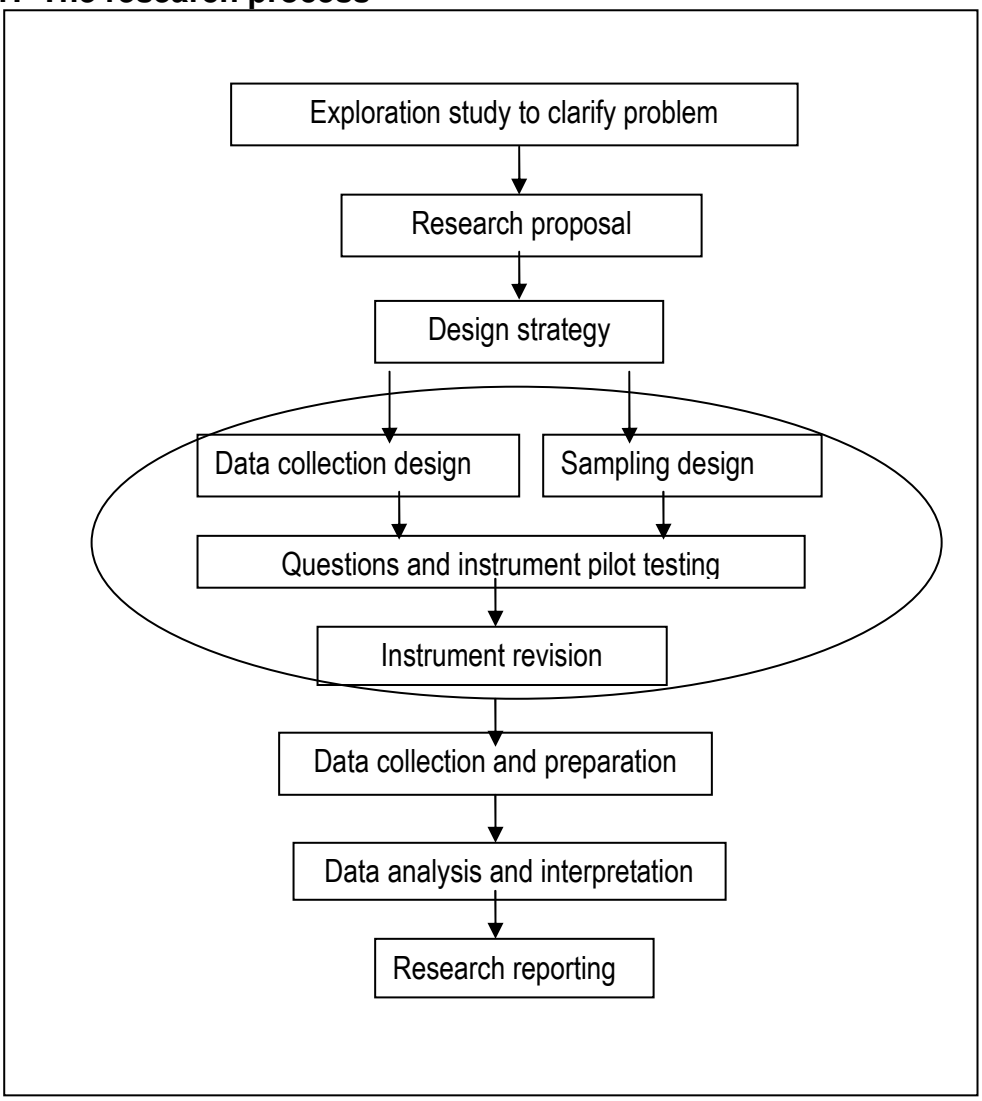
Chapter 5: Research Methodology

5.1 Introduction

This chapter describes the research methodology used to address the objectives of this study. This study uses a two-stage research design comprising of an exploratory study and a formal study to survey SMEs in order to derive owner/manager's opinions on the skills needed in running SMEs successfully. The study focuses on SMEs in the textile and clothing Industry that operate in the municipality of Johannesburg in Gauteng, South Africa. The study follows the research process outlined by Cooper & Schindler (2008:82) as illustrated in Figure 5.1 below:

Figure 5.1: The research process



Source: Cooper & Schindler (2008:82)

This chapter describes the research problem and propositions culminating from the exploratory study. He exploratory study was an extensive literature research towards a detailed and integrated model linking a set of skills and venture performance. The model and the propositions derived are the basis for the research proposal. The research design strategy for the formal study is then described. Issues relating to the design of the measuring instrument to be used, research questions, instrument testing, target population and sample size determination are discussed in detail.

This chapter also outlines the data collection and preparation method. This is followed by an explanation of the methods used to describe, test and analyse the data obtained from the field study against the propositions of the study and for validity and reliability (accuracy and precision) of the constructs. The format of reporting is then summarized and research method chapter concluded by a schematic diagram that links the theory to the research design and instrument.

5.2 Research problem

The research problem was triggered by the need to understand whether training of emerging SMEs (as supported by the training programs of the Department of Labour and the Economic Development Unit of the city of Johannesburg among other stakeholders) has a positive impact on SME success. First the question to be answered was whether there was a set of specific competencies or skills needed for SME success and then to investigate whether these skills are acquired through training or not. So while the study is investigating the nature of skills needed to success in SMEs, its principal aim is to make a valuable contribution to existing government intervention programs that support SME training.

The study sought to answer the following research questions:

- What set of skills are associated with successful entrepreneurs? Can this set of skills be developed and enhanced?
- How important are these skills as perceived by entrepreneurs and SMEs in the Textile and Clothing industry in Johannesburg? Answering this question will show whether these SME owners are aware of the importance of the set of competencies in enhancing business survival and growth.

- How competent are the SME owners in these skills? Answering this question will indicate whether there a link between the skills competence and success of the SME.
- In which of the skills has training been received? Answering this question will illustrate whether competence in the said skills is associated with specific prior training as well as give an indication of current training needs of entrepreneurs working in textile and clothing SMEs.

5.3 Objectives of the study

The aim of this study is to establish which skills, as identified in theory, are perceived as affecting (negatively or positively) the success in the textile and clothing industry within the South African context. The objectives of this study are:

- To review the literature to determine whether there are any common management competencies that contribute to the success of an SME. This will generate a list of specific competencies as identified with international researchers in literature.
- To investigate the importance of these skills as perceived by SMEs in the textile and clothing industry in the city of Johannesburg. This will assess whether SMEs are aware of the most important skills that lead to a competitive SME force in this industry in this area.
- To investigate the levels of competencies of successful and less successful SMEs in the city of Johannesburg. This will give statistical evidence as to whether differences in owner/manager skills result in differences in performance and if there is correlations existing between key skills and success of SMEs.
- To analyse levels of training of SMEs in the textile and clothing industry in Johannesburg in terms of the skills identified as important for SME success. This will map the training of SMEs in the Textile and Clothing industry in Johannesburg and investigate whether training received has any impact on the success of SMEs.
- To suggest areas of improvement in the training of SMEs and in the research needed to help bridge the information gap in addressing problems relating to entrepreneurship and SME development in Africa.

