Chapter 6: Research design and methodology of the study

"In spite of the impressive proliferation of the entrepreneurial education and training courses in the 1980s and 1990s, little is known about the performance and effectiveness of this training or the extent to which it really matches the needs of target groups. Empirical research in this area remains the exception. Without a stringent feedback about the usefulness of the education programmes, the contents and methods of courses stay to be 'gospel' more than theoretically based teaching."
- Klandt (1993: 37)

6.1 Introduction

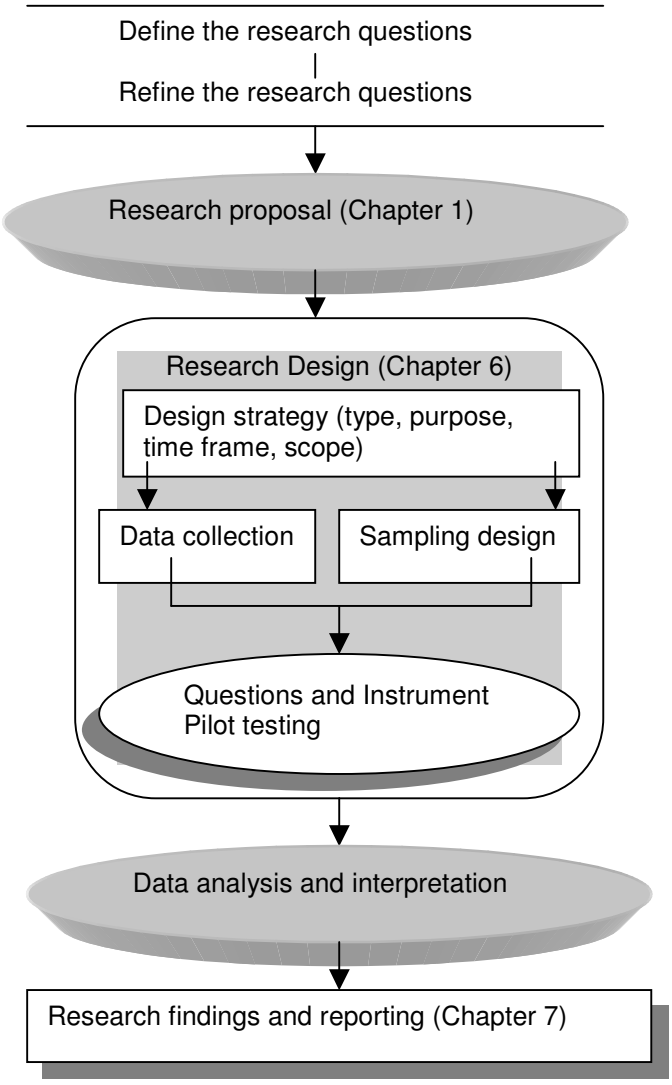
The above statement by Klandt indicates that little is known about the performance and effectiveness of entrepreneurship training programmes. This necessitated a thorough review of the literature. The literature study revealed a need to do empirical research measuring the effectiveness of the WEP as a training intervention for women starting and growing their own businesses. This chapter is concerned with the research methodology and design used to assess the likely success of the WEP as a training intervention for women entrepreneurs in South Africa. Figure 6.1 on the next page illustrates the research process as used throughout this study. The most important objective of the study is to examine whether the WEP is effective in assisting women to start and grow their own businesses.

This study made use of an experimental design, which consists of an independent variable that serves as the manipulated entity. The experimental design of the study involves the treatment of the independent variable as the WEP with an experimental group (women who received entrepreneurial training in the form of the WEP) and a control group (women who did not receive any entrepreneurial training). The time frame when the empirical research was done ranges from 2004 to 2006.

This is a formal study which highlights research problems and hypothesis statements and involves precise procedures and data source specifications. This chapter presents the problem statement, objectives of the study, hypotheses and data collection methods. This chapter also describes how the research questionnaires

were designed and measured to ensure that the researcher obtained valid responses from the respondents. Measurement levels and key performance measures are provided to determine the effectiveness of the WEP. The final section of the chapter concerns the data processing, analysis and characteristics of sound measurement.

Figure 6.1: The research process of the study



Source: Adapted from Cooper and Schindler (2001: 61)

The research proposal was summarised in Chapter 1, in which the research problem and questions were stated. As explained above, this chapter takes an in-depth look at the research design, data collection and sampling design. The data analysis and interpretation of the research findings are presented in Chapter 7.

6.2 The research problem

The research problem was triggered by the study (mentioned in Chapter 1) which was conducted in 2003 by the Chair in Entrepreneurship at the University of Pretoria on 174 women entrepreneurs in South Africa. This study highlighted the lack of training and education as a barrier to women entrepreneurs and how a training programme could solve these problems. As stated at the outset in Chapter 1, while this study is essentially an investigation into the nature and effectiveness of the WEP, its principal aim is to make a valuable contribution to the area of entrepreneurship training programmes.

With this in mind, the study sought to address the following research questions:

- Is the WEP effective in assisting start-up and already established women entrepreneurs to grow their businesses by means of improving their business performance?
- Is the WEP effective in assisting potential women entrepreneurs to start their own businesses?
- Is the WEP effective in assisting start-up and already established women entrepreneurs to start multiple businesses?
- Is the WEP effective in training potential, start-up and established women entrepreneurs?
- Will skills transfer take place after the completion of the WEP?
- Are there significant differences regarding the business performance between the women entrepreneurs who attended and completed the WEP (experimental group) and the women entrepreneurs who did not take part in the WEP (control group)?
- Are there significant differences regarding skills transfer between potential, start-up and established women entrepreneurs?
- Are there significant differences regarding business performance between women entrepreneurs in different provinces in South Africa?
- Does the WEP satisfy the training needs of the experimental group?
- Does the WEP meet the expectations of the experimental group?

6.3 Objectives of the study

The primary and secondary objectives are presented here to illustrate and guide the direction of the research.

6.3.1 Primary objective

The primary objective of the study is to:
Measure the effectiveness of the WEP, as a training intervention, on potential, start-up and established women entrepreneurs in South Africa.

6.3.2 Secondary objectives

The secondary objectives of the study are to:

- Determine whether the training content of the WEP has an effect on women starting their own businesses;
- Determine whether the training content of the WEP has an effect on women entrepreneurs growing their businesses;
- Determine which entrepreneurial, as well as business, skills and knowledge the experimental group learned and gained after they completed the WEP;
- Compare the experimental and control groups approximately six months after the intervention has taken place;
- Determine whether there are significant differences regarding skills transfer between women who already have businesses and those who recently started businesses and those who are potential business owners;
- Determine whether the WEP satisfied the training needs of the experimental group;
- Determine whether the WEP met the expectations of the experimental group;
- Determine whether there are significant differences regarding business performance between women entrepreneurs in different provinces in South Africa.

