CHAPTER 6

RESEARCH DESIGN AND METHODOLOGY

6.1 INTRODUCTION

A research design is a master plan specifying the methods and procedures for collecting and analysing data to get the needed information. It is a framework of the research plan of action (Zikmund, 2000:59). Research methods refer to the various means by which data can be collected and/or analysed. Hussey and Hussey (1997:54) argue that research methodology refers to the overall approach to the research process, from the theoretical underpinning to the collection and analysis of the data.

Dutka (1995:25) argues that qualitative research involves free-format responses in which words and observations are used. It provides in-depth information obtained from a few cases. The qualitative paradigm is identified by a tendency to be process orientated, subjective, based on natural uncoordinated observations and descriptive. In general it is based on "feelings" and not tangibles.

Quantitative research concentrates on numbers to represent viewpoints and opinions. The numbers that are generated from a larger number of cases are manipulated by using arithmetic and statistics. It contains a factual basis, measurable data, is analysis-based, and depends on logic with analytical, mathematical deductions (Dutka, 1995:26).

Most researchers see qualitative and quantitative methodologies as complementary and argue that both should be combined to maximise their individual strengths (Dutka, 1995:26). For example, initial qualitative research can be used to help develop an effective quantitative telephone or mail questionnaire.

In an attempt to be objective in understanding barriers and constraints to tourism entrepreneurs, it was decided to use both qualitative and quantitative methods. Initial qualitative research was used to help determining the critical factors as well as the questions of the questionnaire.

This chapter aims to provide an insight into the practical ways and methods that were employed in gathering the information for the empirical part of this study. The universe and sample frame will be discussed as well as the sample method and size. The method of data collection and questionnaire design is described. The last part of the chapter concerns the data processing, analysis and evaluation of results.

6.2 DEFINITION OF THE RESEARCH PROBLEM AND OBJECTIVES

6.2.1 Problem definition

According to Welman and Kruger (1999:12), a research problem refers to some difficulty which the researcher experiences in the context of either a theoretical or a practical situation and to which he/she wants to obtain a solution.

The focus of the research is to identify the barriers and constraints encountered by tourism entrepreneurs in Gauteng and Mpumalanga. Identifying the barriers and constraints will allow us to take steps to develop and incorporate new strategies to break down these barriers and therefore enhance the role that tourism entrepreneurs play in South Africa.

6.2.2 Objectives of the research study

Primary objective:

The primary objective of the research is to determine the barriers and constraints that entrepreneurs are facing in the tourism industry in South Africa.

Secondary objectives:

The secondary objectives of the study are:

- To analyse the tourism industry in providing a better understanding of how it is functioning in the South African context.
- To analyse the role and structure of the Government in tourism.
- To analyse the role of other stakeholders in the tourism industry.
- To investigate the size of the tourism industry in South Africa.
- To investigate the employment opportunities that exist in the tourism industry.
- To analyse how the tourism sector links and is conducive to SMME development.

6.3 THE POPULATION AND SAMPLING FRAME

6.3.1 The population (universe)

A population (also called the universe) is that set of people, products, firms and markets that is of interest to the researcher. It is the responsibility of the researcher to provide a precise definition of the population of interest (Dillon, Madden and Firtle, 1993:215).

In the case of this particular study, the population refers to the owners of tourism enterprises located in Gauteng and Mpumalanga.

6.3.2 The sample frame

The sample frame is closely related to the population. It is the list of elements

from which the sample is actually drawn. Ideally, it is a complete and correct list of population members only (Cooper and Schindler, 2001:170). Dillon, *et al* (1993:218) say that the sampling frame is the vehicle the researcher uses to assemble eligible sampling units. A sampling frame is therefore a list of the sampling units.

Sudman and Blair (1998:338) identify three ways in which the sample frame may differ from that of the universe:

- The frame may contain ineligibles or elements that are not part of the population;
- The frame may contain duplicate listings; and
- The frame may omit units of the population, which is by far the most serious problem.

It can be assumed that convenience sampling was used in this particular study. Convenience samples involve selecting sampling units on the basis of where and when the study is being conducted. For this reason convenience samples are also referred to as "accidental samples", since sampling units are selected by "accident". Convenience samples provide very little or no control over who is included in the sample. If respondent participation is voluntary or if the interviewer as opposed to the researcher selects sampling units, then convenience samples are produced (Dillon, *et al* 1993:229).

The sample frame that was used in this study were tourism entrepreneurs located in Gauteng and Mpumalanga. Biases could exist between the opinions of members of the sample frame and the population. However, it is assumed that the opinions of the sample frame used represents the opinions of informed tourism entrepreneurs (owners) in South Africa.

6.3.3 Sampling size, method and response rate

Sample size:

According to Lockhart and Russo (1994:144), refer a *sample size* to the number (*n*) of items to be selected from the universe of the population to make up a specific sample. According to Cooper, *et al* (2001:172) some principles that influence sample size include:

- The greater the dispersion or variance within the population, the larger the sample must be to provide estimation precision;
- The greater the desired precision of the estimate, the larger the sample must be;
- The narrower the interval range, the larger the sample must be;
- The higher the confidence level in the estimate, the larger the sample must be;
- The greater the number of subgroups of interest within a sample, the greater the sample size must be, as each subgroup must meet minimum sample size requirements; and
- If the calculated sample size exceeds five percent of the population, sample size may be reduced without sacrificing precision.

Sampling method:

Questionnaires were the main sampling method. Questionnaires were e-mailed to 1 134 entrepreneurs in Gauteng and Mpumalanga. Questionnaires were also faxed to 216 entrepreneurs in Gauteng and Mpumalanga who don't have e-mail facilities.

