

CHAPTER 5

RESEARCH DESIGN AND METHODOLOGY

5.1 INTRODUCTION

This chapter focuses on the research design and methodology procedures used in this study. The chapter begins with a discussion of the quantitative and qualitative research design and methodology; this is followed by a full description of the mixed methodologies (triangulation) approach used in this study. Data analysis by means of univariate, bivariate and multivariate analysis used for the treatment of data in this study is discussed in detail. Included are details of the population selected for the study, a description of respondents, sampling procedures, the variables investigated, quantitative and qualitative instrumentation used, data collection methods and the treatment and analyses of data.

5.2 OVERVIEW OF RESEARCH METHODOLOGIES IN THE SOCIAL SCIENCES

To satisfy the information needs of any study or research project, an appropriate methodology has to be selected and suitable tools for data collection (and analysis) have to be chosen (Mouton, 2001). Primarily there are two distinct approaches that inform the gathering of data in any research project, namely the qualitative approach and the quantitative approach. Qualitative and quantitative methodologies in the social sciences are governed by specific paradigms.

5.2.1 Qualitative approach

The qualitative approach is grounded in the interpretive social sciences paradigm. Qualitative forms of investigation tend to be based on a recognition of the importance of the subjective, experiential 'lifeworld' of human beings. Such reflection is the province of phenomenology reports (Babbie, 1995; Blanche & Durrheim, 1999). Gilbert (1993) notes that qualitative methodologies provide avenues that can lead to the discovery of these deeper levels of meaning. Easterby-Smith *et al.* (1991) describe the task of the qualitative methodologist as to capture what people say and do as a product of how they

interpret the complexity of their world, and to understand events from the viewpoints of the participants. In the domain of tourism specifically, Jennings (2001) notes that the qualitative methodology gathers information as text-based units, which represent the social reality, context and attributes of the phenomenon under study. The methodology is inductive in nature.

Again within the context of tourism research specifically, Finn, Elliot-White and Walton (2000) and Walle (1996) explain that qualitative or inductive research commences in real-world settings, that is, in the empirical social world, where data about the tourism phenomenon are gathered, then analysed, and theoretical constructs are either generated or modified. Research that utilises a qualitative methodology will draw on data collection methods such as participant observation, in-depth interviews and/or focus groups (Jennings, 2001). As a consequence of its underlying, paradigm, qualitative research is subjective, since it relies on the texts and discourses of participants and involves small numbers of participants in the research process as a result of the process of gathering in-depth information (Gilbert, 1993; Walle, 1993; Gunn, 1994). Moreover, qualitative research, because of the small numbers of participants, does not presume to represent the wider population. Qualitative research enables researchers to present detailed snapshots, as it were, of the participants under study (Blanche & Durrheim, 1999; Finn *et al.*, 2000).

Since qualitative reports are not presented as a statistical summation, but rather adopt a more descriptive, narrative style, this type of research is likely to be of particular benefit to the practitioner (Easterby-Smith *et al.*, 1991; Blanche & Durrheim, 1999). However, it is on those grounds that qualitative research has often been described as not being empirical. Nevertheless, this argument does not hold, since the term ‘empirical’ has nothing to do with numbers or the manipulation of variables, but refers instead to whether phenomena are capable of being found in the real world and assessed by means of the senses (Gilbert, 1993; Finn *et al.*, 2000; Jennings, 2001). Perhaps one of the major limitations of qualitative research and evaluation is the time required for data collection, analysis and interpretation. The researcher has to spend a considerable amount of time in

the research setting in order to examine, holistically and aggregately, the interactions, reactions and activities of subjects (Babbie, 1995).

The problem of adequate validity and reliability is a criticism often levelled by quantitative researchers at qualitative methods. Because of the subjective nature of qualitative data and its origin in single contexts, it is difficult to apply conventional standards of reliability and validity (Gilbert, 1993; Creswell, 1994). Contexts, situations, events, conditions and interactions cannot be replicated to any extent, nor can generalisations be made to a wider context than the one studied with any degree of confidence. In short, the richness, individuality and subjective nature of a participant's perspective and understanding are not amenable to the usual scientific criteria. However, Neuman (1994), Walle (1996) and Jennings (2001) argue that this does not make such understanding any less real or valid for that participant. Quantitative researchers expect the qualitative researcher to demonstrate the validity and reliability of claims, to demonstrate the generality of feelings — in short, to meet the same criteria as quantitative research.

5.2.2 Quantitative approach

A quantitative research approach is grounded in the positivist social sciences paradigm, which primarily reflects the scientific method of the natural sciences (Creswell, 1994; Jennings, 2001). This paradigm adopts a deductive approach to the research process. In the tourism context it thus commences with theories, hypotheses or research questions about a particular tourism phenomenon, gathers data from the real-world setting and then analyses the data statistically to support or reject the hypotheses (Veal, 1997; Blanche & Durrheim, 1999; Welman & Kruger, 2001). Researchers who adopt a more deductive approach use theory to guide the design of the study and the interpretation of the results (Neuman, 1994). The overall objective is to test or verify a theory, rather than to develop one. Thus the theory offers a conceptual framework for the entire study, serving also as an organising model for the research questions or hypotheses and for the entire data collection procedure (Veal, 1997; Blanche & Durrheim, 1999; Welman & Kruger, 2001).

