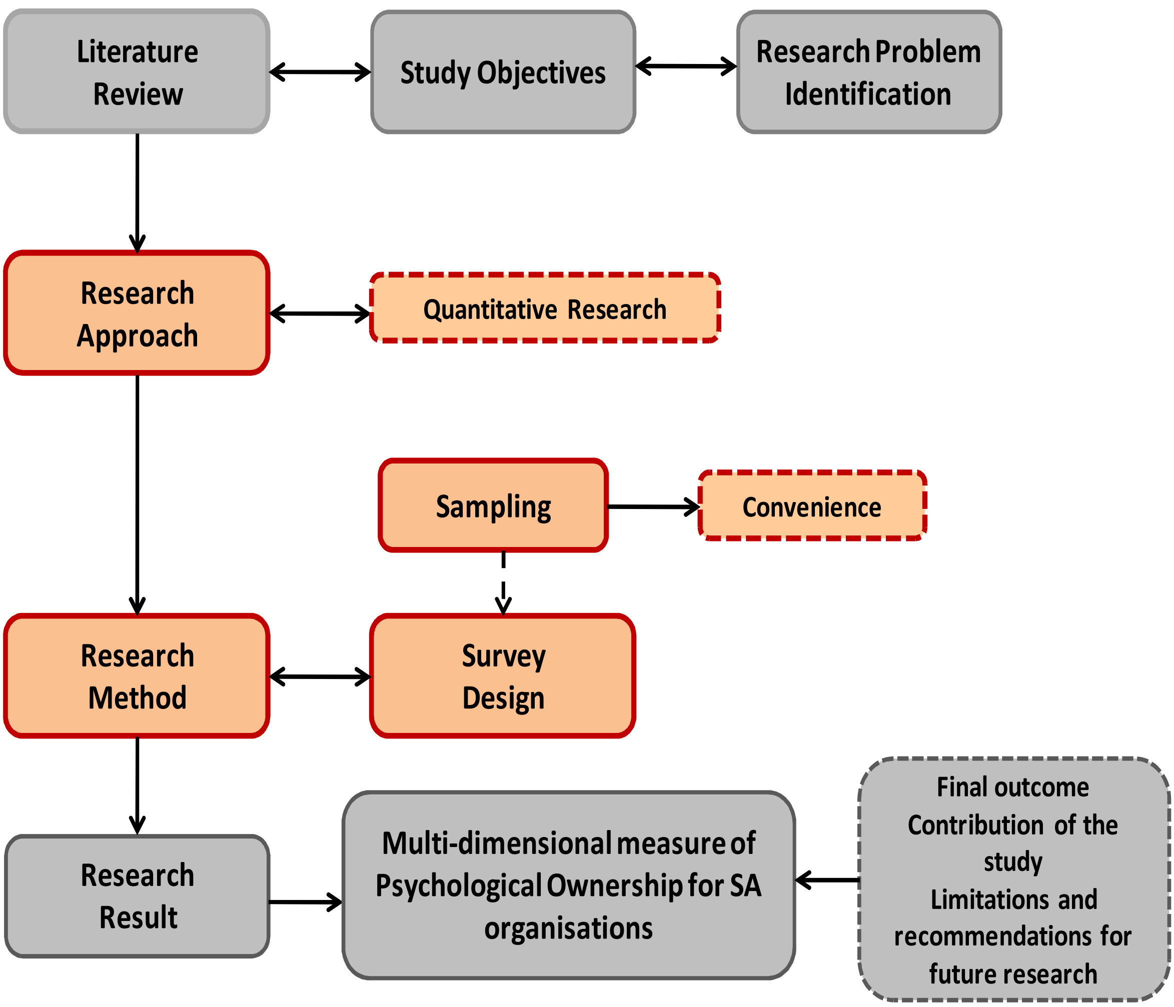
CHAPTER 3

RESEARCH DESIGN AND METHODS

If *the map shows a different structure from the territory represented… then the map is worse than useless, as it misinforms and leads astray.*

- ***Alfred Korzybski***

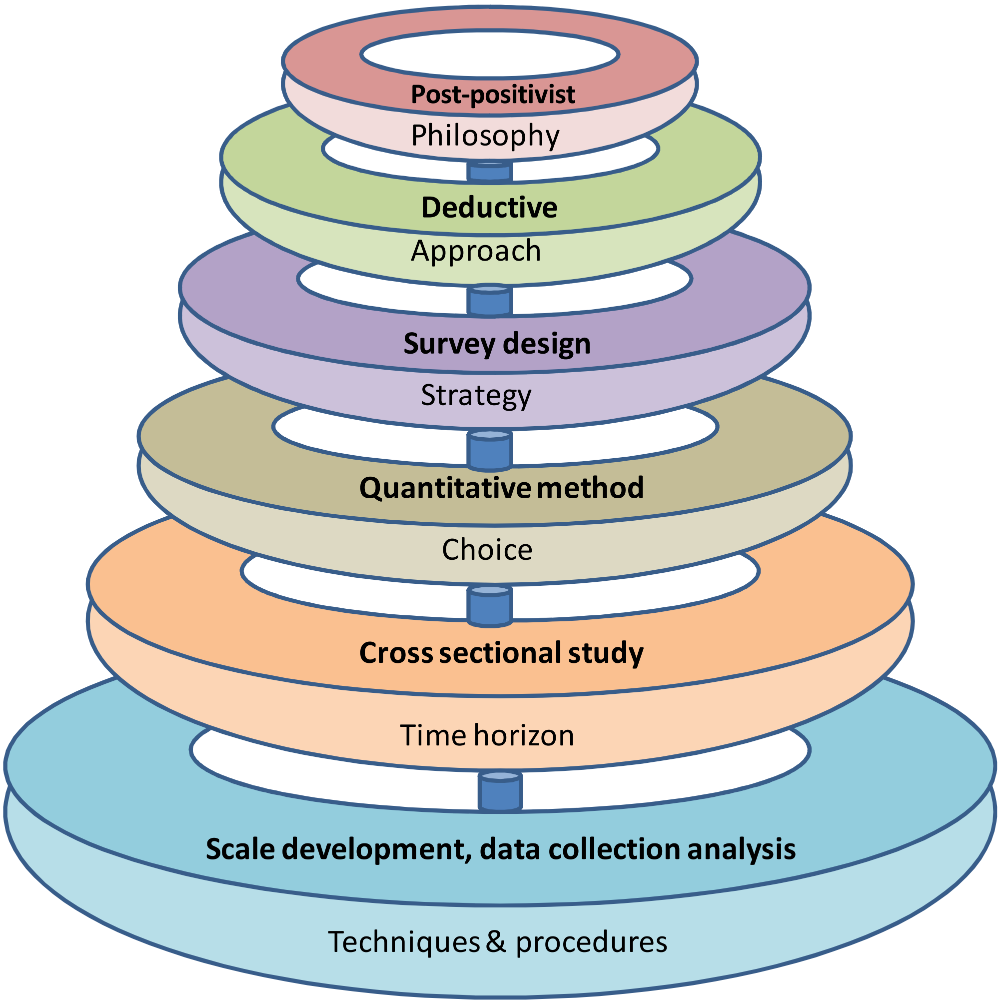
In this chapter...



3.1 INTRODUCTION

This study was conducted in order to develop a multi-dimensional measure whereby psychological ownership of employees could be measured in South African organisations. The purpose of this chapter is to explain the research methodology and strategy that were followed in the study, as displayed in Figure 3.1

Figure 3.1: Research methodology and strategy outline



Source: Adapted from Saunders, Lewis & Thornhill (2007)

3.2 THE THREE WORLDS FRAMEWORK

In order to describe and clarify research problems and other aspects of the logic of research, Mouton (2001) proposes that researchers should make use of a simple structure that he called the Three Worlds Framework. The framework is based on a distinction between three "worlds" in which research can be conducted:

- World 1: The world of everyday life
- World 2: The world of science
- World 3: The world of meta-science

3.2.1 World 1: The world of everyday life

Ordinary people spend most of their lives in World 1 – the ordinary social and physical reality that people exist in. This world consists of multiple worlds and in this world people produce and use knowledge of different kinds referred to as *lay knowledge*. Lay knowledge refers to knowledge that people use in everyday life, which enables them to cope effectively with their daily tasks (pragmatic coping). People apply their lay knowledge to solving problems.

In this study, the questions to be answered with regard to World 1 include the following:

- What is the binding force (apart from money) that keeps an employee with an organisation?
- How do organisations get their employees to stay in their organisations?

3.2.2 World 2: The world of science

In World 2, scientists select phenomena from World 1 and make them into objects of inquiry. The overriding goal of science is to search for the "truth" or "truthful knowledge". Thus, the aim of science is to generate valid and reliable descriptions, models and theories of the world. Mouton (2001) refers to this as the *epistemic imperative*.

Researchers suggest that psychological ownership may be an integral part of the employee's relationship with the organisation. In this study, questions to be answered with regard to World 2 include the following:

- What is meant by the construct of psychological ownership?
- Why it is necessary to measure psychological ownership?
- What motivates psychological ownership?
- What are the constructs that influence and define psychological ownership?
- What are the consequences of psychological ownership?
- Are there any measures of psychological ownership available within the South African context?

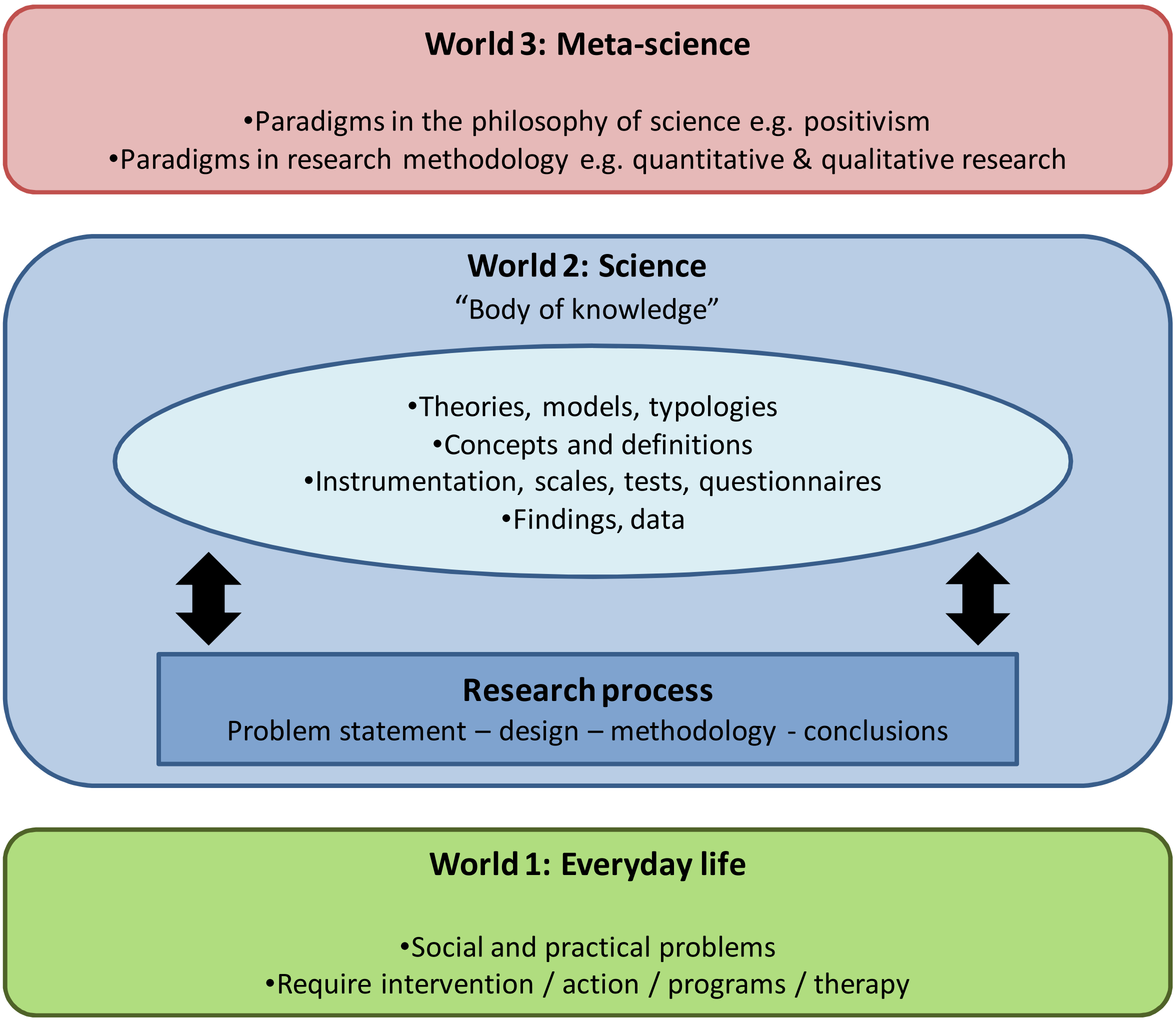
3.2.3 World 3: The world of meta-science

Throughout history scientists have critically reflected on science and scientific practices to develop their various *meta*-disciplines. Scientists have done this by laborious research and reflection to attain truthful and valid research results. Today this is reflected in the disciplines of philosophy and methodology of science, research ethics and the sociology and history of science. In this study the questions to be answered with regard to World 3 include the following:

- If measures of psychological ownership are available, are they comprehensive enough to represent psychological ownership, or is future theory building and research necessary to demonstrate a link between psychological ownership and other related concepts?
- Are these measures of psychological ownership valid and reliable within the South African context?
- If not, what are the steps that are necessary to develop an instrument that will be valid and reliable for the diverse South African context?