5.4 Research design

The research study will use the two-stage design comprising of an exploratory study and a formal study (Cooper & Schindler, 2008:150). The exploratory study was qualitative research which formed part of the first phase of the research, to determine the concepts to be included in the study theory and to support the foundation and background of this study. This was an intensive exploratory study on existing literature and secondary data available on skills linked with SME success and related training (Chapters 1 to 4). A comprehensive number of text books and articles were reviewed. The focus of the literature was those published in leading academic journals and annual conference proceedings in such disciplines as marketing, entrepreneurship, management, social psychology, economics, organization behaviour and organization theory. Furthermore various internal documents of SME development stakeholders in Gauteng were especially examined for specific data, trends, results and conclusions that were relevant to the study. The exploratory study achieved the following:

- It clarified key definitions, concepts and constructs used in the study.
- It identified variables linked with SME development and entrepreneurial performance. This included the variables that measure performance (turnover, number of employees and profit) and those that measure competence in entrepreneurship, business, personal and technical skills.
- It identified previous research studies on SME development and SME training in Gauteng Province, South Africa. This helped to focus the study and avoid duplication of research work plus encourages the building on work already done.
- It assisted with the development, refining and breaking down of study propositions and sub-propositions.
- It refined the research design into the final blueprint that guided this study from the formulation of the propositions to the report about the analysis of the collected data.

The final 20 skills categories chosen were a combination of skills identified in the literature review and listed in table 5.1 below:

Table 5.1: The 20 skills categories used in the final questionnaire

Number	Skill type
1.	Business systems
2.	Business networking
3.	Communication
4.	Computer literacy
5.	Creative opportunity innovation
6.	Financial Management
7.	Human Resource Management
8.	Legal
9.	Life skills
10.	Literacy
11.	Marketing
12.	Operations
13.	Research & development
14.	Risk taking
15.	Role models
16.	Securing and controlling resources
17.	Self motivation
18.	Strategy planning
19.	Supplier management
20.	Technical / vocational ability

Source: Own compilation based on literature and detailed E/P model developed

The second phase starts where the exploration study leaves off. It involved a formal research with the summary of the research design descriptors given in table 5.2.

Table 5.2: Descriptors of the formal research design

Category	Option Used
Research question clarity	Formal
Method of data collection	Communication/interrogative study
Type	Ex post facto
Purpose	Part descriptive / part casual
Time frame	Cross sectional
Scope	Statistical study

Environment	Field setting
Subjects perception	Actual routine

Source: Adapted from Cooper & Schindler (2008:142)

- The formal research begins with the propositions and involves a precise procedure and data source specifications. The goal of the study is to test the propositions and answer the research questions.
- The method of data collection for the study is communication through a combination of personal interviews, telephone interviews and self-administered questionnaires that are formal and interrogative.
- The formal study is ex-post facto, as the investigators have no control over the variables. The study seeks only to report on what has happened and is happening.
- The research purpose is part descriptive in that one of its objectives is to find out what skills are perceived as important by SMEs in the textile and clothing industry in Johannesburg and if these SMEs have been trained in these skills. It is part causal part descriptive as it seeks to find out the link between a set of certain skills and venture performance and why in the same place, same industry and same period, one set of SMEs is more successful than another. The study gathers information about potentially confounding factors and uses such information to make cross-classification comparisons. In this way, it determines whether there is a relationship between the skill, the success and the training.
- The formal study was cross-sectional, taking a snapshot of the perception of SMEs in the textile and clothing industry operating in Johannesburg in terms of how they perceive the importance of certain skills to their businesses, how they rate their competences in those skills, and if they had been trained in those skills.
- The topical scope of the study is statistical aimed capturing the characteristics of the population of the SMEs in the Textile and Clothing industry in Johannesburg by making inferences from sample characteristics. The propositions are tested quantitatively based on empirical data. Generalizations about the results are presented based on the actual data findings, representativeness of the sample and the validity of the design.
- The design was in the fieldwork environment with all the interviews conducted on the business premises with owner/manager.
- Subjects perceive no deviations from everyday route due to the research.

5.5 Propositions

The study tested the following propositions:

Propositions 1: The following skills are not likely to be considered to be key skills:

- **Propositions 1.1** Marketing
- **Propositions 1.2** Finance
- **Propositions 1.3** Human resource
- **Propositions 1.4** Motivation
- **Propositions 1.5** Gathering of resources
- **Propositions 1.6** Opportunity identification
- **Propositions 1.7** Technical

Propositions 2: The following skills are not likely to be considered as supportive skills:

- **Propositions 2.1** Life skills
- **Propositions 2.2** Numeracy and literacy
- **Propositions 2.3** Communication
- **Propositions 2.4** Business systems
- **Propositions 2.5** Business linkages
- **Propositions 2.6** Computer literacy
- **Propositions 2.7** Legal
- **Propositions 2.8** Operations management
- **Propositions 2.9** Research and development
- **Propositions 2.10** Strategy and business planning
- **Propositions 2.11** Supplier management
- **Propositions 2.12** Risk taking
- **Propositions 2.13** Role models

Technical skills

Proposition 3.1: Successful SMEs are not likely to consider technical skills to be more important for business than less successful SMEs.

Proposition 3.2: Successful SMEs are not likely to be more competent in technical skills than less success that less successful SMEs.

Proposition 3.3: Successful SMEs are likely to have been more trained in technical skills than less successful SMEs.

