Chapter 5: Research procedures and methodology

5.1 Introduction:

The study was concerned with the assessment of the likely success of a creativity, innovation and opportunity finding training intervention, within an entrepreneurship training context. The CIO training model (Creativity, Innovation and Opportunity finding) was applied to the curriculum of the second year of *Baccalaureus Commercii* (BCom) degree specialising in Entrepreneurship, at the University of Pretoria, South Africa. The timeframe ranges from 1999 to 2002. An action learning approach was applied within an experiential learning context. Consequently, the action learning set is defined as being applicable to second-year entrepreneurship learners. The training model forms part of a programme that focuses on the acquisition of entrepreneurial and business skills. The study has an experimental design that constitutes the following (as discussed in Chapter 1). Zickmund (1997:307) defines the experimental design as one that exists as a method based on the manipulation of a variable with the sequential testing of causal relationships among variables.

The experimental design consists of an independent variable that serves as the manipulated entity. The experimental design of the study involves the treatment or the independent variable as the CIO training model with an experimental group (entrepreneurship learners) and a control group (business learners not specialising in entrepreneurship). The first experimental group formed part of the specialised degree in Entrepreneurship (B.Com) and received the CIO treatment as a compulsory subject in the second year of study. The size of the group is 22. The second experimental group is B.Com students specialising in various commercial directions that took the subject (CIO) as an elective.

The size of the group is 69. The Control group were students chosen from different fields of commerce, whom do not take any entrepreneurship subjects, but developed new products as part of a marketing elective. The size of the group is 50.

The *Innovator* © (see Addendum), a measurement instrument that tests the likely success of new products, services or processes, serves as the dependent variable or criteria for judgement. The test units are firstly the learners specialising in entrepreneurship and secondly learners not specialising in entrepreneurship but in general business studies.

The treatment (independent variable) was conducted in a controlled research environment (non-laboratory), therefore striving towards a "constancy of conditions". Extraneous variables were limited as far as possible but interference was present. The main interference was non-class attendance, whereby some learners missed out on the process approach as part of the action learning paradigm applied in this programme.

5.1.1 Internal validity

Zickmund (1997:308) categorises six different types of extraneous variables that may influence internal validity negatively: History, maturation, testing, instrumentation, selection and mortality. Internal validity may, to a limited extent, be affected due to the unknown background or experience (history) of the learners (in both cases: treated and control groups). The experimental treatment (training programme) can therefore not be seen as the sole cause of observed changes in the dependent variable. The age distribution of learners ranged between 19 to 22 years and can to a great extent be generalised as limited business experience. Hence it provides relatively high evidence of internal validity.

5.1.2 External validity

The measurement instrument (*Innovator ©)* has been developed on the basis of the needs of the external business environment. It therefore measures the likely commercial success of new products, services or processes (innovations) in the market place. The external validity of the results tends to be positive, while research results can be generalised to the external environment.

5.1.3 Classification of experimental design

This study is based on the Campbell and Stanley symbolisation, in which:

X = exposure of a group to an experimental treatment
O = observation or measurement of the dependent variable

The classification of this study is a “**static group design**”. A static group design implies that an “after-only” design is present. The treated group is measured after treatment (the CIO training programme) took place and the control group is measured without a treatment intervention (the CIO training programme). The experimental symbolisation can be illustrated as follows:

Experimental group:	**X**	**O_1**
Control group:		**O_2**

Where the effect of the experimental treatment equals 0_2 - 0_1

5.2 Statistical Analysis

This section discusses the statistical techniques used to analyze the data and obtain the research results. Two basic types of statistical analysis were performed in the current study namely descriptive and inferential statistics.

5.2.1 Descriptive statistics

Salkind (2001:150) describes descriptive statistics as the characteristics of the sample. The descriptive statistics used were frequencies, means and standard deviations.

Frequencies refer to the actual amount or percentage of responses to a certain question. These are presented in the current study by means of bar charts or tables. The arithmetic mean, measuring the central tendency, was used to determine the average response of respondents towards a test.

5.2.2 Inferential statistics

Cozby (1985:142) indicates that inferential statistics allow researchers to make inferences about the true differences in the population on the basis of the sample data. An integral concept in inferential statistics is statistical significance and this is discussed firstly.

5.2.3 Statistical Significance

Hypotheses cannot be tested directly on the population (due to constraints of time and money and the population being too big in most cases). Differences, which appear to exist between groups (in the sample data), may in reality (in the real population) not exist (Salkind 2000:170).

Cooper and Emory (1995:434) states that the statistical significance of a result is the probability that the observed relationship (i.e. between variables) or a difference (i.e., between means) in a sample occurred by pure chance ("luck of the draw"), and that in the population from which the sample was drawn, no such relationship or differences exist. This probability is computed for each statistical test and expressed as a p-value.

The higher the p-value, the less one can believe that the observed relation between variables in the sample is a reliable indicator of the relation between the respective variables in the population. Therefore we would in fact like a smaller p-value.

The most commonly used p-values are 0.01 or 0.05. The chosen value p- depends on the researcher and the level of risk, of accepting a difference where in fact there is none (Type 1/Type 2 error).

Cooper and Emory (1995:434) argue that statistical significance is an important research tool, but should not be confused with scientific significant findings. The authors also distinguish between statistical significance and practical significance.

A result may be statistically significant yet whether it is important to the scientific community will depend on the nature of the variable being studied. Practical significance of a treatment effect depends on factors other than statistical significance such as cost and validity of the study. Even if the treatment effect is significant it may be to costly to implement and has therefore no practical significance.