6.4 Hypotheses

In this study it was decided to state hypotheses rather than propositions, due to the fact that several business research authors state that a hypothesis is a testable proposition (Cooper & Schindler, 2001: 47; Lewis, Saunders & Thornhill, 1997: 344). Blumberg, Cooper and Schindler (2005: 36) agree, stating that a proposition is a statement about concepts that may be judged as true or false if it refers to observable phenomena. When a proposition is formulated for empirical testing, it is called a hypothesis. Zikmund (2003: 43) agrees that propositions are statements concerned with the relationships among concepts; an assertion of a universal connection between events that have certain properties. The author adds that a hypothesis is a proposition that is empirically testable. It is an empirical statement concerned with the relationships among variables.

Therefore the hypotheses are stated below and the hypothesis testing is presented in Chapter 7, which indicates that the hypotheses will be tested empirically. Furthermore, the null hypothesis (Ho) indicates that there are no differences between groups or no relationship between measured variables. The alternative hypothesis (Ha) indicates that there is a difference or relationship between measured variables.

The following hypotheses were formulated out of the research objectives:

Null hypothesis (H1o): The WEP, as a training intervention, is not effective in assisting start-up and established women entrepreneurs to grow their own businesses.

Alternative hypothesis (H1a): The WEP, as a training intervention, is effective in assisting start-up and established women entrepreneurs to grow their own businesses.

The following secondary hypotheses were stated for the study:

H2o: The WEP, as a training intervention, is not effective in assisting women entrepreneurs to start their own businesses.

H2a: The WEP, as a training intervention, is effective in assisting women entrepreneurs to start their own businesses.

H3o: There are no significant differences regarding business performance between the experimental and control groups six months after the experimental group completed the WEP.

H3a: There are significant differences regarding business performance between the experimental and control groups six months after the experimental group completed the WEP.

H4o: The experimental group has not gained entrepreneurial, as well as business, skills and knowledge after the completion of the WEP.

H4a: The experimental group has gained entrepreneurial, as well as business, skills and knowledge after the completion of the WEP.

H5o: There are no significant differences regarding skills transfer between potential, start-up and already established women entrepreneurs.

H5a: There are significant differences regarding skills transfer between potential, start-up and already established women entrepreneurs.

H6o: The WEP did not satisfy the training needs of the experimental group.

H6a: The WEP satisfied the training needs of the experimental group.

H7o: The WEP did not meet the expectations of the experimental group.

H7a: The WEP met the expectations of the experimental group.

H8o: There are no significant differences regarding business performance between women entrepreneurs in different provinces in South Africa.

H8a: There are significant differences regarding business performance between women entrepreneurs in different provinces in South Africa.

6.4.1 Hypotheses testing

The hypotheses testing procedure will be done in Chapter 7 and 8, where the null or alternative hypothesis will be accepted or rejected. According to Zikmund (2003: 500) the significance level is a critical probability in choosing between the null hypothesis and the alternative hypothesis. The statistical significance is presented

later in this chapter in section 6.9.3. The level of significance determines the probability level (0.05 or 0.001) that is to be considered too low to warrant support of the null hypothesis. On the assumption that the null hypothesis being tested is true, if the probability of occurrence of the observed data is smaller than the level of significance, then the data suggests the null hypothesis should be rejected. Table 6.1 indicates that two types of error can be committed in hypotheses testing. Four possible situations can occur when the null hypothesis can be either true or false, and the statistical decision will be either to accept or to reject the null hypothesis.

Table 6.1: Type I and Type II errors in hypotheses testing

State of null hypothesis in the population	Decision	
	Accept Ho	Reject Ho
Ho is true	Correct – no error	Type I error
Ho is false	Type II error	Correct – no error

Source: Zikmund (2003: 504)

6.5 Research methodology

The research methodology which is presented below specifies the methods and procedures for the collection, measurement and analysis of data that the researcher made use of. The research design, used in this study, is explained in section 6.5.1.

6.5.1 The experimental design

According to Cooper and Schindler (2001:136), an experimental design is appropriate when one wishes to discover whether certain variables produce effects in other variables. Experimentation provides the most powerful support possible for a hypothesis of causation.

The empirical study consisted of quantitative research in which three different research questionnaires were used to obtain information from respondents. In terms of evaluating entrepreneurship programmes, it has been suggested that measuring

only quantitative (i.e. economic) results of a programme will fail to provide a full picture of its true value (Henry *et al.,* 2003: 114). Therefore, measurement of the effectiveness of the WEP was done at various levels, as indicated in section 6.8. The first questionnaire was given to respondents, before the actual training took place, to measure the respondents' level of knowledge and skills as well as training expectations and needs (this is referred to as O_1). The second questionnaire was given to respondents to measure their behaviours and attitudes directly after they completed the WEP (this is referred to as O_2) and the third questionnaire measured the respondents' business performance six months after they completed the WEP (this is referred to as O_3).

6.5.2 Classification of experimental designs

Experimental designs vary widely in their power to control contamination of the relationship between independent and dependent variables. According to Cooper and Schindler (2001: 403), the most widely accepted designs are based on this characteristic of control and can be divided into the following four designs:

- Pre-experimental designs;
- True experimental designs;
- Extensions of true experimental designs;
- Field experiments (quasi-experiment).

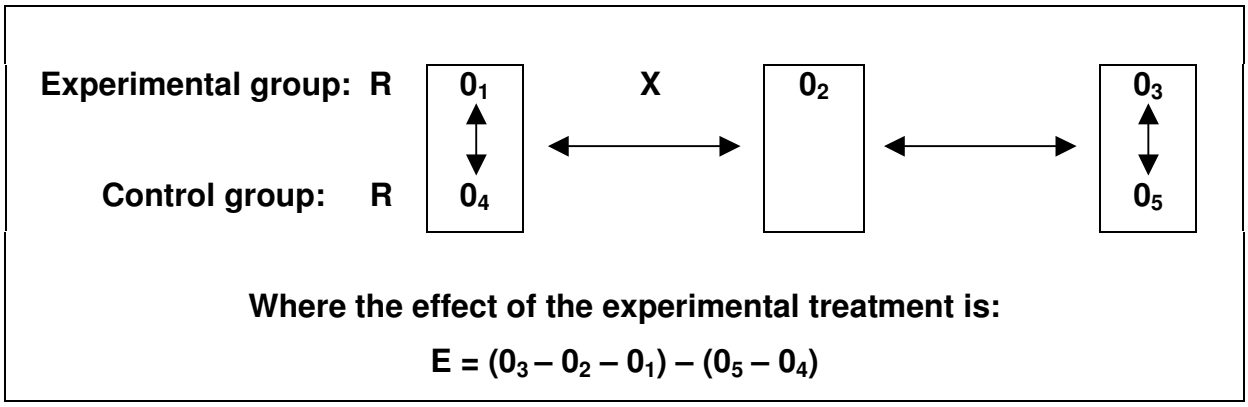
This study was based on a true experimental design which can be classified as the Pretest-Post-Test Control Group Design in which:

R = Randomly assigned group members to a group
X = Exposure of a group to an experimental treatment
0 = Observation or measurement of the dependent variable

Cooper and Schindler (2001: 406) state that the Pretest-Post-Test Control Group design consists of an experimental group and a control group, whereby the experimental group is exposed to a treatment and the control group is not. There was random assignment of women entrepreneurs before the selection and screening

processes, as explained in Chapter 5. Randomisation is the basic method by which equivalence between experimental and control groups is determined (Cooper and Schindler, 2001: 153). In extension of this design, a follow-up observation (six months after the training) was added to strengthen the experimental design and improve the scientific contribution to the field of study.

Figure 6.2: The true experimental design



Source: Own compilation, as adapted from Cooper and Schindler (2001: 406)

The experimental group received the treatment (X) in the form of a training intervention and was observed before the training (O_1), directly after training (O_2) and approximately six months after training (O_3). The control group was observed (O_4) at the same time as the experimental group was observed (O_1) and again approximately six months after that period (O_5). The control group did not receive the treatment (X).

6.5.3 Sampling design and data collection methods

For the selection of the target group out of the population it is important to describe the deciding elements that determine the profile of the selected target group. The determining factors that were taken into consideration when the sample was selected is known as the sampling frame and include the following: Determinant 1 – Already established, start-up or potential women entrepreneurs;

Determinant 2 – Women entrepreneurs with high-growth or potential high-growth ventures;
Determinant 3 – Women whose training needs matched the training content of the WEP.

The sample of the study consisted of 180 women entrepreneurs. The sample included respondents from six different provinces and every ethnic group in South Africa. This was done by running the WEP on six different groups (±20 trainees per group) in the different provinces. One part of the total group consisted of an experimental group (116 respondents), while the other part was the control group (64 respondents). After six months the results of the experimental group were compared with those of the control group. The control group were as far as possible similar to the experimental group in terms of age, experience, skills level and business owners, to name but a few factors. Table 7.1 – 7.10 in Chapter 7 indicates the similarities and differences between the experimental and control groups regarding their personal and business demographics. The *t*-test and chi-square tests were executed to measure the significant differences and similarities between these groups regarding the WEP and are explained in section 7.6.

6.5.3.1 Response rate for the experimental group

Table 6.2 below indicates that a total of 116 women were trained on the WEP from January 2004 to November 2005, all of whom completed and returned the first research questionnaire (O_1). This represents a 100 % response rate. The entrepreneurial learning programme evaluation instrument (O_2) was completed and returned by 106 women after they underwent the six-day WEP. This represents a 91.38 % response rate. Of the 116 women entrepreneurs who were a part of the experimental group, a total of 98 follow-up research questionnaires (O_3) were completed and returned after six months. This represents a 84.48 % response rate. The sample of 98 respondents who completed the follow-up research questionnaires (O_3) was the same respondents that supplied information for questionnaires (O_1 and O_2).

Table 6.2: Experimental group (Time frame: January 2004 – November 2005)

Province	Pre-questionnaire (O_1)	Post-questionnaire (O_2)	Follow-up questionnaire (O_3)
Gauteng group 1	19	18	16
Gauteng group 2	15	13	9
KwaZulu-Natal	19	17	16
Western Cape	19	19	18
Free State, Eastern Cape and Northern Cape	21	18	18
Limpopo and Mpumalange	23	21	21
TOTAL	**116**	**106**	**98**

Source: Own compilation

6.5.3.2 Response rate for the control group

Table 6.3 below indicates that a total of 100 research questionnaires (O_1) were distributed at a women entrepreneurs' meeting in November 2004. A total of 100 questionnaires were returned, of which 64 were usable and could be included in this study, which indicates a 64 % response rate. The 64 women entrepreneurs were contacted by means of personal and telephonic interviews, and 50 completed and returned the follow-up questionnaire (O_3) after the six-month period, in June 2005. This represents a 78.13 % response rate.

Table 6.3: Control group (Time frame: November 2004 – June 2005)

Province	Pre-questionnaire	Post-questionnaire	Follow-up questionnaire
Various Provinces (mainly Gauteng, refer Figure 7.2)	64	-	50
TOTAL	**64**	**-**	**50**

Source: Own compilation

6.5.3.3 Data collection

The method of data collection was based on an interrogation or communication study, in which data was collected by means of personal responses. Primary research was conducted in the form of data collection resulting from the research questionnaires used. Personal and telephonic interviews were conducted with the women entrepreneurs during the follow-up period, after six months. Secondary research was also conducted in the form of a literature review (Chapters 1 to 5) to support the foundation and background of this study.

6.5.4 Sample selection and size

As already indicated, the sample size consisted of 116 (experimental group) and 64 (control group), providing a total of 180 women entrepreneurs. It should be noted that only one respondent from the Eastern Cape Province are included in this sample due to the fact that training was scheduled for that province in 2006. This was unfortunately beyond the time frame of this study and those respondents could not be included in this sample.

Originally the sample was selected by inviting any potential, start-up or established woman entrepreneur to attend the WEP, which meant that each population element had an equal chance of being included in this sample (Probability sampling). Hereafter 40 participants in each province were screened and selected (Chapter 5), therefore non-probability sampling was used. Each population element does not have an equal chance of being included in the sample. For the purposes of this study the sample was restricted, which indicates that each sample element (women entrepreneur) was not drawn individually from the population at large (Cooper & Schindler, 2001: 185). Judgemental sampling was used because the researcher selected sample members who were women entrepreneurs or potential women entrepreneurs in South Africa. According to Zikmund (2003: 392), the advantages of judgemental sampling are that it is useful for certain types of forecasting and the sample is guaranteed to meet specific objectives. A disadvantage may be bias, due to projecting data beyond the scope of the sample. A sample of 180 women entrepreneurs was selected as part of the target population to represent that

population. It can be concluded that the final response rate of this study is 75 % when it is assumed that 40 women out of six different provinces (240 women) were a part of the population at large.