Interviews were conducted with 3 respondents during the pre-testing phase while an additional 12 interviews were conducted during the final stage of data collection.

Mail surveys were not included in the study because of the typical low response rates that are synonymous with the application of this type of survey method. Another factor is that non-respondents to mail surveys could feel that they do not know enough about the topic of survey being conducted (Leedy, 1997:32).

Response rate:

Response rate refers to the percentage of the sample that co-operates and completes the questionnaire (Dillon, *et al* 1993:165).

Therefore, the response rate for this study was 13,63 percent (184 responses from a total of 1 350 questionnaires were received).

6.4 DATA COLLECTION METHODS

Once the research design (including the sampling plan) has been formalised, the process of gathering information from respondents may begin (Zikmund, 2000:65). Blankenship and Breen (1993:122) argue that there is no simple answer to which available method of data collection the researcher should use when collecting primary data. Each method has its advantages and limitations. Therefore, the user has to choose the most appropriate method.

The researcher decided on a questionnaire that comprised of closed-ended questions and 1 open-ended question as the method of data collection.

6.4.1 Questionnaire design

The first step in the questionnaire design was to develop a preliminary questionnaire with questions developed from the initial research problem.

Responses to the questionnaire were anonymous though some researchers have found that this has no significance on the response rate (McDaniël and Rao, 1981:157). With the purpose of obtaining different views and to avoid repetition and imitation, the questionnaires were sent to only 1 person (owner) in a family, group or firm. All questionnaires were attended to, whether completed or not.

Questions of the same kind or that seek the same sort of information were grouped together (Section A of the questionnaire). They fell into a range of categories where some consisted of a small number of related questions. This intended to elicit information of a particular type. The various sections of the questionnaire were as follows:

Section A:

The majority of the questions in Section A were closed-ended questions. It required respondents to choose from a list of options. A number of combination questions were also created, i.e. "other".

The first part of the questionnaire covered the demographic information of respondents. Respondents were required to mark a cross (tick) next to the appropriate box (item) provided. The demographic information covered 3 areas, namely gender, ethic group and age.

The second part of Section A covered the business information in the following areas: the year (and month) when the business was started, the form of business, where the business operates, the location of the business, products/services that are offered by the business, the main primary product/ service that are offered by the business, the annual turnover (sales) of the business and the number of full-time and part-time employees in the firm. Business owners also had to self-evaluate 8 factors on business success, followed by an evaluation of 8 business skills.

Section B:

Section B of the questionnaire dealt with statements on problems, barriers and constraints that tourism entrepreneurs could face. Respondents were asked to scale the extent of each problem, barrier or constraint. Thirty-six (36) statements were developed and they were a mix of positive and negative statements. In order to determine whether they served as problems, barriers, and constraints, each statement had to be classified in a category from which the statement was derived. The classification of statements is presented in Table 6.1.

Table 6.1: Classification of statements

NO.	STATEMENT	VARIABLE	BARRIER/CONSTRAINT CLASSIFICATION
1	The South African economy is favourable for tourism	42	Economy
24	The South African tourism industry creates jobs	65	Economy
2	The Government supports tourism enterprises	43	Government
4	The Government regulates tourism in South Africa	45	Government
15	The Government develops infrastructure for tourism in South Africa	56	Government
25	Government structures exist for tourism in South Africa	66	Government
27	Government policies for tourism in South Africa are coherent	68	Government
30	Government tourism structures are effective in South Africa	71	Government
32	Tourism is monitored by the South African Government	73	Government
35	Government policies exist for tourism in South Africa	76	Government
14	Quality standards exist for tourism products and services	55	Legislation, regulation and standards
23	South African tourism products/services are of a high standard	64	Legislation, regulation and standards
28	Too high prices are charged for tourism products/services in South Africa	69	Legislation, regulation and standards
3	Tourism development is planned in South Africa	44	Tourism development
5	Natural resources are protected in tourism development	46	Tourism development
9	Local communities benefit from tourism in South Africa	50	Tourism development

Table 6.1 continues on the next page

Table 6.1 continued

NO.	STATEMENT	VARIABLE	BARRIER/CONSTRAINT CLASSIFICATION
10	Tourism is negatively affected by urbanisation	51	Tourism development
16	Local communities support new tourism development programmes	57	Tourism development
18	The South African public lacks environmental awareness	59	Tourism development
26	Opportunities exist in South Africa to establish partnerships in tourism	67	Tourism development
29	Large tourism enterprises dominate tourism in South Africa	70	Tourism development
34	South Africans lack the expertise to develop tourism in South Africa	75	Tourism development
7	Developed countries generate tourism demand to South Africa	48	Tourism demand
21	There are sufficient demand/supply for tourism products/services in South Africa	62	Tourism demand
31	South African tourism products/services cater for a wide variety	72	Tourism demand
33	The tourism industry in South Africa keeps abreast of tourism trends/growth	74	Tourism demand
36	Skilled positions in the South African tourism industry are occupied by expatriate (foreign) labour	77	Tourism demand
11	Safety and security hamper tourism in South Africa	52	Social
19	HIV/Aids affect the tourism industry's labour force negatively	60	Social
6	The picture of tourism as created by the media in South Africa, is a true reflection of it	47	Marketing
12	Perceptions about Southern Africa affect tourism in South Africa negatively	53	Marketing
22	Seasonality affects tourism in South Africa negatively	63	Marketing
20	Tourism entrepreneurs lack marketing funds	61	Finance
8	Customer service in the South African tourism industry is of a high standard	49	Education/Training
13	The South African tourism industry provides jobs mainly for the less-skilled people	54	Education/Training

Statement 17 (Section B) is not classified in Table 6.1 as it is not really a barrier or constraint. However, during the pre-testing phase no consensus could be reached on the meaning (interpretation) of this statement. It was then decided to include it in the questionnaire. The statement is: "In South Africa, a tourist is perceived as someone who travels outside his/her own country".