Personal Skills

Proposition 4.1 to 4.4: Successful SMEs are not likely to consider the following personal skills to be more important for business than less successful SMEs:

Proposition 4.1: Motivation skills

Proposition 4.2: Life skills

Proposition 4.3: Literacy and numeracy skills

Proposition 4.4: Communication

Proposition 5.1 to 5.4: Successful SMEs are not likely to be more competent in the following personal skills than less successful SMEs:

Proposition 5.1: Motivation skills

Proposition 5.2: Life skills

Proposition 5.3: Literacy and numeracy skills

Proposition 5.4: Communication

Proposition 6.1 to 6.4: Successful SMEs are not likely to have been more trained in the following personal skills compared to less successful SMEs:

Proposition 6.1: Motivation skills

Proposition 6.2: Life skills

Proposition 6.3: Literacy and numeracy skills

Proposition 6.4: Communication

Business skills

Proposition 7.1 to 7.11: Successful SMEs are not likely to consider the following business skills to be more important for business success than less successful SMEs:

Proposition 7.1: Business systems

Proposition 7.2: Business linkages

Proposition 7.3: Computer literacy

Proposition 7.4: Financial management

Proposition 7.5: Human resource management

Proposition 7.6: Legal

Proposition 7.7: Marketing

Proposition 7.8: Operations management

Proposition 7.9: Research and development

Proposition 7.10: Strategy and business planning

Proposition 7.11: Supplier management

Proposition 8.1 to 8.11: Successful SMEs are not likely to be more competent in the following business skills than less successful SMEs:

Proposition 8.1: Business systems
Proposition 8.2: Business linkages
Proposition 8.3: Computer literacy
Proposition 8.4: Financial management
Proposition 8.5: Human resource management
Proposition 8.6: Legal
Proposition 8.7: Marketing
Proposition 8.8: Operations management
Proposition 8.9: Research and development
Proposition 8.10: Strategy and business planning
Proposition 8.11: Supplier management

Proposition 9.1 to 9.11: Successful SMEs are not likely to have been more trained in the following business skills compared to less successful SMEs:

Proposition 9.1: Business systems
Proposition 9.2: Business linkages
Proposition 9.3: Computer literacy
Proposition 9.4: Financial management
Proposition 9.5: Human resource management
Proposition 9.6: Legal
Proposition 9.7: Marketing
Proposition 9.8: Operations management
Proposition 9.9: Research and development
Proposition 9.10: Strategy and business planning
Proposition 9.11: Supplier management

Entrepreneurial skills

Proposition 10.1 to 10.4: Successful SMEs are not likely to consider the following entrepreneurial skills to be more important for business success than less successful SMEs:

Proposition 10.1: Opportunity identification, creativity and innovation
Proposition 10.2: Risk taking
Proposition 10.3: Role models
Proposition 10.4: Securing and controlling resources

Proposition 11.1 to 11.4: Successful SMEs are not likely to be more competent in the following entrepreneurial skills than less successful SMEs:

Proposition 11.1: Opportunity identification, creativity and innovation

Proposition 11.2: Risk taking

Proposition 11.3: Role models

Proposition 11.4: Securing and controlling resources

Proposition 12.1 to 12.4: Successful SMEs are not likely to have been more trained in the following entrepreneurial skills compared to less successful SMEs:

Proposition 12.1: Opportunity identification, creativity and innovation

Proposition 12.2: Risk taking

Proposition 12.3: Role models

Proposition 12.4: Securing and controlling resources

Demographics variance

Proposition 13.1 to 13.9: Statistically significant variance does not exist between how successful SMEs view the importance of functional skills regarding the following demographics:

Proposition 13.1: Age

Proposition 13.2: Education

Proposition 13.3: Ethnic group

Proposition 13.4: Gender

Proposition 13.5: Work experience

Proposition 13.6: Region

Proposition 13.7: Subsector

Proposition 13.8: Form of business

Proposition 14.9: Place where business is operated

Proposition 14.1.1 to 14.9: Statistically significant variance does not exist between how less successful SMEs view the importance of functional skills regarding the following demographics:

Proposition 14.1: Age

Proposition 14.2: Education

Proposition 14.3: Ethnic group

Proposition 14.4: Gender

Proposition 14.5: Work experience

Proposition 14.6: Region

Proposition 14.7: Subsector

Proposition 14.8: Form of business

Proposition 14.9: Place where business is operated

Proposition 15.1 to 15.9: Statistically significant variance does not exist between how successful SMEs view the importance of enterprising skills regarding the following demographics:

Proposition 15.1: Age

Proposition 15.2: Education

Proposition 15.3: Ethnic group

Proposition 15.4: Gender

Proposition 15.5: Work experience

Proposition 15.6: Region

Proposition 15.7: Subsector

Proposition 15.8: Form of business

Proposition 15.9: Place where business is operated

Proposition 16.1 to 16.9: Statistically significant variance does not exist between how less successful SMEs view the importance of enterprising skills regarding the following demographics:

Proposition 16.1: Age

Proposition 16.2: Education

Proposition 16.3: Ethnic group

Proposition 16.4: Gender

Proposition 16.5: Work experience

Proposition 16.6: Region

Proposition 16.7: Subsector

Proposition 16.8: Form of business

Proposition 16.9: Place where business is operated

Proposition 17.1 to 17.9: Statistically significant variance does not exist between how successful SMEs rate their competence in functional skills regarding the following demographics:

Proposition 17.1: Age

Proposition 17.2: Education

Proposition 17.3: Ethnic group

Proposition 17.4: Gender

Proposition 17.5: Work experience

Proposition 17.6: Region

Proposition 17.7: Subsector

Proposition 17.8: Form of business

Proposition 17.9: Place where business is operated

Proposition 18.1 to 18.9: Statistically significant variance does not exist between how less successful SMEs rate their competence in functional skills regarding the following demographics:

Proposition 18.1: Age

Proposition 18.2: Education

Proposition 18.3: Ethnic group

Proposition 18.4: Gender

Proposition 18.5: Work experience

Proposition 18.6: Region

Proposition 18.7: Subsector

Proposition 18.8: Form of business

Proposition 18.9: Place where business is operated

Proposition 19.1 to 19.9: Statistically significant variance does not exist between how successful SMEs rate their competence in enterprising skills regarding the following demographics:

Proposition 19.1: Age

Proposition 19.2: Education

Proposition 19.3: Ethnic group

Proposition 19.4: Gender

Proposition 19.5: Work experience

Proposition 19.6: Region

Proposition 19.7: Subsector

Proposition 19.8: Form of business

Proposition 19.9: Place where business is operated

Proposition 20.1 to 20.9: Statistically significant variance does not exist between how less successful SMEs rate their competence in enterprising skills regarding the following demographics:

Proposition 20.1: Age
Proposition 20.2: Education
Proposition 20.3: Ethnic group
Proposition 20.4: Gender
Proposition 20.5: Work experience
Proposition 20.6: Region
Proposition 20.7: Subsector
Proposition 20.8: Form of business
Proposition 20.9: Place where business is operated

5.6 Sampling design

Population

The population of the study will be SMEs in the Textile and Clothing Industry in Johannesburg in the Gauteng Province of South Africa. The Gauteng Enteprise Propeller estimates the population of SMEs in the textile and clothing industry in Johannesburg to be approximately 5000 SMEs (South Africa, 2005b).

