UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DETERMINING THE CONSTRUCT VALIDITY OF UDAI PAREEK'S LOCUS OF CONTROL INVENTORY

by

CORNè ENGELBRECHT

Submitted in partial fulfilment of the requirements for the degree

MASTER OF COMMERCII

(INDUSTRIAL PSYCHOLOGY)

in the

FACULTY OF ECONOMIC AND MANAGEMENT SCIENCES

at the

UNIVERSITY OF PRETORIA

Study leader: Mr M.A. Buys

PRETORIA **OCTOBER 2009**

DECLARATION

I Corne Engelbrecht declare that the dissertation that is titled "*Determining the construct validity of Udai Pareek's locus of control inventory*", which I herewith submit for the degree Magister Commercii in Industrial Psychology at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

All the resources used for this study are cited in the text and referred to in the reference list by means of a comprehensive referencing system.

NAME: Cornè Engelbrecht **DATE: October 2009**

____

Signature of student

ABSTRACT

Psychometric instruments are used internationally in organisations to assist companies when making recruitment or development decisions. But the differences that exist on similar psychometric instruments between international groups of people with similar qualifications necessitate the evaluation of the construct validity of these instruments. The objective of this study was to determine the construct validity of the locus of control inventory developed by Udai Pareek in1998.

The instrument consists of 30 items and it was completed by 155 pre-and postgraduate students at the University of Pretoria. Oblimin rotation was used to assist with the interpretation of the factors. In the pattern and structure matrix the highest loadings were highlighted which meant that these items measured the construct that they were supposed to measure.

Items were deleted where there was no clear indication as to what the item was measuring. Item analysis was done on each of the constructs identified to further investigate the appropriateness of each item and even more items were deleted, which also had an impact on the Cronbach's alpha value. The chi-square as well as other fit indices was used to determine the model fit. Confirmatory factor analysis was conducted to establish how well the model fitted the data.

It was clear that the two-factor model fitted the data considerably better than the three-factor model because of a high correlation between two of the factors which indicated that they might be measuring the same construct. The results indicate that the locus of control instrument is not suitable for use in the South African context and should therefore be modified until a satisfactory model fit is found.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

OPSOMMING

Psigometrise instrumente word wereldwyd benut om maatsappye te help met werwing en ontwikkelings besluite. Maar die verskille wat ontstaan wanneer internasionale groepe met soortgelyke kwalifikasies teen mekaar opgeweeg word vereis dat die konstruk geldigheid van die instrumente bepaal word. Die doel van die studie was om die konstruk geldigheid van die lokus van kontrole vraelys, wat deur Udai Pareek ontwikkel is, te ondersoek.

Die instrument bestaan uit 30 items en dit was voltooi deur 155 voor-en nagraadse studente van die Universiteit van Pretoria. Oblimin rotasie was gebruik om die interpretasie van die faktore te vergemaklik. Die items wat werklik meet wat dit ontwikkel is om te meet is ingekleur sodat dit maklik geidentifiseer kan word.

Items is slegs uit die vraelys gehaal waar daar geen duidelike indikasie was wat die item meet nie. Item analise was op elke konstruk toegepas, en die chi-square en ander passings indekse was gebruik om te bepaal hoe die model op die data pas.

Dit was duidelik dat die twee-faktor model die data baie beter pas as die voorgestelde drie-faktor model. Volgens die drie faktor model is daar twee van die voorgestelde drie faktore wat moontlik die selfde konstruk kan meet. Die resultate dui aan dat die lokus van kontrole vraelys nie geskik is vir gebruik in Suid–Afrika nie, en sal daarom aangepas moet word totdat ‘n aanvaarbare model gevind is.

ACKNOWLEDGEMENTS

I would like to extend a special thanks to the following people, who supported me during this strenuous academic year, especially in completing this research project:

- God who gave me the strength and opportunity to complete this study to the best of my ability.
- My Parents who always support me in everything that I do.
- Mr. M.A. Buys for his help and guidance.
- My friends and family for their encouragement and support.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TABLE OF CONTENTS

Chapter 3: The validation of international psychological instruments for application in South Africa

Chapter 4: Steps followed in this study to determine the construct validity of the locus of control inventory

Chapter 5: Methodology

Chapter 6: Results

Chapter 7: Findings and recommendations

LIST OF TABLES

LIST OF FIGURES

APPENDICES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

It has become common practice for psychologists to predict behaviour with the help of psychometric instruments. In certain working situations and now more than ever, this is also used for employee selection during the employment process. But what is mostly overlooked is that such a situation requires ethical responsibilities, which include the psychological instruments being shown to be valid and reliable for their specific use.

The study focused on the construct validity of the locus of control inventory by Udai Pareek (1998). The main reason for the study was to determine if the instrument measures the specific constructs that it is supposed to measure. The instrument was originally developed in India, but the results of the present study can also be used to determine if this instrument can be applied to people in the South African context.

Psychometric instruments are used all over the world in organisations to assist in making recruitment and development decisions which have a significant impact on the future of current employees (Taylor & Radford, 1986; Van der Merwe, 1999). Therefore, it is imperative to determine if the instruments are valid to be used in other countries than the country in which it was developed. Industrial psychologists specialise in the prediction of employee behaviour in the workplace. But the significant mean difference that exists on the same tests between different international groups with similar qualifications brings into question the role of construct validity with regard to psychometrical instruments which are internationally available.

To solve this problem, the construct validity of the specific instruments should be ascertained and published for different international groups. This will ensure that all instruments used will measure what they are supposed to measure in each group.

1.2 PROBLEM STATEMENT

As every human resource manager knows, placing the right person in the right position is a very critical issue for most organisations in South Africa and abroad. Peterson in (Taylor & Radford, 1986) explain that fairness in testing situations is related to a concern for ensuring equality of job opportunities for applicants. According to Cascio (1998), it will be seen as a violation of the Employment Equity Act if psychometric instruments are used in an unfair manner.

According to Cascio (1998) to be a fair instrument it is important that the test battery is designed to discriminate between candidates with higher and lower abilities on certain criteria, and not according to race, gender or age. The instrument was developed in India and the researcher wanted to determine if the test can also be used in South Africa by looking at its construct validity in the South African context.

Most psychometric tests are validated for western cultures, which are mostly based on white middle-class employees. The fact that more and more instruments are being developed all over the world and shared on the internet, can cause a problem because most instruments have not been shown to be valid and reliable in the countries where they are used. Baron and Kenny (1986) state that culture affects behaviour and accordingly the psychological constructs that are being measured. This leads to the important question of the construct validity of psychometric instruments when utilised in different countries.

According to Reynolds (1982), bias in the construct validity of a test exists when it can be indicated that a test measures other psychological constructs in one group than that in another group, or when it measures a similar construct, but with different levels of accuracy. Owen and Taljaard (1996) maintain that construct bias means that the test measures something different in one group from that in another group, even though it assumes that the same construct is being measured. For a test to be unbiased, all the items included in the test must measure the same construct for all the subgroups.

Construct validity of psychometrical instruments is important because the instruments are used in the selection processes of organisations in different countries to identify the best person matching the specific job requirements, which will be measured by certain constructs, and therefore someone who will be able to perform successfully in the company.

In practice, when trying to indicate that a selection tool is valid, the researcher will try to prove its predictive validity because this indicates that the instrument has the ability to predict future performance. Because Pareek's instrument was validated for people living in India, the researcher first had to determine if the locus of control instrument measures the constructs that it was designed to measure in the South African context. Thereafter it would be possible to determine if the instrument can effectively be used in South Africa as opposed to where it was developed.

1.3 RESEARCH OBJECTIVES

The objective of this study was to determine the construct validity of the locus of control inventory to be used in the context of South Africa (Pareek, 1998).

The main objective involved the following:

- to determine what is meant by locus of control;
- to understand why it is necessary to measure locus of control;
- to understand the importance of reliability and validity as statistical concepts;
- to evaluate the locus of control inventory as a valid and reliable instrument especially focusing on the construct validity;

- to understand the meaning and application of construct validity;
- to determine the steps that should be followed when assessing the construct validity of an instrument.

1.4 SIGNIFICANCE OF THE STUDY

In the field of industrial psychology, locus of control is seen as a crucial aspect of personality. The concept was developed initially by Julian Rotter in the 1950s (Rotter, 1966). *Locus of control* is a concept that indicates an individual's perception about the underlying key causes of events in his/her life.

Rotter's (1966) view is that behaviour is mostly led by reinforcements and that through events such as rewards and punishments, individuals come to cling to certain beliefs about why they act the way they do. These beliefs will then direct them to the kinds of attitudes and behaviours they will adopt. This description of locus of control is coherent with the views of Philip Zimbardo, a famous psychologist namely locus of control orientation is a belief about whether the reasons of one's actions are dependent on what one does (internal control orientation) or on events outside one's personal control (Zimbardo, 1985, p. 275).

Thus, *locus of control* refers to a unidimensional continuum, which stretches from external to internal. The fundamental question is whether an internal or external locus of control is sought after. Usually, it seems to be psychologically healthy to feel that one has control over the things which one is able to change (Zimbardo, 1985). In basic terms, an internal locus of control is mostly seen as sought after. Having an Internal locus of control can also be referred to as 'self-agency', 'personal control', 'self-determination', etc. Research has found the following trends:

- males seem to be more internal than females;
- older people tend to become more internal;
- the senior managers in organisations tend to be more internal (Mamlin, Harris, & Case, 2001).

Psychological research has found that people with an internal locus of control are usually better off. These people are more success focused and they get the higher-paying jobs (Mamlin, Harris & Case, 2001). However, thoughts concerning causality have to be remembered here to. Do environmental circumstances cause locus of control beliefs or do the beliefs cause the situation?

Sometimes locus of control is seen as a constant, underlying personality construct, but this may be deceptive, since the theory and research indicate that locus of control can for the most part be learnt. There are some facts which also include the notion that at least to some extent, locus of control is a reaction to circumstances. Some psychological and educational interventions have been found to produce a movement towards internal locus of control (Hans, 2000; Hattie, Marsh, Neill & Richards, 1997).

The study will provide the business community with insight into the importance of construct validity not only for the locus of control questionnaire but also for ensuring that all psychometrical tests are validated before they are used in South African organisations. The study aimed at expanding the knowledge of human resource / industrial psychology practitioners with regard to the validation of psychometric instruments for the South African workplace. Human resource managers have an enormous responsibility with the selection and employment of the most suitable candidates without discriminating unfairly against certain groups. Thus the ultimate aim of the study, form the human resources / industrial psychology perspective, is to ensure that the importance of determining the construct validity of psychometrical instruments is known and understood.

Furthermore, the study should stimulate further research to both explore and empirically establish the construct validity of all psychometrical instruments that are currently in use and also those that are being developed in the near future. The study will firstly focus on the theory, which includes concepts and theories related to the study as well as the findings from research that was previously conducted. The importance and types of reliability and validity that are applicable to the study will also be discussed. Next, the concept *construct validity* as well as steps to determine construct validity will be described in detail because of the vital role that it will play in the interpretation of the results. The methodology will be discussed with regard to the research approach, research design, sampling used as well as the locus of control instrument. Finally, the results will be presented, and a conclusion will be made.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

There are different attitudes regarding the way rewards and outcomes are ascertained. Some people believe that one can neither predict nor influence fundamental events, whereas others believe that one can do both (Pareek, 1982). Issues like this and specifically those relating to prediction and causes of social and personal matters have intrigued philosophers, industrial psychologists and organisational development practitioners alike. According to Pareek (1982), it is therefore necessary that there is an instrument which is able to distinguish between those who have an external locus of control and those who have an internal locus of control. Locus of control will be discussed extensively in the following paragraphs.

People are continually seeking causes for their own and other people's behaviour. According to Schepers (2005), the causes of particular behaviours are called attributions. He explains that the attributions that people make, as well as their understanding thereof, determine their perceptions of the social world to an enormous extent. They can ask questions like: is this a friendly or a threatening world? Do we they have control over certain events through our their abilities or are their lives controlled by influential people?

Schepers (2005) points out that the causes of human behaviour can be divided into two categories. The two categories consist of dispositional and situational causes. Dispositional causes relate to one's natural disposition while situational causes relate to the external world and this includes all the environmental factors (Roediger III, Capaldi, Paris & Polivy, 1991). Theories that are related to locus of control will now be discussed.