The 5-point Likert scale that was applied measured responses that ranged from "disagree strongly" to "agree strongly". A value of 1 was taken as "disagree strongly" with the statement, a value of 5 was taken as "agree strongly" and the value of 3 represented a neutral perception/feeling to the statement. A value of 2 was thus indicative of disagreement with the statement while the value of 4 was indicative of agreement.

The idea of using the Likert scale method was that instead of learning whether or not the respondent was favourably inclined on an issue, one also gets a measure and a reasonably reliable one of the actual position on the continuum. However, this method is frequently criticised because a number of unique response patterns can result in the same total score (Tull and Hawkins, 1980:322).

Section C:

Open-ended questions are included in questionnaires to obtain a description of the respondent's reaction in the respondent's own words. However, the resulting variety of responses makes coding open-ended questions difficult (Dillon, *et al* 1993:342).

The cost of open-ended response questions is also substantially greater than that of fixed-alternative questions, because the job of coding, editing and analysing the data is quite extensive. As each respondent's answer is somewhat unique, there is some difficulty in categorising and summarising the answers. Another disadvantage of the open-ended response question is that interviewer/researcher bias may influence the responses (Zikmund, 2000:311).

Dillon, *et al* (1993:310) give the following reasons for asking open-ended questions, namely:

- Open-ended questions are useful to check and/or corroborate the results of quantitative or closed-ended questions. Along these same lines, open-ended questions also may be used to develop a wider range of

response than is possible when using quantitative or structured questions;

- Open-ended questions may be used to obtain direct comparisons and to specify particular causes for preference or rejection when two or more stimuli (e.g. products or concepts) are involved in a test;

- Open-ended questions are useful in determining whether a particular communication vehicle (e.g. concept) conveys its intended objectives; and

- Open-ended questions elicit respondents' general reactions to or feelings on exposure to specific products or concepts involved in a test.

Section C of the questionnaire contained a single open-ended question where respondents could give their general comments on problems, barriers and constraints that hamper them as tourism entrepreneurs. It enabled the researcher to gain a wider and more dynamic perspective on their unique views. The comments on responses were evaluated and combined to represent a general view from the tourism industry (refer to Appendix B).

Questionnaires versus interviews:

The distribution of questionnaires electronically via electronic mail ensure fast turnaround time, it ensures flexibility and is less paper chasing. It is also been argued that many respondents feel they can be more candid on e-mail than on personal or telephone surveys for the same reasons they are candid on other self-administered questionnaires. It is argued that e-mail questionnaires are successful for two reasons. They arouse curiosity because they are novel and they reach respondents when they are opening their e-mail, which is when they are prepared to interact (Zikmund, 2000:208).

Questionnaires served mainly the purpose to elicit information while the interviews conducted served to communicate verbally and in a face-to-face

situation. Interviews were helpful to verify that the researcher's understanding and interpretation were correct. The researcher could back track to previous questions and could obtained more information and in greater depth. The researcher could also steer the subject towards the next subject of the questionnaire. Because of interview interaction, questions could be clarified. One could also follow up on ambiguous or interesting responses.

However, interviews tended to be time-consuming and interviewees need to be well informed. An advantage of interviews was that some of them were conducted at home. This ensured that interviewees were relaxed. Interviews that were conducted at office were mainly done after-hours. Business owners seemed to be more relaxed if other staff was out of office. The outcomes of these questionnaires were recorded during the interviews.

6.4.2 Measuring instruments

In measurement, a clear distinction is made between different levels of measurement.

According to Churchill (1991:414) consist measurement of "rules for assigning numbers to objects in such a way as to represent quantities of attributes". However, note the following concerning the definition:

- *First*, it indicates that we measure the attributes of objects and not the objects themselves; and

- *Second*, the definition is broad in that it does not specify how the numbers are to be assigned. In this sense, the rule is too simplistic and conveys a false sense of security, because there is a great temptation to read more meaning into the numbers than they actually contain. We often incorrectly attribute all the properties of the scale of numbers to the assigned numerals.

Tull, *et al* (1980:219) distinguish four different types of number or scales of measurement, namely:

- *Nominal scales* comprise of numbers used primarily to categorise objects or events. A nominally scaled number serves only as a label for a class or category. The objects in each class are viewed as equivalent with respect to the characteristic represented by the nominal number.

- *Ordinal scales* represent numbers used primarily to rank items. Items can be classified not only as to whether they share some characteristic with another item, but also whether they have more or less of this characteristic than some other object. However, ordinal scale numbers do not provide information on how much more or less of the characteristic various items possess.

- *Interval scales* represent numbers used to rank items such that numerically equal distances on the scale represent equal distances in the property being measured. However, the location of the zero point is not fixed. Both the zero point and the unit of measurement are arbitrary.

- *Ratio scales* consist of numbers that rank items such that numerically equal distances on the scale represent equal distances in the property being measured and have a meaningful zero. In general, simple counting of any set of objects produces a ratio scale of the characteristic "existence". In this case, the number zero has an absolute empirical meaning - none of the property being measures exists.