The relationship between meta-science, science, and everyday life knowledge is illustrated in Figure 3.2.

Figure 3.2: The relationship between meta-science, science and everyday life knowledge



<u>Source</u>: Mouton (2001, p.140)

3.3 RESEARCH PARADIGM/PHILOSOPHY

Saunders et al. (2007) refer to research philosophy as a term that relates to the development of knowledge and the nature of knowledge, thus the developing of knowledge in a particular field. Creswell (2009) refers to this philosophy/paradigm as a worldview. The researcher's worldview is informed by his or her own assumptions based on their own knowledge, experience and preferences. These elements in combination will influence the philosophy that the researcher will apply in formulating the strategy for a particular research topic.

According to Creswell (2009), there are four different worldviews: post-positivism, constructivism, advocacy/participatory, and pragmatism. The major elements of each position are presented in Table 3.1.

Table 3.1: Four worldviews

Post-positivism	**Constructivism**
• Determination • Reductionism • Empirical observation and measurement • Theory verification	• Understanding • Multiple participant meanings • Social and historical construction • Theory generation
Advocacy/Participatory	**Pragmatism**
• Political • Empowerment issue-orientated • Collaborative • Change-orientated	• Consequences of actions • Problem-orientated • Pluralistic • Real-world practice orientated

Source: Creswell (2009, p. 6)

Positivism is a scientific method that involves the systematic observation and description of phenomena contextualised within a model or theory, the statement of a hypothesis, the implementation of a tightly controlled experimental study, the use of inferential statistics to test the hypothesis, and the interpretation of statistical results in the light of the original theory (Ponterotto, 2005).

The main difference between the positivist and the post-positivist views is that the former stresses "theory verification" and the latter "theory falsification" (Lincoln & Guba, as cited in Ponterotto, 2005, p. 129). Positivists on the one hand believe in an objective, anticipated reality, whereas post-positivists on the other hand admit an objective reality that is only inadequately anticipated. According to post-positivism, human intellectual mechanisms are defective and life's phenomena are basically inflexible, and for this reason, researchers can never fully capture a "true" reality. The primary goal of both a positivist and post-positivist enquiry is a clarification that will ultimately) lead to forecasting and control of phenomena. Positivism and post-positivism form the foundation of quantitative research (Ponterotto, 2005).

Post-positivists are concerned about identifying and assessing the aspects that influence outcomes. Post-positivists base their knowledge on the observation and measurement of the objective reality that exists "out there" in the world. Therefore, it is important for a post-positivist to develop several measures of observations and to study individual behaviour. The world is regulated by specific theories and laws. These theories, however, need to be tested and improved so that people can understand the world. Post-positivists will pursue the following approach to research: firstly, they will start with a theory, secondly they will collect data that either supports or counters the theory, and lastly they will make the necessary adjustments before additional tests are made (Creswell, 2009). The main characteristics of post-positivism are presented in Table 3.2.

Table 3.2: The main characteristics of post-positivism

Definition	Post-positivism is a scientific method that involves systematic observation and description of phenomena contextualised within theory
Ontology (concerns the nature of reality and being)	There is one true reality that is apprehend able, identifiable, and imperfectly measurable (a position known as *critical realism*)
Epistemology (concerned with the relationship between the participant and the researcher)	The researcher, participant and topic are assumed to be independent of one another (dualism), although the researcher may have some influence on that being researched. By following rigorous, standard procedures, the participant and topic can be studied by the researcher without bias (objectivism)
Axiology (concerns the role of researcher values in the scientific process)	The researcher remains emotionally detached from the investigative inquiry – the values, hopes, expectations, and feelings of the researcher have no place in the scientific inquiry. The researcher eliminates or strictly controls any influence he or she might have on the participants or on the research process by using standardised, systematic investigative methods.
Methodology (refers to the process and procedures of the research)	Researchers attempt to simulate, as closely as possible, strict scientific methods and procedures where variables are carefully controlled or manipulated, and where the researcher's emotional stance on the problem under study is irrelevant
Researcher's role	The researcher remains objective, neutral and distant

Source: Adapted from Ponterotto (2005, p. 130-132)

3.3.1 Motivation for choice

The post-positivist paradigm was appropriate to this study because the researcher set out to explore and further define the construct of psychological ownership and was concerned with developing an objective, accurate measure for psychological ownership in South African organisations. To ensure that this instrument was valid, the researcher gathered scientific evidence, which corresponds with the positivist way of thinking. A positivist approach quantifies and measures the observable into scientific evidence. Therefore, the researcher made use of scientific methods that are standardised, valid, and reliable. The aim was to generalise findings from the sample to the population, while being an objective researcher. Quantitative methods were applied and results were presented in an objective manner.

3.4 RESEARCH APPROACH

According to Saunders et al. (2007), there are two approaches that could be followed when conducting research: a deductive or inductive approach. Following the deductive approach, the researcher develops a theory and/or hypothesis and designs a research strategy to test the hypothesis. Following the inductive approach, the researcher would collect data and develop a theory as a result of the data analysis.

3.4.1 Deduction

Deduction involves the development of a theory that is subjected to a thorough test. Deduction is the dominant research approach followed in the natural sciences, where theory presents the basis of explanation, allows the anticipation of phenomena, predicts their occurrence and thus permits them to be controlled (Saunders et al., 2007).

Saunders et al. (2007) list five chronological stages through which deductive research progresses:

(1) A hypothesis (a testable proposition about the relationship between two or more concepts or variables) will be stated from the theory.
(2) An indication will be given of how the concepts are to be measured, which proposes a relationship between two specific concepts or variables.
(3) The hypothesis will be tested.
(4) The specific outcome of the inquiry will be examined (it will either tend to confirm the theory or indicate the need for its amendment).
(5) If necessary, the theory will be adjusted in the light of the findings.

According to Saunders et al. (2007), deduction should possess the following characteristics:

(1) There is a search to explain causal relationships between variables.
(2) Quantitative data is collected.
(3) Controls are applied to ensure the validity of data.
(4) A highly structured methodology is followed to ensure reliability.
(5) The researcher is independent of what is being observed.
(6) Concepts are operationalised in a way that will enable facts to be measured quantitatively.
(7) Samples should be of sufficient numerical size to allow generalisation of conclusions.

3.4.2 Induction

Hinkin (1998) suggests that the induction approach should be followed when it is not easy to identify the conceptual basis of a construct. To get an understanding of the nature of the problem, it is sometimes better for the researcher to ask a sample of respondents to give, for example, descriptions of their feelings about their organisation or to describe some aspect of behaviour. The result of this would be the formulation of theory. According to Saunders et al. (2007), induction should possess the following characteristics:

(1) The researcher tries to gain an understanding of the meanings attached by individuals to certain events.
(2) A less structured approach is followed, which might reveal alternative explanations for the problem at hand.
(3) The approach is likely to be particularly concerned with the context in which the events take place.
(4) A smaller sample may be appropriate.
(5) Researchers collect qualitative data.
(6) Researchers may use a variety of methods to collect data.
(7) Researchers are less concerned with the need to generalise their findings.

The major differences between deductive and inductive approaches to research are summarised in Table 3.3:

Table 3.3: Major differences between deductive and inductive approaches to research

Deduction emphasises	**Induction emphasises**
• Scientific principles • Moving from theory to data • The need to explain causal relationships between variables • The collection of quantitative data • The application of a control to ensure validity of data • The operationalisation of concepts to ensure clarity of definition • A highly structured approach • Researcher independence of what is being researched • The necessity to select samples of sufficient size in order to generalise conclusions	• Gaining an understanding of the meanings humans attach to events • A close understanding of the research context • The collection of qualitative data • A more flexible structure to permit changes of research emphasis as the research progresses • A realisation that the researcher is part of the research process • Less concern with the need to generalise

Source: Adapted from Saunders et al. (2007, p. 120).

3.4.3 Application to this study

Hinkin (1998) suggests that if the theoretical foundation provides adequate information to generate the initial set of items, the deductive approach should be followed. This approach necessitates an understanding of the phenomenon to be investigated and a comprehensive literature review to develop the theoretical definition of the construct under examination. This definition should serve as a guideline for the development of items (Schwab, 1980).

In this study, a thorough review of the literature helped the researcher to explore and further define the construct of psychological ownership in order to develop a multi-dimensional measure of psychological ownership within the diverse South African context.

3.5 DESCRIPTION OF INQUIRY STRATEGY AND BROAD RESEARCH DESIGN

According to Creswell (2009), strategies of inquiry set out unambiguous procedures and guidelines for the research design. Strategies of inquiry are types of qualitative, quantitative, and mixed-methods design that provide specific direction for procedures in research design. A summary of the alternative strategies of inquiry is presented in Table 3.4:

Table 3.4: Alternative strategies of inquiry

Quantitative	Qualitative	Mixed methods
• Experimental designs • Non-experimental designs, such as surveys	• Narrative research • Phenomenology • Ethnographies • Grounded theory studies • Case study	• Sequential • Concurrent • Transformative

<u>Source</u>: Creswell (2009, p. 12)

According to Cresswell (2009), the strategies of inquiry associated with quantitative research are those that invoke the post-positivist worldview, as is the case in this particular study. This study was based on a quantitative research method with a non-experimental research design. Babbie and Mouton (2001, p. 646) refer to quantitative research as the "numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect". The main characteristics of quantitative research, as proposed by Neuman (1997), are as follows:

- The research tests the hypothesis or hypotheses that the researcher starts with.
- Concepts are in the form of well-defined variables.
- Measures are methodically constructed before data is collected and are standardised.
- Data is in the form of numbers from accurate measurement.
- Theory is mainly fundamental and is deductive.
- Standard procedures are followed, and replication is assumed.
- Analysis continues by means of statistics, tables, or diagrams and discussing how what they show relates to the hypothesis.

In this study, a self-administered questionnaire was compiled and data collected by means of distributing the questionnaire as part of a survey research design. A non-experimental, cross-sectional survey design was employed for this particular study.

This research project consisted of two distinct phases, the research design and the data collection. The researcher followed a deductive approach to test whether the theoretical or applied research problem could be supported by empirical measurement and data analysis. The researcher tested whether the variable could represent alternative explanations, and then through statistical analysis eliminated these alternative explanations by measuring them against the control variables. This process is called non-experimental research. On the basis of the theory, the researcher developed a survey questionnaire that was used for measuring psychological ownership. The detailed scale development process will be discussed in more detail below. After the planning phase, data was collected (Neuman, 1997).

This study collected and analysed primary data; thus it represents empirical research. Saunders et al. (2007) describe primary data as data that is compiled for a specific research project being undertaken. In this study empirical data was collected to address the research objectives. Therefore in this study basic research was undertaken to increase the scientific knowledge on psychological ownership experienced by employees in various organisations within the South African context.

In this study, a cross-sectional survey was chosen within the non-experimental design, as measurement occurred at a single time. Individuals were selected to provide a depiction of the overall psychological ownership experienced at a specific time. A cross-sectional method is usually deployed for descriptive studies, as is the case with this study.

The advantage of the cross-sectional design is that it avoids problems related to longitudinal designs, which include being costly and time-consuming, eventually making respondents less interested in taking part in the research. However, disadvantages of the cross-sectional design are that because research is only conducted at one point in time, changes over time are ignored.