2.2 DEFINING LOCUS OF CONTROL

The concept *locus of control* is based on the extent to which people perceive certain events to affect outcomes (Rotter, 1945). Rotter (1945) states that individuals who have low perceptions of such events have an internal locus of control: they believe that their own actions produce outcomes. Those who have high perceptions of the reason for certain events are characterised by an external locus of control: they believe that outcomes are the result of fate rather than the result of their own actions.

The locus of control construct first received attention when Rotter (1966) published his assessment instrument of an individual's generalised expectancies for internal versus external control of reward. The Rotter instrument, said to measure locus of control, was developed on the basis of the social learning theory (Anastasi, 1990, Schepers, 1995). The theory focuses on the important role that reward, respect and gratification play in determining behaviour.

Rotter (1966) proposes the following definition of the concept of *internal-external locus of control*: "When a reinforcement is identified by the subject as following some action of his own but not being completely contingent upon his own action, then, it is characteristically identified as the result of luck, chance, fate, as under the control of an influential other, or as unpredictable because of the great intricacy of the forces surrounding him. When an individual evaluates the event in this way it can be labelled that this is a reaction that can be identified as external control. If the person identifies that the event is in control of his own behaviour or his own stable characteristics, it can be identified as a belief in internal control."

O'Brian (2004) defines *locus of control* as an individual difference that is focused on the degree that individuals ascribe responsibility for outcomes, both positive and negative, to either themselves or to an external cause. It is linked to job outputs, such as job satisfaction and job performance (Spector, 1982) and has been researched in terms of organisational citizenship behaviour and counter productive work behaviour (Paulhus, 1983).

The concept *locus of control* is compared with various economical concepts and organisational behaviour by assorted researchers (Dubinsky, Skinner & Whittler, 1989; Kalechstein & Nowicki, 1994; Kren, 1992). These concepts include: motivation, job satisfaction, rewards and achievement. Studies that examine the relationship between locus of control and work performance (Aijzen & Madden,1986; Kalechstein & Nowicki,1994; Witt, 1988) show that individuals with an internal locus of control usually perform better than those who have an external locus of control. These studies were mostly done on white employees and should therefore be validated before they are used on other cultures.

2.3 CONSTRUCTS MEASURED BY LOCUS OF CONTROL INVENTORY

Internal and external loci of control are represented by the terms *internality* and *externality* respectively. Similarly, people with high internality are called internals: those with high externality, externals (Rotter, 1945). Internality is related to effectiveness and adjustment. When compared with externals, internals have been reported to be more sensitive to new information, more observant, more likely to attend to cues that help resolve uncertainties (Lefcourt & Wine, 1969) and are prone to both intentional and incidental learning (Wolk & DuCette, 1984).

The association of internality with various aspects of learning seems to make good sense. For example, in order to influence or control outcomes, the person with an internal approach must acquire as much information as possible and then process that information as quickly as possible. Evidence supports the assumption that an internal locus of control leads to academic achievement (Crandall, Katkovsky & Crandall, 1965; Harrison, 1968; Lessing, 1969).

Some studies show a high and positive correlation between internality and perseverance, which is defined by extra time spent on work, proceeding involvement in difficult and complex tasks, and willingness to postpone gratification (Franklin, 1963; Mischel, 1966). Lefcourt (1976) summarises the research on the relationship between internality and postponed gratification as follows:

"Involvement in long term goals requires deferment of satisfaction and tenacity in effort requires undivided attention, which is not possible unless the temptation of instant satisfaction is resisted."

Because internals believe that their efforts lead to pleasing results, they can rely on their own understanding and predictability. In contrast, externals - perceiving a lack of personal predictability and being concerned that unforeseen external factors will influence outcomes - may find it more attractive to seek immediate satisfaction than to try to achieve long-term goals (Lefcourt, 1976). Internality was found to be an essential characteristic of people with enhanced achievement motivation (McClelland, 1961). It was further reported that internal locus of control ensures moderate risk-taking, and one study showed that the correlation between achievement motivation and preference for moderate risk was fundamental and positive among internals but almost zero among externals (Wolk & DuCette, 1984).

A study by Mitchell, Smyser and Wood (1975) reveals relationships between internality and certain organisational attitudes and behaviours. For example, internals underwent greater job satisfaction than externals. Internals also preferred a participatory management style, whereas externals preferred a directive style. Even more comparisons indicated that internals considered that hard work was more likely to lead to rewards and that internals had more control over the ways they worked (Mitchell, Smyser & Wood, 1975). The use of rewards, respect and expertise was seen by internally focused supervisors as the most powerful way to influence subordinates, those with an external orientation saw coercion and their formal positions as most effective.

There is some evidence that externals experience higher role stress, especially those who believe that most things are restrained by powerful others. When forty women entrepreneurs completed the Levenson instrument, Surti (1982) found positive correlations between external others and the following role stressors: inter-role distance, role overload, result inadequacy, resource inadequacy, role inadequacy and total entrepreneurial role stress. When Keshote (1989) used the locus of control inventory of Rotter with 212 managers in engineering firms, negative correlations were found between both external others and external chance.

The sum of these findings suggests that internality plays a significant role in human development and purposeful living. Nevertheless, there are also negative consequences for the internal. Those who perceive their own abilities and actions as exclusively responsible for their failures are likely to feel that they have a lot of stress and may become self-punitive (Mitchell, Smyser & Wood, 1975). Ascription of failure or negative conditions to external factors can help people cope with harmful experiences more effectively, perceive social reality in a proper manner and remedy unwanted situations.

Mueller and Thomas (2001) state that in individualistic cultures there is an enhanced likelihood of an internal locus of control. They also found reinforcement for their hypothesis that internal locus of control together with innovativeness is more likely in individualistic low uncertainty avoidance cultures than in collectivistic, high uncertainty avoidance cultures.

Research indicates that negative affectivity and locus of control are likely to distort people's self–reports on a broad range of variables, this could range from work characteristics to well-being and coping behaviours (Parkes, 1991; Siu, Spector, Cooper & Lu Yu, 2002; Spector, 1988). People with elevated levels of negativity, who perhaps already feel depressed or nervous, may disregard the extent to which they do agree with their negative thoughts; likewise, people with an external locus of control may undervalue the degree to which they are able to react in certain situations, especially in the face of unwanted internal events. Such cognitive distortions that centre around people who are underestimating their abilities, and which result from unhelpful personality characteristics, have previously been recognised in psychopathology literature (Beck, Rush, Shaw & Emery, 1979; Beck & Freeman, 1990)

Autonomy is a construct that is closely related to internal control. It can be defined as the "tendency to attempt to be effective in the environment, to impose one's wishes and designs on it" (Wolman, 1973). It is expected that people high on autonomy would look for situations where they can take control of situations that offer possibilities of change. These people would also accept the challenge of solving complicated problems, would take the lead in most situations and they would also prefer working alone.

In summary, internal managers tend to have higher role efficacy, they experience less role stress, and they use problem-solving methods when feeling stressed and when dealing with conflict. These managers will use their motivational behaviour more effectively and use more persuasive bases of power in working with their employees (Keshote, 1989). Externals seem to do the opposite and it was found that they have lower interpersonal trust. Coercive power is an integral part of their managing processes when dealing with employees (Pareek, 1982; Pareek, 1998)

2.4 THEORIES RELATED TO LOCUS OF CONTROL

2.4.1 Social learning theory

The *locus of control* construct was used in the social learning theory developed by Rotter (1966). This theory introduces the word *expectancy* and focuses on three other general classes of variables, namely behaviours, reinforcements and psychological situations. Rotter (1975) ascribes a main role to expectancy, which is one's belief or subjective judgement that, in certain situations, a specific behaviour may lead to reinforcement. He adds that no individual understands an event or situation the same way. For one person, a situation might look rewarding whereas other individuals might interpret the same event completely differently (Hall & Lindsey, 1985)

Anastasi (1990) and Schepers (1995) affirm the above by stating that the social learning theory stresses that reinforcement, respect and gratification play a critical role in determining behaviour. Schepers (1995) describes the perception of locus of control according to the social learning theory as the way in which reinforcement takes place in the social environment, and the effect this has on future behaviour. Rotter, Chance and Phares (1972) also state that the social learning theory explains that rewarding behaviour could lead to a higher expectation that certain behaviours or an event will take place in the future.

When reinforcement follows on behaviour but is not the consequence thereof, it will be seen as the result of something else, which can include fate, the influence of others, something that is not predictable. It could be described as external locus of control. If the person decides that the event took place as a reaction of his/her own behaviour, skills or personal characteristics, this will be defined as internal locus of control (Lefcourt, 1966; Phares, 1976).

2.4.2 Attribution theory

Attribution theory has to do with how individuals take and use information out of the social environment to use as reasons for certain things that happened and to assess the behaviour of people (Collins, 1974). According to Heider (1958), there are three basic principles that are closely related to the attribution theory:

- Individuals try to find the reasons for their own as well as for other individuals' behaviour.
- Individuals give causal explanations for behaviour in a systematic way.
- These attributions have an effect on an individual's behaviour and interactions with others in future.

Schepers (2005) agrees with Collins (1974) that the attribution theory has to do with the way in which a person gathers information about the stable or unchangeable characteristics of others – their motives, intentions and traits – as well as those of the external world.

Phares (1976) is of the opinion that a person that has an internal locus of control attributes his/her achievements to things that are in his/her control while an individual who has an external locus of control attributes his/her achievements to things that are out of his control. Phares (1976) also states that ongoing achievements are dependent on the links that an individual makes to achievements and events. The reasons, according to Phares (1976), why individuals with an internal locus of control are more achievement orientated than individuals with an external locus of control are because a high level of achievement motivation promotes the trust in the individual's own skills and the reasons of events.

2.4.3 Self-determination theory

Since locus of control defines the degree to which an individual ascribes responsibility for a result to either him/herself or to his/her environment, environmental support will possibly moderate the association of locus of control to both organisational citizenship behaviours and counter-productive work behaviour. This can be to such an extent that environmental support will have a significant relationship with normal work behaviours when the locus of control is more external (O'Brian, 2004).

The basic assumption of the self-determination theory, according to O'Brian (2004), is that all humans are generally active, looking for opportunities to learn and develop themselves as well as to fit into a positive sense of self-identity. When the social environment supports a person by satisfying his/her basic psychological needs, the individual will focus on attainment of growth and development. This person might make decisions that will guide him/her towards self-advancement.

O'Brain (2004) confirms that when the environment does not fulfil the basic psychological requirements, the person's attempt to grow and develop is stopped. In this case, people might decide on actions that are not beneficial to themselves or to their environment. Therefore, the basic psychological wants, which can be acknowledged, are autonomy, affiliation and perceived competence.

The self-determination theory is unlike previous motivational need theories, which include the achievement theory (McClelland, Atkinson & Clark, 1953) and Maslow's (1943) needs hierarchy, stating that people are determined to take part in activities in order to fulfil certain needs. The self-determination theory states that humans have essential psychological desires which, when satisfied, enable them to attain some other goal. In this theory, the fulfilment of needs is the way, not the end (O'Brian, 2004).

2.4.4 Self consistency theory

From qualitative reviews of experimental evidence, there are theoretical facts to anticipate a positive relationship between traits such as locus of control and job satisfaction. Locke, McClear and Knight (1996) notes that a person with a high self-esteem will view a demanding job as a deserved chance which he/she can master and take advantage from, whereas a person with a low self-esteem is most likely to see it as an unfair opportunity or a possibility to fail.

In fact, research suggests that people with high self-esteem uphold optimism during situations of failure, which makes future achievement more likely (Dodgson & Wood, 1998). An additional theoretical mechanism connecting these traits to job satisfaction is provided by Korman's (1970) self-consistency theory. Korman's theory foresees that individuals with high self-esteem mostly prefer occupations that are in line with their interests. Because of this choice they usually experience higher levels of job satisfaction. As Tharenou (1979) notes, Korman's (1970) hypothesis has been widely acknowledged with regard to occupational choice.

But more specifically, Korman's theory proposes that individuals with a high self-esteem will be more prone to take part in a broad array of behaviours and cognitions that supports their self-concept. Spector (1982) agrees with Korman (1970) in that individuals with an internal locus of control should be more content with their jobs because they will not stay in an unsatisfying position and are more prone to be successful in organisations. Therefore, internal locus of control is positively connected to job satisfaction.

2.4.5 Control theory

The Control theory (Lord & Hanges, 1987) predicts that when individuals do not perform as expected, they put forth extra effort to achieve the performance goal, or reduce their level of effort and withdraw from the task completely. Research indicates that when individuals who possess an internal locus of control are faced with differences between satisfactory standards of performance and their performance in reality, they tend to enhance their efforts to match their real performance to the standards (Weiss & Sherman, 1973). On the other hand, people who have low self-esteem or external locus of control will be inclined to either lower their standards or totally withdraw from the task when they do not receive positive feedback.