A quantitative methodology abstracts data from the participants into statistical representations rather than textual pictures of the phenomenon. The entire research process is objectively constructed and the findings are usually representative of the population being studied. The main strengths of the quantitative approach lie in precision and control. Control is achieved through the sampling and design, and precise and reliable quantitative measurement. A further strength is that experimentation leads to statements about causation, since the systematic manipulation of one variable can be shown to have a direct causal effect on another when other variables have been eliminated or controlled (Babbie, 1995; Blanche & Durrheim, 1999). Furthermore, hypotheses are tested through a deductive approach, and the use of quantitative data permits statistical analysis (Welman & Kruger, 2001). The method thus provides answers which have a much firmer basis than a lay person's common sense, intuition or opinion.

One of the limitations of quantitative research reported by critics is that many researchers are concerned that the scientific quantitative approach denigrates human individuality and the ability to think (Walle, 1996; Massey, 2003). Gilbert (1993) argues that its mechanistic ethos tends to exclude notions of freedom, choice and moral responsibility. Quantification can become an end in itself rather than a human endeavour seeking to explore the human condition. It fails to take account of people's unique ability to interpret their experiences, construct their own meanings and act on these (Gilbert, 1993; Massey, 2003). It is worth noting, however, that a scientific approach cannot in fact be totally objective, since subjectivity is involved in the very choice of a problem as worthy of investigation and in the interpretation of the results.

5.3 TOWARDS TRIANGULATION AS AN APPROACH

In practice, both quantitative and qualitative approaches are frequently appropriate within a single investigation. It is up to the researcher to choose specific methodologies that will permit a clear understanding of the topic to emerge.

Within academia it has been common practice to associate particular epistemologies with distinctive methodologies (Neuman, 1994; Decrop, 1999). For example, in the discipline of geography, positivism has traditionally been linked with quantitative methods, whilst more recent epistemological perspectives such as humanism and postmodernism have been linked with qualitative methods (Creswell, 1994). Faced with the quantitative-qualitative dichotomy, researchers have often been forced to choose one paradigm over the other, and, as a result, there are few accounts of integrating quantitative and qualitative approaches in the social science methods literature.

However, more recently, social science researchers have exhibited a growing recognition of the benefits of a multiple methods approach to research, especially as positivism has been discredited and new approaches such as postmodernism have emerged (Blaikie, 1991; Bowen, 2003; Massey, 2003). Also, while in the past policy makers have tended to display a preference for quantitative research, they have gradually begun to demonstrate a heightened awareness of the role of qualitative research in informing policy formulation (Decrop, 1999).

Triangulation implies that techniques are used in a parallel sense, thus providing overlapping information, making it possible to check results from more than one viewpoint. Easterby-Smith *et al.* (1991) distinguish between four distinct types of triangulation:

- ? Data triangulation: where data is collected at different times, or from different sources, in the study of a phenomenon
- ? Investigator triangulation: where several different researchers collect data relating to the same phenomenon independently and compare findings
- ? Methodological triangulation: where different methods of data collection, commonly both quantitative and qualitative, are combined in the study
- ? Triangulation of theories: where a theory derived from a new discipline is used to explain a phenomenon in another discipline

Bowen (2003) contends that a combination of quantitative and qualitative approaches should be viewed as an acceptable methodological approach for research occupying a variety of epistemological positions and concerning a wide range of substantive research areas in tourism. Both Bowen (2003) and Massey (2003) report that the multiple methods approach represents a poly-vocal approach to research, where employing a range of methodological strategies means that the researcher does not necessarily privilege one particular view of the social world over another. In recognition of these and other such arguments, many social science researchers are increasingly rejecting the automatic association of particular methodologies with particular epistemologies (Bowen, 2003; Massey, 2003). Instead, they are exhibiting flexibility in selecting the method or methods most appropriate to a particular research project.

Despite the fact that they are often presented as a dichotomy, quantitative and qualitative methods are not mutually exclusive and they do indeed share common ground, for example overlapping in the processes of logical enquiry by which they are underpinned (Decrop, 1999). Indeed, some of the key arguments in favour of quantitative methods (for instance, arguments citing the objective nature of quantitative research as opposed to the subjective nature of qualitative research) have been increasingly discredited (Creswell, 1994; Decrop, 1999). Many positivists would in fact agree that no research is entirely objective and error free, as researchers choose their subject and the manner in which they will conduct their investigation. Furthermore, whilst quantitative methods have been regarded as deductive (associated with the formulation and testing of hypotheses), qualitative methods are associated with an inductive approach. However, as Bowen argues, “in all research we move from ideas to data to ideas” — in other words, researchers continually move between research questions and evidence, regardless of the methods adopted to carry out the research (Bowen, 2003).