However because the demographic characteristics of the SME population are not accurately known in South Africa (Dockel & Ligthelm; 2005:56; Strydom & Tustin; 2003:5), this study uses another method to estimate the population. Research statistics indicate the number of SME over the nine provinces in SA and operating within diverse sectors to be 1,065,494 as at 2001 (South Africa, 2004; Berry et al, 2002:13). The SME sector population in Gauteng is between 33% and 40% of the countries SMEs which is between 352,250 to 426,1976 (Baard & van den Berg, 2004:8; Rogerson, 2004:769). Johannesburg has 70% of SMEs in Gauteng (South Africa, 2006). Thus the population sample of Johannesburg SMEs was estimated to be at least 246,575. In particular the study focused on SMEs in the textile and clothing sector which account for 5% of SMEs in Johannesburg (Rogerson, 2004b:117) resulting in population sample being estimated at 12,090 SMEs.

So this study assumes that the population of SMEs in the textile and clothing industry in Johannesburg is between 5,000 and 12,090 SMEs.

Sampling frame

Owing to the fact that an official register of SMEs in the Textile and Clothing industry in Johannesburg, was not available (i.e. there was no sampling frame) the research started by compiling a list from various sources and SME agencies including:

- SME database for SMEs that have been trained by the Department of Labour in all its skills development programs in the Gauteng province from 2001 to 2004.
- SME database of various departments of the city of Johannesburg.
- SME database at the Premier's Office and with the SMME desk department of economic affairs and finance in Gauteng.
- SME database in SETAs for Textiles & Clothing and related services.
- SME database from Non Government Organizations (NGO's), Community based organizations (CBOs) and development agencies.
- SME database from organized local business associations in Johannesburg.
- SME database from organized labour (South African Workers in the clothing and textile union - SAWCTU) and bargaining councils.
- SME database from the South African Revenue Services (SARS) statistics.
- SME database from the department of trade and industry (the DTI) and its registrar of companies.
- SME database from professional sources like www.brabys.com; ezeedex, monitor, bee gees.
- SME database from industrial indexes and speciality magazines.

Type of sample

A probability sample was used to ensure that each member of the SME population is given a known non-zero chance of selection. The sample was also stratified to ensure that adequate representativity in all the industry groups and in both sub-samples (successful and less successful SME). Simple random sampling was utilized to identify the respondents. This increased accuracy and precision of the sample in representing the characteristics of the population of SMEs in the Textile and Clothing industry in that province.

Some of the businesses interviewed were identified through snowball sampling where the SME participants referred another SME in the industry. This was used because the frame list was not exhaustive and it was difficult to get all the names of all the SMEs and the projects in the Textile and Clothing industry.

Sample size

Sample size determination is one of the most crucial aspects of any empirical research. Too small a sample size undermines the power of the statistical tests of significance (Hair, Anderson, Tatham and Black, 1998). In order to be acceptable a sample must be representative of the entire target population. Due to cost and time constraints, the sample did not include all SMEs in the Textile and Clothing industry in Johannesburg. As it is desired that to have a high degree of confidence in the data obtained in the fieldwork, care was taken when determining the sample size, to limit the standard error of the mean and thus to increase precision levels.

The sample size acceptable is 5% of the total population (Cooper & Schindler, 2008:409). Given this study's estimate of a population of between 5000 and 12,090 means that the targeted sample will be between 250 and 616 respondents.

Furthermore for a sample to be ready for factor analysis, the sample size (or the number of responses) should be at least 5 times the number of variables (Dykman, 2005:147; Urban & van Vuuren, 2006:110). Thus the ideal sample size for this study was calculated to be at least 175 (5 X 35) per sample. Since this study further subdivides the sample into those who were successful and those who were less successful, the initial sample size targeted for the worst case scenario was at least 350 respondents, with 175 respondents per sub-sample (successful and less successful).

The number of respondents interviewed came to a total of 570 SMEs. This covered both sample size criterion in terms of factor analysis of each sub-sample as well as the acceptable sample size per the two population estimations. One sub-sample had a total of 373 respondents who were from less successful SMEs and another had a total of 197 respondents who were from successful SMEs.

This sample size doesn't take into consideration that these SMEs could be further divided into the four SME size categories namely informal; micro; very small, small and medium enterprises and that they were operating across 8 different subsectors of the textile and clothing Industry in the 11 regions of the city of Johannesburg and from every ethnic group in South Africa.

Sample elements

Targeted respondents included SME owners, owner-managers, managers of small ventures and emergent entrepreneurs. Screening procedures eliminated those individuals who do not run their businesses. Thus target respondents have to meet these criteria:

- They operate in the geographic proximity of Johannesburg.
- Their business is within the textile and clothing industry.
- They are actively running the SME.
- Successful SMEs must have been in existence for more than 3 years and have turnover of more than R150 000 and have more than 3 employees.
- Less Successful SMEs must have been in existence for less than 3 years or have turnover of less that R150 000 or have less than five employees.

5.7 Instrument and questions

A structured research instrument (a questionnaire) was used as the instrument to collect data through self administration, face-to-face and telephone interviews. The self designed questionnaire was adapted from previously used and tested instruments with the questions designed using constructs and variables that have been identified in the exploratory study described in chapters 2, 3 and 4.

First investigative questions were developed from the list of refined research questions and information needs coming from the above exploratory study. From here measurement questions were induced by selecting the data type and aligning it to the communication approach. The measurement questions were all aimed at showing or not showing the casual relationship between SME success with a certain set of skills and related training factors.

Only questions that contributed to meaningful answers of whether having certain skills and the related training does contribute significantly to the success of SMEs were included. The nice to have questions were eliminated. Each of the questions was checked for whether they should be asked, if the question is of proper scope and coverage, if the respondent could adequately answer the question and if the respondent will be willing to answer the question. All the questions were checked to ensure that they were asking questions relevant to the propositions. They were also checked to ensure

there was no double meaning, bias and that the respondent would not mistake the meaning of the question or what the interviewer was trying to say.

A fully structured technique was employed to promote objective and efficient scoring and analysis. The structure of the instrument was such that the first section of the questionnaire asked for general characteristics of the firm and demographic background on owner/manager. The firm factors considered included product and service; turnover; number of employees; age of business; form of business; type of industry sub-sector and business location. The owner/manager factors considered included age, gender, levels of education, ethnic group and home language, previous work experience and place of origin indicating family background. The demographics section is important as statistical significance variance can be checked for all demographic variables that impact business performance (Kangasharju, 2000:37; Stewart et al, 2003:35). These demographic factors were measured with mainly closed multiple choice single response questions.