2.5 CULTURAL ISSUES WITH REGARD TO LOCUS OF CONTROL

A popular criticism is that psychometric instruments, which are based largely on middle–class white values and knowledge, are culturally biased and less valid for other population groups (Van Zyl & Visser 1998). Increased debate and widespread attention have been given to the testing of persons with dissimilar cultural backgrounds (Anastasi & Urbina, 1997; Gregory, 1996; Holburn, 1992).

Gaa and Shores (1979) conducted research and concluded that locus of control is not only reliant on culture but also on explicit components or domains of locus of control. According to their research findings domain-specific locus of control measures portray distinct, but not reliable, differences in culturally different populations. The idea that culture is identified as just one feature with regard to particular domains of locus of control is apparent in the research done by Krampen and Weiberg (1981). Differences were identified with regard to the internality and externality of American, Japanese and German students.

Rieger and Blignaut (1996) found a positive correlation between the following constructs: individualism, internal locus of control and autonomy, but there was no correlation between collectivism and external locus of control. Anastasi (1990) argues that there are only cultural differences between cultures or subcultures and that each culture reinforces the development of behaviour that is adapted to the specific culture's values and demands. For maximum utilisation of human resources, a need exists for cross-cultural testing.

2.6 HISTORY OF LOCUS OF CONTROL INVENTORIES

Rotter (1966) developed the first instrument to measure internality and externality with regard to locus of control. Although Rotter's instrument has been used extensively in research and training, his unitary concept of internality has been challenged. On the basis of factor analysis of the responses to Rotter's instrument, several studies found multidimensionality in Rotter's instrument, which seemed to contain items related to control ideology, personal control, system modifiability, and race ideology (Gurin, Gurin, Lao, & Beattie, 1969; Guttentag, 1972; McDonald & Tseng, 1971; Minton, 1972; Mirels, 1970).

Another problem that was identified with regard to Rotter's I-E scale is the fact that the forced choice item format led to ipsative measurement, while the user of the instrument wanted to use it in a normative way (Schepers, 2005). In the opinion of Schepers (2005), there is nothing wrong with ipsative measures, but the users of such instruments should be aware of the limitations that are associated with ipsative measures. Ipsative measures are successfully used to determine the relative strength of drives intra-individually, but not to determine inter-individual differences. For this purpose, normative measures are required (Clemans, 1966).

There were various attempts to analyse Rotter's I-E scale but it was doomed to failure because of the inappropriate factor analytical techniques that were used. Some of these include studies of Rotter (1966), Franklin (1963), Mirels (1970), Abramowitz (1973) and Erwee (1986). Collins (1974) was more successful with his efforts. He converted the 23 pairs of the scale to 46 items with a Likert format. He also added 42 items with the connotation that "it depends on the situation".

Levenson (1972) questioned putting three external factors (chance, fate and powerful others) together. Levenson also proposed a new scale to measure internality and externality; instead of viewing these elements along a continuum, Levenson's stance is to measure both internality (I) and externality (E). Furthermore, he supported two subscales for externality; one to measure perceived influence of chance (EC) and the other to measure perceived influence of powerful others (EO). Gutkin, Robbins, and Andrews (1985) reported through their factor analysis that the results of a health locus of control scale also revealed internal and external factors.

Another type of locus of control scale that was developed by Spector was named the work locus of control scale (Spector, 1988). This 16–item measure was used to measure the degree to which people anticipate that rewards, reinforcements and other reactions in the work domain are influenced either by one's own actions or by others. Reactions to each item were scored on a scale that ranged from "disagree very much" to "agree very much". Higher scores were a sign of a greater external locus of control. Research provided evidence that this measure predicted work outcomes in a more superior way than Rotter's (1966) general locus of control scale (Spector, 1988).

Udai Pareek was a professor in the late 1960s at IIMA in India. According to various statements, it is clear that human resource management (HRM) came to India through his efforts and the L & T Corporation in India. Before Udai Pareek introduced human resource (HR) to India, no company in India used to have HR practices and he is therefore a world renowned personality. Udai Pareek can be identified as the father of the human resource development (HRD) movement in India. He developed the locus of control inventory in 1998 and it was put to use in India soon afterwards.

Another psychologist that made a contribution in the field of instrument development specifically focusing on locus of control is Prof J.M. Schepers. He is the author of the revised edition (1995) of the locus of control inventory. The locus of control inventory of Schepers (1999) was initially standardised on a sample of first-year university students at the Rand Afrikaans University (RAU). The items of the locus of control inventory were endorsed on a seven point scale anchored at the extreme values of 1 and 7.

The three-factor structure of the locus of control inventory was later confirmed by Schepers (2004). The obtained factors were interpreted as external control, internal control and autonomy. The resulting scales yielded reliabilities of 0,841, 0,832 and 0,866, respectively. As far as validity is concerned, statistically significant relationships were found between the matriculation marks and several measures of the Sixteen Personality Factor Questionnaire and the Personal, Home, Social and Formal Relations Questionnaire.

2.7 THE LOCUS OF CONTROL INVENTORY

The locus of control Inventory is a psychometric instrument that was developed by Udai Pareek (1998). The construct validity of this instrument was the focus of the research study. The specific instrument will be discussed in detail in the following paragraphs.

The instrument links locus of control to seven areas:

1. General
2. Success of effectiveness
3. Influence
4. Acceptability
5. Career
6. Advancement
7. Rewards

Table 2.1: Distribution of Items in Locus of control Inventory

	Internality	Externality (Others)	Externality (Chance)
General	1, 27	4, 30	7, 24
Success or effectiveness	3, 10, 16	6, 19, 22	9, 13, 21
Influence	28	17	26
Acceptability	25	29	18
Career	2	5	8
Advancement	23	11	14
Rewards	20	15	12

Source: Pareek (1998: 7)

2.7.1 Purpose

The locus of control inventory can be used for both research and training purposes in human resource development, organisation development or training packages. It was mainly developed for training purposes. Levenson's (1972) concept of locus of control was used to develop the locus of control inventory of Udai Pareek to measure internality and externality in the organisational context. Levenson's (1972) instrument was not developed specifically for organisations, therefore the need arose for the locus of control inventory of Udai Pareek.

2.7.2 Description of the instrument

The locus of control inventory was based on the Levenson (1972) concept of locus of control. An earlier version of the locus of control inventory contained a six-point scoring system and 24 items (parallel to Levenson's instrument). The current five-point system appears to be a superior measure and the 30–item version contains 10 statements for each construct that is measured. Thus 10 statements of which each is used to determine the following: internality (I) externality - others (EO) and externality – chance (EC).

Likert-type scaling was used in the locus of control inventory. This is evident when the questionnaire response categories include words like “strongly agree”, “agree”, “disagree”, and “strongly disagree”. The value of this format is the unambiguous normality of response categories. The Likert format solves the problem where the researcher would find it impossible to judge the relative strength of agreement intended by various respondents (Babbie & Mouton, 2006). The Likert scale is one of the most commonly used ones in contemporary questionnaire design.

2.7.3 Administration of the instrument

The administration of a psychological instrument usually progresses through three stages:

- Stage 1: Preparation
- Stage 2: Assessment administration
- Stage 3: Aftermath

A complete discussion of the administration of an instrument is beyond the scope of this research study. Foxcroft and Roodt (2005) can be consulted for a detailed discussion on this subject. The basic administrative guidelines, which are specific to this instrument, will be discussed next.

Good assessment practices require the assessment practitioner to prepare thoroughly for the assessment session. The respondents complete the instrument by evaluating each statement according to a five-point scale ranging from zero “seldom or never agree” to five “strongly agree”. When the assessment is finished it is the administrator’s job to collect and secure assessment materials. The safekeeping of these documents forms an important part of the ethical use of psychometric measures. These assessments should be locked away and only be available to the designated persons (Foxcroft & Roodt, 2005).

The responses must be transferred to the scoring sheet, which presents three scores: internality, externality – others and externality – chance. If possible, the scoring sheets should be completed in advance, so that the mean and standard deviation can be calculated prior to the discussion of the scores. Norms can also be created, for more information on this refer to 8.8 Norms.

2.7.4 Scoring and interpretation

The numbers that respondents have assigned to the instrument items are transferred to the scoring sheet and a total is computed for each column. Scores will range from one to 40 for each of the three columns (Internality, Externality – other and Externality – chance).

After scoring the instrument, it is firstly important to remember that if one is not happy with the outcome of one's scores, one can create an action plan that will help one to change the way in which one looks at certain things (Pareek, 1998). Next one can select the column with the highest total. The section that is provided in the interpretation sheet should be read. The sheet is attached as Appendix C. During the feedback session a facilitator can lead a discussion based on the concepts and findings included in this study. Respondents can be asked to predict their own levels, namely high, medium or low on each of the three dimensions. In very open group feedback sessions, each member of a trio can estimate the levels of the other two members. Completed scoring sheets should be distributed to the respondents, as well as copies of the interpretation sheet. Trios can be formed to discuss discrepancies between actual scores and observed behaviour (Pareek, 1998).

The facilitator could present implications of internality for employee effectiveness and lead a discussion on how to increase internality and reduce externality. The discussion could include which organisational practices promote I, EO and EC. Another important discussion could deal with how to increase internality among the employees (Pareek, 1982). Material that would help the facilitator lead this discussion includes Baumgartel, Rajan, & Newman (1985), Reichard (1975), Mehta (1968) and DeCharms (1976).

2.7.5 Norms

Many human characteristics that are measured in psychology are assumed to be normally distributed in the population. This is referred to the normal distribution that has a mean of 0 and a standard deviation of 1 (Foxcroft & Roodt, 2005). In order to make the raw scores obtained from testees more meaningful, the raw scores are converted to a normal score through statistical transformation. According to Foxcroft and Roodt (2005), the definition of a *norm* is a measurement against which the individual's raw score is evaluated so that the individual's position relative to that of the normative sample can be determined.

Based on data from more than 300 managers, mean and standard deviation (SD) values are presented in Table 1. High and low scores were calculated by adding or subtracting one-half SD value to or from the mean. Similarly, very high and very low scores were obtained by adding or subtracting one SD value to or from the mean. Such norms can be worked out for specific organisations for interpretation purposes.

Table 2.2: Norms - Mean and Standard Deviation Values

	Mean	SD	Very High	High	Low	Very Low
I	25	8	33	29	21	17
EO	25	9	34	29.5	20.5	16
EC	19	9	28	23.5	14.5	10

Source: Pareek (1998, 4)

2.7.6 Previous research on the reliability and validity of the instrument

The reliability and validity of a questionnaire should be taken into account when one uses the results to draw conclusions. Reliability can be determined by looking at the questionnaire's internal consistency.

- **Reliability**

Levenson (1972) reported moderately high internal consistency, with Kuder-Richardson reliabilities (coefficient alpha) of .64, .77, and .78 and split half reliabilities of .62, .66 and .64 for I, EO and EC respectively. Retested reliability for a one-week period for the three subscales were 0.64, 0.74 and 0.78, respectively. Reliabilities of the Levenson instrument were also moderately high in another study in India (Sen, 1982). Split-half reliability coefficients for the earlier version of the locus of control inventory were .43, .45 and .55 and even-odd reliability coefficients were .41, .48 and .54 for I, EO and EC subscales, respectively. The current version has similar reliability coefficients.

- **Validity**

With regard to previous and related research, there was a high correlation (.89) between Levenson's instrument and the locus of control inventory in a sample of 26 bankers. This finding indicates the validity of the locus of control inventory. Using Levenson's scale, Surti (1982) reports a highly significant coefficient of correlation (.70) between EO and EC in a sample of 360 professional women and correlation values of .00 and .06 between I and EO and between I and EC, respectively. This finding shows the validity of Levenson's two-factor concept. These reliabilities were also used in the construction of the locus of control inventory of Udai Pareek.

2.8 CONTROL OVER THE USE OF PSYCHOLOGICAL ASSESSMENT MEASURES

2.8.1 Employment Equity Act 55 of 1998

The reason for the development of the Employment Equity Act, No 55 of 1998, was to provide rules and regulations for employment equity. The Employment Equity Act of South Africa, No 55 of 1998 places all test developers and users under an obligation to consider the impact of psychometric assessments on different groups as carefully as they consider other technical psychometric issues. The importance of including this requirement in the design of psychometric instruments cannot be overemphasised. The fact that some tests may be biased against certain groups has become a primary concern in South Africa (Schaap, 2001).