A summary of the survey research design that was conducted in this study is presented in Table 3.5.

Table 3.5: Survey research design

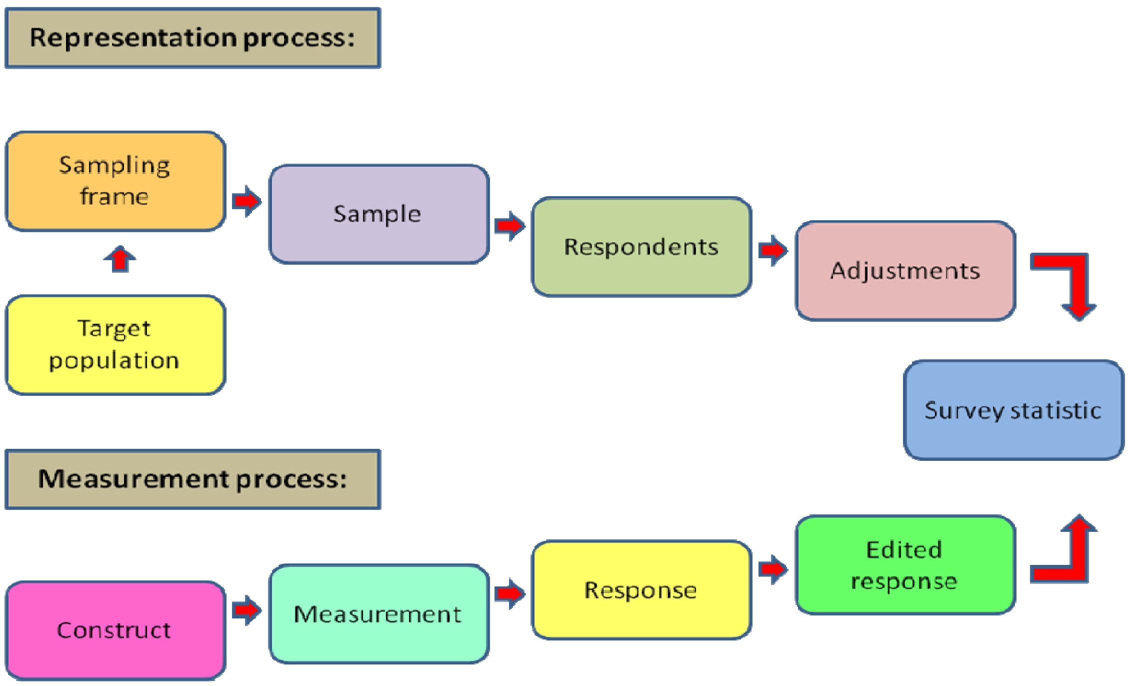
Description	It is quantitative in nature and aims to present a sample that is representative
Design classification	Empirical research was conducted that collects primary data that is numeric and allows the researcher medium control
Key research questions	The research was exploratory and descriptive
Design types	A cross-sectional survey was conducted in that data was only collected at one point in time
Application	The survey was conducted in various organisations within the South African context

Meta-theory	This approach is associated with the post-positivist meta-theory and variable analysis
Conceptualisation	The survey was theory-driven
Sampling	Non-probability sampling was used.
Mode of observation	Data was collected by means of structured electronic questionnaires that were e-mailed or delivered by hand
Analysis	Descriptive and inferential statistics were used
Sources of error	Sampling error; questionnaire error; high refusal rates; high non-response; respondent effects; data capturing error; use of inappropriate statistical techniques

Source: Mouton (2001:152)

Umbach (2005) notes that new technologies, such as the world-wide web and advanced scanning software, have made the use of surveys easy and inexpensive. Umbach views survey methodologies as a survey life cycle process. The survey life cycle is illustrated in Figure 3.3.

Figure 3.3: The survey life cycle



Source: Umbach (2005, p. 92)

In the measurement process, the researcher starts by identifying a construct (in this case psychological ownership) or understanding of what she wants to measure. The researcher then develops measurements (in this case the psychological ownership measure) and considers the concrete ways in which she will gather information about the constructs. Responses are the information provided by the survey participants (in this study, employees from various South African organisations) through the survey measurements. The researcher then edits such responses.

In the representation process, the researcher starts by identifying the target population from which the sample frame (units of analysis) and sample are to be drawn. In this study, employed individuals in various South African organisations were approached. Respondents are those participants from the sample who completed the questionnaires. Post-survey adjustments were made to the respondent data; frequent adjustments made to the data are weighting and imputation of missing data.

3.6 SAMPLING

The full data set that is gathered in the research project is called the research population. However, in order to test or verify the research theory, the researcher utilises a subset of information from the research population called the sample. Various sampling techniques exist in order to ensure that the sample group remains representative and valid for the total population.

3.6.1 Target population, context and unit of analysis

The target population concerns the population from which data will be gathered. The target population from which the sampling frame was chosen in this study consists of a diverse group of professional, high-skilled and skilled individuals employed in both the private and public sector. The reason for choosing employed individuals was that psychological ownership towards their organisation can only develop in individuals who are employed, and the research objective for the proposed study was to develop a measure of psychological ownership. A skilled employee is defined as one who possesses some special skill, knowledge, or ability in his or her work. A skilled worker may have attended a university, college or technical school, or may have learned the skills on the job. A highly skilled worker is any worker who is capable of working efficiently, of exercising substantial independent judgment, of carrying out duties with responsibility and who usually efficiently supervises the work of skilled employees. A professional is defined as a individual who typically possess a large body of knowledge derived from extensive, specialised educational training (usually tertiary), who earns a comfortable salary, who usually exercises autonomy in the workplace, is frequently engaged in intellectually and creative challenging work, and is expected to make use of independent judgement and professional ethics in carrying out his or her responsibilities (Mattes & Richmond, 2000). Various organisations within the private and public sector were approached that allowed for the gathering of a diverse dataset. The criteria for the organisation were that it must have a diverse workforce from a broad range of occupations, and that the individuals should have access to electronic mail.

The unit of analysis for this study was therefore professional, highly-skilled and skilled individuals employed in various types of organisations in both the private and public sector from the South African population.

3.6.2 Sampling methods

There are two types of sampling techniques: probability and non-probability sampling. Probability sampling is a sampling technique in which the chance, or probability, of each case being selected from the population is known and is not zero. Probability sampling is often associated with survey and experiential research strategies. Non-probability sampling is a sampling technique in which the chance or probability of each case being selected is not known (Saunders et al*.,* 2007). The different types of sample are illustrated in Table 3.6.

Table 3.6: Types of sample

Non-probability	**Probability**
Convenient: select anyone who is convenient	*Simple*: select people based on a true random procedure
Quota: select anyone in predetermined groups	*Systematic*: select every *n*th person (Quasi-random)
Purposive: select anyone in a hard-to-find target population	*Stratified*: randomly select people in predetermined groups
Snowball: select people connected to one another	*Cluster*: take multistage random samples in each of several levels

Source: Neuman (1997, p. 205)

This particular study employed non-probability sampling. The reason for this sample type being chosen is that the researcher was not aware of the probability of each case being selected. The specific type of non-probability sampling employed was heterogeneity sampling, which is a type of purposive sampling. Purposive sampling enables the researcher to use her judgment in selecting cases, or it selects cases with a specific purpose in mind (Saunders et al., 2005). In this particular study the purpose was to collect data from a diverse, employed South African sample. Heterogeneity sampling is concerned with getting data from a wide range of diverse individuals and allows the researcher to collect data about the key themes. The population from which data was to be obtained was employed professional, high-skilled and skilled individuals in various organisations in both the private and public sector. The population was accessible to the researcher and the researcher proposed to provide individuals with an equal opportunity to be selected for the survey. Non-probability sampling also has an advantage of convenience and economy. Due to an insufficient response rate, the researcher followed a form of *quota sampling* in that additional respondents were reached to ensure that the sample size was sufficient for quantitative research. The result was to create a somewhat representative sample that would increase the population validity.

3.6.3 Sample size

Hinkin (1998) recommends that the sample used for the subsequent data collection should be of ample size and be representative of the population of interest. According to Tabachnick and Fidell (2007), correlation coefficients tend to be less reliable when estimated from small samples. Correlation of the sample can only be reliably estimated if the sample size is large enough and representative of the research population. However, the size of the research population is a key factor in determining the size of the sample; similarly, the stronger the correlations, the smaller the allowable sample will be. The number of research factors has a similar impact on sample size.

Tabachnick and Fidell (2007) calculate that it is “comforting” to have at least 300 cases for factor analysis. Krzysofiak, Cardy and Newman (1988) suggest that if the initial sample is large enough, the sample could randomly be split in half and parallel analyses for scale development could be conducted, using exploratory and confirmatory factor analysis. According to Hinkin (1998), both exploratory and confirmatory factor analysis are particularly susceptible to sample size effects. It was the intention of the researcher to conduct both exploratory and confirmatory factor analysis as part of the scale development process.

3.6.4 Application to this study

For this study a non-probability convenience sample of 712 was collected from employed professional, high-skilled and skilled individuals in various organisations in both the private and public sector in South Africa. The sample was collected over a three-month period and sample selection came to an end when the researcher arrived at a sufficient sample. According to Tabachnick and Fidell (2007), at least 300 cases would be needed to perform the statistical methods described in this chapter. Following the recommendation made by Tabachnick and Fidell, an arbitrary criterion was set such that persons with 5% or more missing values were not included for further analysis. In a case where fewer than five missing values per respondent occurred, the missing value was replaced with the mean of that particular respondent's responses to other items. Based on a suggestion made by Krzysofiaket al. (1988), the sample was randomly split in half, such that a sample of 356 respondents was used for the Exploratory Factor Analysis (EFA) and a sample of 356 for the Confirmatory Factor Analysis (CFA)

Detailed information about the demographic characteristics of the sample is presented in Table 3.7

Table 3.7: Demographic information on the respondents

Variable	Category	Frequency	Percentage	Cumulative percentage
Gender	Male	287	40.59	40.59
	Female	420	49.41	100
	Total	707		
	Omitted data	5		
Ethnic group	African	226	32.19	32.19
	Coloured	24	3.42	35.61
	Indian	34	4.84	40.45
	White	418	59.54	100
	Total	702		
	Omitted data	10		
Age	Younger than 29	137	19.2	19.1
	30-39	213	28.87	49.07
	40-49	240	33.66	82.73
	50+	122	17.13	100
	Total	712		
Education	Grade 12	60	8.6	8.6
	Diploma	223	31.95	40.55
	Bachelor's degree	135	19.34	59.89
	Postgraduate degree	280	40.11	100
	Total	698		
	Omitted data	14		

Variable	Category	Frequency	Percentage	Cumulative percentage
Sector in which organisation operates	Financial services	103	14.84	14.84
	Chemical/Petroleum	186	26.80	41.64
	Government	172	24.78	66.42
	Manufacturing and production	110	15.86	82.28
	Other	123	17.72	100
	Total	694		
	Omitted data	18		
Operating level in organisation	Operational level	221	31.94	31.94
	Junior management	150	21.68	53.62
	Middle management	201	29.05	82.66
	Senior management level	120	17.34	100
	Total	692		
	Omitted data	20		
Years working in current organisation	Less than 5 years	307	43.67	43.67
	6-10 years	138	19,35	63.02
	11-20 years	163	22.44	85.46
	21+ years	104	14,61	100
	Total	712		
Years working in current job	Less than 5 years	471	66.06	66.06
	6-10 years	102	14.31	80.37
	11-20 years	99	13.88	94.25
	21+ years	40	5.62	100
	Total	712		
Registered at Professional Body	Yes	324	47.58	47.58
	No	357	52.42	100
	Total	681		
	Omitted data	31		

The sample consisted of 40.59% (*n* = 287) males and 49.41% (*n* = 420) females. Of the sample, 59.54% (*n* = 418) were White respondents, 3.42% (*n* = 24) were Coloureds, 4.84% (*n* = 34) were Indian and 32.19% (*n* = 226) were Africans. Of the respondents, 19.2% (*n* = 137) were younger than 29 years of age, 28.87% (*n* = 213) were between 30 and 39 years of age, 33.66% (*n* = 240) were between 40 and 49 years of age and 17.13% (*n* = 122) were above 50. Many respondents in the sample had obtained a postgraduate degree, representing 40.11% (*n* = 280) of the sample. The least represented category was respondents who had obtained Grade 12, only 8.6% (*n* = 60) respondents falling in this category. Employees who had obtained a diploma constituted 31.95% (*n* = 223) of the sample, while employees who obtained a bachelor's degree represented 19.34% (*n* = 135) of the sample.