The second section asked three investigative questions aimed at exploring what correlations may or may not have existed between success in entrepreneurship and the said set of competencies (Gartner et al, 1999:219; Klofsten & Spaeth, 2004:1). The investigative questions were:

- To indicate which skills the SMEs perceived to be important for business survival, success and growth. The study also compares differences between successful and less successful SMEs in terms of their perception of which skills are important for successfully running their businesses. These questions aimed at showing or not showing the relationship between important of skill awareness and SME success. Four-point Likert scale questions were used.
- To self evaluate themselves on their abilities on the said skills set. The study will also indicate which skills successful SMEs are more competent in compared with less successful SMEs. The questions aimed at showing or not showing the relationship between sets of competencies and SME success. Four-point Likert scales were used for rating SME ability/knowledge/skill/competencies.
- To indicate for which of these skills they had undergone training and for which they had not been trained. The study will compare training received by successful SMEs with that received by less successful SMEs. The result will show whether the successful SMEs are more trained in the key skills set than less successful SMEs. The questions aimed at showing or not showing the relationship between training in

certain skills sets and SME success. Dichotomous yes/no questions were used to indicate training received.

For each of the investigative questions, the respondents were asked to answer the question evaluating each of the 20 skills categories listed above, meaning each of the three main questions is a cluster of twenty questions; one question for each skills category.

The third and last section of the questionnaire asks the respondent to indicate if the training had impact on performance factors of their business using a five-point Likert scale where 1 = strongly disagree and 5 = strongly agree. This section ends with a free response open-ended question that invited respondents to provide additional unstructured remarks on skills necessary to run an SME; give details of the actual training the SME had received and general comments on the training intervention for SMEs.

One question per time linked the one variable to a proposition. Closed questions facilitated data analysis. The inclusion of a limited number of open-ended questions and the 'other – please specify' multiple choice format option, ensured the receipt of qualitative and quantitative data.

Once developed, the questionnaire was pilot-tested by running pilot interviews with a sample of 10 SMEs, followed by a group discussion to strengthen the validity of the questionnaire and to identify unclear or ambiguous formulated items; to observe nonverbal behaviour signifying discomfort in responding to a particular section or question, to detect flaws in measurement procedures and to investigate the reliability of the selected instrument. This served as a check on whether the planned measurement questions met the data needs of the research question. The questions were then revised and the instrument refined accordingly. Table 5 shows the thesis design flow chart.

Table 5.3: Study design flow chart - from constructs to questionnaire

Propositions, constructs & elements		Questions 1 - 12 = Yes-No Questions 13-48 = FOUR point Likert Scale statements
C1: Background	General human capital	Q1, Q2, Q4, Q5: gender, age, ethnic group, language
	Specific human capital	Q13 , Q14: education & previous experience
	Company demographics	Q11, Q12, Q13: SME age, annual turnover, employees Q7, Q8, Q10: sector, form of business, products Q3, Q9: region and business location
Q15: Importance	Personal Skills	S10, S9: numeracy, literacy, problem solving, decision, time S3, S17: communication, motivation
	Business Skills	S1, S2, S4: business systems, networks, ICT S6, S7, S11: finance, human resources, marketing S8, S12, S19: legal, operations, value chain S13, S18: R&D, strategy and business planning
	Entrepreneurial Skills	S16, S14, S15: ability to gather resources, risk, role models S5: opportunity recognition, creativity and innovation
	Technical Skills	S20: technical and vocational skills
Q16: Competence	Personal Skills	S10, S9: numeracy, literacy, problem solving, decision, time S3, S17: communication, motivation
	Business Skills	S1, S2, S4: business systems, networks, ICT S6, S7, S11: finance, human resources, marketing S8, S12, S19: legal, operations, value chain S13, S18: R&D; strategy and business planning
	Entrepreneurial Skills	S16, S14, S15: ability to gather resources, risk, role models. S5: opportunity recognition, creativity and innovation
	Technical Skills	S20: Technical and vocational skills
C17: Training	Personal Skills	S10, S9. numeracy, literacy, problem solving, decision, time S3, S17: communication, motivation
	Business Skills	S1, S2, S4: business systems, networks, ICT S6, S7, S11: finance, human resources, marketing S8, S12, S19: legal, operations, value chain S13, S18: R&D, strategy and business planning
	Entrepreneurial Skills	S16, S14, S15: ability to gather resources, risk, role models S5: opportunity recognition, creativity and innovation
	Technical Skills	S20: technical and vocational skills
Q18.	Impact analysis	I1-6: quality, productivity, operations, sales, motivation, skills

Source: Own compilation based on literature and detailed E/P model developed

5.8 Data collection methods

The method of data collection for this study was communication or interrogation where data was collected using a combination of personal interviews, telephone interviews and self-administered questionnaires. Most of the interviews were conducted during the period October 2005 to February 2006. In total 10 interviewers (students) were recruited to assist with the collection of data. All were trained before and a guideline document given to them.

The research questionnaire was completed by both successful SMEs (total 270) and less successful SMEs (327). All the interviews were conducted on the business premises with owner/manager. Some SMEs asked for the questionnaire to be either mailed or left at their premises for later collection. Follow-up telephone reminders after 15 days contributed a high response rate. About 30% of the questionnaires were filled using telephone interviews in case of non-respondents. This was consistent with Dahlqvist (2000:7) who relied heavy on telephone interviews to minimize non-response.

Once the data was captured in the questionnaires, a sub-sample of a reasonable number of SMEs and projects was identified for a closely matched comparison of respondents from a wide variety of personal backgrounds such as: age, gender, ethnic origin, educational achievements, as well as organizational characteristics that included, amongst others, size, location, market orientation and economic activity. These in-depth interviews solicited qualitative data that allowed for a comparative analysis.

5.9 Analysis methods

The study questionnaire incorporated all the four data types, namely nominal, ordinal, interval and ratio scales data types.

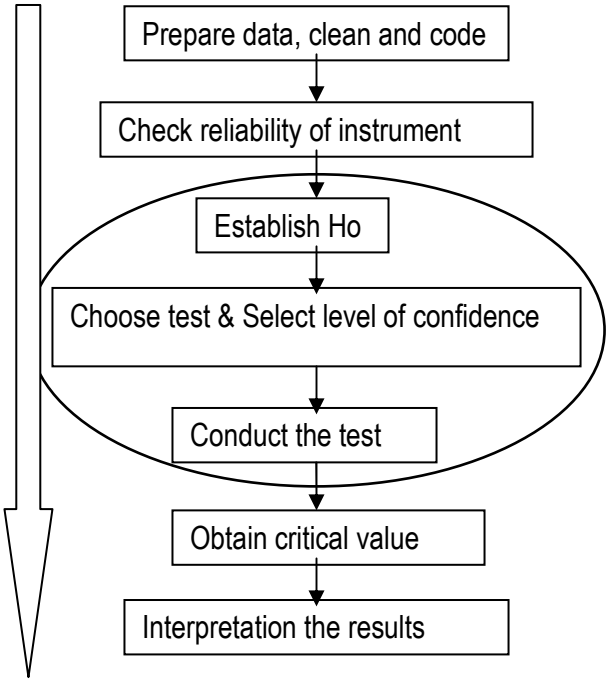
Table 5.4: Types of data and their measurement characteristics

Type of data	Characteristic of data	example	Analysis to determine
nominal	classification	gender (male or female)	equality
ordinal	classification and order	Well, medium, raree	greater or lesser value
Interval	classification, order and distance	Temperature, Likert scale	Equality of intervals
ratio	classification, order, distance & origin	Age, employees	Equality of ratios

Source: Cooper & Schindler (2008:282)

Once the data has been collected the following steps will be applied for analytical purposes as adapted from Cooper & Schindler (2008:476).