The Employment Equity Act, No 55 of 1998, has the twofold purpose of ensuring that only valid and reliable assessments are used and that assessments are used in a fair manner, which does not contain bias. Psychological testing and other comparable assessments of an employee are forbidden unless the test or assessment that is being used:

a) has been scientifically revealed to be valid and reliable;

b) can be used in a fair manner on all employees;

c) is not biased in opposition to any employee or group.

2.8.2 Code of practice for psychological assessment in the workplace

The primary objective of the code is to guarantee that psychological assessments are used appropriately, professionally and ethically, by identifying the needs and rights of those who are part of the assessment process, the reasons for the assessment and the wider context in which the assessment will operate. To attain these outcomes, it is believed that the practitioner or consultant should have the relevant competence as well as knowledge and understanding of psychological tests and other assessment actions that inform and underpin this (SIOPSA, 2006).

The code adds value to the professional codes and laws, which ensure effective test use and employment conditions, the code also ensures that tests being used are in line with the professional codes and laws. One of the most important topics that is covered in the code of practice and which is related to this study is to advance the professional and ethical use of assessment measures. Assessment practitioners should have a say by giving their opinions on how to improve standards of ethical assessment. Some of the actions already identified can include any of the following:

- ensuring that information about assessment methods is circulated to all users as it becomes accessible;
- encouraging the ethical use of assessment procedures and actively trying to stop unethical use;
- publishing information for circulation to members of the public and especially stakeholders in assessment of the workplace, which explains the use of assessment methods;
- encouraging and facilitating exposure of abuses;
- encouraging research of the cross-cultural validity of all assessment instruments;
- encouraging and facilitating the access of data from such research;
- encouraging the addition of topics covered in the code in the basic training and ongoing professional education of assessment practitioners;
- encouraging the access of information and training to people who are the decision-makers with regard to the use of assessment methods, or those who make decisions on the basis of the outcomes of such instruments.

2.8.3 Health professions Act No 56 of 1974

The purpose of the Health Professions Act was to launch the Health Professions Council of South Africa. The Council should offer control over the training, registration and practices of practitioners and health professionals.

According to the Health Professions Act No 56 of 1974 a psychological act with regard to assessment is described as being: the use of measures to assess mental, cognitive or behavioural processes and functioning, intellectual or cognitive ability or functioning aptitude, interest, emotions, personality, psycho-physiological functioning or psychopathology.

In South Africa, the use of psychological instruments is under legislative control. Foxcroft and Roodt (2005) explain that this means that a law has been promulgated restricting the use of psychological assessment measures to suitably registered psychology professionals. According to the Health Professions Act No 56 of 1974, the following are defined in Section 37 (2) (a), (b), (c), (d) and (e) as specifically pertaining to the profession of a psychologist:

- The evaluation of behaviour or mental processes or personality adjustments or adjustments of individuals or groups, by making use of the interpretation of tests for the purpose of identifying intellectual abilities, aptitude, interests, personality make-up or personality functioning and the analysis of personality and emotional functions and mental functioning deficiencies according to a standard scientific system for the categorisation of mental deficiencies.

- The use of any method focused on helping persons or groups of persons in the modification of personality, emotional or behavioural problems or at the support of positive personality change, growth and development, and the classification and evaluation of personality dynamics and personality functioning according to psychological scientific methods.

- The assessment of emotional, behavioural and cognitive processes or alteration of personality of individuals or groups of persons by making use of and interpretation of questionnaires, tests, projections or other techniques, whether of South African origin or international, for the purpose of identifying intellectual abilities, aptitude, personality make-up, personality functioning, psycho-physiological functioning or psychopathology.

- The focusing of control over prearranged questionnaires or tests or prescribed techniques, or instruments for the identification of intellectual abilities, aptitude, personality make-up , personality functioning, psycho-physiological functioning or psychopathology;

- The improvement of and control over the development of questionnaires, tests techniques or instruments for the identification of intellectual abilities, aptitude, personality make-up, personality functioning, psycho-physiological functioning or psychopathology.

2.8.4 International guidelines for test use

The International Test Commission (ITC) compiled international guidelines for test use as there was a need for international guidelines. The focus of the International Test Commission projects was on rules for good test use and for promoting best practice in assessment. The work done by the ITC was also to encourage good practice in test adaptions (Hambleton, 1994; Van de Vijver & Hambleton, 1996). It was a significant step towards ensuring uniformity in the quality of tests, which were modified for use across various continents, cultures and languages.

There are various reasons why rules with regard to test use are required at international level (ITC, 2000). Some of these are: countries vary significantly in the level of legislative control that they can implement with regard to the use of testing and its effect on those who are tested. Access to the instruments in terms of rights to purchase or use test materials fluctuate to a great extent from country to country. A number of well-known instruments have been published on the internet in violation of copyright and without acknowledgement of the test authors or publishers. Within the work-related testing arena, the greater international mobility of labour has enlarged the demand for instruments to be used on job applicants from diverse countries.

The most important part of the guidelines for this specific study is no 2.9, which states that competent test users will (ITC, 2000):

- monitor and review changes in the populations of individuals being tested and any criterion measures being used at regular intervals over time;

- observe tests for confirmation of adverse impact;
- be alert of the need to re-evaluate the use of a test if changes are made to its form, contents or way of administration;
- recognise the need to re-evaluate the facts of validity if the rationale for which a test is being used is changed;
- where achievable, try to validate tests for the use to which they are being put;
- give support in the updating of information concerning the norms, reliability and validity of the test by providing applicable test data to the test developers, publishers or researchers.

2.8.5 Categories of psychology professionals

Foxcroft and Roodt (2005) state that there are currently five categories of assessment practitioners within the profession of psychology in South Africa who may use psychological measures, namely psychologists, registered counsellors, psychometrists (independent practice), psychometrists (supervised practice) and psycho-technicians. All these psychology professionals should, apart from completing modules in assessment, be in possession of a degree in which they studied psychology in depth as well as related disciplines. This provides them with a broader knowledge base on which to draw during the assessment process.

2.9 SUMMARY

The literature review covered the definition as well as a whole range of theories which are related to locus of control. The history of locus of control inventories as well as the legislation applicable to psychometric testing was also discussed. In the next chapter, the various types of reliability and validity concepts that are applicable to the study will be discussed. The main focus will be on construct validity, which forms part of the research objectives of this study, and therefore the focus will mostly be on the description and explanation of construct validity. The concepts convergent and discriminant validity, which are essential in determining construct validity, will also be discussed.

Next, the researcher will focus on the threats that may have an impact on construct validity. The nomological network and the multitrait-multimethod matrix are two methods which can be used to determine the construct validity of an instrument, therefore the basics of these two methods will also be explained in the chapter that follows. Finally, other types of validity that can have an influence on the research study will also be discussed.

CHAPTER 3

THE VALIDATION OF INTERNATIONAL PSYCHOLOGICAL INSTRUMENTS FOR APPLICATION IN SOUTH AFRICA

3.1 INTRODUCTION

Psychometrics has an impact on the lives of thousands of South Africans (Sehlapelo & Terreblanche, 1996) as psychological tests are used for selection purposes, placement, promotion, transfers, training and development (Van der Merwe, 1999; Kemp, 1999; Shaw & Human, 1989; Taylor & Radford, 1986). The important thing to remember is that these psychometric tests, which are used for all the above-mentioned purposes, have to be reliable, valid, fair and non-biased especially those that have been developed in other countries and then used in South Africa (Bedell, Van Eeden & Van Staden, 1999; Van der Merwe, 1999). After a brief discussion of the development of psychometrics in South Africa, prerequisites for psychometric instruments will be discussed.

3.2 HISTORY AND DEVELOPMENT OF PSYCHOMETRICS IN SOUTH AFRICA

Although a lot of employees have little trust in psychometric tests and the test process, these tests and the process are commonly used as aids in occupational decisions, which include the selection and classification of human resources (Kemp 1999; Sehlapelo & Terreblanche, 1996). According to Shaw and Human (1989) the black population mostly see psychometric testing as being biased, irrelevant and unfair therefore psychometrics is seen as racism and unfair discrimination (Sehlapelo & Terreblanche, 1996).

Instruments that do not respect the diversity of this country lead to unfair discrimination against a lot of employees, especially against the previously disadvantaged individuals (Erasmus and Schaap, 2003). Because of the above-mentioned reasons, there is a need to develop alternative methods of assessing potential (Shaw & Human, 1989). This need emerged out of issues of fairness and equal opportunities for all employees. If these tests are proved to be scientifically valid and culturally fair, there will not be a need to develop new methods and tests (Sehlapelo & Terreblanche, 1996; Bedell, et al., 1999).

Even though not all of the psychometric tests have been proved to be culturally fair, it will probably go on to be widely applied in South Africa (Foxcroft in Bedell et al, 1999). The Employment Equity Bill as well as the policy of the Health Professions Council of South Africa (HPCSA) for the classification of psychometric tests are putting pressure on test developers and users to validate existing psychometric instruments. By doing this, they will ensure that psychometric tests are valid and reliable for their specific uses.

The concept of unfair labour practice was introduced in South Africa through the Labour Relations Act 66 of 1995. An unfair labour practice takes place when any individual or group is paced at an unfair disadvantage. Psychometrical testing can be seen as an unfair labour practice when different cultural groups obtain significantly different mean scores (Taylor & Radford, 1986). One of the major stumbling blocks in creating tests for South African psychological testing as identified by Huysamen (1996) is the diverse linguistic and cultural backgrounds, for which these tests must be suitable, as only culture-fair tests may be used for comparisons between races and different countries. This need arose because of companies establishing international branches all over the world, which led to the need of South African people to go and work abroad and vice versa. Van der Merwe (1999) and Kemp (1999) also state that a definite need for culture–fair tests has been expressed.

In general, studies support the view that South African tests are reliable and valid for the groups for which they were developed and standardised. This means that comparison of scores of individuals within these groups may be justified. However the fact that cross-cultural validity has not been openly determined for a number of tests implies that cross- group comparison of scores could yield information that is not valid and which could lead to discrimination (Bedell, et al., 1999).

3.3. RELIABILITY AND VALIDITY

3.3.1 Reliability

Price (1997) explains *reliability* as the extent to which a measure produces the same results when used repeatedly. Test reliability refers to the instrument's degree of accuracy and consistency (Foxcroft & Roodt, 2005). According to Erasmus and Schaap (2003) reliability simply tells one how confident one could be that the scores obtained with the instrument are consistent and accurate. Salkind (2006) also agrees with the above statements by stating that reliability occurs when a test measures the same thing more than once and results in the same outcomes.

The reliability of the instrument can be influenced by the group to whom the instrument is administered. When used in the appropriate formula, reliability will indicate the relation between true scores and observed scores. This is known as the standard error of measurement. Standard error of measurement provides a direct indication of the degree of inconsistency or error one could expect with individual scores (Erasmus & Schaap, 2003).

The validity of a scale is limited by its reliability, and therefore unreliable measurements will hamper efforts to predict behaviour (Erasmus & Schaap, 2003). A factor that needs consideration because it will influence the size of the reliability coefficient is the range of individual differences in the group. According to Erasmus and Schaap (2003), a more heterogeneous group usually results in a higher reliability coefficient than that of a homogeneous group. Variance due to content heterogeneity can also influence the reliability coefficient. More homogeneous items lead to a higher coefficient.

The best way to establish the reliability of a measure is to use it repeatedly on the same object (Smith & Robertson, 1986). Smith and Robertson also state that there are several means by which a measure of reliability can be estimated.

Salkind (2006) distinguishes between the following methods of obtaining the reliability coefficients:

- Test-retest reliability: It measures how stable a test is over time. The same test is given to the same group of people at two different points in time usually after a short time interval (Salkind, 2006).

- Parallel form reliability: According to Salkind (2006), it is when different forms of the same test are given to the same group of participants. Then the two sets of scores are correlated with each other. The tests are said to be equivalent if the correlation is statistically significant, therefore the relationship is due to something shared between the two forms, not some chance occurrence (Salkind, 2006).

- Internal consistency: It measures how united the items are in a test or assessment (Salkind, 2006). The similarity of the test items can be determined by the split-half reliability method, the Kuder Richardson's formula or Cronbach's alpha. The Kuder Richardson's formula is generally used when the items are scored in terms of pass or fail whereas Cronbach's alpha is used when items are scored on a continuum; 0.70 is seen as minimum acceptable level for alpha (Price, 1997).