6.5.5 Purpose of the study

The main purpose of the study is causal in nature; there is at least one independent variable and one dependent variable in a causal relationship. In this study the independent variable (IV) is the WEP and the dependent variables (DV) are "starting own businesses" and "growing start-up or established businesses". The purpose of the study is to measure the effectiveness of the WEP on women entrepreneurs and whether it can equip them with the knowledge and skills to start and grow their own businesses. This study will also make agencies, government, financial institutions and other role players aware of the WEP. The main purpose, however, is to inform other women entrepreneurs about the programme which could provide them with training.

6.5.6 The time dimension

The study was based on a mixture of a cross-sectional study and a longitudinal study. Blumberg *et al.* (2005: 130) define a cross-sectional study as one that is carried out once and represents a snapshot of one point in time. The respondents were not measured only once and therefore it is also suggested that this study was based on a longitudinal study, which is one repeated over an extended period. The researcher studied the same women over a period of six months, measuring them at three different points in time. The advantage of this type of study is that it can track changes over time (Cooper & Schindler, 2001: 136). Therefore, longitudinal studies can provide data about past attitudes of women entrepreneurs with reference to entrepreneurial training as well as their future expectations.

6.5.7 The topical scope

The topical scope of this study was based on a statistical study in which the researcher attempts to capture a population's characteristics by drawing inferences

from a sample's characteristics. According to Cooper and Schindler (2001: 137), generalisations about the findings of a statistical study are based on the representativeness of the sample and the validity of the design.

6.5.8 Subjects' perceptions

Cooper and Schindler (2001: 139) emphasise that the usefulness of a design may be reduced when people in a disguised study perceive that research is being conducted. The women entrepreneurs who completed the questionnaires might have perceived deviations as researcher-induced, as they knew research was being conducted.

6.6 Questionnaire design, validity and measurement

According to Sudman and Blair (1998: 300), there is always a chance that some questions could cause problems and questionnaire testing is needed to identify and eliminate these problems. Therefore the next section provides a discussion of the validity of the research questionnaires used in this study.

6.6.1 Validity of the research questionnaires

The research questionnaires (O_1, O_2 and O_3) were first-level pretested on three fellow instrument designers in the Chair of Entrepreneurship in the Department of Business Management at the University of Pretoria. Each specialist examined the questionnaires individually and their comments were then used and the questionnaires were adapted accordingly. This is an example of face (content) validity which refers to the subjective agreement among professionals that a scale logically appears to accurately measure what it is intended to measure (Zikmund, 2003: 302). Face (content) validity and other forms of internal validity are further explained in section 6.7.1.1. The research questionnaires O_1 and O_2 were also pretested in the pilot phase as discussed in Chapter 5, section 5.2.1, by selecting women entrepreneurs from the target population and by simulating the procedures and protocols that had been designated for data collection. The questionnaires were then adapted and some unclear statements were changed or replaced. The validity of the sample was based on the accuracy and precision of the questionnaires.

Although the questions presented in the research questionnaire did not lean in one direction more than another, some respondents did not respond to certain questions asked. These questionnaires were not, however, discarded as they did not affect the validity of the response that the researcher obtained from the respondents. The design of each questionnaire is now explained.

6.6.2 Research questionnaire (O_1) design

This questionnaire (refer Annexure A) was used for the pre-testing of respondents before the actual training took place, and consists of 93 variables. These variables include 15 items dealing with measuring the effectiveness of the WEP. The research questionnaire can be divided into five sections. The first section of the questionnaire collected demographic data on the respondents: age, education, province where their businesses operate, race, language and marital status. The second section of the questionnaire collected data on the respondents' business information: ownership, form of ownership of their businesses, type of industry, annual sales/turnover, value of capital assets, number of employees and customers. The third section collected data on respondents' business success, including profitability, satisfaction of customers, break-even point and business growth indications. The purpose of the fourth section of the questionnaire was to collect data on respondents' entrepreneurial knowledge and skills; they had to rate themselves on several topics. The fifth and final section collected data on the WEP, in terms of what their expectations were, how they rated themselves on certain topics that were going to be covered during the programme, and their entrepreneurial performance. Finally, the last sub-section deals with the business plan, which as seen in earlier chapters plays an important role in the WEP. This research questionnaire (O_1) was completed by both the experimental group (n = 116) and the control group (n = 64); Total = 180.

6.6.3 Entrepreneurial learning programme evaluation instrument questionnaire (O_2) design

This questionnaire (O_2) was distributed directly after the respondents received training (refer Annexure B). This questionnaire consists of 76 variables, including 40 items dealing with the entrepreneurial and management skills and knowledge gained

after the WEP. The research questionnaire can be divided into four sections. The first section of the questionnaire collected data on the respondents' business information: motivational factors to start or grow an own business, and growth plans and strategies. The second section of the questionnaire collected data on respondents' entrepreneurial knowledge and skills, in which they had to rate themselves on several topics. The third section collected data on the WEP, in terms of the respondents' level of satisfaction and whether their expectations of the WEP had been met. The final section of this questionnaire gave the respondents the opportunity to evaluate their level of satisfaction with the facilitator or lecturer. This research questionnaire (O_2) was completed only by the experimental group (n = 106).

6.6.4 Follow-up research questionnaire (O_3) design

The final research questionnaire used in this study is known as the follow-up research questionnaire (O_3) which was used after a six-month period, to measure the effect that the WEP had on the respondents' businesses (refer Annexure C). This questionnaire consists of 75 variables, including 33 items dealing with the measurement of the effectiveness of the WEP. The research questionnaire can be divided into three sections. The first section of the questionnaire collected data on the respondents' business information: annual sales/turnover, value of capital assets, number of employees and customers after six months. The second section collected data on respondents' business success, including profitability, satisfaction of customers, break-even point and business growth indications. The third and final section collected data on the WEP, in terms of how the respondents rated themselves on certain topics covered during the course and what their knowledge of certain business concepts was before and after the WEP. The last sub-section deals with the business plan and what they achieved with the business plan that they prepared during the WEP. This research questionnaire (O_2) was completed by both the experimental group (n = 98) and the control group (n = 50).