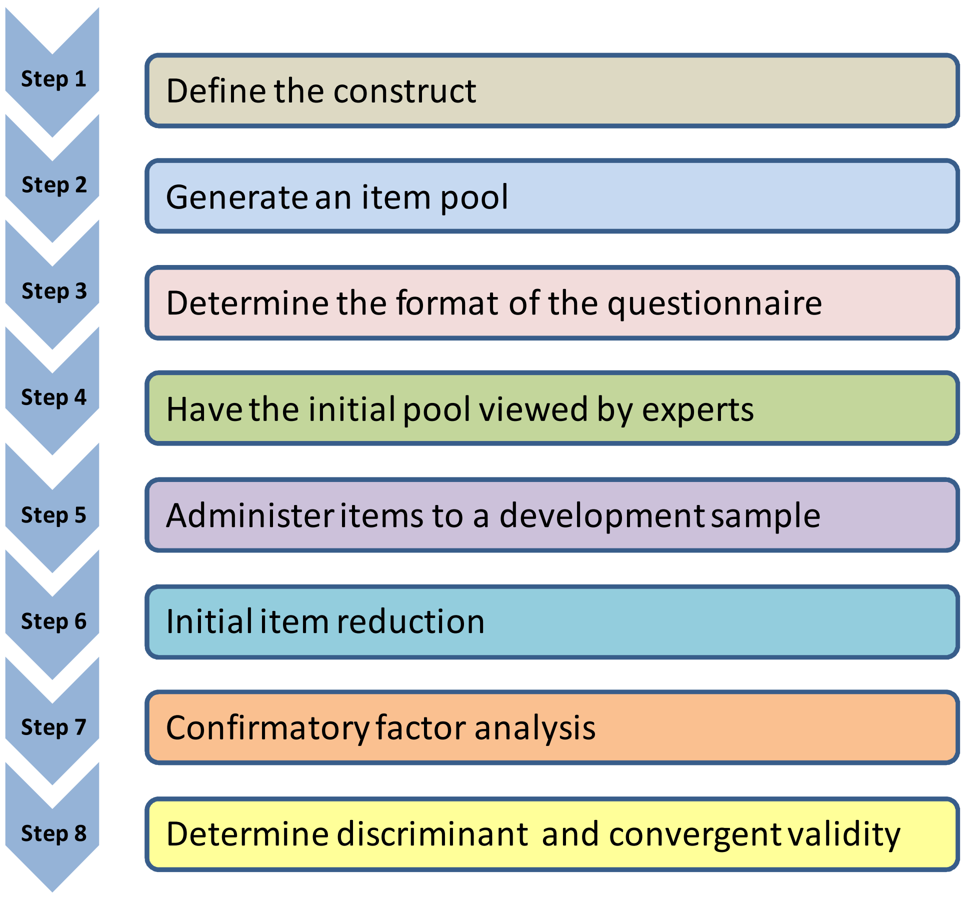
Of the sample, 14.84% (*n* = 103) respondents were employed in the financial service sector, 26.8% (*n* = 186) in the chemical/petroleum industry, 15.86 (*n* = 110) in manufacturing and production, and 17.72% (*n* = 123) in "other" sectors, which include for example professional services, information technology and telecommunications. Of the respondents, 31.94% (*n* = 221) are functioning on operational level, 21.68% (*n* = 150) on junior management level, 29% (*n* = 201) on middle management level, and 17.34% (*n* = 120) on senior management level.

Table 3.17 also indicates that most of the sample respondents, 43.67% (*n* = 307) had been working in their current organisation for less than 5 years, with 19.35% (*n* = 138) having worked between 6 and 10 years, 22.44% (*n* = 163) between 11 and 20 years, while 14.61% (*n* = 104) of the respondents had been working in their current organisation for more than 21 years. The majority of the respondents, 66% (*n* = 471) had worked in their current job for less than 5 years, 14.31% (*n* = 102) between 6 to 10 years, 13.88% (*n* = 99) between 11 and 20 years, and 5.62% (*n* = 40) had worked in their current job for more than 21 years. There were 324 (47.58%) respondents registered with a professional body, but 52.42% (*n* = 357) respondents were not registered with any professional board or body.

3.7 SCALE DEVELOPMENT

As Spector (1992) points out, researchers use various factors to formulate summated rating scales. The factors included in the scales in this study are emotional state, personality, personal needs and job description. According to DeVellis (2003), researchers achieve their goal by quantifying specific phenomena through scale development. In this study, a measure of psychological ownership was developed by using a combination of the steps suggested by DeVellis (2003), Hinkin (1998), and Spector (1992), as indicated in Figure 3.4.

Figure 3.4: Steps in the scale development process



Source: Adapted from DeVellis (2003), Hinkin (1998), and Spector (1992)

3.7.1 Step1: Defining the construct

DeVellis (2003, p. 60) states that clarifying exactly what the scale is intended to measure seems "deceptively obvious." Spector (1992, p. 7) argues that: "This may seem to be a simple-minded requirement, but it is at this step that many scale development efforts go astray". DeVellis adds that theory is a great aid to clarity and recommends that social science theories should always be considered before developing a scale. If there is no theory available to guide the research, a conceptual framework must be developed before developing the scale instrument.

In contrast to this, Du Plessis and Hoole (2006) suggest that a tentative theoretical model be developed with a specific set of defined measures that the researcher has applied his or her mind to. This is not a process of trial and error, but a process wherein the researcher has taken an active view on the number of dimensions, the content domains, the population being researched and applicable environmental factors. The old adage of less is more is proposed, whereby the researcher is firm as to what is within the scope of the research and what is not. DeVellis (2003, p. 62) supports this by stating that "Scale developers should ask themselves if the construct they wish to measure is distinct from other constructs". Therefore, researchers should make sure that the underlying construct is well defined and is focused on the main purpose.

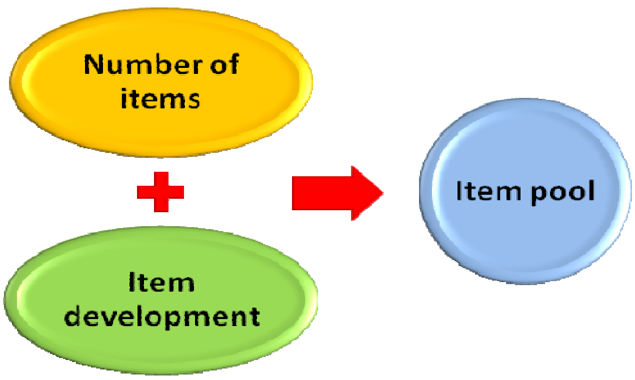
3.7.2 Step 2: Generating an item pool

The key to successful item generation, according to Hinkin (1998), is a well-articulated theoretical underpinning, specifying the content area for the new measure. The aim of the researcher, at this point, is to develop items that will result in measures that sample the theoretical area of interest to reveal content validity. Ghiselli, Campbell, and Zedeck (1981) emphasise that the sample of items drawn from the potential items should adequately represent the construct under examination. DeVellis (2003) supports this by stating that each content area must be well represented in the initial item pool. He also suggests that multiple items will constitute a more reliable test than individual items, but each must still be sensitive to the true score of the latent variable.

Hinkin (1998) suggests that if the theoretical foundation provides enough information to generate the initial set of items, the deductive approach should be followed. This approach requires an understanding of the phenomenon to be investigated and a thorough review of the literature to develop the theoretical definition of the construct under examination. This definition is then used as a guide for the development of items (Schwab, 1980).

Hinkin (1998) discusses several advantages and disadvantages of the deductive approach. An advantage of the deductive approach to scale development is that if properly conducted, it will help to assure content validity in the final scales. Through the development of adequate construct definitions, items should capture the domain of interest. The disadvantages of the deductive approach are that it is very time-consuming and requires that researchers have a working knowledge of the phenomena under investigation. In most situations in which theory does exist, the deductive approach would be most appropriate. The two aspects that need to be taken into consideration when generating the item pool, namely the number of items and the item development process, are displayed in Figure 3.5

Figure 3.5: Generation of the item pool



(Author's own)

- **Number of items**

Worthington and Whittaker (2006) state that scale development studies sometimes include as many as three to four times the number of items that will eventually end up on the instrument. However, Schmitt and Stults (1985) and Schriesheim and Eisenbach (as cited in Hinkin, 1998) state that limiting the number of items in the instrument reduces response bias caused by boredom and fatigue. Additional items result in more analysis and increase the analytics applied to the instrument, while not adding more to scale reliability (Carmines & Zeller, 1979). Harvey, Billings, and Nilan (1995) suggest that at least four items per scale are needed to test the homogeneity of items within each latent construct. Cook, Hepworth, and Warr (1993), however, found that reliable internal consistency can still be achieved with fewer than four items for each latent construct. Churchill (1979) posits that it is essential to ensure that the research domain has been sufficiently sampled, as inadequate sampling is a major source of measurement error. Thurstone (1947) points out that scales should possess simple structure, or parsimony. Not only should any one measure have the simplest possible factor structure, but any scale necessitates the contribution of a minimum number of items that adequately tap the domain of interest. These findings would suggest that the ultimate goal will be the retention of four to six items for most constructs, but the final determination must be made only with accumulated evidence in support of the construct validity of the measure.

- **Item development**

Hinkin (1998) and DeVellis (2003) provide scale developers with a number of guidelines that they should follow in writing items. Statements should be simple and as short as possible. Scale developers should avoid using exceptionally lengthy items, as length usually increases complexity and diminishes clarity. The language used should be familiar to target respondents. Items that determine behaviour and items that determine affective responses should not be combined, therefore all items should be kept uniform in terms of perspective (Harrison and McLaughlin, 1993). Items should focus on a single issue; a "double-barrelled" item may lead to misunderstanding on the part of the respondents because it represents two constructs. Little variance will be generated if items are structured in such a manner that respondents answer the items in the same way, therefore these types of question should be avoided as well. Avoid leading questions, as they may lead to biased responses. The reason for wording items both positively and negatively (reversed-score items) within the same scale is usually to avoid an affirmation, acquiescence or agreement bias (i.e., a respondent's tendency to agree with items, irrespective of their content). Some researchers, such as Price and Mueller (1986) argue that the use of reversed-score items may reduce bias in responses. Others, such as Harrison and McLaughlin (1993), however, have found that the use of a few of these items randomly interspersed within a measure may have a detrimental effect on the psychometric properties of the measure. Reversals in item polarity may be confusing to respondents, so if the researcher does choose to use reversed-score items, they must be very carefully worded, and careful attention should be paid to factor loadings and communalities at the factor analytical stage of scale development (Schriesheim, Eisenbach & Hill, 1989).

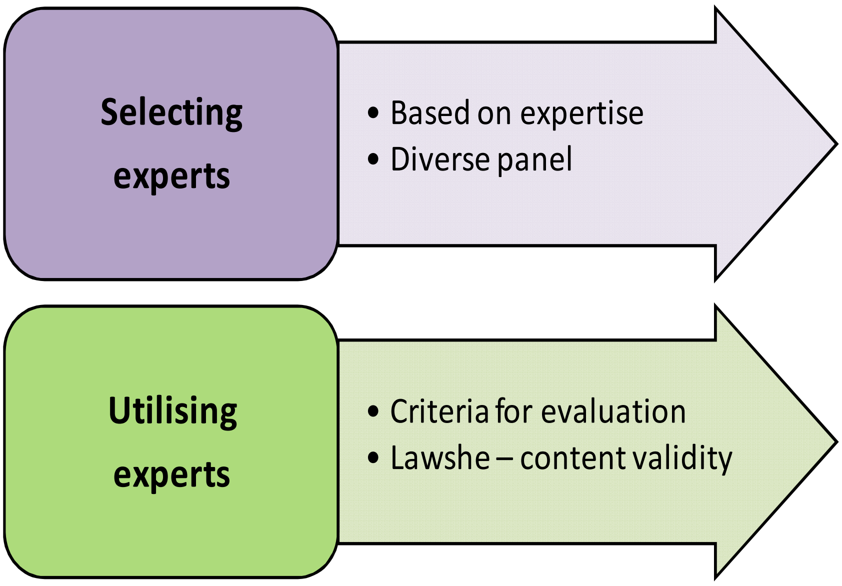
3.7.3 Step 3: Determining the format for measurement

Stone (1978) argues that the effectiveness of statistical analysis is a direct result of the variance among the respondents in a research population. Scaling techniques have been extensively developed by writers such as such as Guttman, Thurstone, and Likert. DeVellis (2003) points out that Likert-type scales are the most frequently used in instruments measuring opinions, beliefs, and attitudes. They are also most suitable for use in factor analysis. Six possible responses are frequently included, such as *strongly disagree, moderately disagree, mildly disagree, mildly agree, moderately agree* and *strongly agree.* The use of this five-point scale increases reliability to a certain point before it starts to level off, according to Lissitz and Green (1975). However, mid-range options can result in respondents' choosing the middle options, whereas equal number options can result in respondents' falling on one side.