5.2.4 t-test

DeFusco et al. (2001:327) shows that the appropriate inferential test when comparing two means obtained from different groups of subjects is a t-test for independent groups. The t for independent groups is defined as the difference between the sample means divided by the standard error of the mean difference

The authors indicate furthermore that the p-level reported with a t-test represents the probability of error involved in accepting our research hypothesis about the existence of a difference. The null hypothesis is that of no difference between the two categories of observations (corresponding to the groups).

Some researchers suggest that if the difference is in the predicted direction, you can consider only one half (one "tail") of the probability distribution and thus divide the standard p-level reported with a t-test (a "two-tailed" probability) by two. Others, however, suggest that you should always report the standard, two-tailed t-test probability.

As the two-tailed p-values in the current study is all highly significant (below 0.001), it was not considered necessary to divide them even though the differences are in the expected direction.

5.2.5 ANOVA

In general, the purpose of analysis of variance (ANOVA) is to test for significant differences between means. ANOVA tests the null hypothesis that all the population means are equal:

H_0: $\mu_1 = \mu_2 = ... = \mu_a$

Two estimates are derived. One estimate (called the Mean Square Error or "MSE" for short) is based on the variances within the samples and the other (Mean Square between or "MSB" for short) is based on the variance of the sample means.

(http://davidmlane.com/hyperstat/intro_ANOVA.html)

In ANOVA we can test each factor while controlling all others; this is actually the reason why ANOVA is more statistically powerful (i.e., we need fewer observations to find a significant effect) than the simple t-test.

When the F-test in the one-way analysis of variance proves significant at the 5% level of significance, it shows that there are statistically significant differences. Yet often this simply says that 3 or more groups are different with respect to their mean scores and to understand between which of the 3 groups the differences actually lies, a so-called post-hoc test procedure is applied to test which pair-wise group difference are significant. The procedure used in the present study is the Bonferroni.

5.2.6 Reliability and Validity

Salkind (2000:105) states that reliability and validity is the hallmarks of good measurement. Reliability and validity is a researcher’s first line of defence against spurious and incorrect conclusions. Reliability is when a test measures the same thing more than once and results in the same outcome.

The present study calculated the Cronbach Alpha coefficient as a measure of the internal consistency reliability of each of these scales.

Cronbach's alpha measures how well a set of items (or variables) measures a single unidimensional latent construct. When data have a multidimensional structure, Cronbach's alpha will usually be low. (http://www.ats.ucla.edu/stat/spss/faq/alpha.html)

Alpha coefficient ranges in value from 0 to 1 and may be used to describe the reliability of factors extracted from dichotomous and/or multi-point formatted questionnaires or scales.

The more items there are in a scale designed to measure a particular concept, the more reliable the measurement instrument will be.

The alpha is calculated as follows:

$$\alpha = \frac{N \cdot \bar{r}}{1 + (N-1) \cdot \bar{r}}$$

The higher the Alpha is, the more reliable the test. There is not a generally agreed cut-off. Usually 0.7 and above is acceptable (Nunnally 1978). It is a common misconception that if the Alpha is low, it must be a bad test. Actually the test may measure several attributes/dimensions rather than one and thus the Cronbach Alpha is deflated. (http://www.ats.ucla.edu/stat/spss/faq/alpha.html)

5.2.7 Effect sizes

The effect size tells you something very different from the p-value. A result that is statistically significant is not necessarily particularly important as judged by the magnitude of the effect, a highly significant result should therefore not automatically be interpreted as reflecting large effects (Rosnow and Rosenthal 1996:276). Sometimes also a significant effect might be missed as the result failed to be significant (a Type 2 error). The effect size is also not dependent on the sample size, where the significance p-value is linked to the size of the sample.

Many indices of effect size have been formulated, each test corresponding with a separate index or formula. As the effect size measure the strength of a result/relationship, Rosnow and Rosenthal (1996) explains that the product moment r (correlation) is a good index of effect size.

The following guideline to the interpretation of the size are set by the Authors of “Conceptual introduction to power and effect size” :

Test	**Small**	**Medium**	**Large**
t-test (d)	0.2	0.5	0.8
Correlation r	0.1	0.3	0.5
Chi-square (e)	0.05	0.1	0.2

Four of the commonly used measures of effect size in ANOVA are:

- Eta squared, η^2
- partial Eta squared, η^2 p
- omega squared, w^2
- the Intraclass correlation, rI

In the current study eta squared ($\eta^{2)}$ was calculated by the program SPSS as part of an ANOVA process.

η^2 varies from 0 to 1 the bigger the number the greater effect (just like a correlation coefficient).

η^2 is the proportion of the total variance in the dependent variable that is associated with a particular effect.

The calculation is:

$$\eta^2 = \frac{SS_{school}}{SS_{total}}$$

5.2.8 Discriminant analysis

Discriminant function analysis is applied to determine which variables discriminate between two or more naturally occurring groups. For example, an educational researcher may want to investigate which variables discriminate between high school graduates who decide (1) to go to college, (2) to attend a trade or professional school, or (3) to seek no further training or education. For that purpose the researcher could collect data on numerous variables prior to students' graduation. After graduation, most students will naturally fall into one of the three categories. *Discriminant Analysis* could then be used to determine which variable(s) are the best predictors of students' subsequent educational choice.

The groups in this study are obviously not naturally occurring, therefore the one group is manipulated into what is now their results. If one obtains predictors, it is acquired for people that already know/have the knowledge that the experimental group has after the intervention/training.