6.6.5 Measurement of research questionnaires

The research questionnaires (O_1 and O_3) completed by both the experimental and control groups were adapted to a slight extent when given to the control group due to the fact that they did not receive training in the form of the WEP.
The process structure (response strategies) that was used in the three research questionnaires consisted of the following questions (Table 6.4):

- Dichotomous questions (closed, structured questions);
- Free-response questions (open-ended, unstructured questions);
- 5-point and 4-point Likert scale summated rating question;
  (An even-numbered scale was mostly used in the research questionnaires to avoid the average rating and mid-scores that suggest neutral, average or "don't know" concepts. This enabled the researcher to get usable responses).
- Multiple-choice, single-response questions.

The last mentioned type of question includes multiple options for the respondent, but only one answer is sought and one question can be classified using the multiple-choice, multiple-response scale (also called a checklist).

There are four different types of scales of measurement: nominal scales, ordinal scales, interval scales and ratio scales, as seen in Table 6.4. The nominal data type is used to collect information on gender that naturally or by design can be grouped into male or female categories that are mutually exclusive and collectively exhaustive. This study, however, focused solely on female respondents. According to Diamantopoulos and Schlegelmilch (2002: 25), an ordinal scale establishes an ordered relationship between persons or objects being measured. In ordinal scaling, numbers are used to indicate whether a person, object etc., has more or less of a given characteristic than some other person or object. These authors add that an interval scale possesses all the characteristics of an ordinal scale and, in addition, is characterised by equality of intervals between adjacent scale values. The last scale is the ratio scale, which has all the features of an interval scale plus an absolute zero point (also known as true or natural zero). All of these scales were incorporated in the research questionnaires.

Table 6.4: Characteristics of response strategies used in the questionnaires

Characteristics	Dichotomous	Multiple Choice	Checklist	Free Response	Likert Scale
Type of data generated (measurement scales)	Nominal	Nominal and ordinal	Nominal	Nominal	Interval (pragmatic view) and ordinal (purist view)
Usual number of answer alternatives provided	Two	Three to ten	Ten or fewer	None	Three to seven
Characteristics of data	Classification	Classification and order	Classification	Classification	Classification, order and distance

Source: Cooper and Schindler (2001: 351)

6.7 The characteristics of sound measurement

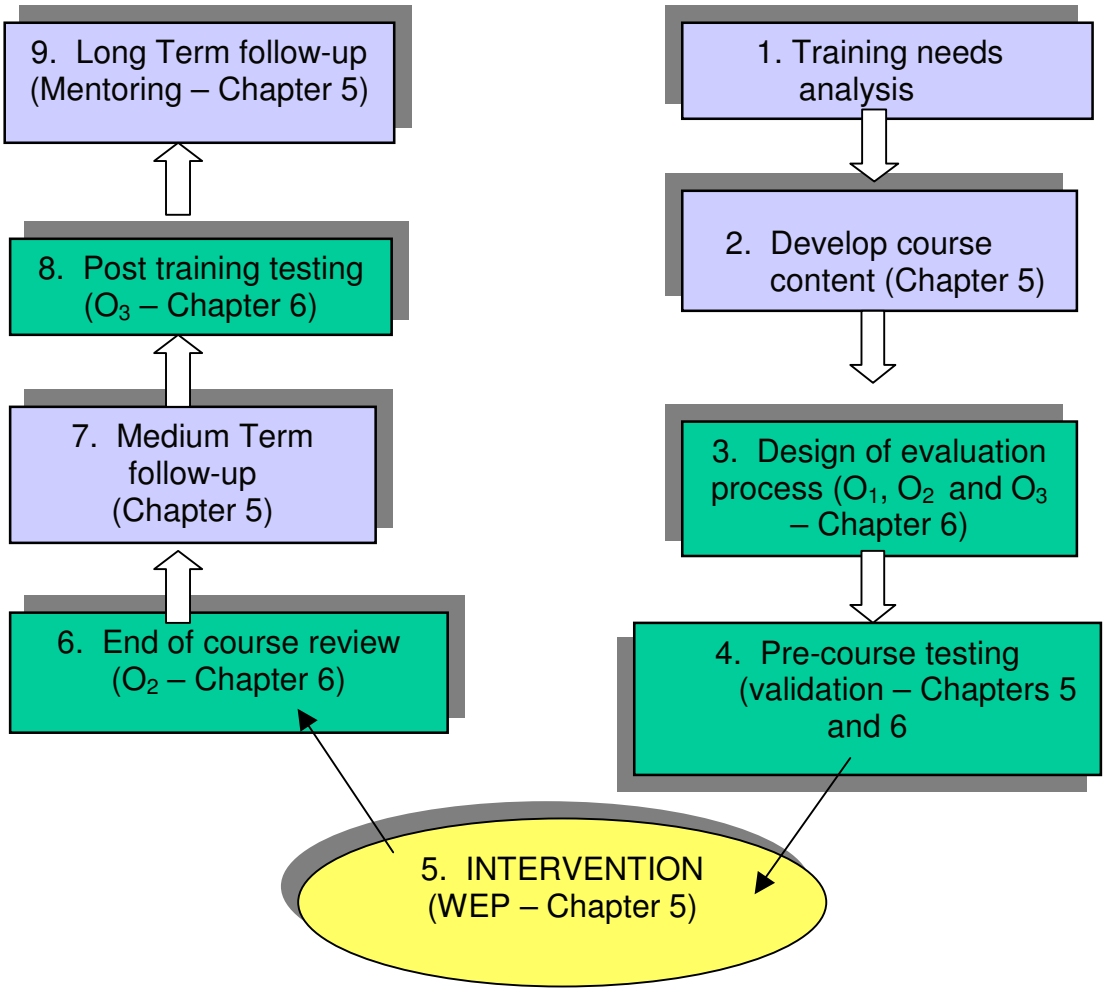
According to Diamantopoulos and Schlegelmilch (2002: 33), the ensuring of validity and reliability is a prerequisite for research data in order to circumvent possible shortcomings and pitfalls in research results. Cooper and Schindler (2001: 210) agree and identify validity and reliability as two characteristics of sound measurement of a research study.

6.7.1 Validity of the training intervention

According to Rae (2002: 88), validation and evaluation are concerned with identifying the change which takes place from the state existing before the training event to that evolving after the training. A number of models of validation and evaluation of training interventions have been put forward. The three principal ones are those

attributed to Kirkpatrick (1967: 98), Hamblin (1976) and Kalleberg and Leicht (1991: 148). Although differing in a number of respects, the first two models are very similar and restrict their coverage to the actual acts of validation and evaluation rather than a complete approach to evaluation. These evaluation models were used in this study to determine the effectiveness of the WEP and will be explained further in section 6.8 of this chapter. Rae's (2002: 6) model of validation and evaluation, shown in Figure 6.2, was used to graphically explain how the WEP was designed as well as provide guidance towards the flow of this study. This model was adapted and used due to the fact that it has been developed, over a number of years of experience in training and development and the practical evaluation of the programmes, as part of a more comprehensive training cycle model (Rae, 2002: 5).