An instrument can be reliable but not valid, but a test cannot be valid without first being assessed as reliable. This means that reliability is a needed, but not adequate, condition of validity (Salkind, 2006; Smit, 1991). Just as reliability does not refer to only one attribute of a test, the concept of *validity* is also multifaceted.

3.3.2 Validity

All psychometric tests are developed to make inferences about people. According to Murphy and Davidshofer (2001), *validity* can be seen as the accuracy of these inferences. Anastasi and Urbina (1997) state that *validity* concerns what is measured and how well it is measured. *Validity* is the degree to which a measure reflects the concept it is designed to measure (Price, 1997). Adams (1966) confirms this by stating: "Validity is always validity for a specific purpose to aid in making a specific type of decision concerning members of a specific group."

These definitions are important because they focus on the fact that the essence of validity is the accuracy of the conclusions that may be made from it. Thus, it is not the measure that have validities, it is the decision made from the results that is important when determining validity (Adams, 1966). Foxcroft and Roodt's (2005) understanding of validity is that the validity of a measure has to do with what the test measures and how accurately it does so. They also add that validity is not an explicit characteristic of a measure, but that a psychological measure is valid for a specific purpose, this means that it has a high or low validity for a specific purpose.

Salkind (2006) affirms this and states that it should be kept in mind that the validity of an instrument is often defined within the situation of where and how the test is being used. There are three aspects of validity:

- Validity refers to the results of a test, not to the test itself.
- Just as with reliability, validity is never a question of all or none. The results of a test cannot just be interpreted as valid or invalid. This progression occurs in degrees from low validity to high validity.
- The validity of the results of a test must be interpreted within the context in which the test occurs.

There are two different approaches to use when making decisions about an instrument. The first one concerns the attribute being measured by the instrument and the second will affect decisions made about the testee. Here one must remember that tests are not only developed to make inferences about individuals, they are also used to make decisions (Murphy & Davidshofer, 2001). According to the results, the individual may or may not be appointed in a specific position.

3.4 CONSTRUCT VALIDITY

Construct validity refers to the level at which decisions can be made from the operationalisations in one's study to the theoretical constructs on which those operationalisations were based (Trochim, 2006). Trochim (2006) is of the opinion that equal to external validity, construct validity is also linked to generalising. But where external validity involves generalising from one's context to other people, places or times, construct validity can be thought of as a labelling issue.

Construct validity will now be discussed under various headings so that a comprehensive understanding of the term can be reached. This will aid the researcher in identifying the steps that can be followed when determining the construct validity of various instruments.

3.4.1 Defining a construct

To understand the traditional definition of construct validity, it is first necessary to understand what a construct is. Garson (2008) states that a high-quality construct has a theoretical basis, which is transcribed through clear definitions involving indicators that can be measured. He defines a poor construct as follows: "A poor construct may be identified by lack of theoretical agreement on its content, or by flawed operationalization such that its indicators may be seen as measuring one thing by one researcher and another thing by another researcher".

A *construct* is a way of defining something, and if the construct does not conform to the existing literature then its construct validity might be invalid (Garson, 2008). Garson (2008) is of the opinion that the more a construct is used by researchers in different contexts with results that are consistent with theory, the more there is construct validity. Therefore, researchers should establish both of the two major types of construct validity, convergent and discriminant validity, for their constructs.

According to Chronbach and Meehl (1955), a *construct* is an attribute of people, supposed to test performance. In test validation, it is the attribute about which we make statements when scoring and interpreting a test. Chronbach and Meehl (1955) explain that a person may at any time have a qualitative attribute or structure or possess some level of a quantitative attribute. They also state that a construct has certain meanings related to its general character. Persons who possess this attribute will, in situation X, act in a manner Y (with a stated probability).

3.4.2 The meaning of construct validity

Garson (2008) suggests that construct validity determines the logic of items which include measures of social concepts. In more simple and practical terms, Foxcroft and Roodt (2005) agree with this statement of Garson by saying that the *construct validity* of a measure is the degree to which it measures the theoretical construct or trait that it should measure. It is vital to determine the construct validity especially when the test user wants to determine the degree to which a specific trait or construct that should be reflected in the test is in fact present in the test (Owen & Taljaard, 1996).

Salkind (2006) describes *construct validity* as a time-consuming and difficult type of validity to establish, but also the most attractive. A definition will be the extent to which the outcomes of a test are associated with an underlying psychological construct. It associates the practical components of a test score with some underlying theory of behaviour. As an example, a personality test actually measures personality. In psychometrics, *construct validity* can be defined as the degree to which a scale measures a theorised construct. It is linked to the theoretical ideas behind the trait which is being analysed a non-existent concept in the physical sense may be suggested as a method of organising how personality can be viewed (Pennington, 2003).

Welman and Kruger (2001) define *construct validity* as: the degree to which procedures which are supposed to produce the independent variable which are being studied indeed succeed in generating this variable rather than something different. In simpler terms, it indicates whether a scale correlates with a theoretical psychological construct. After looking at all the different validity concepts, Trochim (2006) came to the conclusion that a definition of *construct validity* is the estimated truth of the conclusion that one's operationalisation exactly reflects its construct.

He also made a distinction between two different types: translation validity (a term that he created as an overarching term) and criterion-related validity. According to Trochim (2006) in translation validity, the focus is on whether the operationalisation is an accurate indication of the construct. This approach is definitional in nature. It assumes one has a high quality definition of the construct and that the operationalisation can be verified against it.

In criterion-related validity, Trochim (2006) states that items should be examined to determine if the operationalization reacts the way it should when compared against the theory of the construct. This is a more relational approach to construct validity. It assumes that one's operationalisation should react in predictable ways in relation to other operationalisations based upon one's theory of the construct.

Regardless of how construct validity is defined, there is no single best way to study it. In most cases, construct validity should be confirmed from a number of perspectives. Salkind (2006) points out that the more strategies used to reveal the validity of a test, the more confidence test users will have in the construct validity of that test, but only if the facts provided by those strategies are persuasive.

It is important to remember that a measure is only valid for the particular purpose for which it was designed. Foxcroft and Roodt (2005) insist that the validity of a measure should be determined when it is used in a different way or in a different context. When interpreting assessment results, it is important to establish the construct validity of the measure for its particular purpose otherwise the results may be invalid.

3.4.3 Convergent and discriminant validity

3.4.3.1 Convergent validity

Trochim (2006) states that convergent validity is present when it is shown that measures that are theoretically supposed to be highly interrelated are, in fact, highly interrelated. Garson (2008) confirms the above and notes that there are a few ways to measure convergent validity. He suggests a few alternatives, which include the following:

- it is assessed by the correlation between items which make up the instrument measuring a construct;
- by the correlation of the given scale with measures of the same construct using scales and instruments identified by other researchers and which are preferably already accepted in the field;
- also by correlation of relationships, which involves the given scale across samples or methods. It is expected that these correlations should at least be moderate to be proof of external validity.

Convergent validity, can be understood as the degree to which the operationalisation is similar to other operationalisations that it should be similar to as stated in the theory (Foxcroft and Roodt, 2005; Trochim, 2006). High correlations should be confirmation of convergent validity.

Trochim (2006) and Salkind (2006) maintain that to establish convergent validity, there exists a need to indicate that measures that should be related are actually related. It is theorised that some items in the instrument should reflect the idea of internal locus of control. It should readily be seen that the item intercorrelations for those item pairings are very high. This provides evidence that the theory is correct in that all items are related to the same construct.

Trochim (2006) also advises that it is important to notice that while the high inter-correlations demonstrate that the specific items are probably related to the *same* construct, which does not automatically mean that the construct is internal locus of control. It could be that there is some other construct that all those items are related to. But it can be assumed from the pattern of correlations that the items are converging on the same thing.

Testing for internal consistency

Internal consistency is a type of convergent validity which seeks to assure that there is at least moderate correlation among the indicators for a concept. Poor convergent validity among the indicators for a construct may mean the model needs to have more factors. According to Pallant (2005), Cronbach's alpha is commonly used to determine internal consistency. Important numbers to look out for when working with Cronbach's alpha are the following:

- .60 is considered acceptable for exploratory purposes;
- .70 is considered adequate for confirmatory purposes;
- .80 is considered good for confirmatory purposes.

Rasch models, which can also be named one-parameter logistic models, are an internal stability test used in item response theory for binary items. Rasch models, like Guttman scales, ensure that items measure a construct from an ordered relationship (Rasch, 1960; Wright, 1977). It is vital to note that a group of items may have well-organised internal consistency even though they do not correlate very highly.

Ordered internal consistency reflects a complexity factor, whereby answering a difficult item predicts responses on easier items but not vice versa (Wright, 1977). Wright (1977) explains that when factor analysis is used to validate the addition of a set of indicator variables in the scale for a construct, the researcher will take a stance which identifies a linear, additive model. Linearity is assumed as part of correlation, which is the basis for grouping indicator variables into factors. Wright (1977) further states that additivity is also assumed, meaning that items will be evaluated to be internally consistent only if they are equally highly correlated.

3.4.3.2 Discriminant validity

With reference to discriminant validity, Garson (2008) and Trochim (2006) write that it is the second main type of construct validity, which refers to the belief that the indicators for diverse constructs should not be so highly correlated as to lead one to conclude that they measure the same thing. Trochim (2006) believes that this would happen if there is a definitional convergence between constructs. Discriminant validity analysis refers to testing statistically whether two constructs differ. The results should then indicate that discriminant validity was achieved by presenting that the relationship indicates very low scores between measures of different constructs.

Campbell and Fiske (1959) notes that *discriminant validity* is the degree to which the concept is different from other concepts that it should not be similar to. Campbell and Fiske (1959) present the concept of discriminant validity within their debate on evaluating test validity. They highlight the necessity of using both discriminant and convergent validation techniques when assessing new tests. They also declare that a successful evaluation of discriminant validity shows that a certain test that measures a specific concept is not highly correlated with other tests that are destined to measure different concepts.

3.4.4 Methods used to determine discriminant validity

- **Correlational methods**. In developing scales, researchers often eliminate an indicator if it correlates more highly with a construct different from the one which it was anticipated to measure. Some researchers use r = .85 as a rule-of-thumb cut-off f, because of apprehension that correlations above this level could signal definitional overlap of concepts (Trochim, 2006). Other researchers use the criterion that two constructs significantly vary if the correlations between a given scale and a specific criterion measure are larger in magnitude than the correlations between that same scale and criterion measures used for other scales that are not related.
- **Average variance extracted (AVE) Method.** An alternative factor-based procedure for determining discriminant validity was identified and developed by Fornell and Larcker (1981). In this method, the researcher understands that constructs are diverse if the AVE for one's constructs is larger than their

shared variance. This means that, the square root of the average variance extracted for a given construct should be higher than the value of the standardised correlation of the construct with any other construct in the analysis.

- **Structural equation modelling (SEM) Methods**. Confirmatory factor analysis within structural equation modelling, discussed in the next chapter, is a general method of determining discriminant validity. If goodness of fit measures for the measurement model in SEM are sufficient, it can be concluded that the constructs in the model are diverse.
- **Nested models.** A more thorough SEM-based approach to discriminant validity is to run the model unconstrained but then also constraining the correlation between the constructs to 1.0. If the two models do not diverge significantly on a chi-square difference test, the researcher will fail to conclude that the constructs differ (Bagozzi, Yi, & Phillips, 1991). In this procedure, if there are more than two constructs, one must use a comparable analysis on each pair of constructs, constraining the constructs to be correlated and then releasing the constraints. This method is considered more accurate than either the SEM measurement model approach or the AVE method.

3.4.5 Interpretation of convergent and discriminant validity

To determine the degree to which any two measures are linked to each other, the correlation coefficient is usually used (Pallant, 2005). That means that the patterns of inter-correlations among measures should be looked at. Correlations among theoretically similar measures should be high while correlations between theoretically different measures should be low.

The main problem that Trochim (2006) discussed with the convergent-discriminant idea has to do with his use of the terms high and low. A popular and very important question is: How high do correlations need to be to provide sufficient evidence for convergence and how low do they need to be to provide sufficient evidence for discrimination? Salkind (2006) states that in general convergent correlations should be as high as possible and discriminant correlations should be as low as possible, but there is no specific rule that exactly prescribes this. According to Trochim (2006), a basic rule that can be followed is that convergent correlations should always be higher than the discriminant correlations.