6.4.3 Questionnaire testing

A pilot study is an abbreviate version of a research project in which the researcher practices or tests the procedures to be used in the subsequent full-scale project (Dane, 1990:43). No survey data can be trusted unless one can be sure the respondents understood the instrument and provide appropriate responses (Dane, 1990:127). Pre-testing questionnaires should be taken

seriously, as stated by Dillon, *et al* (1993:321): "No amount of intellectual exercise can substitute for testing an instrument designed to communicate with ordinary people".

The questionnaire was pre-tested with responses from the owners of 2 travel agencies, 2 tour operators, 1 guesthouse owner and 1 lodge owner. All 6 pre-tested respondents were located in Gauteng. Three interviews were conducted during this pilot testing trial, while 3 questionnaires were e-mailed.

The main objective during this phase was to learn whether the questionnaire was correctly worded and for respondents to easily understand all elements of it. The questionnaire was adapted and adjustments were made to some instructions and statements.

6.4.4 Editing and coding of data

Cooper, *et al* (2001:423) state that the customary first step in data analysis is to edit the raw data. Editing detects errors and omissions, corrects them when possible and certifies that minimum data quality standards are achieved. The editor's purpose is to guarantee that data are:

- Accurate;
- Consistent with intent of the question and other information in the survey;
- Uniformly entered;
- Complete; and
- Arranged to simplify coding and tabulation.

Dillon, *et al* (1993:340) argue that response consistency and accuracy are the primary concerns in editing. Editing instructions are written to include as complete a check as possible for consistency. When 2 answers are inconsistent, it may be possible to determine which, if either of the 2 is correct. When this is not possible, both answers should be discarded. With respect to accuracy the person responsible for editing concentrates on signs of interviewer

bias or cheating. For example, common patterns of responses across different questionnaires for the same interviewer can signal potential problems.

In addition to accuracy and consistency, the person who conducts the editing function also is concerned with:

- Response legibility;
- Response clarity, and
- Response completeness.

Once we have collected data, we have to make sense of it. In order to do this, we must organise and code it so that we can analyse it. Therefore, coding means that we have to identify the variable that we want to analyse statistically and decide on the different code values such a variable level presents (Welman, *et al* 1999:208).

Churchill (1991:687) mentions that through coding, raw data are transformed into symbols (usually numerals) that may be tabulated and counted. The transformation is not automatic, but it involves judgement on the part of the coder. Coding closed questions and most scaling devices are simple because the coding is established (for all practical purposes) when the data collection instrument is designed. On the other hand, coding open-ended questions can be very difficult and is often much more expensive than coding closed questions. The coder has to determine appropriate categories on the basis of answers that are not always anticipated.

In the case of this particular study, the editor did not complete incomplete answers, as it could bias the responses of the study. In some cases questionnaires were discarded as unreliable and were consequently totally ignored.

6.4.5 Record-keeping

The researcher kept record of each respondent, questionnaire and comment irrespective if complete or incomplete. Record was also kept of the dates to send follow-up responses to respondents. The researcher reviewed responses on individual questionnaires and stored the 184 original questionnaires for the purpose of transferring the information from questionnaires to a format for statistical analysis.

6.4.6 Data processing

The questionnaires were coded prior to data capturing. The responses were directly captured from the questionnaire using software packages at the Department of Statistics at the University of Pretoria. Some basic calculations were made to check the reliability of the data. Finally the data was imported into the SAS statistical software programme where the final analyses were made. The Iteman programme of the Assessment Systems Programme was used to determine the Cronbach alpha.

6.5 DATA ANALYSIS AND INTERPRETATION

Data analysis is the application of reasoning to understand and interpret the data that have been collected about a subject. Data analysis may involve determining consistent patterns and summarising the appropriate details revealed in the investigation (Zikmund, 2000:66).

Data analysis usually involves reducing accumulated data to a manageable size, developing summaries, looking for patterns and applying statistical techniques. Scales responses on questionnaires and experimental instruments often require the analyst to derive various functions, as well as to explore relationships among variables (Cooper, *et al* 2001:82).

6.5.1 Distribution of questionnaires

Questionnaires were distributed to tourism entrepreneurs in Gauteng and Mpumalanga. All questionnaires were collected either by e-mail, fax or collected personally (during interviews - Gauteng). A number of questionnaires were eliminated in the case of incomplete questionnaires and questionnaires with an excessive amount of missing data. After elimination, 184 questionnaires were coded for data analysis.

The forms of businesses that were selected were:

- Accommodation providers;
- Food and beverage providers;
- Transport providers;
- Tour operators;
- Travel agents;
- Tour guides;
- Conservation and wildlife;
- Shops/retail;
- Conferences, events and meeting providers;
- Art and craft; and
- Adventure, recreation and entertainment providers.

Contact details (e-mail addresses, fax and telephone numbers) of potential respondents were obtained from booklets, pamphlets, brochures, tourism trade shows as well as from the researcher's personal contacts and collection.

The researcher allowed 2 weeks for responding to the questionnaire. Unsuccessful responses where then followed up by either e-mail, fax or telephone call. A total of 3 reminders (follow-ups) were sent to those who did not respond to the questionnaire.

Limitations of the survey method used:

The survey method (e-mail) used in this study had certain limitations, namely:

- Respondents were representative of different ethnic groups. A large majority of especially black tourism entrepreneurs did not have e-mail facilities.

- E-mail addresses were both personal and business addresses. The questionnaire was addressed to the owner of the business. It was assumed that the e-mailed responses were coming from the owner of the business as a number of individuals in a business setting may have access to the same e-mail address.

- Some respondents indicated responses in the spaces that were allocated "for office use only" (refer to Appendix A for questionnaire layout). Furthermore some respondents indicated responses outside of the boxes allocated, or selected one response when more were asked for or vice versa.