Figure 6.3: Model of validation and evaluation of the study



Source: Own compilation from Rae (2002: 6)

Figure 6.3 illustrates that the WEP and the design thereof took place in several steps:

Step 1: The WEP was designed based on a training needs analysis that was done on the target population (refer to Chapter 4).

Step 2: The programme content was developed based on the training needs of women entrepreneurs (refer to Chapter 5).

Step 3: The design of the evaluation process was done by means of three research questionnaires measuring the respondents at different time periods.

Step 4: The pilot programme took place in 2003 as mentioned in Chapter 5.

Step 5: Six interventions took place in several provinces from January 2004 until November 2005.

Step 6: The respondents were measured directly after the intervention to determine the level of skills transfer that had taken place.

Step 7: The respondents were continuously monitored and mentored after the intervention took place.

Step 8: The respondents were measured again six months after the intervention to determine the effect of the WEP on their businesses.

Step 9: Continuous mentoring and business advising took place.

6.7.1.1 Internal validity

Zikmund (2003: 270) states that in an experimental design, internal validity indicates whether the independent variable was the sole cause of the change in the dependent variable. In other words, was the WEP the sole cause of changes in the women entrepreneurs' businesses?

One widely accepted classification consists of three major forms of validity: content validity, criterion-related validity and construct validity (see Table 6.5). As already mentioned, in this study the principal one that the researcher and trainer encountered is known as content (face) validity – does the research questionnaire reflect the material that is included in the WEP and is it representative of the skill, knowledge or attitude presented in the programme? Rae (2002: 74) indicates that a high content validity for a test will be one in which the majority of items included in the well-balanced programme are included in the test/questionnaire.

Six types of extraneous variables that may influence internal validity negatively are identified by Zickmund (1997: 308). These variables are: History, maturation, testing, instrumentation, selection and experiment mortality. Internal validity may, to a limited extent, be affected due to the unknown background or experience (history) of the delegates (both experimental and control groups). The experimental treatment (training programme) can therefore not be seen as the sole cause of observed changes in the dependent variable. Maturation may also be possible in this study due to the six-month time lapse during which respondents might have forgotten skills gained during the WEP. During this study the same observer was used for all measurements. According to Cooper and Schindler (2001: 401), there can be an instrumentation problem if different observers or interviewers are used for all measurements, yet at the same time they state that using the same observer can also threaten validity.

Table 6.5: Summary of 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.	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. Concurrent: description of the present, data is available at same time as predictor scores Predictive: prediction of the future, after a passage of time	Correlation
Construct	Attempts to identify the underlying construct(s) being measured and determine how well the test represents it (them).	Judgemental, correlation of proposed test with established one, convergent-discriminant techniques, factor analysis and multitrait-multimethod analysis

Source: Cooper and Schindler (2001: 211)

6.7.1.2 External validity

Cooper and Schindler (2001: 403) state that external validity is concerned with the interaction of the experimental treatment with other factors, and the resulting impact on the ability to generalise to (and across) times, settings, or persons. In other words external validity is concerned with whether the research findings indicate a generalisation of results in this study in order to accept or reject the hypotheses stated in this chapter. Among the major threats to external validity are the following interactive possibilities:

- The reactivity of testing on the experimental stimulus (X). The "before" measurement of the respondents' knowledge about entrepreneurship and management concepts of the WEP might have sensitised the respondents to various experimental communication efforts.
- Interaction of selection and X. The process by which respondents were selected to be included in the experimental study may be a threat to external validity. However in this study this was not the case for the experimental group, due to the fact that respondents were screened to do the WEP and be a part of the study (see Chapter 5, Figure 5.2). For the control group selection this might be a threat to external validity.

6.7.2 Reliability of the measuring instruments

Reliability is concerned with whether the measure is reliable to the degree that it supplies consistent results. In this study, however, when looking at the basic definition of reliability, if a measuring instrument/questionnaire is applied at the start of the WEP and the same constructs are measured at the end, with very similar results, the conclusion could be made that the training had failed by not producing the essential change. According to Rae (2002: 74), in practice the reliability is demonstrated by a significant change, provided the same test/questionnaire is administered under the same sort of conditions to the same group that has followed a common programme. According to Diamantopoulos and Schlegelmilch (2002: 34), if a measure is not reliable then it cannot be valid, but if it is reliable it may or may not be valid; put differently, a measure that is valid is also reliable but the reverse is not necessarily true. Reliable instruments are robust, they work well at different times

under different conditions. This distinction of time and condition is the basis for frequently used perspectives on reliability – stability, equivalence and internal consistency as seen in Table 6.6.

Table 6.6: Summary of reliability estimates

Type	Coefficient	What is measured	Methods
Test-retest	Stability	Reliability of a test or instrument inferred from examinee scores. Same test is administered twice to same respondents.	Correlation
Parallel forms	Equivalence	Degree to which alternative forms of the same measure produce the same or similar results.	Correlation
Split-half, Kuder-Richardson formula 20 (KR20) and Cronbach's alpha	Internal consistency	Degree to which instrument items are homogeneous and reflect the same underlying construct(s).	Specialised correlational formulas

Source: Cooper and Schindler (2001: 216)

Cooper and Schindler (2001: 218) state that reliability can be improved by:

- Minimising external sources of variation;
- Standardising conditions under which measurement occurs (During this study this was attempted by the researcher as the measurement was done under the same conditions and circumstances for both the experimental and control groups);
- Broadening the sample of measurement questions used by adding similar questions to the data collection instrument.

Factor analysis was furthermore executed to confirm the validity and reliability of the measuring instruments (questionnaires) used in this study and is explained below.

6.7.2.1 Factor analysis

The main application of factor analysis techniques is, firstly, to reduce the number of variables and, secondly, to detect structure in the relationship 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 in 1931 (Terre Blanche & Durrheim, 2002: 117). Factor analysis is furthermore executed on variables to strengthen the reliability of the research questionnaires used in this study.

One of the most 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 can compute as follows:

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

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 (Visser, 2002: 195).

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

- Confirmatory factor analysis allows researchers to test specific hypotheses about the factor structure for a set of variables, in one or several samples (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 the 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 of categorical variables included in the table.

6.8 Determining and measuring the effectiveness of the WEP

The literature review suggested how the effectiveness of a training programme could be determined and measured. The description of the research design, methodology and the design of the research questionnaires has indicated how they were formulated to measure the effectiveness of the WEP.