Figure 5.2: Steps in analysing data



Source: Adapted from Cooper & Schindler (2008:476)

In quantitative research, data analysis is normally used to refer to the process of breaking down collected data into constituent parts in order to obtain answers to research questions (Terre Blanche & Durheim, 2002:105). Data analysis involves reducing accumulated data to a manageable size, developing summaries, looking for patterns and applying statistical techniques. Scales responses or questionnaires require the analyst to derive various functions as well as to explore relationships among variables. Furthermore, researchers must interpret the results in the light of the research questions or determine if the results are consistent with their propositions and theories; and make recommendations based on the interpretation of the data (Cooper & Schindler, 2001:93).

Preparation of the data

Data analysis begins with editing and coding of the data. Editing included checking of data collection forms for omission, legibility and consistency in classification as well as discarding of completed responses that have missing data; identifying potential error in data collection and discussing its implications (Zikmund, 2003:74). The coding of the data included code development, coding of data, accommodating "Don't know" responses. Transcripts were analysed using content analysis, a method used for coding

both words and phrases dependent upon participants' responses which allows open ended questions to be analysed systematically.

Thereafter data was entered into a user friendly and retrievable database or spreadsheet and SAS statistical software. For this study the questionnaires were processed by the Department of Statistics at the University of Pretoria. The SPSS statistical package of SAS was used to compile descriptive and inferential statistics.

Validity

Validity measures the degree to which a study succeeds in measuring intended values and the extent to which differences found reflects true differences among the respondents (Bateman et al, 2002:79; Cooper & Schindler, 2008:289). There are three types of validity tests namely content, construct and criterion-related validity tests (Cooper & Schindler, 2008:289).

Table 5.5: Summary of validity estimates

Type of validity	What is measured	Methods
Content	Degree to which the content of the items adequately represent the universe of all relevant items under study	Judgemental or panel evaluation with content validity ratio
Criterion related	Degree to which the predictor is adequate in capturing the relevant aspects of the criterion	Correlation
Construct	Identifies the underlying constructs being measured and determine how well the tests represents them	Judgemental; correlation of proposed test with established one; Factor analysis; Multitrait-multi-method analysis and Convergent-discriminant techniques

Source: Adapted from Cooper & Schindler (2008:289)

Reliability

Reliability refers to the degree of reliability of a measurement or low variation between results of different samples of the same population (Bateman et al, 2002:78). Tests of reliability aim to show if the survey can be relied upon to provide the same values if the survey were to be administered repeatedly under similar conditions.

Table 5.6: Summary of reliability estimates

Coefficient	What is measured	Methods and type
Stability	Reliability of a test or instrument is inferred from examinee scores. The same test is administered twice to same subject over a period of less than six months.	Correlation; test-retest
Equivalence	Degree to which alternative forms of the same measure produce same or similar results. Administrated simultaneously without a delay.	Correlation, parallel forms
Internal consistency	Degree to which instrument items are homogenous and reflect the same underlying constructs.	Specialized correlation: split half, KR20 and Cronbach's alpha

Source: Cooper & Schindler (2008:293)

Factor analysis

Factor analysis was executed to confirm the validity and reliability of the measuring instruments (questionnaires) used in this study. The term factor analysis was first introduced by Thurstone in 1931 (Terre Blanche & Durrheim, 2002:117). Factor analysis looks for patterns among the variables to discover whether an underlying combination of the original variables (a factor) can summarize the original set. Factor analysis is used to reduce the number of variables and second to detect structure in the relationship between variables as well as to discover the underlying constructs that explain the variance (Cooper & Schindler, 2008:292).

Factor analysis has two main purposes. First it is used for data reduction and secondly for detection of structure (underlying dimensions) in a set of variables (Zikmund, 2003:586). Factor analysis looks for patterns among the variables to discover whether an underlying combination of the original variables (a factor) can summarise the original set (Cooper & Schindler, 2008:562).

One of the commonly used measures of reliability is Cronbach's alpha coefficient (α); which provides a measure of internal consistency. It 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. It is computed as follows:

$$\alpha = (k/(k-1)*[1- \Sigma (S^2_1)/S^2_{sum}]$$

If there is no true score but only error in 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 the coefficient alpha will be equal to zero. If all items are perfectly reliable and measure the same thing (true score), then the coefficient alpha is equal to 1 (Visser, 2002:195).

Confirmatory factor analysis and corresponding analysis are further emphasized and mentioned below (Statsoft.com, 2006):

- Confirmatory factory analysis allows researchers to test specific propositions about the factor structure for a set of variables, in one or several sampled (for example comparing factor structures across samples).
- Correspondence analysis is a descriptive/exploratory technique designed to analyse two-way and multi-way tables containing some measure of correspondence between two rows and columns. The results provide information which is similar in nature to that produced by factor analysis techniques, and allows one to explore the structure if categorical variables included in the table.

A good factor solution should show invariance in the structure when the factor loadings are derived from various solution techniques. Generally factor analysis is a mathematical procedure not a statistical one, and often misused under this guise. The factor loadings are produced by sampling information, but they cannot be easily tested for significance. Factor analysis assumes that all the variables are caused by the underlying factors Factor analysis can be used to check out the meaning of a particular variable or element to see if it fits the construct. If it does not fit, the element may be dropped (Kim & Mueller 1988:78).

For factor analysis to be reliable, the number of responses should be equal or greater than five times the variables (Brigant & Yarmold, 1995:100). This was achieved by securing 197 responses from the successful group and 373 from the less successful group covering at least 175 responses for each question with 35 variables.

Frequencies

This analysis was used to determine the frequencies distribution and the percentages for categorical variables.

Descriptive statistics

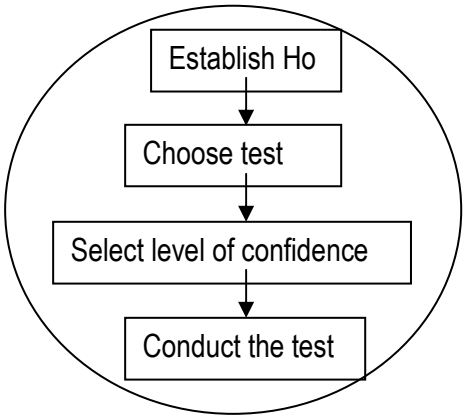
In order to have a broader appreciation of the data collected, descriptive statistical techniques was used to describe characteristics of the population or samples. Descriptive statistics is aimed describing the data by investigating the distribution of the scores for each variable by determining whether the scores on different variables were related to each other (Terre Blanche & Durheim, 2002:105). This reduced the data set and allowed for easier interpretation. It was important to carry out this analysis because it provided a broad biography of the data under study. This enabled the contextualizing of the results.