Blaikie (1991), Easterby-Smith *et al.* (1991), Creswell (1994), Decrop (1999), Bowen (2003), and Massey (2003) have emphasised the following benefits of combining qualitative and quantitative methods:

- ? While the quantitative design strives to control for bias so that facts can be understood in an objective way, the qualitative approach strives to understand the perspective of the programme stakeholders, looking to first-hand experience to provide meaningful data (Easterby-Smith *et al.*, 1991).
- ? The accumulation of facts and causes of behaviour are addressed by the quantitative methodology, whereas the qualitative methodology addresses concerns with the changing and dynamic nature of reality (Bowen, 2003).
- ? Quantitative research designs strive to identify and isolate specific variables within the context of the study (seeking correlation, relationship, causality), while the qualitative design focuses on a holistic view of what is being studied (via documents, case histories, observations and interviews).
- ? Quantitative data is collected under controlled conditions in order to rule out the possibility that variables other than the one under study may account for the relationships identified, while qualitative data is collected within the context of its natural occurrence (Massey, 2003).
- ? Both quantitative and qualitative research seek reliable and valid results. Data that is consistent or stable, as indicated by the researcher's ability to replicate the findings, is of major concern in the quantitative arena, while the validity of qualitative findings is paramount so that data is representative of a true and full picture of the constructs under investigation (Blaikie, 1991; Bowen, 2003).

When methods are combined, the advantages of each methodology complement those of the other, making a stronger research design that will yield more valid and reliable findings (Decrop, 1999). The inadequacies of individual methods are minimised, and more threats to internal validity are recognised and addressed.

In selecting an approach for the present study, the benefits and shortcomings of the various methodologies were considered, and an integrated approach combining elements of both qualitative as well as quantitative data was decided upon, thus making triangulation possible. Both qualitative and quantitative methods would make it possible

to gather the most needed data to address the research problem and to ensure that the objectives of the study were successfully met.

5.4 METHODS OF DATA COLLECTION USED IN THE STUDY

5.4.1 Primary research methods for data collection

- ? A Likert scale questionnaire survey was the main instrument providing quantitative data, and was designed around opinion statements as a means of exploring respondents' perceptions of a wide range of socio-cultural impacts. Questionnaire household surveys using the Likert scale have been used widely by researchers measuring perceptions of the impacts of tourism on residents (Allen *et al.*, 1988; Ap, 1992; Ap & Crompton, 1993; Getz, 1994; Lankford, 1994; McCool & Martin, 1994).
- ? Semi-structured personal interviews were conducted, providing qualitative insights and illuminations.
- ? Participant observation was conducted by the researcher's going on township tours.

5.4.2 Secondary research methods for data collection

Secondary research is research based on secondary resources that already exist (Veal, 1997; Jennings, 2001). Secondary research methods in the current study included Soweto tourism brochures, leaflets, photographs, videos, newspaper and magazine articles, government publications, conference proceedings, reports, academic journals, books, diaries, visitor record books, unpublished manuscripts, statistics and the World-Wide Web (Internet).

5.5 DESCRIPTION OF THE MAIN MEASUREMENT INSTRUMENT USED IN THE STUDY: THE LIKERT METHOD

Resident perceptions of tourism development have been well documented, and in tourism impact studies, the development of a tourism impact assessment scale has received considerable attention (Allen *et al.*, 1988; Ap, 1992; Ap & Crompton, 1993; Getz, 1994; Lankford, 1994; McCool & Martin, 1994). A standard scale could provide researchers and tourism planners with a tool for measuring resident perceptions of tourism in different townships and on different occasions, thus providing a basis for adequate

comparative analysis. The development of a scale of this kind responds to the call for the establishment of standardised instrumentation for use in tourism research. Likert in 1932 proposed a method of attitude measurement (Likert, 1967); the same method remains in use today, and is appropriate to the current context, since Likert scale questionnaire surveys have been widely used for measuring perceptions and attitudes of the host community towards socio-cultural impacts (Ap & Crompton, 1993; Lankford, 1994; McCool & Martin, 1994).

About two decades ago, with the introduction of tourism impact attributes by Pizam *et al.* (1978), researchers began using various resident-attitude-related attributes to quantify perceived tourism impacts. Liu and Var (1986) and Liu, Sheldon and Var (1987) further distilled these attributes into a smaller number of identical impact domains. From their development of a standardised impact scale, Lankford (1994) identified two impact factors from a 27-item tourism impact scale. The variables pertaining to residents' perceptions of the socio-cultural impacts of township tourism in Soweto used in this study mirror those used by Lankford (1994). Although a relatively large number of perceived impact studies have been reported, there is a need to develop better measures of perceived impacts of tourism. To date, the only reasonably generic, reliable and valid perceived impact scale to have appeared in the literature is that developed by Lankford (1994), which provides a conceptual framework for perceived tourist impact.

A Likert scale instrument was therefore developed for the purposes of this study to assess residents' perceptions of the socio-cultural impacts of township tourism. The research variables were measured on a 5-point Likert-type scale, with a score of 1 representing 'strongly disagree' and a score of 5 representing 'strongly agree'. The scale was designed to elicit respondents' opinions on a range of issues relating to the socio-cultural impacts of township tourism. In such scales no judges are used to rank the scale statements: it is assumed that all subjects will perceive 'strongly agree' as expressing greater favour towards the attitude statements than 'moderately agree' and 'strongly disagree' (Likert, 1967; Lankford 1994). Some of the item statements should be expressed positively and some negatively to encourage respondents not to respond

automatically, but to think about every item. Ideally there should be roughly equal numbers of positively and negatively worded items (Lankford, 1994). Individual items can be, and normally are, analysed by counting how many respondents gave a particular response to the item. A subject's score is tabulated by assigning a numerical value to each of the answers, ranging from 1 for the alternative at one end of the scale to 5 for the alternative at the other, and then calculating the sum of the numerical values of the answers to all questions (Jennings, 2001). However, the principal objective, which is not uncontroversial, is to arrive at an overall score for all the items combined together.