- A number of respondents also altered the format of the questionnaire. In most cases it was possible to interpret the responses. In instances where responses were not clearly visible, responses were carried over on an original questionnaire and attached to the questionnaire from the respondent. Some responses even had to be verified through a follow-up e-mail, fax or telephone call.

- Some entrepreneurs responded from e-mail addresses to which the questionnaires were not sent to. This made the tracking of responses very difficult. It was assumed that in these instances, the respondent might have received the questionnaire at the business e-mail address, but opted to complete the questionnaire at home or from an alternate address and may have forwarded it to the home/alternate address from

where the response was sent. However, some respondents also forwarded their contact details on their responses, which could then be followed up.

- A number of respondents e-mailed responses with a note saying "questionnaire attached" where there was none. Quite a large number of questionnaires were also received blank. Furthermore, a number of respondents who had completed the questionnaires experienced difficulties in returning them electronically. Only a few respondents resorted to fax it through. All "failed" responses were followed up with either a telephone call or e-mail note where they were requested to re-send their responses including instructions to save and attach the completed questionnaire. Some respondents did, while the majority was discouraged. Quite a large number of potential responses were lost in this way.

- Some respondents don't run Microsoft word processing software, their mailboxes were full, and some had corrupt e-mails or were too afraid of viruses to open the questionnaire.

- Additionally, a number of respondents preferred not to complete the questionnaire, but rather forwarded information on their establishments such as information regarding its location, facilities available, rate structures, pictures, etc.

6.5.2 Validity and reliability

According to Martins, Loubser and van Wyk (1996:26) researchers often neglect to point out possible shortcomings and pitfalls in research results. Reliability and validity are a prerequisite for research data to be useful and is therefore important to be able to proof reliability and validity.

- ***Validity***

Bagozzi (1994:18) says that a measure or a scale is valid to the extent that it measures what it is intended to measure. While this definition serves to convey the general meaning of validity, it is too vague to be of much theoretical or operational use. A better way to define validity is as a "law-cluster concept" (Putnam, 1962:379). That is, as a group of interrelated ideas whose totality captures the essence of the phenomenon under consideration. Specifically, we define validity as the degree to which a concept and its measures achieve theoretical *and* empirical meaning within the overall structure of one's theory (Bagozzi, 1980:1984).

Forms of validity:

Content validity

Content validity (or face validity) refers to the subjective agreement among professionals that a scale logically appears to reflect accurately what it purports to measure. The content of the scale appears to be adequate. When it appears evident to experts that the measure provides adequate coverage of the concept, a measure has content (face) validity (Zikmund, 2000:282).

Criterion-related validity

Criterion validity is an attempt by researchers to answer the question "Does my measure correlate with other measures of the 'same' construct?" (Zikmund, 2000:282). Criterion validity may be classified as either *concurrent validity* or *predictive validity*, depending on the time sequence of associating the "new" measurement scale and the criterion measure.

Concurrent validity

If the new measure is taken at the same time as the criterion measure, the method is called concurrent validity (Zikmund, 2000:283). Concurrent validity is

the extent to which one measure can be used to estimate an individual's current score on another variable. In its most common form, concurrent validation involves comparing the results of two different measures of the same characteristic in the same objects at the same point in time. Concurrent validity is concerned with the relationship between the predictor variable and the criterion variable when both are assessed at the same point in time (Churchill, 1991:489).

Predictive validity

Predictive validity is established when an attitude measure predicts a future event. The two measures differ only on the basis of a time dimension, that is, only if the criterion is separated in time from the predictor measure (Zikmund, 2000:283). Predictive validity focuses on the usefulness of the measuring instrument as a predictor of some other characteristic or behaviour of the individual (Churchill, 1991:489).

Construct validity

Construct validity refers to the accuracy with which a variable represents a theoretical concept. Validity, including construct validity, is assessed through consensus (Dane, 1990:34). It involves determining the extent to which a measure represents concepts it should represent and does not represent concepts it should not represent. Construct validity involves both making comparisons between a new measure and existing, valid measures of the same concept and contrasting the new measure with existing, valid measures of a different concept (Dane, 1990:259).

Cronbach (1970:143) (Tull, *et al* 1980:234) lists three important steps in construct validation, namely (a) derive constructs that could account for test performance, (b) derive hypothesis from the theory involving the construct and (c) to test the hypothesis empirically. Thus, in construct validation, both the measure and the theory relating the construct to other constructs are being evaluated. Table 6.2 summarises the validity estimates.

Table 6.2: Validity estimates

TYPE	WHAT IS MEASURED	METHODS
Content	Degree to which the content of the items adequately represents the universe of all relevant items under study	Judgmental or panel evaluation with content validity ration
Criterion-related	Degree to which the predictor is adequate in capturing the relevant aspects of the criterion	Correlation
Concurrent	Description of the present; criterion data are available at the same time as predictor scores	-
Predictive	Prediction of the future; criterion data are measured after the passage of time	-
Construct	Answers the question, "What accounts for the variance in the measure?" Attempts to identify the underlying construct(s) being measured and determine how well the test presents it (them)	Judgmental Correlation of proposed test with established one Convergent-discriminant techniques Factor analysis Multitrait-multimethod analysis

Source: Cooper, *et al* (2001:211)

Internal versus external validity:

If a study is well controlled, it is said to be internally valid. A study is internally valid to the extent that important variables are controlled so that the actual relationship between the independent variable and the dependent variable can be observed (Vadum and Rankin, 1998:72). For a study to have internal validity, it must clearly demonstrate that a specific factor causes an affect. The logic of establishing internal validity is simple. *First*, one determines that changes in a treatment are followed by changes on an outcome variable. *Second*, one determines that the treatment is the only factor responsible for the effect. Extraneous factors are ruled out - factors other than the treatment. The most direct way to rule out the possibility that the results are due to the effects of extraneous factors is to eliminate all extraneous factors from the study. If there are no extraneous factors in the study, extraneous factors obviously can't be responsible for the results (Mitchell and Jolley, 2001:144).