3.7.4 Step 4: Having the initial item pool reviewed by a pool of experts

After items have been generated, they should be subjected to an assessment of content validity. The validity of research content is a critical element in the development of instrument measurement, according to Grant and Davis (1997). Hinkin (1998) states that this process serves as a pre-test, permitting the deletion of items that are deemed to be conceptually inconsistent. This judgment quantification process entails asking a specific number of experts to evaluate the validity of items individually, as well as the entire instrument. Grant and Davis state that it is important to report on the characteristics and qualifications of the experts as well as on the process they are asked to use to assess validity. According to them, the soundness of the validation process is significantly influenced by how content experts are chosen and utilised for instrument development. The two aspects that play an important role in this step, namely, how to select the panel of content experts and how to utilise them, are displayed in Figure 3.6.

Figure 3.6: Panel of experts



(Author's own)

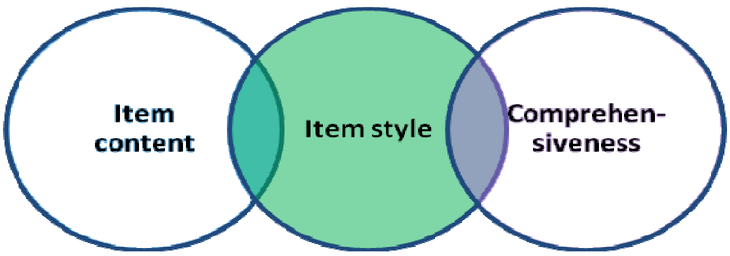
- **Selecting a panel of content experts**

Grant and Davis (1997) stipulate the necessity for content experts to have the relevant training, qualifications, and experience. According to Grant and Kinney (1992), the following could serve as criteria in selecting content experts: publications in refereed journals, national presentations, and research on the phenomenon of interest. Davis (1992) suggests that if a clear theoretical basis for the instrument is provided, panel members might be selected based on their expertise related to the conceptual framework. It is often difficult to find individual content experts who meet all criteria. Therefore, various content experts should be asked to judge the instrument on the different aspects. Authorities differ on the number of content experts needed for a panel. According to Lynn (1986), a minimum of three content experts are needed, but others (Gable & Wolf, 1993; Waltz, Strickland, & Lenz, 1991) recommend from two to twenty panel members.

- **Utilising the panel of experts**

Experts should be provided with the conceptual theoretical basis for the instrument. The relevant dimensions of the constructs to be used in the instrument should be included in the conceptual definitions. Grant and Davis (1997) suggest that the pool of experts validate the initial pool of items in terms of the criteria displayed in Figure 3.7.

Figure 3.7: Expert evaluation criteria



(Author's own)

Item content Experts should judge how representative individual items are of the content area and whether the content area sufficiently measures all dimensions of the construct. Lynn (1986) suggests that content experts should be asked to suggest amendments of items that are not consistent with conceptual definitions or are not representative of the content.

Item style Experts can be asked to evaluate the items' clarity and conciseness. The content of an item may be relevant to the construct, but respondents may provide incorrect data because directions for using the instrument, the items, or response scale are ambiguous. Reviewers could be asked to indicate complex or confusing items and should suggest substitute wordings.

Comprehensiveness Lastly, the panel of experts should be asked to evaluate the total instrument for comprehensiveness. If the instrument is comprehensive, all the dimensions of the desired content area of the concept will be included. In judging the whole instrument, the panel of experts evaluate whether the entire set of instrument items is sufficient to represent the total content area. This will allow the scale developer to identify items needed to be added to the content area, or deleted since they do not present the content area.

Lawshe's (1975) quantitative approach to the content validity of items will be applied in this study. The judgment of content experts is regarded as the highest authority to test the alleged content validity of an instrument.

The content validity ratio (CVR) is an item statistic that is functional in the rejection of specific items from the initial item pool and the calculation of the content validity index (S-CVI/Ave – the mean of the CVR values retained in the test) for the entire item pool. Polit, Beck and Owen (2007) recommend that unless only minor revisions are needed based on the first round results, a second round of expert review should be conducted. A smaller group of experts (3-5) can be used to evaluate the relevance of the revised set of items in the second round. According to Lynn (1986), these second round of experts can be drawn from the same pool as in the first round, or they can also be a new panel.

3.7.5 Step 5: Administering items to a development sample

Items should then be presented to a sample representative of the actual population of interest. The goal should be to analyse how well the items confirm the psychometric property expectations of the newly developed measure. Hinkin (1998) suggests that it is important to examine the relationship between the newly developed scale and other established measures to assess the "nomological network". Therefore, the newly developed instrument should be administered along with other existing measures. Theory should serve as a guideline, indicating those variables which the new measures should correlate with, or which they should be independent of. These other measures will be employed later in successive analyses to provide initial evidence of criterion-related, convergent, and discriminant validity, and thus the construct validity of the new scales.

- **Sample size**

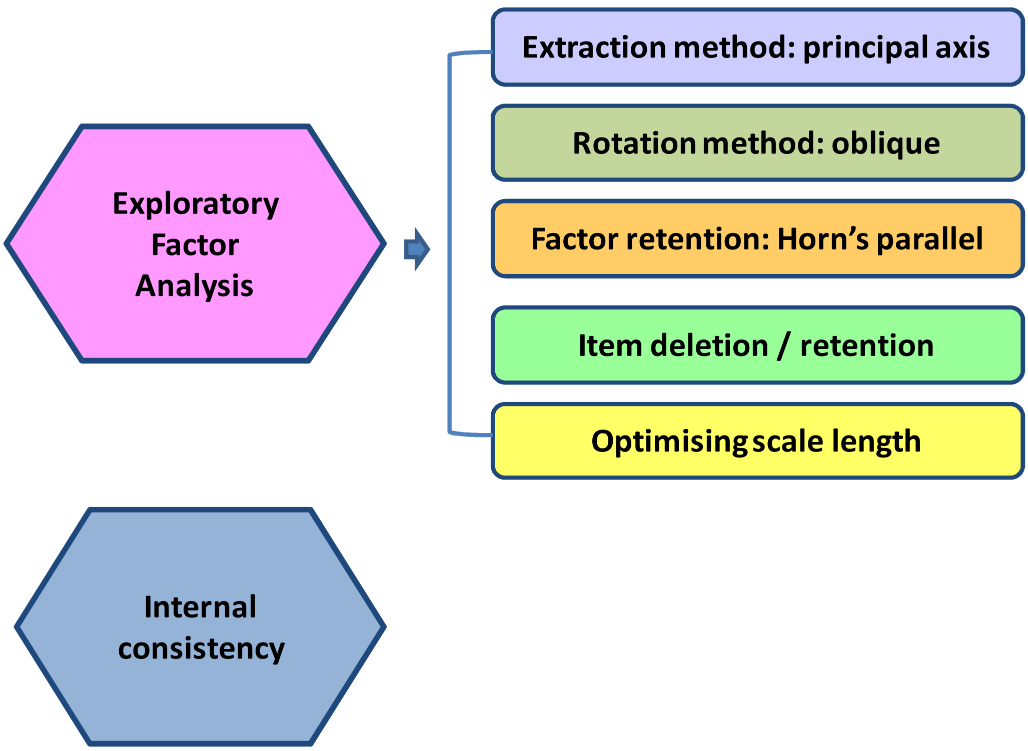
The selection of a suitable sample is crucial, to guarantee enough variance in responses and prevent the effects of an irregular context. Hinkin (1998) recommends that the sample used for the subsequent data collection should be of ample size and be representative of the population of interest. This sample should also be clearly described. According to DeVellis (2003), there are risks involved when using too small a sample. Firstly, patterns of covariance may not be stable, as likelihood may considerably influence correlations between items when the ratio of participants to items is fairly low; and secondly, the development sample may not adequately represent the intended population. Worthington and Whittaker (2006) offer the following guidelines with regard to sample size:

1. Sample sizes of at least 300 are usually sufficient in most situations.
2. Sample sizes between 150 and 200 are expected to be sufficient with data sets containing communalities higher than .50, or with 10:1 items per factor with loadings at around .4.
3. Smaller samples may be sufficient if all communalities are .60 or greater, or with at least 4:1 items per factor and factor loadings greater than .6.
4. Sample sizes of less than 100 or fewer than 3:1 participant-to-item ratios are normally insufficient.

3.7.6 Step 6: Initial item reduction

The two aspects that play an important role in item reduction, namely exploratory factor analysis (EFA) and internal consistency are displayed in Figure 3.8.

Figure 3.8: Initial item reduction



(Author's own)

- **Exploratory factor analysis**

Exploratory factor analysis (EFA) is associated with theory development. According to Worthington and Whittaker (2006), EFA is a technique used to identify a smaller number of factors from a large number of observed variables (or items). EFA is used to assess the construct validity of an instrument during the initial development phase. After the development of an initial set of items, EFA is employed to explore the underlying dimensionality of the item set. This will allow researchers to group a large item set into meaningful subsets that measure different factors. EFA allows items to be related to any of the factors underlying examinee responses. Consequently, this allows the scale developer to identify items that do not measure an expected factor or that simultaneously measure multiple factors, in which case they could be poor indicators of the desired construct and should be eliminated from further consideration (Worthington & Whittaker, 2006).

EFA is particularly susceptible to sample size effects. Bartlett's (1950) test of sphericity can be utilised to estimate that correlations in a matrix are zero. However, according to Tabachnick and Fidell (2007), because of its sensitivity and dependence on sample size, this test is likely to be significant for large samples with relatively small correlations. Therefore, they recommend using this test only if there are fewer than about five cases per variable. Worthington and Whittaker (2006) recommend that researchers provide additional evidence for scale factorability in studies with cases-per-item ratios higher than 5:1.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is also useful for evaluating factorability. This measure of sampling adequacy shows the degree to which a correlation matrix in fact contains factors or simply chance correlations between a small subset of variables. Values of 0.60 and higher are required for good factor analysis (Tabachnick & Fidell, 2007).

Extraction methods The two most commonly known and studied factor-extraction methods are principal-components analysis (PCA) and common-factors analysis (FA). The purpose of PCA is to reduce the number of items while retaining as much of the original item variance as possible. FA assesses the latent factors of constructs that are indicative of the shared variance between items. Therefore, Worthington and Whittaker (2006) suggest that the purpose of FA is more closely aligned with the development of new scales. There are several techniques of FA, including principal-axis factoring, maximum likelihood, image factoring, alpha factoring, and unweighted and generalised least squares. Gorsuch (1997) recommends principal axis factoring.

Criteria for determining rotation method There are two types of FA rotation methods, namely orthogonal and oblique. Orthogonal rotations are used when the set of factors underlying a given item set are assumed or known to be uncorrelated. Oblique rotations are used when the factors are assumed or known to be correlated (Worthington & Whittaker, 2006). Byrne (2005) provides three reasons why oblique rotation is considered preferable to orthogonal rotation. Firstly, in psychological research overall, the constructs and/or their multiple dimensions are expected to be correlated based on theoretical and empirical grounds. As such, a simple structure based on oblique rotation will yield inter-factor relations that are more practical than those from one based on orthogonal rotation. According to Byrne (2005), a factor analytic solution that is based on orthogonal rotation when, in fact, total independence among the factors is unnecessary, will cause estimates that are rigorously deceptive. Secondly, Floyd and Widaman (1995) suggest, given that an optimally derived sample structure shows that the factors are truly orthogonal, an oblique rotated factor solution will still reflect these independent factor relations. Thirdly, as mentioned by Fabrigar, Wegener, MacCullum, and Straham (1999), the information generated by oblique rotations is broader than is the case for orthogonally rotated solutions, given that oblique rotations routinely generate factor correlation estimates. Furthermore, Byrne (2005) notes that evidence of such factor correlations suggest the possibility of one or more higher-order factors. Since factor solutions based on orthogonal rotation do not yield factor correlation estimates, they are unable to reveal the possibility of a hierarchical factor structure.