5.5.1 Advantages of the Likert method

Likert (1967), Lankford (1994) and Veal (1997) list the advantages of the Likert method as including:

- ? the fact that the method is based entirely on empirical data regarding subjects' responses rather than the subjective opinions of judges;
- ? the fact that this method produces more homogeneous scales and increases the probability of a unitary attitude being measured; as a result, validity (construct and concurrent) and reliability are reasonably high; and
- ? greater ease of preparation.

5.5.2 Construction of the Likert scale questionnaire used in the study

The basic procedures employed in developing the scale for measuring resident perceptions of tourism impacts for the purposes of this study followed the procedures recommended by Likert (1967), Churchill (1979) and Lankford (1994). Four main steps were taken in developing the instrument.

- ? Identification and generation of socio-cultural tourism impact variables, i.e. long-term impact variables derived from three source of information: (a) a review of a pool of impact items used in research on resident perceptions of the socio-cultural impacts of tourism, as mentioned earlier in chapter 2, (b) a review of questionnaires utilised for attitude and perception studies relating to socio-cultural tourism impacts and (c) insights gained from an examination of information relating to township tourism from

secondary sources such as newspapers, magazines, tour operator web sites and leaflets.

- ? These variables were then used to formulate as 57 attitudinal statements. The survey questionnaire (see Appendix A) consists of three parts. The first section (Part A, consisting of 6 questions) was structured in such a way as to elicit demographic information regarding the respondents' gender, age, educational level, income, and years of residence. Part B consisted of 57 impact variable items incorporating a 5-point Likert scale to measure respondents' level of agreement or disagreement with a statement about the impact of township tourism on a given economic, social or cultural aspect. Respondents were asked to rate items on an ordinal scale of 1–5. In Part C, respondents were asked to provide any additional positive or negative comments they wished to make regarding township tourism development and planning as a way of identifying other impacts and problems not included in the questionnaire.
- ? The pre-test was conducted using a convenience sample of 25 Unisa staff and students resident in Soweto, primarily to ensure the clarity of the questions and to measure whether the questionnaire could be completed within a reasonable period of time (about 20 minutes), and secondly, to elicit some comments about the content validity, as respondents were asked to describe any difficulties they had in completing the questionnaire accurately.
- ? The survey instrument was modified on the basis of comments and suggestions made by the pre-test subjects. Re-wording of the questionnaires was necessary in order to remove any jargon, inconsistencies or leading questions. The final questionnaire appears as Appendix A.

5.5.3 Pool of items (impact variables) derived from the literature and fieldwork

The following scales were developed using impact variables derived from the literature; these were then used to formulate statements specific to township tourism in Soweto.

The statements were subsequently divided into four scale categories, and these were then contextualised within the Likert scale questionnaire.

CATEGORY 1		SOCIAL IMPACTS
NO	ITEMS	
1	Township tourism will encourage an increase in street children along the tourist route	
2	The current level of township tourism has significantly improved the standard of living of Soweto residents	
3	Family life of local residents has been disrupted by the presence of tourists	
4	Community life has become disrupted as a result of the development of tourism in Soweto	
5	Local residents view foreign tourists as intruding into their community	
6	Residents feel that their safety is affected as more tourists are encouraged to visit Soweto	
7	Local people are being exploited because of the growth of township tourism	
8	Further growth in Soweto tourism will result in overcrowding of local amenities by tourists	
9	An increase in tourists into Soweto will lead to resentment between residents and tourists	
10	The Soweto community should take steps to restrict tourism development	
11	Local resentment is generated because of the inflated prices for the tourist market	
12	Contact with tourists may introduce health risks to the host community	
13	Township tourism has resulted in a greater demand for female labour	
14	The number of tourists on township tours should increase significantly	
15	Township Tourism will gradually result in an increase in municipal rates and taxes	
16	Soweto residents have been consulted and made aware of the tourism development plan for township tourism	
17	The current level of township tourism has significantly improved the local community's hospitality toward strangers	
18	Local residents oppose the presence of township tourists in the Soweto region	
19	The benefits of township tourism outweigh the negatives	
20	Township tourism increases the rate of organised crime in the Soweto community	
21	Government should restrict further development of township tourism in Soweto	
22	Township tourism has increased traffic problems in Soweto	
23	The noise levels caused by township tourism are not appropriate for a residential community	
24	Tourists taking photographs of local people can cause great offence to locals	
25	Locals are barred from using tourist facilities in Soweto	
26	Tourists who are seen to be wealthier than the majority of the residential population are more likely to generate resentment.	
27	Tourism development increases the development of recreational facilities and amenities for residents	
28	Local residents are the ones who really suffer from living in an area popular with tourists	
29	Local people are treated equally, rather than as inferiors by tourists	
CATEGORY 2		PHYSICAL/ENVIROMENTAL IMPACTS
NO	ITEMS	
30	Township tourism has made residents more conscious of the need to maintain and improve the appearance of the area	
31	There are better roads (infrastructure) due to township tourism development	
32	Residents are satisfied with the manner in which township tourism development and planning is currently taking place	
33	The development of township tourism has generally improved the appearance of Soweto	
34	Township tourism in Soweto has led to more litter in the streets	