External validity is the ability to generalise findings from a specific setting and small group to a broad range of settings and people. It addresses the question: If something happens in a laboratory or among a particular group of subjects, can the findings be generalised to the "real" world or to the general public? High external validity means that the results can be generalised to many situations and many groups of people. Low external validity means that the results apply only to a very specific setting (Neuman, 2000:172).

Vadum, *et al* (1998:74) claim that a study is externally valid to the extent that the relationship observed between the independent and dependent variables generalises to circumstances other than those in the study. This would include generalising to different types of subjects, different settings and different types of measures.

- **Reliability**

Reliability is the extent to which a measure produces stable, consistent scores. Measures are able to produce such stable scores if they are not strongly influenced by random error. A measure can be reliable, but not valid. However, if a measure is not reliable, it cannot be valid (Mitchell, *et al* 2001:115). Vadum, *et al* (1998:15) claim that the reliability of a measuring instrument is a numerical index of the extent to which it yield consistent results from one occasion to the next. To quote Anastasi (1988:109): "Reliability refers to the consistency of scores obtained by the same persons when re-examined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions".

Reliability is a prerequisite for validity (Mitchell, *et al* 2001:96). Test-retest reliability tells the total extent to which random error is influencing the measure. If test-retest reliability is low, one may want to calculate other types of reliability coefficients to find out the main source of the measure's unreliability. The main limitation of reliability is that it does not guarantee validity. Key points to remember about reliability are:

- Two *major* avoidable sources of unreliability are:
  - Random fluctuations in the measurement environment;
  - Random fluctuations in how observers interpret and code observations;
- All reliability coefficients are not the same;
- Test-retest reliability tells the total extent to which random error is influencing the measure;
- Other types of reliability can help to find the source of the measure's unreliability;
- Reliability is necessary for validity. Valid measures are reliable;
- Reliability does not guarantee validity. Reliable measures are not always valid;
- Unreliability weakens validity, but does not introduce systematic bias into the measure; and
- Reliability is an important, but not all-important consideration in choosing a measure.

Reliability is necessary for validity and is easier to achieve than validity. Although reliability is necessary in order to have a valid measure of a concept, it does not guarantee that a measure will be valid. It is not a sufficient condition for validity. A measure can produce the same result over and over (i.e., it has reliability), but what it measures may not match the definition of the construct (i.e. validity).

To ensure validity and reliability, questionnaires that indicated "no response" were not rejected. They applied mainly to the demographic and business information of the questionnaire (Appendix A). Some respondents did not feel comfortable to state their age and/or annual turnover.

6.5.3 Tabulated data

Though informative, descriptive statistics of central tendency and variability are frequently not sufficient for a researcher to fully understand the relationships among a set of variables. Dillon, *et al* (1990:466) say for this reason, we often

turn to tabulation procedures that can provide additional insight into the data before we consider what specific statistical analyses to perform. A number of techniques that are reflected in tabulation form are referred to.

- ***Percentages***

Percentages are the most widely used statistic for reporting and summarising data. One of the most useful features of the percentage is that it can reduce everything to a common base and thereby allow meaningful comparisons to be made (McGown, 1979:214).

- ***Frequencies***

Typically, the initial step is to tabulate responses on a question-by-question basis. At the cleaning stage, these frequencies are sometimes called *marginals* because they give the number and percentage of respondents who chose each alternative (Dillon, *et al* 1993:345). Frequency distributions are therefore useful for summarising responses to specific questions as well as for data cleaning (Dillon, *et al* 1993:377).

Clover and Balsley (1979:186) claim that frequency distribution lends itself to statistical analysis since there are formulas permitting the calculation of all of the averages and the measures of dispersion from frequency distributions. The *arithmetic mean* and the *standard deviation* are the two most widely used measures of average and dispersion.

- ***Averages (measure of central tendency)***

Averages come in three forms, namely the *mode*, the *mean* and the *median*. The *mode* represents the point in the array showing the greatest response level. The *mean* is readily understood as the sum total of values divided by the number of cases. However, it has one shortcoming. It is affected by a few large or small numbers (skewing) since it is based on all the values in the array. The *median* is the value of the middle case in a series. There are as

many cases on the higher side as on the lower side. The measure offers the advantage of being unaffected by extreme cases at one end or the other (Blankenship, *et al* 1993:259).

The *mean* is a particularly informative measure of the "central tendency" of the variable if it is reported along with its confidence intervals. Usually we are interested in statistics (such as the *mean*) from our sample only to the extent to which they can infer information about the population. The *confidence intervals* for the *mean* give us ranges of values around the *mean* where we expect the "true" (population) *mean* is located.

The width of the *confidence interval* depends on the sample size and on the variation of data values. The larger the sample size, the more reliable its *mean*. The larger the variation, the less reliable the *mean*. The calculation of *confidence intervals* is based on the assumption that the variable is normally distributed in the population. The estimate may not be valid if this assumption is not met, unless the sample size is large, say n = 100 or more (StatSoft, 2002).

- ***Item analysis and scaling***

Item analysis involves calculating the *mean* scores for each scale item among the low scores and high scorers. The item *means* between the high-score group and the low-score group is then tested for significance by calculating *t* values (Cooper, *et al* 2001:240).