Criteria for factor retention According to Hinkin (1998), the two components for determining the number of factors to be retained must include underlying theory and qualitative results. The theory will examine the loading on latent factors which will confirm the researcher's expectations. The most widely known approaches for estimating the number of factors for a given item set were recommended by Kaiser (1960) and Catell (1966) on the basis of the eigenvalues, which will determine the importance of a factor and indicate the amount of variance in the entire set of items accounted for by a given factor. Kaiser believed that eigenvalues of less than 1.0 reflect potentially unstable factors, because the variance that each standardised variable contributes to principal component extraction is 1.0. According to Ledesma and Valero-Mora (2007) and Zwick and Velicer (1982), the K1 method proposed by Kaiser (1960) demonstrated a tendency to often overestimate the number of factors.

Cattell's scree test (1966) estimates the correct number of factors to be examined during factor analysis. This is done by using the descending value of eigenvalues to establish a change in the size of eigenvalues prior to horizontal levelling. According to researchers (Hayton, Allen & Scarpello (2004); Ledesma & Valero-Mora (2007); Zwick & Velicer (1982)), the scree test is criticised for its subjectivity and ambiguity, given that there is no objective definition of what the cut-off points between the important and insignificant factors should be. The graph may be difficult to interpret, especially if cases present various drops and possible cut-off points (Ledesma & Valero-Mora, 2007).

Velicer (as cited in Ledesma & Valero-Mora, 2007)), suggests the MAP (Minimum Average Partial) test – a method based on the partial correlations matrices. In this method the EFA concept of "common" factors is employed to determine the number of components to extract. The method seeks out what components are common, and is proposed in general to find the best factor solution, rather than to find the cut-off point for the number of factors. According to Zwick and Velicer (1986), although the MAP has proved to be accurate under many conditions, under certain conditions it seems to reveal a tendency to underestimate the number of factors.

Zwick and Velicer (1986) compared the five different methods (Bartlett's Chi-Square test, Kaiser's greater-than-one rule, Cattell's Scree, Velicer's MAP and Horn's parallel analysis) for determining the number of factors to retain, and came to the conclusion that Horn's parallel method was the most accurate (92% of the time) of the five methods evaluated. Therefore, the researcher decided to employ Horn's parallel analysis (1965) for the purposes of this study. This method randomly orders the participants' item scores and conducts a factor analysis on both the original data set and the random ordered scores. The number of factors to retain is determined by comparing the eigenvalues determined in the original data set and in the randomly ordered data set. A factor should be retained if the original eigenvalue is larger than the eigenvalue from the random data.

Worthington and Whittaker (2006) suggest that researchers should retain a factor only if they can interpret it in a meaningful way, no matter how solid the evidence is for its retention, based on the empirical criteria earlier stated.

The larger the number of items on a factor, the more confidence the researcher will have that it will be a reliable factor in future studies. Tabachnick and Fidell (2007)counsel against retaining factors with fewer than three items. According to them, it is possible to retain a factor with two items only if the items are highly correlated (i.e., $r > .70$) and relatively uncorrelated with other variables.

Criteria for item deletion or retention Parsimony and simple structure are desired for scales; therefore, researchers should only retain those items that clearly load on a single appropriate factor. According to Hinkin (1998), the objective is to identify those items that most clearly represent the content domain of the underlying construct. Hair, Black, Babin, Anderson, and Tatham (2006) suggest that factor loadings of .30 to .40 are minimally acceptable, but that values greater than .50 are generally considered necessary for practical significance. Worthington and Whittaker (2006) suggest that researchers should delete items with factor loadings of less than .32, or items with cross-loadings less than a .15 difference from an item's highest factor loading. In addition, items that contain absolute loadings higher than a certain value (e.g., .32) on two or more factors should also be deleted. Hinkin (1998) suggests that it is also important to report the percentage of the total variance that is explained; the larger the percentage the better. According to Hayton et al. (2004), as many common factors as possible should be kept to explain at least 50% of the variance in the data set. At this stage, items with inappropriate loadings were deleted, and the analysis repeated, until a clear factor structure matrix that explained a high percentage of total item variance was obtained.

Optimising scale length According to DeVellis (2003), a scale's alpha is influenced by two characteristics, namely the extent of covariation among the items and the number of items in the scale. Longer scales tend to be more reliable, but Converse and Presser (1986) recommend that questionnaires take no longer than 50 minutes to complete. Worthington and Whittaker (2006) notice that scales that take longer than about 15 to 30 minutes to complete might become problematic, depending on the respondents, the intended use of the scale, and the respondents' motivation regarding the purpose of administration. Shorter scales are good in the sense that they place less of a burden on respondents. Maximising one of these assets reduces the other. Scale developers must try to optimise the length of subscales by deleting items that have low factor loadings or high cross-loadings. Similarly, items that contribute little to internal consistency should also be removed. The challenge in scale development optimisation is not to degrade the quality of the factor structure, inter-correlations, loadings and cross-loadings. A final EFA must be performed to confirm factor solution has not changed due to the deleted items.

- **Internal consistency assessment**

Kerlinger (1986) defines reliability as the accuracy or precision of a measuring instrument that is a necessary condition for validity. Gerbing and Anderson (1988) suggest that the reliability of a measure should be assessed after unidimensionality has been established. There are a number of ways of calculating reliability, but the most appropriate accepted measure in field studies is internal consistency reliability using Cronbach's alpha (Price & Mueller, 1986). Cortina (1993) recommends the use of this statistic, especially when used in conjunction with factor analysis. Hair et al. (2006) argue that items with an alpha correlation of .70 and higher are viewed as acceptable, but indicate that alpha correlations of .60 are also acceptable in exploratory research. Hinkin (1998) suggests that since the unidimensionality has been confirmed, the researcher can remove items that do not contribute to, or devalue the reliability of the developed scales. The sensitivity of alpha to the number of items in the measure, according to Cortina, will remain high, regardless of the low inter-correlation and multi-dimensionality. In contrast to Hair et al., Cortina suggests that an alpha correlation of .70 should serve as an absolute minimum for newly developed measures, and that through appropriate use of factor analysis, the internal consistency reliability should be considerably higher than .70. According to Hinkin, reporting of internal consistency should be considered necessary.

3.7.7 Step 7: Confirmatory factor analysis

Confirmatory factor analysis (CFA) is often used during the scale development process to help support the validity of a scale following an EFA. A CFA should be performed in the first place to confirm that observed variables sort themselves into factors corresponding to the latent variables and in the second place to examine in general the quality of the solution and the specific factor loadings that represent the measurement model (Kelloway, 1998). Hinkin (1998) recommends that CFA from an independent sample should be conducted using the item variance-covariance matrix computed from data collected from the independent sample. Krzysofiak et al. (1988) suggest that if the initial sample was large enough, the sample could randomly be split in half and parallel analyses for scale development could be conducted, using exploratory and confirmatory factor analysis. The purpose of the analysis is twofold. Firstly, the analysis assesses the goodness of fit of the measurement model, comparing a single common factor model with a multi-trait model with the number of factors equal to the number of constructs in the new measure (Jöreskrog & Sörbom, 1989). The multi-trait model restricts each item to load only on its appropriate factor. Secondly, the analysis examines the fit of individual items within the specified model using the modification indices and *t* values.

The chi-square statistic allows the assessment of fit of a specific model as well as the comparison between two models. The smaller the chi-square, the better the model fit. Carmines and McIver (1981) suggest that an acceptable chi-square is two to three times as large as the degrees of freedom, but the fit is considered better the closer the chi-square value is to the degrees of freedom for a model. However, Ullman (2001) suggests that two or less reflects good fit, while Kline (1998) notes that three or less is acceptable. A non-significant chi-square is desired. This will indicate that differences between the model-implied variance and covariance and the observed variance and covariance are small enough to be due to sampling fluctuation. Chi-square is very sensitive to sample size, with the result that a model with a large chi-square may still have a good fit if the fit indices are high. Therefore, the chi-square statistic must be used with caution, and other multiple-fit indices should be used to assess a model's goodness-of-fit. Bentler (2007) proposes that the standardised root mean square residual (SRMR) should be reported, accompanied by at most two other indices of fit, such as the comparative fit index (CFI).

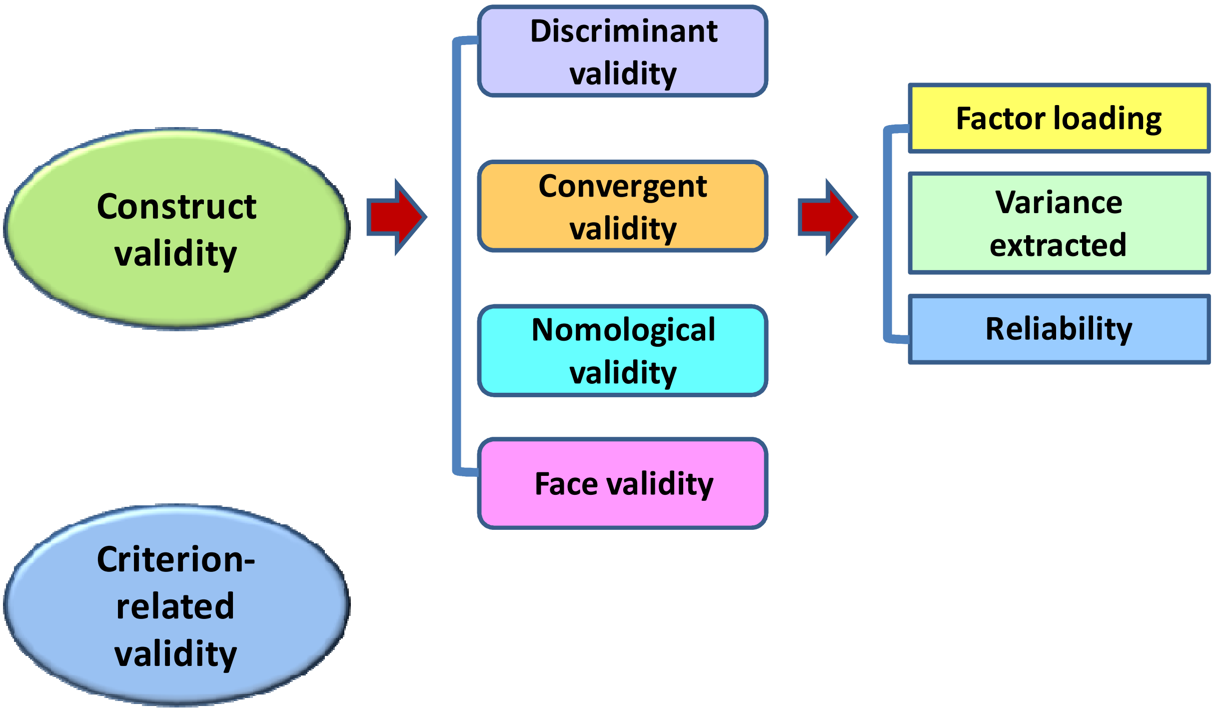
Values for the SRMR less than .10 are generally indicative of acceptable fit. Values for the root mean square error of approximation (RMSEA) at or less than .05 indicate close model fit, which is customarily considered acceptable (Hu & Bentler, 1999). Brown and Cudeck's (1993) general guideline for RMSEA values states that a result of .05 and smaller confirms that the model developed and the research data have an acceptable correlation. The CFI values should be equal to or greater than 0.95, according to Hu and Bentler (1999), for the model to be accepted. RMSEA and CFI are less sensitive to sample size; however, SRMR tends to be lower due to larger sample size (Garson, 2002).

In summary, CFA allows the researcher to quantitatively assess the quality of the factor structure, providing further evidence of the construct validity of the new measure. Results should include at the minimum the chi-square value and the associated degrees of freedom, and the recommended goodness-of-fit indices used for each competing model. One of the biggest advantages of CFA is its ability to assess the construct validity of a proposed measurement theory. The CFA must not only provide acceptable fit, but must also show evidence of construct validity (Hair et al., 2006).