It is best when the researcher is trying to argue for construct validity to show that both convergent and discriminant validity are supported. Trochim (2006) advises that it is best to put both principles together into a single analysis and then to examine both at the same time. By providing more constructs and measures it will improve the researcher's ability to assess construct validity using approaches like the multitrait-multimethod matrix and the nomonological network.

Another approach to determining the construct validity is the idea of pattern-matching. Instead of seeing convergent and discriminant validity as differences of kind, pattern-matching views them as differences in degree. This was identified as being a more reasonable idea and helps to steer clear of the problem of how high or low correlations need to be; and rather to say that convergence or discrimination has been established (Trochim, 2006).

3.4.6 Threats to construct validity

Cook and Campbell (1979) and Mehta (2004) made a list of all the threats to construct validity that are important to look at when trying to determine the construct validity of a specific instrument. These threats will be discussed in short in the following paragraphs.

3.4.6.1 Inadequate preoperational explication of constructs

In this case, *preoperational* means before changing constructs into measures, and *explication* in more simple terms, means explanation, therefore the researcher did not explain what he/she meant by the construct in the best way possible. This means that the researcher can be accused of not thinking through his/her constructs well before using them (Cook & Campbell, 1979). Some possible solutions are:

- to ensure enough time to think through the concepts;

- to make use of different methods to express the concepts;
- to ask experts in the relevant field to criticise the identified operationalisations.

Mehta (2004) also observes inadequate explication of constructs as a threat to construct validity. She describes this concept as failure to adequately explain a construct that may lead to wrong conclusions made about the association between the operation and the construct.

3.4.6.2 Mono-operation bias

Mono-operation bias has to do with the independent variable, cause, program or treatment in a research study (Cook & Campbell, 1979). If only a single version of an instrument is used in a specific place at a certain point in time, the whole range of the concept might not be captured. Every instrument is imperfect relative to the construct on which it was developed. According to Mehta (2004) mono-operation bias can be described as any operationalisation of a construct that does not accurately represent the construct of interest and measures constructs that are not relevant. All of this will then lead to the inference to be more complicated.

If it can be decided that the instrument accurately reflects the construct of the instrument, the critics might dispute that the results of the study only show the peculiar version of the instrument that the researcher implemented, and not the real construct that should have been reflected (Cook & Campbell, 1979). A solution might be to try and implement a lot of versions of the instrument.

3.4.6.3 Monomethod bias

Monomethod bias includes one's measures or observations but not one's programmes or instruments, because this might be the same issue identified as mono-operation bias. In the view of Mehta (2004) *Monomethod bias* means that when all operationalizations apply the same method, that specific method forms part of the construct which is being studied. A solution that is proposed by Mehta (2004) is to try to implement various measures of the most important constructs and to try to demonstrate that the measures used behave as expected, based on the relevant theory.

3.4.6.4 Interaction of different treatments

By using an example, this threat will be explained in more understandable terms. For the purposes of remedying a certain unwanted effect, there might be two or more intervention programmes running simultaneously, and by looking at the effects of all of these programmes, it might be concluded that they are all designed to have the same effects. When looking at the wider context, the researcher should be careful not to label the effect as the sole consequence of his/her intervention and should therefore be aware of the effects of the other programme by also taking that into account when interpreting the findings (Trochim, 2006).

3.4.6.5 Interaction of testing and treatment

It is vital to determine if the measurement actually make the groups more accessible to the treatment that are being administered. If it does, then the testing is actually a part of the treatment, it is inseparable from the effect of the treatment. This can also be seen as a labeling issue because using the label programme to refer to the programme alone is not accurate, but in fact the word programme also includes the measurement (Cook & Campbell, 1979).

3.4.6.6 Restricted generalisability across constructs

Cook and Campbell (1979) states that restricted generalizability can also be referred to as the unintended consequences threat to construct validity. The researcher did a study and came to some conclusions, but the researcher could have failed to anticipate some negative consequences of the study. When it is stated that something is effective, one has to clearly define what one means by effective and also state the conditions for effectiveness. This threat is focused on reminding the researcher that care should be taken when the observed effects are generalised to other potential outcomes.

3.4.6.7 Confounding constructs and levels of constructs

Mehta (2004) describes the term in the following way: operations usually include more than one construct, but by not describing all the constructs, it may result in imperfect construct inferences. But it also describes levels of constructs in terms of inferences about constructs that best characterise study operations. It may also fail to describe the restricted levels of the construct studied.

3.4.6.8 Construct-irrelevant variance

In the views of Messick (1989), there might be some disparate sub-dimensions that creep into measurement and contaminate it. These sub-dimensions are not relevant to the focal construct and in fact there is no need to measure them, but their addition in the measurement is to be expected. They produce dependable variance in test scores, but it is not relevant to the construct. Construct irrelevant variance may arise in two different forms. Firstly, it is construct-irrelevant easiness and secondly construct-irrelevant complexity (Messick, 1989).

As the term imply, *construct-irrelevant complexity* means inclusion of some tasks that make the construct difficult or more complex and result in very low scores for some people. *Construct-irrelevant easiness*, on the other hand, lowers the difficulty of the test. For instance, construct-irrelevant easy items include items that are vulnerable to 'test-wise' solutions, giving an advantage to 'test-wise' examinees who obtain scores which are unsustainably high for them (Messick, 1989).

3.4.7 The nomological network

The nomological network is a concept that was designed by Cronbach and Meehl in 1955. This formed part of the American Psychological Association's efforts to expand the standards for psychological testing. The term nomological was taken and developed from a Greek word which means "lawful", so the nomological network can be understood as the “lawful network”. The nomological network was Cronbach and Meehl's (1955) vision of construct validity. This means that in order to ensure that there will be evidence that the measure includes construct validity, Cronbach and Meehl (1955) argue that a nomological network had to be designed for the measure. This network can include the basic framework for what should be measured, a framework for how it will be measured, and the indication of the linkages among and between these two frameworks.

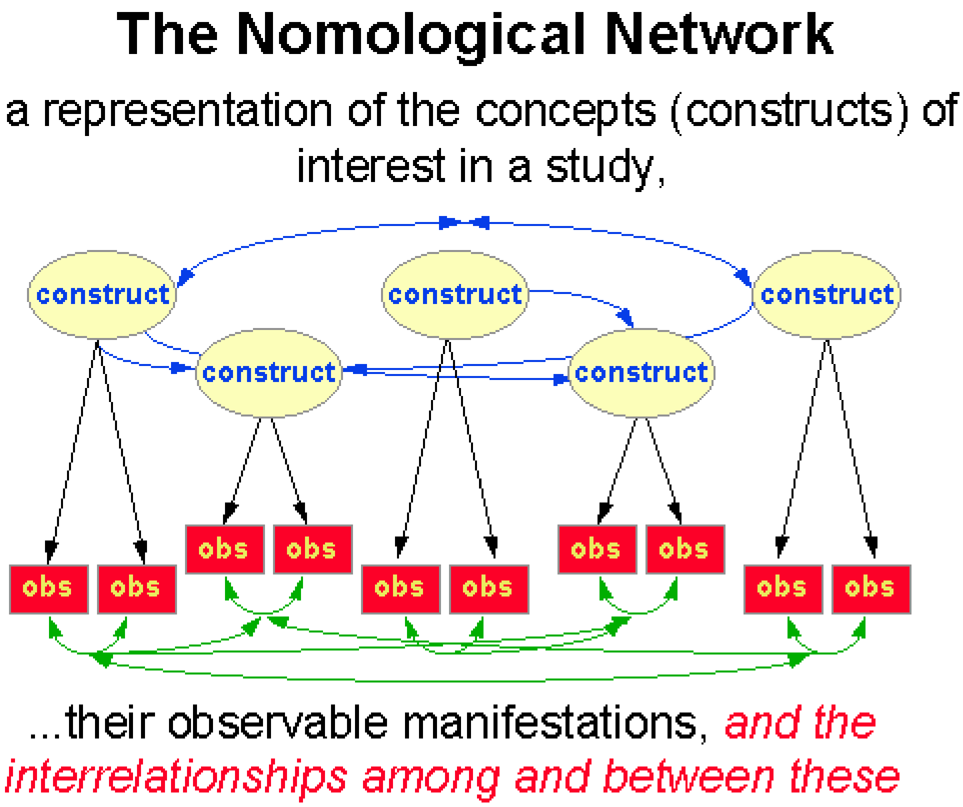


Figure 3.1: A representation of the constructs of interest in a study

Source: Trochim (2006: 10)

The nomological network is based on a number of principles that guide the researcher when trying to determine construct validity. In Cronbach and Meehl's (1955) views they are:

- To make scientifically clear what something is or means, so that it can be dealt with in the appropriate manner.
- To relate the laws in a nomological network to:
    - visible characteristics or quantities to each other;
    - various theoretical constructs to each other;

  - theoretical constructs to what can be observed.
- To at least have a few laws in the network that involves observables.
- To know a construct is a matter of elaborating on the nomological network in which it appears or increasing the confidence of its components.
- To know that the ground rule for adding a new construct to a theory is that it must produce laws confirmed by examination or decrease the number of nomologicals needed to predict some observables.
- To identify operations as 'different', 'overlapping' or 'measuring the same thing' if their positions in the nomological net link them to the same construct variable.

Cronbach and Meehl (1955) tried to associate the conceptual/theoretical realm with the observable one, because this is the central concern of construct validity. While the nomological network idea may work as a philosophical foundation for construct validity, it does not provide a realistic and working methodology for essentially assessing construct validity. The next phase in the development of the concept of construct validity was the development of the multitrait-multimethod matrix.

3.4.8 The multitrait-multimethod matrix

3.4.8.1 Describing the multitrait-multimethod matrix

The multitrait-multimethod matrix (MTMM) is an approach to determine the construct validity of a range of measures in a study. It was developed by Campbell and Fiske (1959) in part as an effort to provide a realistic methodology that a researcher can make use of.

Along with the MTMM, Campbell and Fiske (1959) defined two new types of validity called convergent and discriminant Validity as subcategories of construct validity. The concepts of convergent and discriminant validity were discussed extensively at the beginning of this chapter. Both convergent and discriminant validity can be tested by making use of the MTMM. In order to argue that the measures have construct validity, both convergence and discrimination should be indicated.

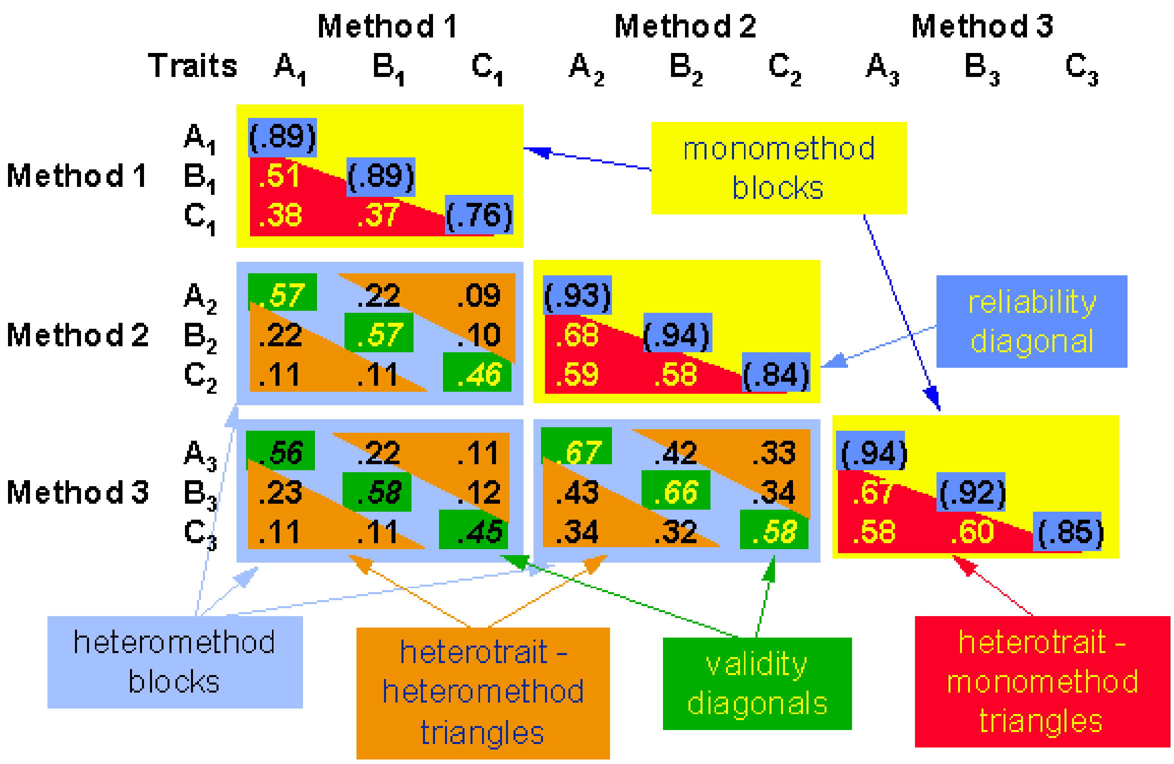


Figure 3.2: Example of the MTMM matrix

Source: Trochim (2006: 13)

Campbell and Fiske (1959) explain that the MTMM is simply a matrix of correlations set to facilitate the explanation of the assessment of construct validity. The MTMM assumes that the researcher will measure each of several concepts by each of several methods, for example, a paper-and-pencil test, a direct observation and a performance measure. The MTMM is a very restrictive methodology and ideally every concept by every method should be measured (Babbie & Mouton, 2006).