This statistical method provided information that helped in deciding whether the central location value could be regarded as a reliable representative value of all observations in the data. According to Cooper & Schindler (2008:436) descriptive statistics were used to point out location tendency (mean, median, mode), spread (variance, standard deviation, range, interquartile range) and shape (skewness and kurtosis). The arithmetic average or mean (X) comprised a point which coincided with the sum of the scores divided by the number of scores. The standard deviation (S) showed the variation about the average of the data (Dimantopolous & Schlegelmilch, 2002:97). Calculating the standard deviation of the theoretical distribution of the sample reflected how far the sample means could be derived from the population mean.

Inferential Statistics Tests

Inferential statistics is the method used to draw conclusions about the population itself. While the descriptive analysis allows the researcher to generalize from the sample to the population; inferential statistics allows the research to draw conclusions about the population on the basis of data obtained from samples (Terre Blanche & Durheim, 2002:105). The process for testing statistical significance is illustrated in figure 5.3 below:

Figure 5.3: Process for establishing statistical significance



Source: Cooper and Schindler (2008:503)

Based on the results, several conclusions were reached to reject or accept the propositions. A method of testing checks was performed indicating what percentage of the sampling distribution lies beyond the sample statistics on the curve by comparing the probability values (p values) with the significant level (a). Based on the results, the proposition was either rejected or not rejected. If the p value was less than the significant level, the proposition was rejected. If the p value was greater or equal to the significance level, the proposition was not rejected.

The following techniques were used to perform the inferential analysis: chi-square, t-test and the one way analysis of Variance (ANOVA) as per the table 5.7 below:

Table 5.7: Analysis done

Measurement level	**Independent 2 samples**
Nominal	chi-square 2 sample test
Ordinal	Median test,
Interval and ratio	T test, n way ANOVA

Source: Own compilation

Any appropriately performed test of statistical significance indicates the degree of confidence one can have in accepting or rejecting a proposition.

Chi-square test

The chi-square test is probably the most widely used non-parametric test of significance that is useful for tests involving nominal data. The binomial test is appropriate for situations in which a test for differences between samples is required especially where the population is viewed as only two cases such as successful and less successful and

all observations fall into one or the other of these categories (Cooper & Schindler, 2001:484).

The chi-square test was used to test for significant differences between observed distribution of data among categories and the expected distribution based on the null propositions (Cooper and Schindler, 2001:485). Typically the proposition tested with chi-square is whether or not two different samples are different enough in some characteristic or aspect of their behaviour to allow for the generalization that the population from which the sample drawn is also different in behaviour or characteristic.

A non-parametric test is a rough estimate of confidence. It accepts weaker less accurate data as input than parametric tests and therefore has less status in the pantheon of statistical tests. The chi-square also has its strengths as it is more forgiving in the data it will accept and can thus be used in a wide variety of research contexts. The chi-square test was found to be appropriate for the following questions:

- Is there a relationship between any two variables in the data?
- How strong is the observed relationship in the data?
- What is the direction and shape of the observed relationship in the data?
- Is the observed relationship due to some intervening variables in the data?

While the issue of theoretical or practical importance of a statistically significant result cannot be quantified, the relative magnitude of statistically significant relationship can be measured. A statistically significant result in a chi-square test symbolize the degree of confidence in that the relationship between variables described in the results is systematic in the larger population and not attributable to random error.

The chi-square is a series of mathematical formulae that compare the actual observed frequencies of some phenomenon with the frequencies expected if there were no relationship at all between the two variables in the larger (sampled) population. The chi-square test used is calculated as following:

$$\mathbf{X^2 = SS\ (O_{ij} - E_{ij})^2/E_{ij}}$$

Where:

- O_{ij} is the observed number of cases categorized in the ij^{th} cell
- E_{ij} is the observed number of cases under Ho to be categorized in the ij^{th} cell

For The chi-square test to operate smoothly it requires the following (Blumberg et al, 2005:671; Cooper & Schindler, 2008:482):

- The sample must be randomly drawn from the population or have multinomial distributions.
- Data must be reported in raw frequencies (actual counts and not percentages)
- Measured variables must be independent
- Values / categories on independent and dependant variables must be mutually exclusive and exhaustive
- Observed frequencies cannot be too small. The tradition is that the expected frequency below 5 should not compose more than 20% of the cells and no cell should have an E_i of less than 1.
- As a rule one should perform a chi-square on the data in its un-collapsed form; and if the chi-square value achieved is significant then one may collapse categories to test subsequent refinements of the original proposition.

The chi-square test was used in this study for two independent samples (successful and less successful) to test for difference between the samples in terms of the three main investigative questions.

t-test

According to Zikmund (2003:524) the t-test may be used to test a proposition stating that the mean scores on some variable will be significantly different for two independent sample groups. To test the t-test for difference of means, it is assumed that the two samples are drawn from normal distributions. The null proposition about the differences between groups is normally stated as $u_1 = u_2$ or $u_1 - u_2 = 0$. In most cases comparisons are between two sample means ($X_1 - x_2$). The formula for t is as follows:

$$\mathbf{T = (X_1 - X_2)/(S_{X1-x2})}$$

Where:

- x_1 means for group 1.
- x_2 mean for group 2.
- S_{X1-x2} is pooled or combined standard error of difference between means.

The t test has more tail area that that found in the normal distribution. This is compensation for lack of information about the population standard deviation (Cooper & Schindler, 2008:490).

The t-test was executed to measure the significant differences and similarities between these two sub-samples regarding their perceptions of skill importance and own competence.

One-way Analysis of Variance

The statistical method for testing the null proposition such that the means of several populations are equal, is called the analysis of variance (ANOVA). The testing of two independent variables calls for the introduction of ANOVA. ANOVA is used to test the main and interaction effects of categorical variables on a continuous dependent variable, controlling for the effects of selected other continuous variables which co-vary with the dependent. ANOVA has been used for three purposes:

- In quasi experimental (observational) designs, to remove the effect of variables which modify the relationship of the categorical independents to the dependents.
- In experimental designs to control for factors which cannot be randomized, but which, can be measured on an interval scale.
- In regression models to fit regressions where there are both categorical and interval independents.