Table 5.1 FOUR CATEGORIES WITH IMPACT ITEMS GENERATED FROM LIKERT SCALE

CATEGORY 3		CULTURAL IMPACTS
NO	ITEMS	
35	Tourist interest in culture has resulted in a strengthening of traditional activities and cultural pride	
36	Local people alter their behaviour in an attempt to copy the style of tourists	
37	Township tourism has stimulated the locals' interest in participating in traditional art forms	
38	Local culture is being renewed as a result of township tourism	
39	Interacting with tourists leads to a deterioration of local languages	
40	Township tourists show respect for the cultural lifestyle of the local people	
41	Traditional African culture in Soweto is being commercialised (sold) for the sake of tourists	
42	Tourism encourages a variety of cultural activities by the local population	
43	Township tourism helps to conserve the cultural identity and heritage of the host population	
44	Locals often respond to tourist needs by adapting traditional practices to enhance their commercial value	
45	Township tourism causes changes in the traditional culture of local residents	
46	Meeting tourists promotes cross-cultural exchange (greater mutual understanding and respect for one another's culture)	

CATEGORY 4		SOCIO-ECONOMIC IMPACTS
NO	ITEMS	
47	Only a small minority of Soweto residents benefit economically from tourism	
48	Income-generating opportunities created by township tourism development are evenly distributed across the community	
49	By creating jobs and generating income, township tourism promotes an increase in the social wellbeing of residents	
50	Township tourism has led to more people leaving their former jobs for new opportunities in tourism	
51	In addition to payment to tour operators for tour costs, tourists should be advised to make donations for the benefit of the local community	
52	Township tourism provides many worthwhile employment opportunities for Soweto residents	
53	Township tourism holds great promise for Soweto's economic future	
54	Tourism has already improved the economy of Soweto	
55	The development of township tourism in Soweto benefits the visitors more than the locals	
56	Soweto residents have been adequately consulted in participating in entrepreneurial initiatives in township tourism	
57	Township tourism in Soweto is in the hands of a few operators only	

Table 5.1 FOUR CATEGORIES WITH IMPACT ITEMS GENERATED FROM LIKERT SCALE

5.6 PROFILE OF RESPONDENTS

According to Krippendorff (2001) opinions about and expectations of tourism can be very different, depending on which population or occupational groups are considered. This needs to be taken into account when the sample is chosen.

For both quantitative and qualitative data collection methodologies, the sample was selected from the following categories:

Type 1: Residents who are in constant and direct contact with township tourists; because they depend on township tourism and would perhaps be unemployed without it, they welcome visitors.

Type 2: Township residents who have no contact with tourists or see them only in passing and whose household income is not derived from township tourism.

The rationale behind selecting different categories of respondents was to allow key comparisons to be made.

5.7 SAMPLING DESIGN AND SAMPLING METHODS USED IN THE STUDY

The main purpose of sampling is to achieve representativeness; the sample should be assembled in such a way as to be representative of the population from which it is taken (Gilbert, 1993; Jennings, 2001). To achieve this, the sampling units are randomly selected. This is the commonest approach to sampling, but it is by no means the only one, nor is representativeness — in a numerical sense — the only aim of sampling procedures.

5.7.1 Population and sampling frame

Jennings (2001:136) defines population as “all the study subjects (tourists, visitors, hosts, family, friends, employees, managers) or study units (attractions, transport providers, accommodation facilities) that are the focus of the research project”. In this study the target population consists of Soweto residents living around the 14 main tourism hubs or visiting points in Soweto. The 14 tourism hubs are Dobsonville, Protea South, the Oppenheimer Tower, Meadowlands, Dhlamini, Freedom Square, Kliptown, Phefeni, the Hector Peterson Memorial, Orlando West, Power Park Informal Settlement, Informal Settlement, Taxi Rank and Diepkloof Hostel. These are the destinations to which tour operators take tours, visiting political landmarks, museums, homes, informal settlements, traders, restaurants, shebeens, museums and so on. Residents living in and around these hubs are divided into those who earn an income from tourism, and those who are not directly involved in tourism.

5.7.2 Sample size

A sample of 350 households living around the 14 main tourism hubs of Soweto was selected for the quantitative part of the study. A combination of systematic and stratified random sampling approaches was used for sample selection.

5.7.3 Sampling technique followed for quantitative research design

According to Central Statistics South Africa (2003), the 2001 Census revealed that the population of the areas in which the 14 tourism hubs are located numbers 352 054. However, the researcher observes that the actual population size could well be far higher, since there is considerable difficulty attached to conducting population counts in informal settlements. The population of Soweto is estimated at between 1,3 million and 4 million people. There is no official figure available for the number of households and persons living in the two informal settlements in and around the main tourism hubs. With the assistance of a statistician, it was decided to approximate a 0,1% sampling minimum for the study area as a whole.