3.7.8 Step 8: Convergent and discriminant validity

Construct validity indicates the degree to which a set of measured items truly represents the theoretical latent construct those items are designed to measure. According to Hair et al. (2006), construct validity is made up of four important components, namely convergent validity, discriminant validity, nomological validity and face validity, which are displayed in Figure 3.9.

Figure 3.9: Construct validity and criterion-related validity



(Author's own)

Convergent validity The items that are indicators of a specific construct should share a high proportion of variance, known as convergent validity. The relative amount of convergent validity among item measures can be estimated by studying the factor loadings, the amount of variance extracted and reliability.

According to Garson (2002), it is important to take the size of the factor loading into account. In the case of high convergent validity, high loadings would indicate that the items meet at some common point. Standardised loadings should be .5 or higher, ideally .7 or higher.

With CFA, the average percentage of variance extracted (VE) among a set of construct items is a summary indicator of convergence (Fornell & Larcker, 1981). VE should be .5 or greater to suggest adequate convergent validity.

Reliability is also an indicator of convergent validity. In this study, the most common form of internal consistency reliability, Cronbach's alpha, was used to ascertain the reliability, as recommended by Gregory (2004). Reliability should be .7 or higher to indicate adequate convergence or internal consistency. Reliability between .6 and .7 may be acceptable, provided that other indicators of a model's construct validity are good (Hair et al., 2006). In order to improve the reliability coefficients of the analysis and to prevent cancelling out variables with positive and negative loadings, a reversal of original negative items was done.

Discriminant validity Discriminant validity is the extent to which a construct is truly distinct from other constructs. In this particular study, discriminant validity was determined by comparing the variance-extracted percentages for any two constructs with the square of the correlation estimate between those two constructs. If the variance extracted estimates were greater than the squared correlation estimate (Fornell & Larcker, 1981), discriminant validity was illustrated.

Face validity Face validity is an evaluation of the extent of equivalence between the items selected to form a summated scale and their conceptual definition (Hair et al., 2006). According to Hair et al., when CFA is done, face validity must be established prior to any theoretical testing. A measurement theory cannot be correctly specified without a clear understanding of each item's content or meaning. This content validity was already established in step 4 of the scale development process, where the items were subjected to the scrutiny of a pool of experts.

Nomological and criterion-related validity According to Hair et al. (2006, p. 138), nomological validity "determines whether the scale demonstrates the relationships shown to exist based on theory or prior research". Cronbach and Meehl (1955) suggest that the relationships between the newly developed scale and other measures with which they could be hypothesised to relate to develop a nomological network should be examined and criterion-related validity should be established. These relationships should be based on existing theory and may be examined using correlation or regression analysis. Evidence of criterion-related validity will be provided if the hypothesised relationships attain statistical significance. In this study, psychological ownership will be correlated with organisational commitment, job satisfaction and intention to stay. Empirical evidence for a positive relationship between psychological ownership and job satisfaction were found by Buchko (1993), Vande Walle et al. (1995), Van Dyne and Pierce (2004), Mayhew et al. (2007) and Avey et al. (2009). O'Driscoll et al. (2006), Mayhew et al. (2007) and Avey et al. (2009) have confirmed a strong association between affective organisational commitment and psychological ownership of the organisation. Avey et al. (2009) found a positive relationship between psychological ownership and employees' intention to stay in the organisation.

Construct validity To further investigate evidence of construct validity independent sample *t*-tests and the analysis of variance (ANOVA) technique were conducted to assess whether employees varying in biographical variables (age, gender, ethnic group, education, the sector in which their organisation operates, level in the organisation and registration with a professional board) differ significantly with regard to their feelings of psychological ownership for the organisation, as well as with regard to the specific dimensions underlying the concept of psychological ownership. Independent sample *t*-tests were used to test whether significant differences exist between the means of two groups, and where several independent variables were compared, the analysis of variance (ANOVA) technique was used.

Thompson (as cited in Vacha-Haase & Thompson, 2004, p. 473) states that statistical testing cannot evaluate result importance, although statistical significance "evaluates the probability or likelihood of the sample results, given the sample size, and assuming that the sample came from a population in which the null hypothesis is exactly true". Lately, more emphasis has been placed on the reporting of effect-sizes. According to Vacha-Haase & Thompson (2004), the fifth edition of the *Publication Manual* of the American Psychological Association, 2001, accentuates the fact that researchers should provide readers not only with information about statistical difference but should also give adequate information to calculate the size of the observed effect of the relationship. Therefore, it is important to report effect-sizes, which will give an indication of the practical significance of study results.

In this study, effect-sizes were measured by calculating Cohen's *d* (1992) for two independent groups, and the most widely-used effect-size measure for ANOVA, partial eta-squared, was calculated. Cohen's *d* (1992) is defined as the difference between the means divided by the standard deviation of either group. Partial eta-squared is described by Pierce, Block and Aguinis (2004) as the proportion of total variation in the dependent variable accounted for by the variance between groups formed by the independent variables. Partial eta-squared values range from 0 to 1. Cohen (1992) provided the following regression benchmark for effect-sizes: d = .20 is a minimal solution and therefore has a small effect; d = .50 is a medium effect; anything equal to or greater than .80 is a large effect-size.

In sum, according to Cronbach and Meehl (1955), the demonstration of construct validity of a measure is the ultimate objective of the scale development. Therefore, attempts to demonstrate convergent, discriminant, nomological, face, and criterion-related validity should be sufficiently and clearly reported.

3.8 DATA COLLECTION

Data was obtained by means of self-administered questionnaires. According to Babbie and Mouton (2001, p. 646), a questionnaire is a document that contains "questions and other types of items designed to solicit information appropriate to analysis". After the compilation of a measure for psychological ownership, it was electronically distributed to the participants via the organisation's intranet or e-mailed to employees. In some instances, questionnaires were administered on hard copy to maximise the response rate. This survey method of collecting data was deemed appropriate because it is relatively inexpensive, it is not too time consuming, it eliminates the need for assistants, and data entry is automated (Saunders et al., 2005). An additional benefit of electronically self-administered questionnaires is that respondents can easily be reached via electronic communication, regardless of their geographical position, and they can be sent, completed and returned immediately. This electronically self-administered questionnaire allowed respondents to complete the questionnaire anonymously. The questionnaire was developed with Lime Survey software and respondents were provided with a web-link to complete the questionnaire via the internet. The completed questionnaire was saved on the web-provider server without their e-mail addresses, thus ensuring respondents' anonymity. Anonymity will increase participants' honesty, and the researchers' bias based on the respondents' personal information will be eliminated. The features of the electronically self-administered questionnaire are described in Table 3.8.

Table 3.8: Features of an electronic self-administered questionnaire

Feature	Internet and intranet questionnaire
Population's characteristics for which suitable	Computer-literate individuals who can be contacted by e-mail or intranet
Confidence that right person has responded	High if using e-mail
Probability of contamination or distortion of respondent's answer	Low
Size of sample	Large, can be geographically dispersed
Feasible length of questionnaire	Fewer 'screens' seems to be better
Likely response rate	Variable, 30% reasonable within organisations/via intranet
Suitable question types	Closed, but not too complex questions that are of interest to the respondent
Completion time	2-6 weeks from distribution
Financial resource implications	Low costs
Role of interviewers or field workers	None
Data input	Usually automated

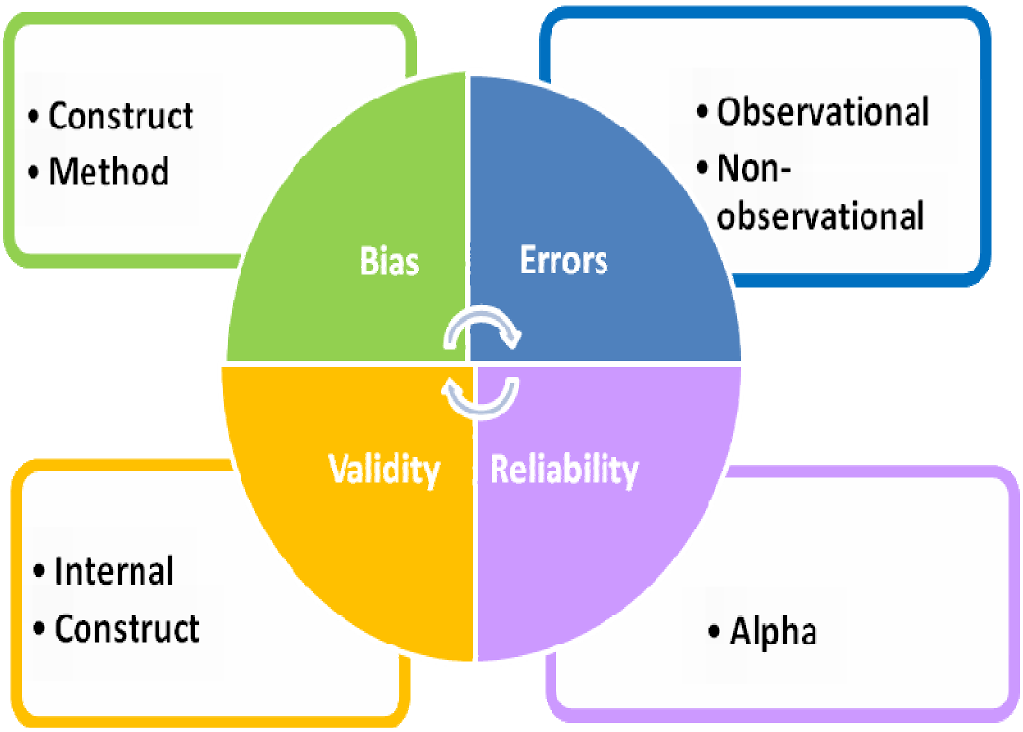
Source: Saunders et al. (2005, p. 358)

Some disadvantages of this approach are that the researcher is not available to implement quality control relating to the answering of questions or the quality of the responses presented. Other potential limitations are that the questionnaire should be carefully designed, that the asking of open questions is usually not practical and that the possibility of prompting and exploring issues in further detail is not possible (Saunders et al., 2005).

3.9 ASSESSING AND DEMONSTRATING THE QUALITY AND RIGOUR OF THE RESEARCH DESIGN

The research process should be correctly described and transparent, to demonstrate that rigour is present and that the research results are accurate and that they can be generalised. This section will subsequently describe the various aspects of bias, errors in human inquiry, validity and reliability that were taken into consideration when conducting the research, as illustrated in Figure 3.10.

Figure 3.10: Rigour of the research design



(Author's own)

3.9.1 Bias

Bias is described as any form of influence, condition or set of conditions that separately or mutually influence the credibility of data (Van de Vijver & Tanzer, 1996). Bias can occur unnoticed and attacks the truthfulness of the results and findings of the research study. Three kinds of bias can be distinguished, namely: *construct bias, method bias* and *item bias*.

- **Construct bias**

According to Van de Vijver and Leung (1997), when the construct does not have the same meaning across cultural groups and when only a partial overlap in the definitions of the constructs across cultures exists, construct bias will probably be present. Construct bias can also occur due to poor sampling of all relevant behaviours associated with the construct. Bias in terms of construct validity probably exists when an instrument measures different attributes for one group than for the other, or when the instrument is measuring the same attribute but with different degrees of accuracy. The employment of factor analysis, followed by target rotation and an evaluation of the factorial agreement across the samples, could determine the presence of construct bias in this study. Tucker's coefficient of agreement (Tucker's phi) is most frequently employed where values larger than.90 are often taken to indicate equivalent factors.

- **Method bias**

There are three types of method bias that can occur: *sample, administration* and *instrument bias*.

Sample bias According to Van de Vijver and Tanzer (1996), sample bias occurs when the samples used differ in a variety of relevant characteristics other than the target constructs. Sample bias may be caused by a lack of comparability of samples or differential stimulus familiarity.

Administration bias All source of bias that is caused by the form of administration is called administration bias (Van de Vijver & Tanzer, 2004). Different situations can lead to administration bias, including differences in physical conditions, social conditions, environment administration conditions, differential expertise of administrators, ambiguous instructions for respondents, communication problems between the respondent and the tester, and halo effects of the tester. The researcher tried here to limit administration bias by using electronically distributed questionnaires. The researcher ensured that all the instructions were clear and concise. Respondents were asked for their consent (the consent form can be viewed in Annexure B). The questionnaire was set in such a manner that the respondents should respond to all the items.