To construct an MTMM, it is Campbell and Fiske's (1959) stance that the researcher needs to position the correlation matrix by concepts within methods. Figure 3.2 shows an MTMM for three concepts (Traits A, B and C), each of which is measured with three different methods. It is imperative to note that the layout of the matrix is in blocks by method. Basically, the MTMM is a correlation matrix between measures, with the exception that instead of 1's along the diagonal, an estimate of the reliability is substituted for each measure.

To be able to interpret an MTMM, it should be understood how to identify the different parts of the matrix (Campbell & Fiske, 1959; Chronbach & Meehl, 1955). Firstly, it should be noted that the matrix consists only of correlations. It is a square symmetrical matrix, therefore there is a need to only look at half of it. Secondly, these correlations can be divided into three different kinds of shapes: diagonals, triangles and blocks.

Campbell and Fiske (1959) as well as Chronbach and Meehl (1955) claim that the specific shapes are:

- **The reliability diagonal (monotrait, monomethod)**

These are the estimates of the reliability of each measure in the matrix. The researcher can estimate reliabilities in a number of different ways. There are as many correlations in the reliability diagonal as there are measures. In the figure 3.2 above, there are nine measures and nine reliabilities. The first reliability is the correlation of Trait A, Method 1 with Trait A, Method 1. This is mainly the correlation of the measure with itself. Such a correlation should always be perfect (i.e., r=1.0).

- **The validity diagonals (monotrait, heteromethod)**

The validity diagonals are the correlations between measures of the same trait measured using diverse methods. Since the MTMM is ordered into method blocks, there is one validity diagonal in each method block. In the figure 3.2 above, the A1-A2 correlation is 0.57. This is the correlation between two measures of the same trait (A) which are measured with two diverse measures. Because the two measures are of the same concept, they are expected to correlate highly. It can also be considered to be monotrait, heteromethod correlations.

- **The heterotrait, monomethod triangles**

These are the correlations among measures that share the same method of measurement. For example in figure 3.2, A1-B1 = 0.51 in the upper left heterotrait monomethod triangle. These correlations have the method in common, not trait or concept. If these correlations are high, it is because measuring different things with the same method results in correlated measures. Therefore, it can be seen as a strong method factor.

- **Heterotrait, heteromethod triangles**

These triangles indicate the correlations that vary in both the trait and method. For instance, A1-B2 is 0.22 in the example. It can be expected to be the lowest in the matrix because the correlations do not share trait or method.

- **The monomethod blocks**

These blocks include all of the correlations that share the same method of measurement. The number of blocks is the same as the number of methods of measurement.

- **The heteromethod blocks**

These blocks include all correlations that do not share identical methods. There are (K(K-1))/2 such blocks, where K = the number of methods. In figure 3.2, there are three methods and so there are (3(3-1))/2 = (3(2))/2 = 6/2 = 3 such blocks.

3.4.8.2 Principles of interpretation

Campbell and Fiske (1959) and Babbie and Mouton (2006) believe that MTMM requires that the researcher uses his/her own judgement, even though some of the principles may not be met. In an MTMM, a somewhat strong construct validity might still be interpreted. Perfect conformances might not always be achieved with regard to the principles in different research settings, even when there is enough reason to support construct validity. Campbell and Fiske (1959) point out that a researcher with superior knowledge of MTMM can use it to locate weaknesses in measurement and can use it to determine its construct validity.

Campbell and Fiske (1959) are of the opinion that the ground rules for the MTMM are:

- Coefficients in the reliability diagonal should always be the highest in the matrix. This should indicate that a trait is more highly correlated with itself than with the other coefficients.
- Coefficients in the validity diagonals should be totally different from zero and large enough to permit further investigation. This is the main evidence of convergent validity.
- A Validity coefficient should be more elevated than values lying in its column and row in the same heteromethod block.
- A validity coefficient should be larger than all the other coefficients in the heterotrait-monomethod triangles. This mainly emphasises that trait factors should be stronger than method factors
- A similar pattern of trait interrelationship should be identified in all triangles.

3.4.8.3 Advantages and disadvantages of MTMM

The MTMM idea ensures an operational methodology for determining construct validity. There is the possibility of examining both convergent and discriminant validity all at once in one matrix by using MTMM. Campbell and Fiske (1955) stress the importance of searching for the effects of how one measures in addition to what one measure. MTMM also provides a thorough framework for assessing construct validity.

Despite these advantages, MTMM has received little use since its introduction in 1959. There are several reasons:

- First, in its original form, MTMM requires a fully crossed measurement design. This means that each of the various traits should be measured with each of the various methods. In some research contexts, it is not possible to measure every trait with as many methods as possible. In most applied social research, it is not feasible to make methods a definite part of the research design.

- Secondly, the critical nature of the MTMM may have caused opposition to its wider adoption. Many researchers searched for a test to measure construct validity that could lead to a single statistical coefficient, namely the equivalent of a reliability coefficient. But it was not possible to use the MTMM strategy to measure the grade of construct validity in a study (Campbell & Fiske, 1955).
- Finally, Campbell and Fiske (1955) were of the opinion that the judgement needed to make a decision about the nature of MTMM meant that different researchers could lawfully come to a variety of different conclusions.

3.4.8.4 A modified MTMM

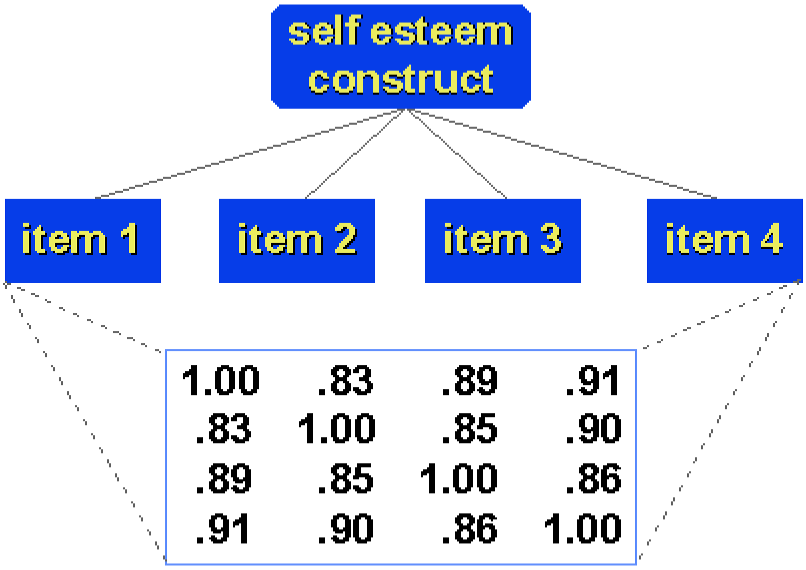


Figure 3.3: Evidence of the correlations of a similar construct

Source: Trochim (2006: 15)

As mentioned above, one of the most complex aspects of MTMM from an execution point of view is that it needs a design that includes all combinations of traits and methods. But to determine convergent and discriminant validity, the method factor is not needed. If every item accurately reflected the construct of internal locus of control, then the items would be expected to be highly intercorrelated. These strong intercorrelations are proof of convergent validity.

Discriminant validity, on the other hand, is the standard by which measures of theoretically dissimilar constructs should not correlate highly with each other (Campbell & Fiske, 1955). Therefore, two different instruments that measure two diverse constructs, cross-construct correlations, should be low, as shown in Figure 3.4 (Trochim, 2006). Low correlations like this are evidence of construct validity.

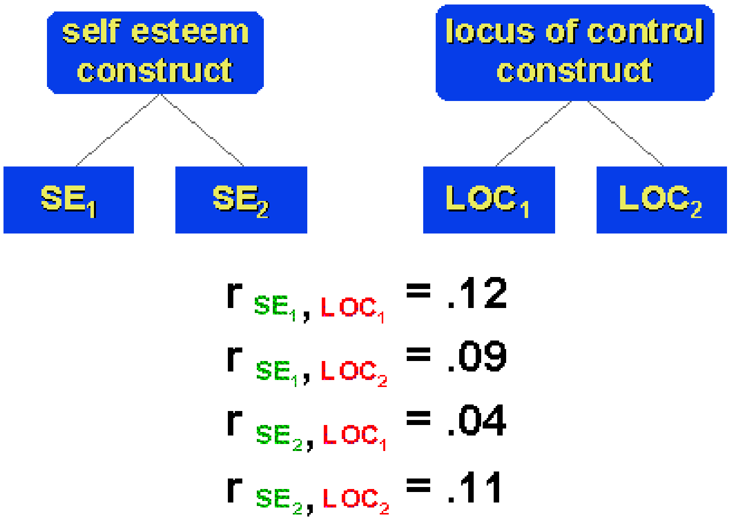
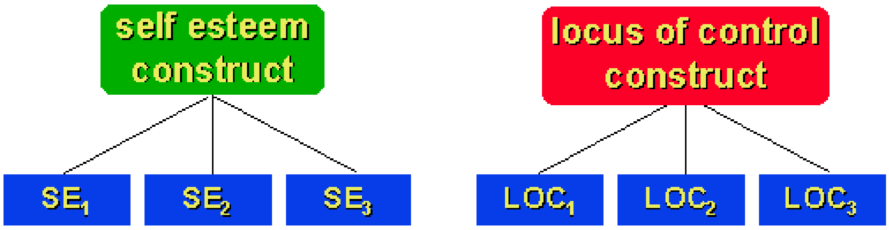


Figure 3.4: Evidence of correlations for different constructs

Source: Trochim (2006: 16)

Trochim (2006) provides evidence that this can all be put together to see how both convergent and discriminant validity can be dealt with at the same time.

The matrix examined both convergent and discriminant validity but it only explicitly looked at construct intra- and interrelationships. Trochim (2006) pointes out that it can be seen that the MTMM concept has two main themes. The first is the idea of looking concurrently at the pattern of convergence and discrimination. The second idea in MTMM is the emphasis on methods as a factor possessing the potential to be a confounding factor.



	SE_1	SE_2	SE_3	LOC_1	LOC_2	LOC_3
SE_1	1.00	.83	.89	.02	.12	.09
SE_2	.83	1.00	.85	.05	.11	.03
SE_3	.89	.85	1.00	.04	.00	.06
LOC_1	.02	.05	.04	1.00	.84	.93
LOC_2	.12	.11	.00	.84	1.00	.91
LOC_3	.09	.03	.06	.93	.91	1.00

Figure 3.5: Evidence of two different constructs

Source: Trochim (2006: 17)

3.4.9 Pattern-matching for construct validity

3.4.9.1 Idea of a pattern

When carefully examining the thoughts about construct validity, construct validity can be seen as the foundation of the nomological network as well as the MTMM, and every other method where statistics should be interpreted. One of the major themes identified in these types of methods is the idea of a pattern. When the researcher claims that the programmes or measures have construct validity, he/she is actually claiming that he/she as a researcher realises how the constructs or theories of the programmes and measures work in theory and it can be claimed that he/she can provide evidence that it behaves in practice the way he/she think it should (Trochim, 2006).

A researcher develops a theory of how the measures are connected to each other. The researcher usually provides evidence by making use of observation of the statistics that the measures actually react in a specific way in reality. In other words, this will be an observed pattern. When construct validity is identified, the researcher is essentially claiming that the observed pattern correlates with the theoretical pattern (Trochim, 2006).