A study sample of 350 households was therefore drawn from the resident household population of the 14 main tourism hubs listed above. The first step involved stratified random sampling; accordingly the population of Soweto was first subdivided into sub-groups (tourism regions). Systematic sampling was then applied through the selection of every third household at each of the tourism hubs. A decision on the number to sample at each tourism hub or visiting point was based on the following formula:

$$(\text{Visiting points}) = 14 = k$$

$$n = 350 \text{ households}$$

$$n/k = 350/14 = 25$$

$$25 \times 14 = 350$$

Therefore it was necessary to interview at least 25 households at each tourism hub or visiting point, of which 12 households derived an income from tourism and 13 households were not dependent on tourism. This was essential in order to obtain true representativeness and to allow for comparisons to be made. Every third household was targeted at each hub to arrive at the 25 households.



Figure 5.1 SOWETO STUDY AREA INDICATING TOURISM HUBS

5.8 DATA COLLECTION PROCEDURE (FIELDWORK) IN THE STUDY

In order to gather information on residents' perceptions of the socio-cultural impacts of township tourism, 350 household surveys (constituting the basis of the quantitative component of the study) together with unstructured personal interviews and participant observation (constituting the qualitative component of the study) were carried out in the Soweto region (14 tourism hubs) over an eight-week period. Two fieldworkers, who knew the study area very well, were employed on a part-time basis and trained by the researcher in both approaching the respondents to elicit their participation and monitoring the completion of questionnaires. The purpose of the training was to make sure the fieldworkers understood the importance of their role in the research project and what they could expect in the field. They were taught the skills required for approaching the respondents and were familiarised with the various sections of the questionnaire. This training was both relevant and necessary, as it was the fieldworkers' first experience of such a situation.

The researcher assumed a supervisory role in monitoring the fieldworkers daily. This helped to ensure that only respondents who were targeted participated. The survey team, together with the researcher, visited the 14 tourism hubs in Soweto and identified those respondents to be interviewed. The survey team then distributed one questionnaire to each of the selected households. The unstructured personal interviews were conducted by the researcher.

5.9 DATA ANALYSIS OF LIKERT SCALE QUESTIONNAIRE

The quantitative data collected was initially coded into numerical representations, so that a series of statistical analyses could be performed using the software package Statistical Analysis System (SAS), version 8. According to Jennings (2001), the software package enables researchers to:

- ? enter and store data
- ? utilise retrieval strategies
- ? engage in statistical analyses
- ? generate graphs and reports

- ? manage research projects
- ? write reports

For analysis purposes, the respondents were asked to rank their responses to the questions according to the Likert scale format. Coding in quantitative analysis differs from qualitative coding in that the raw data is turned into numerical representations to allow statistical analyses to be conducted on the aggregated data (Veal, 1997; Jennings, 2001). A code was developed for each of the response sets in the questionnaire, and numerical codes were assigned for each response (Veal, 1997). These responses were then turned into a series of numbers for capture using SAS for further statistical analysis. The researcher checked and cleaned the data by examining the coded data for any incorrectly assigned codes and correcting these errors by reviewing the original data (Veal, 1997; Jennings, 2001). Finally univariate, bivariate and multivariate data analysis was performed.

5.9.1 The analysis of univariate data used in the study

Univariate data analysis is the analysis of single variables. According to Mouton (2001), descriptive statistics organise and summarise the data to render it more comprehensible. Descriptive statistics enable the researcher to describe trends in the data and also to determine whether relationships exist between variables.

For this study, the researcher made use of the following descriptive statistics, with a range of aims:

- ? The calculation of frequencies and percentages expressed as tables, charts and graphs (see Appendix C)
- ? Measurements of central tendency, namely the mean. The mean is the sum of the individual scores in a distribution, divided by the number of scores (Vogt, 1993). The mean can be found for ordinal and interval variables.
- ? Measurements of dispersion of the data using the standard deviation (SD) of the measurements. The SD is a measure of how much, on average, the scores in a distribution deviate from the mean score (Vogt, 1993). This provides a reflection of

how homogeneous or heterogeneous a population is. The SD further provides an indication of the average distance from the mean (Vogt, 1993). A low SD would mean that most observations cluster around the mean. A high SD would indicate considerable variation in the responses (see Appendix D).

5.9.2 The analysis of bivariate data used in the study

The analysis of two variables, or bivariate analysis, was required for this study. Correlations as instruments of bivariate analysis and testing the significance of a difference between means were used.

Chi-square a test for assessing the statistical significance of cross-tabulated variables was used in the study (Welman & Kruger, 2001). Chi-square indicates whether or not a relation exists between or among variables. Most of the statistical tests report a significance level (or one can be obtained). Difficulty is sometimes experienced in interpreting these values. Generally, the smaller the significance level reported, the more conclusive the results. Social scientists usually establish a cut-off point at $p = 0,05$, i.e. the 5% level. This implies that there is a 5% chance that the results obtained were a result of chance (or sampling error).

Analysis of variance (ANOVA), or a t-test, was employed to test the theoretical framework on which the study was based so as to determine whether there were significant relationships or differences among group mean totals, item mean scores, and independent variables (Jennings, 2001). Independent variables considered in the analysis were residents' income from tourism, gender, and period of residence (see Appendix G).