To measure the effectiveness and impact of a training programme such as the WEP, Kirkpatrick (1967: 98) suggested measurements on four different levels: Reaction measures (do trainees like the programme?); learning measures (do they understand concepts in the programme?); behaviour measures (can they apply skills gained?); and results/success measures (does it make a difference?). In this study assessing training effectiveness was done in line with these suggestions, as indicated in Table 6.7 (Refer Annexures A, B and C).

Table 6.7: Measurement levels used to determine the effectiveness of the WEP

Type of measurement level	Description	Time of evaluation	Research questions (Q) and questionnaires (O)
Reaction measures	Respondents' satisfaction with the WEP (if not satisfied, reasons were asked for).	This evaluation was done directly after the completion of the WEP.	The entrepreneurial learning programme evaluation instrument (O_2) was used (Q6, 10 and 12).
Learning measures	Used to assess whether the respondents gained	This evaluation was done before and directly after the	The research questionnaire (O_1) was used (Q23

Type of measurement level	Description	Time of evaluation	Research questions (Q) and questionnaires (O)
	entrepreneurial and business specific skills, whether their knowledge increased and if there were any changes in their attitudes.	completion of the WEP.	and 25). The entrepreneurial learning programme evaluation instrument (O_2) was used (Q5, 7, 11).
Behaviour measures	Assessing the effect of the training on the businesses of the respondents. Finding out whether the participants were able to apply these skills to their own businesses.	Six months after the completion of the WEP.	The follow-up research questionnaire (O_3) was used (Q1 – 47).
Post training success measures	Used to measure training outcomes in terms of economic and growth factors such as profits, costs, productivity, turnover, increases in customers and employees.	Six months after the completion of the WEP and measured against the findings of the control group.	The follow-up research questionnaire (O_3) was used (Q4 – 11).

Source: Own compilation

To measure the effectiveness of the WEP even further, this study also made use of the key performance measures that were adopted from Kalleberg and Leicht (1991: 148). Table 6.8 illustrates how these key performance measures were used to determine the effectiveness of the WEP in this study (Refer Annexures A, B and C).

Table 6.8: Key performance measures used to determine the effectiveness of the WEP

Key performance measures	Indicators	Time of evaluation	Research questions (Q) and questionnaires (O)
Primary performance measures	Growth in number of employees, number of customers, sales turnover and value of capital assets. Growth in innovation and listing strategic objectives.	This measurement was done before commencement of training and six months after the completion of the WEP.	The research questionnaire (O_1) was used (Q11 – 21). The follow-up research questionnaire (O_3) was used (Q4 – 11).
Proxy performance measures	Geographical range of markets – national versus international markets, formal business and VAT registration.	This measurement was done before the respondents attended the WEP.	The research questionnaire (O_1) was used (Q1, 9 and 10).
Subjective measures	The ability of the business to meet business and domestic needs – confidence in running a business.	This measurement was done before, directly after and six months after the completion of the WEP.	All three research questionnaires [O_1 (Q15, 17 and 22), O_2 (Q11 and 12) and O_3 (Q8, 10, 12 – 28)] were used.
Entrepreneurial performance	The desire to start a business or the	This measurement was done before	The research questionnaire (O_1) was

Key performance measures	Indicators	Time of evaluation	Research questions (Q) and questionnaires (O)
measures	desire for growth, the ownership of multiple businesses.	and directly after the completion of the WEP.	used (Q7, 19 – 21). The entrepreneurial learning programme evaluation instrument (O_2) was used (Q1 – 4).

Source: Own compilation

Chapter 8, Tables 8.2 and 8.3, will highlight the findings of the above measures and determinants of effectiveness of the WEP.

6.9 Data processing and analysis

Data processing generally begins with the editing and coding of data. According to Zikmund (2003: 72), editing involves checking the data collection forms for omissions, legibility and consistency in classification. Thereafter the questionnaires were processed by the Department of Statistics at the University of Pretoria. The SPSS statistical package of the SAS was used to compile the descriptive and inferential statistics. 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 & Schindler, 2001: 82).

6.9.1 Descriptive statistics

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. Descriptive statistics is the method used to describe characteristics of a population or a sample. It therefore aims at describing data by

investigating the distribution of scores for each variable and by determining whether the scores on different variables are related to each other (Terre Blanche & Durrheim, 2002: 105).

6.9.2 Inferential statistics

Inferential statistic is the method used to draw conclusions about the population itself. In other words, while the descriptive analysis allows the researcher to generalise from the sample to the population, inferential analysis allows the researcher to draw conclusions about the population on the basis of data obtained from samples (Terre Blanch & Durrheim, 2002: 101). Based on the distribution of the descriptive statistics obtained from the study, the following techniques were used to perform the inferential analysis: frequency distribution, cross-frequency tabulation, factor analysis, chi-square test, *t*-test, Wilcoxon test, Mann-Whitney test and the Kruskal-Wallis (K-W) One-Way Analysis of Variance (ANOVA). Table 6.9 illustrates the statistical techniques by measurement level as used during this study. During this study two sample cases (related and independent samples) and k-sample cases (independent samples) were measured.

Table 6.9: Statistical techniques by measurement level and testing situation

Measurement level	Two-Sample Case		K-Sample Case
	Related Samples	Independent Samples	Independent Samples
Nominal	-	X^2 Two-samples test	-
Ordinal	Wilcoxon matched-pairs test	Mann-Whitney test	Kruskal-Wallis one-way ANOVA
Interval and ratio	*t*-test for paired samples	*t*-test	-

Source: Adapted from Blumberg *et al.* (2005: 664)

6.9.2.1 Chi-square test

The chi-square test (x^2) is used to test for significant differences between observed distribution of data among categories and the expected distribution based on the null hypothesis (Cooper & Schindler, 2001: 499). The chi-square was used in this study for two independent samples (experimental and control groups) to test for differences between the samples. The chi-square test used in this study can be calculated as follows:

$$x^2 = \sum_i \sum_j \frac{(Oij - Eij)^2}{E_{ij}}$$

in which

0_{ij} = Observed number of cases categorised in the ijth cell.

E_{ij} = Expected number of cases under Ho to be categorised in the ijth cell.

According to Blumberg *et al.* (2005: 671), for chi-square to operate properly, data must come from random samples of multinomial distributions and the expected frequencies should not be too small. The traditional caution is that expected frequencies below 5 should not compose more than 20 % of the cells, and no cell should have an E_i of less than 1.