In this study the ANOVA analysis is used to prove whether a particular independent factor is positively correlated with the success of the SME. One way ANOVA uses a single factor, fixed effects model to compare the effects of one factor on a continuous dependant variable. In a fixed effects model the levels of the factor are established in advance and the results are not generalized to other levels of treatment. To use ANOVA certain conditions must be met:

- The samples must be randomly selected from normal populations.
- The populations must have the equal variances.
- The distance from one value to its groups means should be independent of the distances of other values to that mean (independence of error).

The ANOVA method is inappropriate when the relationship between the covariates and the responses is not the same in each group. By using the one way ANOVA, dichotomous data have been accommodated in the analysis of significant differences between observations. ANOVA is reasonably robust and minor variations from normality

and equal variance are tolerable. ANOVA breaks down or partitions total variability into components parts. ANOVA uses squared deviations of the variance so computation of distances of the individual points from their own mean or from the grand mean can be summed. In an ANOVA model, each group has its own mean ad values that deviate from the mean. Similarly all the data points from all of the groups produce an overall grand mean. The total deviation is the sum of the squared differences between each data point and the overall grand mean.

The total deviation of any particular data point may be partitioned into between-group variance and within-group variance. The between-group variance represents the effect of the treatment or factor. The differences of between-group means imply that each group was treated differently and the treatment will appear as deviations of the sample means from the grand mean. Even if this were not so, there would still be some natural variability among subjects and some variability attributable to sampling. The within-group variance describes the deviations of data points within each group from sample mean. This results from variability among subjects and from random variation. It is often called error. It is concluded that when the variability attributable to the treatment exceeds the variability arising from error and random fluctuations, the viability of the null proposition begins to diminish. And this is exactly the way the test statistic for analysis of variance works.

The test statistic for ANOVA is the *F* ratio. The *F* distribution determines the size of ratio necessary to reject the null proposition for a particular sample size and level of significance (Cooper & Schindler, 2008:493). The *F*- ratio or translate p-value compares the variance from the last two sources (Cooper & Schindler, 2008:494). To compute the *F*-ration the sum of the squared deviations for the numerator and denominator are divided by their respective degrees of freedom as illustrated below:

$$F = \frac{\text{Between-groups variance}}{\text{Within-groups variance}} = \frac{\text{Mean square}_{\text{between}}}{\text{Mean square}_{\text{within}}}$$

where:

$$\text{Mean square}_{\text{between}} = \frac{\text{Sum of squares}_{\text{between}}}{\text{Degrees of freedom}_{\text{between}}}$$

$$\text{Mean Square}_{\text{within}} = \frac{\text{Sum of squares}_{\text{within}}}{\text{Degrees of freedom}_{\text{within}}}$$

Dividing computes the variance as an average or mean, thus the term mean square. The degrees freedom for the numerator, the mean square between groups, is one less than the number of groups (k-1). The degrees of freedom for the denominator, the mean square within groups are the total number of observations minus the number of groups (n-k).

If the null proposition is true there should be no difference between the populations and the ration should be close to 1. If the population means are not equal, the numerator should manifest this difference, and the F ration should be greater than 1. The f-distribution determines the size of ratio necessary to reject the null proposition for a particular sample size and level of significance.

ANOVA is a versatile statistic which tests for the significant differences between two or more groups of means and additionally breaks down the variability of a set of data into its component sources of variation. ANOVA is carried out in order provide a more in-depth analysis of the data. As with the correlations, some of the study's propositions are built on the significant differences between variables and factors. ANOVA is therefore used to prove or disprove some of the study's propositions.

In this study one-way ANOVA (Analysis of Variance) was used to test the propositions that there was no statistical differences in demographic variables in terms of how successful and the less successful SMEs view the importance of the two factors (functional and enterprising skills) and how the rate their competence in those factors.

Scheffe's multiple comparison procedure

The analysis was concluded by conducting Scheffe's multiple comparison procedure. This investigates the source of variance between the variables. In order to control Type 1 Error (where a true proposition is wrongly rejected) further tests are carried out using Scheffe's multiple comparison procedure. According to Schindler and Cooper (2008:497) Scheffe's test is a further test used after a proposition is rejected. It helps the researcher find the sources of differences within the different levels of the variable. Scheffè's comparison procedure will be used where p-values indicate the direction of the statistical significance.

Probability Values (*p* values) measuring statistical significance

Results are regarded as significant if the *p*-values are smaller than 0.05 (Cooper & Schindler, 2008:470) because this value presents an acceptable level on a 95% confidence interval ($p \leq 0.05$). The *p*-value is the probability of observing a sample value as extreme as, or more extreme than, the value actually observed, given that the proposition is true. The *p*-value is compared to the significance level (α) and on this basis the proposition is either rejected or not rejected. If the *p* value is less than the significance level, the proposition is rejected (if *p* value< α, reject null). If *p* is greater than or equal to the significance level, the proposition is not rejected (if *p* value> α, don't reject null). If the *p* value is less than 0.05, the proposition will be rejected. Results are regarded as significant if the p-values are smaller than 0.05, because this value is used as cut-off point in most behavioural science research. A p-value of 0.05 is typical threshold used in industry to evaluate the null proposition.

P-values do not simply provide one with a Yes or No answer. They rather provide a sense of the strength of the evidence against the null proposition. Small p-values suggest that the null proposition is unlikely to be true. The smaller the p-value is, the more convincing is the rejection of the null proposition. A p-value close to zero signals that the null proposition is false and typically that a difference is very likely to exist. On the other hand, large p-values closer to 1 imply that there is no detectable different for the sample size used.

Using the sampling theory approach the study accepts or rejects a proposition on the basis of sampling information alone. The propositions that were stated earlier in this chapter will be tested and be accepted or rejected in chapter 6.

However it must be recalled that statistical significance also does not ensure substantive significance. A large enough sample may demonstrate a statistically significant relationship between two variables, but that relationship may be a trivially weak one. Statistical significance only means that the pattern of distribution and relationship between variables which is found in the data from a sample can be randomly drawn. By itself it doesn't ensure that the relationship is theoretically or practically important or even very large.

5.10 Reporting research findings

The research findings and its analysis will be described in chapter 6 with the report, containing the study conclusion, the interpretations, conclusions, limitations and recommendations, will be detailed in chapter 7. After the results from the analyses, the study will be able to make predictions and recommendations on a model that will deduce a theory on the optimal combination of the five aspects of training needed to help SMEs do well and succeed in growing their businesses. Visual displays will be done using Excel and PowerPoint graphics.

5.11 Conclusion

This chapter provided a description of the methodology applied in this study. In summary this chapter focused on the research questions posed whether there is a certain set of competencies that allows SMEs to start, run and grow successfully. The data collected was primarily based on personal responses and was collected through research questionnaires. The data processing and analysis attempt to answer the research question through the research findings, which are presented in the following chapter.