5.9.3 The analysis of multivariate data used in the study

For the analysis of multivariate data, a range of more complex parametric tests was conducted on the data. Factor analysis and item mean analysis as a multivariate grouping procedure were applied to data (Reese & Lochmuller, 2003).

Factor analysis is a statistical approach that can be used to analyse interrelationships among large number of variables and to explain these variables in terms of their common underlying dimensions (factors) (Wikipedia, 2003). The statistical approach involves finding a way of condensing the information contained in a number of original variables into a smaller set of dimensions (factors) with a minimum load of information (Rummel, 2003; Wikipedia, 2003). Factor analysis has been widely used, especially in the behavioural sciences, to assess the construct validity of a test or a scale.

A review of factor analysis reveals that it entails four basic steps (University of Texas at Austin, 1997; Reese & Lochmuller, 2003; Rummel, 2003; Wikipedia, 2003):

- ? Data collection and generation of the correlation matrix
- ? Extraction of initial factor solution
- ? Rotation and interpretation
- ? Construction of scales of factor scores to use in further analyses

Reese and Lochmuller (2003) conclude that the main applications of factor analytic techniques are: (1) to reduce the number of variables and (2) to detect structure in the relationships between variables, that is, to classify variables. Therefore, factor analysis is applied as a data reduction or structure detection method. Factor analysis with varimax rotation and item analysis was used to determine the underlying perception patterns or dimensions associated with township tourism in Soweto.

Scale purification through the use of exploratory factor analysis and item analysis was used to assess dimensionality of scale and to delete items with either low or multiple loading factor coefficients (Rummel, 2003; Wikipedia, 2003). Item analysis is a detailed method for estimating the internal consistency of the instrument (Rummel, 2003). Here the researcher is interested in finding out how well the responses to each item correspond to the responses to the other items and to the test as whole. This helps the researcher to identify those items within an instrument that do not provide useful information about the subjects or that are actually confusing the data (Reese & Lochmuller, 2003). The researcher can then remove these troublesome items from the instrument (replacing them

with better items if necessary) so as to increase the overall reliability of the instrument. In the case of this study, this allowed for the possible reduction of the 57 impact variables into categories of factors.

Factor loadings, eigenvalues and percentages of variance using principal components extraction with varimax rotation for total composite scores are usually reported (Wikipedia, 2003). The factor loadings, also called component loadings, are the correlation coefficients between the variables (rows) and factors (columns) (Rummel, 2003; Wikipedia, 2003). The eigenvalue for a given factor measures the variance in all variables that is accounted for by that factor (see Appendix E).

Part C of the questionnaire survey provided qualitative data and therefore content analysis as a way of systematically analysing the content of text and converting it to numerical variables to permit quantitative analysis (see Appendix C).

The findings were displayed in the form of tables and graphs, and documented in accordance with the stated aims of the study. Conclusions were drawn based on the findings, followed by recommendations for further research.

5.10 DESCRIPTION OF QUALITATIVE METHODOLOGY FOLLOWED

5.10.1 Semi-structured personal interviews

The personal interviews were semi-structured in nature and were conducted on an individual basis. The interviews made it possible to explore other themes and nuances, further enriching the data. In cases where respondents consented, tape recordings were made, which enabled the interviewer to pay close attention to discussions; transcriptions of recordings were made later. In many instances, however, the researcher found the respondents reluctant to allow the use of a tape-recorder, and in those instances the researcher resorted to note-taking. Each unstructured interview lasted approximately 30 minutes. The researcher explained the purpose of recording the discussion to the respondents in detail. The respondents were assured that the information recorded would be used only for the purpose of the study.

Semi-structured interviewing is based on the use of an interview guide (Veal, 1997; Jennings, 2001), which is a written list of questions and topics that need to be covered in a particular order. The interviews were broadly guided by the following four questions:

- ? How does tourism in Soweto personally affect your way of life?
- ? What do you like best about tourism in Soweto?
- ? What do you like least about tourism in Soweto?
- ? What are residents' concerns about tourism in Soweto?

The respondents were free to expand on the topic as they saw fit, and to relate their own experiences. The interviewer intervened only for clarification or further explanation (Veal, 1997). The interviewer used probing questions for clarification of concepts and ideas. Blanche and Durrheim (1997) and Veal (1997) conclude that the benefits of an unstructured interview include the opportunity it affords the interviewer to interact with respondents in a conversational setting so as to reach the heart of the subject under investigation. Semi-structured interviews are generally the most useful, in that they allow full exploration of the topic and yet retain a degree of structure, which ensures that most of the information obtained is relevant and manageable (Veal, 1997).

5.10.2 Participant observation

In participant observation the researcher becomes a participant in the social process being studied. For instance, the researcher who is studying the use of a park or resort may spend periods there as a user of the facility (Veal, 1997). In this study, the researcher assumed the role of tourist by taking six township tours with registered tour operators guiding tours in Soweto. Since the researcher was the data gatherer and recorder, the process of analysis was ongoing. In agreement with Veal (1997), the researcher constantly related his observations to the objectives of the study, drawing interim conclusions. The very act of deciding what to view, what to say and what to record during these township tours involved choices that were influenced by the researcher's evolving understanding of the phenomenon being studied.