Instrument bias Predictive validity is one of the most crucial forms of validity in relation to instrument bias**.** According to Van de Vijver and Tanzer (2004), instrument bias may be caused by differential awareness of stimulus material, response procedures and differential response styles. The researcher aimed to limit instrument bias by consulting a diverse group of experts.

3.9.2 Errors in human inquiry

Groves (as cited in Umbach, 2005) classifies errors into two general categories: observational and non-observational. Observational errors occur in the measurement process of the survey life cycle and are deviations in a respondent's answers from the true values on a measure; they include measurement error and processing error. Non-observational errors occur in the representation process of the survey life cycle and arise when survey researchers do not take measurements on the part of a population. There are four types of non-observational error in survey research, namely coverage, sampling, non-response and adjustment error. The types of error that occur are summarised in Table 3.9.

Table 3.9: Error in survey research

Type of error	Definition	Suggestions for reduction
Observational error		
Measurement error	When a respondent's answer to a survey question is inaccurate, imprecise, or cannot be compared in any useful way to other respondents' answers. Often the result of poor wording of questions and construction of questionnaires	• Have a defined objective • Pay attention to question wording • Evaluate the survey questions by consulting with experts and pre-testing the questionnaire
Processing error	Is introduced after the data has been collected and prior to the analysis. Most common ones are coding, data entry, and outliers	• Avoid open-ended questions • Search for outliers
Non-observational error		
Coverage error	Occurs when the sampling frame does not match the population because some members of the population did not have a chance to be included in the sample	• Maintain good data • Take careful consideration in the mode of selection
Sampling error	Arises when the sample does not match the sample frame. Is present in every sample survey because statistics calculated on the survey data are only a subset of the population	• Determine an appropriate sample size • Report confidence intervals
Non-response error	*Unit non-response*: when a member of the sample does not respond to the survey *Item non-response*: when the respondent does not answer one or more survey questions	• Use multiple contacts • Draw a sample • Know the population • Ensure respondents match the population • Keep the questionnaire short
Adjustment error	Arises from efforts to reduce errors of non-observation (coverage, sampling, and non-response). Similar to processing error	• Consider the use of weights • Be thoughtful in the handling of missing data

Source: Adapted from Umbach (2005, p. 94)

Umbach (2005) notes that no survey study is error free therefore researchers must be conscious of the possibility of error and bias at every stage of the survey life cycle and consider ways to reduce and eliminate them without incurring large costs. The researcher in this study aimed at increasing the credibility and integrity of the study at each step of the research process.

It is also necessary for the researcher to consider aspects of validity and reliability pertaining to the study, to analyse and determine the appropriateness of the proposed research design and methods. This will be discussed in the subsections to follow.

3.9.3 Validity

Validity is defined as the extent to which an empirical measure accurately reflects the concept it is intended to measure. The researcher needs to verify to what extent the development measurement reflects the concepts being measured. It is not an exact science but rather a measure of relative validity. A summary of the different types of validity is given in Table 3.10.

Table 3.10: Different types of validity

Type of validity	**Definition**	**How determined**
Construct validity	"Extent to which a set of measured items actually represents the theoretical latent construct those items are designed to measure" p.707	Estimated by looking at convergent validity, discriminant validity, nomological validity and face validity.
Convergent validity	"Extent to which a set of measured variables actually represents the theoretical latent construct those variables are designed to measure" p. 137	Estimated by looking at the factor loadings, the amount of variance extracted and reliability.
Discriminant validity	"Extent to which a construct is truly distinct from other constructs" p. 138	Comparing the variance-extracted percentages for any two constructs with the square of the correlation estimate between these two constructs. The variance extracted estimates should be greater than the squared correlation estimate.
Nomological validity	""Determines whether the scale demonstrates the relationships shown to exist based on theory or prior research" p. 138	Looking at the correlations between the factor scores for each construct. Constructs should positively relate to one another.
Face or content validity	"Extent to which the content of the items is consistent with the construct definition" p. 771	A specific number of experts should be asked to evaluate the validity of items individually as well as the entire instrument.
Criterion-related validity	Examines the relationship between existing measures and the newly developed scales. Theory should dictate those	Evidence of criterion-related validity will be provided if the hypothesised relationships attain statistical significance.

Type of validity	Definition	How determined
	variables with which the new measures should correlate or be independent	
External validity	Concerned with the generalisability of results	Generating similar relationships with different samples

Source: Hair et al. (2006)

- **Internal validity**

Threats to internal validity for the proposed study could include the following: history, selection bias, and other third variable problems (Creswell, 2009). The aforementioned threats to internal validity are defined and described in Table 3.11, along with the researcher's proposed counter actions according to each threat.

Table 3.11: Threats to internal validity

Type of threat	Description	Counter actions
History	Events may affect the dependent variable. Participants could be influenced by uncontrolled variables e.g., respondents present themselves better	The rationale of the study were explained and the fact that there are no right or wrong answers was emphasised
Selection bias	Participants with certain characteristics are selected which influence results and findings	A diverse pool of participants was selected
Third variable problems	Nuisance/intervening variables that influenced the proposed research	The researcher tried to determine the possible external third variables that might influence results and considered action plans to establish control over such variables

Source: Creswell (2005, p. 163)

- **Construct validity**

According to Hair et al. (2006), construct validity is the extent to which a set of measured items actually represents the theoretical latent construct those items are designed to measure. Threats that have been identified that may have a possible influence on the study include the subject effect and the experimenter effect (Welman, Kruger & Mitchell, 2005).

The subject effect Participants may act differently and change their behaviour because they are aware that the researcher is assessing their perceptions and experiences, in this case with regard to psychological ownership. In this study, to counter the subject effect the researcher emphasised that the participants should be as honest as possible, and assured them that their responses were anonymous and would be treated as confidential.

The experimenter effect The researcher verified formulated expectation of the research results by testing the hypothesis through data analysis, manipulation and interpretation. The researcher counteracted the experimenter effect by taking cognisance of the influence of personal characteristics, by selecting objective data collection and data analysis methods, and by keeping in mind the research objectives of obtaining diverse data and providing an objective picture of the data and results of the study.

3.9.4 Reliability

Kerlinger (1986) defines reliability as the accuracy or precision of a measuring instrument, which is a necessary condition for validity. In this study reliability was determined by means of Cronbach's alpha (Price & Mueller, 1986). As has been mentioned, Hair et al. recommend that items with an alpha correlation of .70 and higher are viewed as acceptable, but indicate that alpha correlations of .60 are also acceptable in exploratory research. According to Cortina (1993), alpha is very sensitive to the number of items in a measure, and alpha can be high in spite of low item inter-correlations and multidimensionality. As already mentioned, in contrast to Hair et al., Cortina suggests that an alpha correlation of .70 should serve as an absolute minimum for newly developed measures, and that through appropriate use of factor analysis, the internal consistency reliability should be considerably higher than .70. According to Hinkin (1998), reporting of internal consistency should be considered absolute necessary and therefore this was determined in the study.

3.10 ETHICAL CONSIDERATIONS

Approval from the University of Pretoria's ethical committee was sought before the commencement of the study. Saunders et al. (2007) define research ethics as the researcher's appropriateness in terms of his or her behaviour regarding the rights of human beings affected by the work of the researcher. For a research study to be ethical, the research design should be methodologically sound and morally acceptable for the people who are involved.

Researchers should thoroughly consider the ethical implications of the study they propose to conduct, especially when the focus of investigation of the research involve human beings. In this particular study respondents were asked to fill out a questionnaire that measured their psychological ownership toward their organisation.

The Health Professions Council of South Africa, along with the Professional Board of Psychology, requires that any research conducted should be in compliance with and guided by the "Ethical Code of Professional Conduct" (Babbie & Mouton, 2001). This "Ethical Code of Professional Conduct" specifies what is deemed acceptable and unacceptable conduct, from the research planning phase through to the publication of research findings.

Based on the International Research Test Commission's Guidelines for Test Use (2000), fair assessment practices could be defined as entailing:

- The appropriate, fair, professional, and ethical use of assessment measures and assessment results
- Taking into account the needs and rights of those involved in the assessment process
- Ensuring that the assessment conducted closely matches the purpose to which the assessment results will be put

- Taking into account the broader social, cultural, and political context in which assessment is used and the ways in which such factors might affect assessment results, their interpretation and the use to which they are put (Foxcroft & Roodt, 2001).

There are several very important ethical issues in the research that were attended to in the study, displayed in Figure 3.11. Each issue will be discussed in greater detail.

Figure 3.11: Ethical issues in the research

(Author's own)

- **Informed consent** According to Babbie and Mouton (2001), the anticipated consequences, rights and responsibilities as well as the nature and purposes of the research should be communicated as fully as possible to the individuals likely to be affected. Applicants should be totally informed about all the features of the research project that might influence their decision to participate. Informed consent can be ensured by telling participants what the researcher wishes them to do and asking them for their written permission (See Appendix A).

- **Anonymity** Anonymity refers to the principle that the identity of a participant should be kept secret. Participants have the right to remain anonymous. Neither the names nor any identifiable background information of participants may be disclosed. Information on (in this case) their psychological ownership is of a private nature and not for public information and should be respected and regarded as such by ensuring that the information provided by the participants cannot be personally linked to them (Mouton, 2001). Anonymity was ensured due to the fact that respondents completed the questionnaire via the internet using a web-link. The competed questionnaire was saved on the web provider server without respondents' revealing their e-mail addresses.

- **Confidentiality** Researchers are responsible for protecting the security and confidentiality of obtained information. The researcher should not discuss or share any personal information related to the participants without their consent. If conducting survey research, the researcher should ensure that the data collected will be used only for the stated purposes of the specific study (Mouton, 2001). (See Annexure B)

- **Privacy and voluntary participation** Participation in survey research should be voluntary and participants may refuse to reveal certain information about themselves and may have the opportunity to withdraw from the research study at any time. The researcher made every effort to reiterate that participation was entirely voluntary and that participants had the right to withdraw their consent at any time (See the cover letter in Annexure B).

- **Accountability** Researchers may be held accountable for the manner in which survey data is used and interpreted as well as for protecting the confidentiality and security of obtained information (Voskuijl & Evers, 2007).

- **Harm** It will be unethical for researchers to expose participants to unnecessary physical or psychological harm. According to Babbie and Mouton (2001), researchers need to take cognisance of the impact their research will have on participants so as not to expose them to any unreasonable risks and harm to their emotional well-being. In this study a pilot test was done with a small group of participants with whom the experience of participating in the research was discussed, and changes were made to the questionnaire to limit potential unreasonable risks. Participants also had the opportunity to communicate via email with the researcher if they had any concerns resulting from their participation in the research. The researcher also minimised the possibility of harm to participants by explaining that this was not a test with any right or wrong answers, and that no judgements would be made about them as individuals, nor could the results in any way be linked to them.

- **Research conducted in a socially responsive and responsible manner** The researcher ensured that participants were treated in a socially responsive and responsible manner by consulting numerous published journal articles and taking note of how research in similar contexts had been conducted. Voskuijl and Evers (2007) advise researchers to treat participants with respect and consideration, to acknowledge them as persons in specific contexts with specific needs, to protect them from possible negative consequences of the research, and to demand of them only to produce relevant and reasonable information.

- **Plagiarism** Plagiarism is a concern that needs to be prevented at all cost and researchers have to ensure all references have been properly documented and listed throughout the research report.

- **Ethical reporting** It is the responsibility of the researcher not to falsify, distort or leave out any findings. The researcher attempted to report results in an honest and accurate manner. Results that contradicted previous research or which were in conflict with predominant literature were reported and linked to the body of literature.

3.11 CONCLUSION

In this study a post-positivist paradigm was followed because the researcher explored and further defined the construct of psychological ownership and was concerned with the development of an objective, accurate measure of psychological ownership in South African organisations. The study was based on a quantitative research method with a non-experimental, cross-sectional survey design. The research group represented a non-probability convenient sample consisting of 712 professional, highly-skilled and skilled respondents employed in various types of South African organisations in both the private and public sectors.

The chapter set out the various steps to be followed in scale development. These steps have been followed and will be explained in detail in the next chapter; they resulted in a questionnaire that could measure the psychological ownership of employees in organisations. Data was collected by means of an electronically self-administered questionnaire; in some cases hard copies were employed. The chapter described the various aspects of bias, errors in human inquiry, validity and reliability that were taken into consideration when conducting the research to demonstrate that rigour was present. It described several important ethical issues that are applicable to the study. The following chapter will describe the application of the chosen methods and the resultant statistical findings.