6.9.2.2 *t*-test

According to Zikmund (2003: 524), the *t*-test may be used to test a hypothesis stating that the mean scores on some variable will be significantly different for two independent samples or groups. To use the *t*-test for difference of means, it is assumed that the two samples are drawn from normal distributions. The null hypothesis about differences between groups is normally stated as $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$. In most cases comparisons are between two sample means ($\overline{X}_1 - \overline{X}_2$). A verbal expression of the formula for *t* is:

$$t = \frac{\text{Mean 1} - \text{Mean 2}}{\text{Variability of random means}}$$

Thus, the t-value is a ratio with the information about the difference between means (provided by the sample) in the numerator and the random error in the denominator. To calculate *t*, the following formula is used:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

Where:

$\overline{X}_1$ = mean for group 1

$\overline{X}_2$ = mean for group 2

$S_{\bar{x}_1 - \bar{x}_2}$ = pooled, or combined, standard error of difference between means

Zikmund (2003: 525) states that a pooled estimate of the standard error is a better estimate of the standard error than one based on the variance from either sample. It requires the assumption that variances of both groups (populations) are equal. The following formula is used to calculate the pooled standard error of the difference between means of independent samples:

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Where:

S_1^2 = variance of group 1

S_2^2 = variance of group 2

n_1 = sample size of group 1

n_2 = sample size of group 2

The Mann-Whitney test was also carried out in this study and is an alternative to the t-test for two independent samples when assumptions about normality are violated and/or the sample sizes are small (Diamantopoulos & Schlegelmilch, 2002: 181).

The paired sample *t*-test is a parametric test where two related samples are tested and concerns those situations in which persons, objects or events are closely matched or the phenomena are measured twice (Blumberg *et al.*, 2005: 673). This test is done when persons were randomly assigned to groups and given pre-tests and post-tests. In the following formula, the average difference, $\overline{D}$, corresponds to the normal distribution when the α difference is known and the sample size is sufficient. The statistic *t* with (n – 1) degrees of freedom is defined as:

$$t = \frac{\overline{D}}{S_D / \sqrt{n}}$$

Where

$$\overline{D} = \frac{\sum D}{n}$$

$$S_D = \sqrt{\frac{\sum D^2 - \frac{\left(\sum D\right)^2}{n}}{n-1}}$$

The observed significant level for the calculated *t* value is 0.005.

6.9.2.3 Wilcoxon matched-pairs test

Blumberg *et al.* (2005: 577) state that when both direction and magnitude of difference between carefully matched pairs can be determined, the Wilcoxon matched-pairs test must be used. This test has excellent efficiency and can be more powerful than the *t*-test in cases where the latter is not particularly appropriate. Cooper and Schindler (2001: 740) agree and mention that the mechanics of the calculation are also quite simple. Find the difference score (d_i) between each pair of values and rank-order the differences from smallest to largest without regard to sign. The actual signs of each difference are then added to the rank values and the test statistic *T* is calculated. *T* is the sum of the ranks with the less frequent sign.

According to Zikmund (2003: 542), a common situation is the "before/after" experiment, where the same subjects are measured twice.

The formula for the test is:

$$z = \frac{T - \mu_T}{\sigma_T}$$

Where:

$$\text{Mean} = \mu_T = \frac{n(n+1)}{4}$$

$$\text{Standard deviation} = \sigma_T \sqrt{\frac{n(n=1)(2n=1)}{24}}$$

6.9.2.4 Kruskal-Wallis (K-W) One-Way Analysis of Variance (ANOVA)

According to Diamantopoulos and Schlegelmilch (2002: 183), the K-W One-way ANOVA tests the same null hypothesis as the Mann-Whitney *U* test but across three or more independent groups rather than two groups. Zikmund (2003: 544) states that this test may be thought of as a nonparametric equivalent of analysis of variance. However, as with all nonparametric tests, the assumptions are less restricting: if there are three groups, the null hypothesis is that population 1 equals population 2, which equals population 3. In other words, the Kruskal-Wallis test is a technique to determine if the three populations have the same distribution shape and dispersion. Blumberg et al. (2005: 582) agree and suggest that the K-W one-way ANOVA is used to rank all scores in the entire pool of observations from smallest to largest. The rank sum of each sample is then calculated with ties being distributed. The value of *H* is computed as follows:

$$H = \frac{12}{N(N+1)} \sum_{j=l}^{k} \frac{T_j^2}{n_j} - 3(N+1)$$

Where

T_j = Sum of ranks in column *j*

n_j = Number of cases in *j*th sample

$N = \alpha w_j$ = Total number of cases

K = Number of samples

When there are a number of ties, it is recommended that a correct factor (*C*) be calculated and used to correct the *H* value as follows:

$$C = 1 - \left\{ \frac{\sum_{i}^{G} \left(t^3{}_i - t_i \right)}{N^3 - N} \right\}$$

Where

G = Number of sets of tied observations

T_i = Number tied in any set *i*

H = H/C

6.9.3 Statistical significance

The hypotheses that were stated earlier in this chapter will be tested in Chapter 7 and will be accepted or rejected in Chapter 8. Since any sample will almost certainly vary somewhat from its population, it must be judged whether these differences are statistically significant or insignificant (Cooper & Schindler, 2001: 486). A method of presenting the results of a statistical test reports the extent to which the test statistic disagrees with the null hypothesis. This method has become very popular because analysts want to know what percentage of the sampling distribution lies beyond the sample statistic on the curve and most report the results of statistical tests as probability values (*p* values). The *p* value is compared to the significance level (α) and on that basis the null hypothesis is either rejected or not rejected. If the *p* value is less than the significance level (0.05 or 0.001), the null hypothesis is rejected. If p is greater than or equal to the significance level, the null hypothesis is not rejected (refer section 6.4.1).

6.10 Conclusion

This chapter provided a description of the methodology applied in this study. In summary this chapter focuses on the research question posed: Whether the WEP is effective in training potential, start-up and established women entrepreneurs to start and grow their own businesses. The data collection was primarily based on personal responses and was conducted in the form of the research questionnaires (O_1, O_2 and O_3). The data processing and analysis attempt to answer the research question through the research findings which are presented in the following chapter. The measurements and determinants of the effectiveness of the WEP are highlighted in the chapter which provides a foundation for the descriptive statistics discussed in the next chapter.

The explanation of the statistical techniques preceded the actual tests carried out and presented in Chapter 7. These techniques included the t-test, chi-square test, Wilcoxon matched-pairs test, Mann-Whitney U test and Kruskal-Wallis One-Way ANOVA. The next chapter explains and interprets the most significant results as found by executing the above techniques.