5.10.3 Qualitative sampling

Easterby-Smith *et al.* (1991) and Neuman (1994) describe four commonly used qualitative research data collection methods, namely the case study, and ethnographic, phenomenological and grounded theory methods. Since the researcher was interested in understanding individual respondent perceptions of the socio-cultural impacts of township tourism, the phenomenological method was deemed the most suitable for qualitative data collection. The phenomenological method suggests that respondents are chosen specifically because of their knowledge of the topic under investigation (Easterby-Smith *et al.*, 1991; Neuman, 1994; Veal, 1997). In keeping with this method, respondents in this study were selected specifically because of their township tourism experience as a result of living near or working within a major tourism hub in Soweto.

Purposive sampling, which is a type of non-probability sampling, was applied when conducting in-depth personal interviews (Veal, 1997; Jennings, 2001). Purposive sampling is also referred to as judgmental sampling, since it involves the researcher making a decision about who or what study units will be involved in the study. Welman and Kruger (2001) describe non-probability sampling as a method in which the researcher has no way of forecasting, estimating, or guaranteeing each element in the population to be represented in a sample.

5.10.4 Sample size — qualitative study

In using purposive sampling, the researcher decides when enough participants or units have been sampled. This occurs when there is redundancy with regard to data. In this study this was achieved after conducting 38 personal interviews. The cut-off is not predetermined, but emerges from the research process and concurrent data analysis (Babbie, 1995).

5.10.5 Data analysis — qualitative study

According to Blanche and Durrheim (1999), qualitative data analysis tends to be primarily an inductive process of organising data into categories and identifying patterns. For this study a content analysis method was followed to understand the procedures and

importance of qualitative data analysis (Gunn, 1994). This entailed a literature review to learn about the various techniques of analysing qualitative data. Babbie (1995) states that the most general guide to analysing qualitative data involves looking for similarities and dissimilarities. The focus must be on those patterns of interactions and events that are generally common to what the researcher is studying (Babbie, 1995). This type of analysis formed the core of analysing the qualitative data collected during this study. Themes were identified, and the data was then classified into categories and themes.

Qualitative data collected through interview schedules is coded, and repeated themes (responses) or concepts recorded until saturation is achieved (Jennings, 2001; Veal, 1997). Recorded interviews are transcribed and coded into themes already established in the interview schedule. The role of theory-building remains the responsibility of the researcher. The essence of the analysis procedure will be to return to the terms of reference, the research problem and questions of the research, and begin to sort and evaluate the information gathered in relation to the questions posed (Finn *et al.*, 2000). In this way an explanation of the actual meaning of the data and logical reasoning can be achieved (Babbie, 1995). The goal is to integrate the themes and concepts into a theory that offers an accurate, detailed, yet subtle interpretation of the research arena (Veal, 1997; Jennings, 2001). The analysis will be complete when the researcher feels that his or her interpretation can be shared with tourism policy-makers and tourism development planners in both the public and private sectors and will make a meaningful contribution to theory. The above analysis was favoured for its potential to assist the researcher to describe trends in the data and also determine whether there were relationships between variables.

As previously noted, where consent was given, a tape recorder was used during the interviews. The researcher clearly explained the purpose of recording the discussion to the respondents, and respondents were assured that the information recorded would only be used for the purpose of the study.

5.11 VALIDITY OF DATA

Welman and Kruger (2001) describe validity as a mechanism that ensures that the process implemented to collect data has collected the intended data successfully. Validity refers to extent to which an empirical measure adequately reflects the real meaning of the subject under investigation (Babbie, 1995). To ensure that the data acquired was valid in this study, the following steps were taken:

- ? An extensive literature review was undertaken to understand how personal in-depth interviews and household surveys should be conducted (Blanche & Durrheim, 1999; Gunn, 1994; Finn *et al.*, 2000; Jennings, 2001). Interview guidelines were generated in conjunction with the fieldworkers. This ensured that the interviews focused on the topic under investigation.
- ? The purpose of the study was clearly explained to the respondents and issues of concern were resolved satisfactorily. The procedure of the interview was explained to the respondents. Lastly, respondents were assured of anonymity and confidentiality. This encouraged frankness during the interviews.

The above steps ensured that the interviews were conducted under conditions and in an environment acceptable to the respondents, and therefore ensured that the process was trustworthy.

5.12 RELIABILITY OF THE DATA

Babbie (1995) describes reliability as a condition in which the same results will be achieved whenever the same technique is repeated to do the same study. This was achieved by the following means:

- ? The anonymity and confidentiality of the respondents was ensured so that they were able to provide information for use strictly for the purpose of the study. A rapport with the respondents was successfully established during the preliminary fieldwork

stage. Here the researcher began to build a relationship of trust with the respondents, and the credibility of the study was reinforced.

- ? The utilisation of trained fieldworkers ensured that the discussion level was high where necessary, and relevant to the study.

5.13 SUMMARY

Chapter 5 furnished a discussion of the choice of methodology used to conduct the present research. The range of methods and approaches that were applied fall within the paradigms of both quantitative and qualitative research. The researcher supported the choice of approach with a detailed description of the use of triangulation and its benefits, and the way in which this approach was customised to suit the requirements of the study as a whole. The method of sampling, data analysis (univariate, bivariate, multivariate) and the choice of statistics and data analysis used were described in detail.