The Likert-scale procedure has many advantages. It is easy and quick to construct. Each item that is included has met an empirical test for discriminating ability. Since respondents answer each item, it is probably more reliable and it provides a greater volume of data than many other scales (Cooper, *et al* 2001:241).

In Likert-scale procedure a large number of statements are generated and then an *item analysis* is performed. The purpose of the *item analysis* is to ensure that final items evoke a wide response and discriminate among those with

positive and negative attitudes. Items that are poor because they lack clarity or elicit mixed response patterns are eliminated from the final statement list. Hence, a disadvantage of the Likert-type summated rating method is that it is difficult to know what a single summated score means. Many patterns of response to the various statements can produce the same total score. Thus, identical total scores may reflect different "attitudes" because of the different combinations of statements endorsed (Zikmund, 2000:292).

- ***Standard deviation***

The *standard deviation* is a measure of the distance from the observations in a data collection to the middle of that collection. The *variance* of a collection of measurements is the same as *standard deviation* except that the square root is not taken (Hanke, Reitsch and Dickson, 1984:51).

The *standard deviation* is just like the average deviation except that, instead of taking the absolute value of the numbers to eliminate the negative signs, the *standard deviation* takes the square of the deviations of individual values from the arithmetic *mean*. The magnitude of the numbers are altered by this process, so the result is then brought back into line by taking the square root of the resulting sum. Despite the increased difficulty of computation caused by the squaring process, the *standard deviation* is the generally preferred measure of dispersion (McGown, 1979:222).

- ***Correlation***

Correlation is a measure of the relation between two or more variables. The measurement scales used should be at least interval scales, but other correlation coefficients are available to handle other types of data. Correlation coefficients can range from -1,00 to +1,00. The value of -1,00 represents a perfect *negative* correlation while a value of +1,00 represents a perfect *positive* correlation. A value of 0,00 represents a lack of correlation.

The most widely used type of correlation coefficient is *Pearson r*, also called *linear* or *product-moment correlation*. *Pearson correlation* assumes that 2 variables are measured on at least interval scales and it determines the extent to which values of the 2 variables are "proportional" to each other. The value of *correlation* (i.e., correlation coefficient) does not depend on the specific measurement units used. For example, the correlation between height and weight will be identical regardless of whether inches and pounds, or centimetres and kilograms are used as measurement units (StatSoft, 2002).

In order to evaluate the correlation between variables, it is important to know this "magnitude" or "strength" as well as the *significance* of the correlation. The *significance level* calculated for each correlation is a primary source of information about the reliability of the correlation. Therefore, the *significance* of a correlation coefficient of a particular magnitude will change depending on the size of the sample from which it was computed. The test of *significance* is based on the assumption that the distribution of the residual values (i.e. the deviations from the regression line) for the dependent variable *y* follows the normal distribution, and that the variability of the residual values is the same for all values of the independent variable *x* (StatSoft, 2002).

An alternative way of computing the reliability of a sum scale is to divide it in some random manner into two halves. If the sum scale is perfectly reliable, we would expect that the two halves are perfectly correlated (i.e. $r = 1{,}00$). Less than perfect reliability will lead to less than perfect correlation (StatSoft, 2002).

- ***Cronbach alpha***

Cronbach alpha is the mean reliability coefficient calculated from all possible split-half partitions of a measurement scale (Dillon, *et al* 1990:823).

We can estimate the proportion of true score variance that is captured by the items by comparing the sum of item variances with the variance of the sum scale. Specifically, we can compute:

$$\alpha = (k/(k-1)) * [1 - \Sigma (s^2_i)/s^2_{sum}]$$

This is the formula for the most common index of reliability, namely Cronbach's coefficient *alpha* (α). If there is no true score but only error in the items (which is esoteric and unique, and therefore uncorrelated across subjects), then the variance of the sum will be the same as the sum of variances of the individual items. Therefore, coefficient *alpha* will be equal to zero. If all items are perfectly reliable and measure the same thing (true score), then *coefficient alpha* is equal to 1 (StatSoft, 2002).

6.6 FACTOR ANALYSIS, DESIGNS AND LOADING

6.6.1 *Factor analysis*

Factor analysis attempts to describe the relationships that exist among a set of variables by searching for correlation patterns within the data. All variables are given equal treatment. *Factor analysis* is a label that is given to a family of statistical techniques used in the study of such inter-relationships. While the techniques vary from one another in many important ways, they all have one thing in common. They attempt to identify a *factor*, which is defined as a linear combination of the variables being studied. The major difference among techniques is how the weights that are used to compute the linear combination are developed (Hanke, *et al* 1984:458).

The main applications of factor analytic techniques are:

- To *reduce* the number of variables; and
- 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 (the term *factor analysis* was first introduced by Thurstone, 1931) (StatSoft, 2002).

The purpose of factor analysis in this study is to reduce the number of variables and to classify the problems, barriers and constraints into factors.

Closely inter-linked with factor analysis is *eigenvalues*. These equal the sum total of the squares loading for the variables on a factor, provided that a measure of the percentage of variance is contributing variables that are explained by the factor. *Eigenvalues* can be found for square symmetric matrices. There are as many *eigenvalues* as there are rows (or columns) in the matrix. A realistic description of an *eigenvalue* demands a sound knowledge of linear algebra. However, conceptually they can be considered to measure the strength (relative length) of an axis (derived from the square symmetric matrix). *Eigenvalues* are also known as Latent variables (Zikmund, 2000:347).

If we have a measure of how much variance each successive factor extracts, we can turn to the question of how many factors to retain. By its nature this is an arbitrary decision. However, there are some guidelines that are commonly used and that in practice, seem to yield the best results; namely:

- *The Kaiser criterion*:

We can retain only factors with *eigenvalues* greater than 1. In essence this is like saying that, unless a factor extracts at least as much as the equivalent of one original variable, we drop it. This criterion was proposed by Kaiser (1960) and is probably the one most widely used (StatSoft, 2002).

- *The scree test:*

A graphical method is the *scree test* first proposed by Cattell (1966). We can plot the *eigenvalues* in a simple line plot. Cattell suggests finding the place where the smooth decrease of *eigenvalues* appears to level off to the right of the plot. To the right of this point, presumably, one finds only "factorial *scree*" - "*scree*" is the geological term referring to the debris that collects on the lower part of a rocky slope (StatSoft, 2002).

Theoretically, one can evaluate these criteria by generating random data based on a particular number of factors. One can then see whether the number of factors is accurately detected by those criteria. Using this general technique, the first method (*Kaiser criterion*) sometimes retains too many factors, while the second technique (*scree test*) sometimes retains too few. However, both do quite well under normal conditions, that is, when there are relatively few factors and many cases. In practice, an additional important aspect is the extent to which a solution is interpretable. Therefore, one usually examines several solutions with more or fewer factors, and chooses the one that makes the best "sense" (StatSoft, 2002).

6.6.2 Factorial designs

Factorial designs are used to measure the effect of two or more independent variables at various levels. They are particularly useful when there is some reason to believe that the various levels of the independent variables might interact to produce results that neither could produce alone (Tull, *et al* 1980:182).

Churchill (1991:813) argues that there are three very good reasons why one might want to use a *factorial design*. *First*, it allows the interaction of the factors to be studied. *Second*, a factorial design allows a saving of time and effort because all the observations are employed to study the effects of each of the factors. *Third*, the conclusions reached have broader application, since each factor is studied with varying combinations of the other factors. This result is much more useful than it would be if everything else had been held constant.

6.6.3 Factor loading

A *factor loading* is a measure of correlation between a variable and a factor (Hanke, *et al* 1984:461). Many procedures used to rotate the matrix of *factor loadings* do so in a manner to achieve a simple structure (Dillon, *et al* 1990:573). The major characteristics of simple structure are the following:

- Any column of the factor loading matrix should have mostly small values, as close to zero as possible;
- Any given row of the matrix of factor loadings should have non-zero entries in only a few columns; and
- Any two columns of the matrix of factor loadings should exhibit a different pattern of high and low loadings.

6.7 DIFFICULTIES EXPERIENCED DURING THE STUDY

A number of limitations were referred to in 6.5.1 (Chapter 6) concerning the main survey method (e-mailed questionnaires) that was used. Additional difficulties experienced by the researcher include:

- Access to libraries and library material (SAT and the DEAT).
- Difficulties in obtaining databases and statistics from tourism associations, professional bodies, tourism Authorities, Government departments and tourism quality assurance providers.
- The researcher experienced difficulties concerning the definition of the tourism industry. Some conference, meeting and event organisers, bodies and staff felt that they are not part of the tourism industry. It also applied to some accommodation, adventure, recreation and entertainment providers.
- The "Government-related" questions that were asked in Section B of the questionnaire created a problem, as some respondents doubted the aim and purpose of the researcher's study.
- Some respondents indicated that some of the questions were too personal. Therefore, they were discouraged to complete the questionnaire.

- The researcher experienced difficulties in locating the owners of businesses as many of them are abroad, on safari or operate mainly in other provinces or countries (e.g. North West Province, Botswana, Namibia, Mozambique, etc.).

- Some respondents indicated that they don't deal with the public at all, but handle only bookings from travel agents for overseas destinations. Therefore, they were not prepared to assist with the questionnaire.

- Responses from black tourism entrepreneurs were disappointing. The researcher even approached some of them at a popular tourism trade show, but could not convince them to complete the questionnaire. Irrespective of numerous follow-up attempts, the researcher was unable to get their favourable response.

- Complaints raised by (mainly) travel agents and tour operators were that they are too busy to complete questionnaires, that they work under extreme pressure and that they are bombarded by questionnaires. A large number of responses were lost in this way.

- Respondents (mainly tour operators) who make use of answering machines and don't return calls.

- Some accommodation establishments use "old-fashioned" fax machines. Respondents could easily use excuses such as fax machines that were without paper, not working, etc.

- Incorrect telephone numbers, fax numbers and e-mail addresses forced the researcher to make additional telephone calls.

- Some travel agents and guesthouse / bed and breakfast owners felt that their businesses are too small to assist with the questionnaire. In these cases, they had less than 5 full-time employees. On the other hand,

some travel agents and tour operators felt that their firms are too big to complete the questionnaire (more than 20 employees).

- Personal expenses of the researcher mainly because of a large number of e-mails, faxes and telephone calls.

6.8 CONCLUSION

In this chapter only a few methods and techniques in order to create scientifically obtained knowledge by using objective methods and procedures, were discussed. However, research involves different methods for different studies due to different aims. In this particular study relevant and applicable methods and research techniques were used to investigate and test the research problem and objectives.

In order to be able to investigate the research problem and objectives, a thorough and formal methodology should be used. The universe was identified from which the sample frame was derived. A preliminary questionnaire and then the final questionnaire were drawn up, tested and distributed. The results were captured on computer, edited, coded and analysed. Cronbach alpha tests were conducted for reliability of the questionnaire and factor analyses was done.

In the following chapter the research findings of the study are presented.