3.4.9.2 The theory of pattern-matching

A pattern is any range of objects or entities. The term *arrangement* can be used here to indicate that a pattern is non-random and has potential to be described (Trochim, 1985). Marquart (1989) agrees with a lot of Trochim's (1985) ideas in his study, which mainly focuses on a pattern-matching approach to determine the construct validity of an evaluation instrument. Marquart (1989) explains that for each major area of an evaluation, programme, participants, measures and outcomes, there are theoretical and observed patterns, and the level of correlation between the theoretical and observed patterns can be described as a pattern match.

Trochim (1985) explains that the theory might either start from a formal tradition of research, or it might be the ideas or 'hunches' of the researcher. Trochim (2006) and Babbie and Mouton's (2006) stance is that all research use pattern-matching principles, although this is rarely done consciously. He notes that in survey research, pattern-matching is the basis of generalisations across various concepts or populations. In qualitative research, pattern-matching forms the basis of any attempt to carry out thematic analyses.

3.4.9.3 Advantages and disadvantages of pattern-matching

The most evident disadvantage is that pattern-matching requires that the description of the theory of the constructs should be as accurately as possible. This is not done in applied social research, but perhaps it should be done (Trochim, 2006). Possibly the more restraining assumption is that the theoretical and observed patterns can be arranged in the same way so that it can be correlated directly. Advantages of the pattern-matching approach include the following (Trochim, 2006):

- Firstly, it is more general and flexible than MTMM. There is no need to measure each construct with various methods.
- Secondly, it treats convergence and discrimination as a continuum. The concepts are somewhat similar and so their interrelations would be more or less convergent or discriminant.
- Thirdly, the pattern-matching approach enables the estimation of the overall construct validity for a group of measures in a specific context.

3.5 EXTERNAL VALIDITY

External validity involves the extent to which the results of a study can be generalised beyond the sample. In other words, can one relate what one found in one's study to other people (population validity) or settings (ecological validity) (Bracht & Glass, 1968)?

- **Ecological validity**

  Ecological validity is the degree to which the results of research can be generalised or transferred to other contexts or settings (Welman & Kruger 2001). Bracht and Glass (1968) describe *ecological validity* as the degree to which the results of an experiment can be generalised from the set of environmental circumstances created by the researcher to other environmental circumstances or settings.

  Subsequently, ecological validity cannot be ascertained because no form of field or realistic research can be done or is afforded by the chosen research design. Therefore, the threats to internal validity, namely subject and experimenter effects, will affect the extent to which the results may be generalised to the target population and thus have an adverse effect on ecological validity. However, the sampling strategy, which is discussed in following sections, may alleviate this to some extent. It should be cautioned that due to the proposed sample size, it would still be unwise to generalise the results obtained to the population at large.

- **Population validity**

The other side of external validity requires the deliberation of the population that is under study. It can be interpreted to be population validity when the results obtained for the sample of the study can be generalised to the population that is relevant in terms of the research study (Welman & Kruger, 2001; Bracht & Glass, 1968). Involvement in the study was totally voluntary and therefore the use of volunteers may affect the population validity results.

This is because of the fact that the characteristics of individuals who volunteer for research studies may not be the same as for the target group (Welman & Kruger, 2001). Subsequent generalisation to the rest of the population will not be possible because all the participants in the study did not come from the same population and even more participants had to be approached so that the minimum number of participants could be reached. External validity might thus be low.

3.6 CRITERION VALIDITY

Criterion validity indicates how well a test either determines present performance, called concurrent validity, or how well it predicts future performance, called predictive validity. *Criterion validity* is a measure of the degree to which a test is related to some criterion. In both types of criterion validity, a criterion is used as a confirmatory measure (Salkind, 2006; Foxcroft & Roodt, 2005). Foxcroft and Roodt (2005) describe it as a quantitative process, which entails the determination of a correlation coefficient between a predictor or more than one predictor and a criterion.

According to Trochim (2006), in criteria-related validity, the results of one's operationalisation are tested against some criterion. It might be tricky to distinguish this from content validity. In content validity, the criterion is the construct definition itself, it is a specific assessment. In criterion-related validity, an estimation will be made about how the operationalisation will perform based on the theory of the construct. The differences between the various criterion-related validity types are in the criteria used as the standard for making the decision. In the views of Salkind (2006), criterion validity determines the degree to which an instrument is related to some criterion.

3.7 PREDICTIVE VALIDITY

In predictive validity, the operationalisation's ability to forecast something it should be able to predict is determined (Trochim, 2006). A high correlation could provide proof of predictive validity. Salkind (2006) states that a *predictive criterion* is an assessment of how well a test predicts a criterion or in other words how well it predicts future performance.

3.8 CONCURRENT VALIDITY

Salkind (2006) explains that concurrent validity measures how well an instrument estimates a criterion. Foxcroft and Roodt (2005) agrees with Salkind and they point out that it involves the correctness with which a measure can recognise or diagnose the current behaviour or status regarding detailed skills or characteristics of an individual. In concurrent validity, Trochim (2006) states that one tries to determine the item's ability to differentiate between groups that it should be able to differentiate between. As in any discriminating test, the results are more dominant if it can be shown that there is a difference between two groups that are very similar.

3.9 CONTENT VALIDITY

Content validity refers to the degree to which the content of a test covers all the aspects of a specific trait (Foxcroft & Roodt, 2005; Salkind, 2006). It should be established whether the items used in the measure symbolise a reasonable sampling of the total items that make up the content of the concept (Price, 1997). Content validity refers to the degree to which the items that are included in a specific measurement instrument measure the same content or how well the content material was sampled in the measure (Rubio, Berg-Weger, Tebb, Lee & Rauch, 2004).

According to Murphy and Davidshofer (2001), a comprehensive description of the content domain of a test provides the basis for assessing the content validity. In order to understand content validity, it is vital to know what is meant by the term, *content*. Content usually has limits and the domain is always structured. Content validity is also very difficult to measure statistically, because it represents an opinion regarding the level to which a test provides a sufficient sample of a certain content domain (Murphy & Davidshofer, 2001).

In content validity, the operationalisation is checked against the appropriate content domain for the construct (Trochim, 2006). This approach assumes that one has an accurate and comprehensive description of the content domain, which is not always true. According to Salkind (2006), content validity shows the extent to which a test represents the group of items from which it is drawn, and it is helpful when evaluating the effectiveness of instruments that sample a particular area of knowledge.

3.10 FACE VALIDITY

Face validity is not validity in psychometric terms. It does not refer to what the test measures but to what it seems to measure. According to Foxcroft and Roodt (2005), it has to do with whether the measure looks valid to testees who have to undergo testing for a specific purpose. In face validity, the researcher can look at the operationalisation and see whether it seems like a good explanation of the construct. This is probably the weakest way to try to display construct validity.

It would clearly be feeble evidence because it is essentially a subjective decision. Trochim (2006) advises that the value of the face validity method can significantly improve by making it more systematic. It would be more credible if it is sent to a selected sample of experts with superior knowledge on the subject and they can then all report back with their feedback with regard to the measure, if it appears to be a strong or a weak measure of whatever the researcher is trying to measure.

3.11 SUMMARY

While the nomological network idea may work as a philosophical foundation for construct validity, it does not provide a realistic and working methodology for essentially assessing construct validity and the researcher would therefore not make use of the nomological network. When looking at the disadvantages discussed with regard to the MTMM, especially the time and effort needed, it will also not be used to determine the construct validity of the locus of control inventory.

Even though neither the nomological network nor the MTMM will be used, it is important to note that these are methods that one can use depending on the specific type of problem and the amount of time and resources that one has at one's disposal. Pattern-matching will be used to a certain degree while interpreting the results of the statistical analysis and it can therefore be seen as playing an important part of the research process. Next, the steps in determining construct validity will be discussed in detail. This is important because it will allow other researchers to make use of these steps when determining the construct validity of their instruments.

CHAPTER 4

STEPS FOLLOWED IN THIS STUDY TO DETERMINE THE CONSTRUCT VALIDITY OF THE LOCUS OF CONTROL INVENTORY

4.1 INTRODUCTION

In the view of Salkind (2006), there are a variety of ways in which construct validity can be determined. But there is no solitary best way to study construct validity. In most cases, it should be confirmed from a number of perspectives. Therefore, the more strategies used to demonstrate the validity of a test, the more confidence test users will have in the construct validity of that test, but only if the results provided by those strategies are persuasive.

In short, the construct validity of a psychometric test or questionnaire should be confirmed by the gathering of evidence. Salkind (2006) proposes the following: taking the unified definition of construct validity, it could be demonstrated by making use of content analysis, correlation coefficients, factor analysis, analysis of variance (ANOVA) studies demonstrating differences between differential groups or pre-test - post-test intervention studies, factor analysis, multitrait / multimethod studies, etc.

This chapter is dedicated to the steps that should be followed to determine the construct validity of the locus of control inventory. Some of the ideas that will be covered in this chapter will be discussed in short. Descriptive statistics will be performed followed by Cronbach's alpha for determining the reliability. Next, internal reliabilities will be determined for the item analysis. This will be followed by exploratory factor analysis to see how many factors should be extracted and finally, confirmatory factor analysis will be done to determine if the number of factors and the loadings of measured (indicator) variables on them conform to what is expected on the basis of pre-established theory. These steps can be used as a guideline and can be used in a different order depending on the research requirements.

4.2 STEP 1: DESCRIPTIVE STATISTICS

Salkind (2006) advises that the first step in the analysis of data is to describe the data. Describing data generally means computing a group of descriptive statistics, so called because they describe the general characteristics of a distribution of scores. In effect, they allow the researcher to form an accurate first impression of what the data looks like.

Salkind (2006) advises that it is good practice when it is ensured that the descriptive statistics for the specific instrument is set out in a table. The following statistics could be reported on: the variability of the mean, standard deviation, skewness and kurtosis, which reflects how the participants responded to the different items of the questionnaire. The standard deviation is used as a measure of variability and is the average quantity that each of the separate scores varies from the mean of the set of scores. If the size of standard deviation is larger, the scores will be more variable (Salkind, 2006).

If there are any negatively worded statements, it should be reversed. A typical difference that can be observed with regard to the negatively worded statements is that the two sets will differ in the mean values (Pallant, 2005). The reversed items do not have an effect on the standard deviation, skewness and kurtosis values. The variability will show that data collected and analysed should be spread normally. Skewness values should be less than 1, which means that the distribution does not differ to a significant extent from a normal distribution (Pallant, 2005).

4.3 INTERNAL CONSISTENCY COEFFICIENT (ALPHA)

Internal consistency refers to the degree to which the items that comprise the scale 'hang together'. In other words, do they all measure the same underlying construct? One of the most frequently used indicators of internal consistency is Cronbach's alpha coefficient. Ideally, the Cronbach alpha coefficient of a scale should be above 0.7. Cronbach alpha values are very sensitive to the number of items in the scale. With short scales, it is common to find low Cronbach values. In this case, it may be more suitable to report the mean inter-item correlation for the items (Pallant, 2005).

The alpha coefficients for each of the constructs will be tabulated. It should be able to differentiate between the highest and the lowest coefficient while others will conform to the guideline (Chao, Leary-Kelly, Wolf, Klein & Gardner, 1994; Klein & Weaver, 2000).

4.4 STEP 2: ITEM ANALYSIS

Foxcroft and Roodt (2005) point out that an item analysis phase adds value to the item development and the development of the measure in general. The purpose of item analysis is to examine each item to see whether it serves the purpose for which it was designed. Certain statistics are computed to evaluate the characteristics of each item. The statistics are then used to guide the final item selection and the organisation of the items in the measure. They also state that item analysis helps to determine the characteristics of the items, which include: how difficult an item is, whether it discriminates between good and poor performers, and what the shortcoming of an item is.

Item statistics can include the following: item-total correlations, item means and variances. The occurrence of item bias should also be tested and if detected, the source of bias should be eliminated (Van de Vijver & Leung, 1997).

4.5 STEP 3: FACTOR ANALYSIS

Factor analysis takes a large group of variables and looks for a way that the data may be minimised using less items. It does this by looking for sets of intercorrelations among a set of variables (Pallant, 2005). Factor analytic techniques are used by most researchers with the design and testing of instruments and scales. The scale developer starts with a lot of individual scale items and questions, then by using factor analytic techniques, it can be refined and lessened to form a smaller number of coherent subscales. Pallant (2005) also states that factor analysis can be used to decrease a large number of linked variables to a smaller more usable number, prior to using them in other analyses such as multiple regression or multivariate analysis of variance. The researcher makes use of this method to group those items that seem to measure the same construct.

According to Pallant (2005), there are two main approaches to factor analysis: exploratory and confirmatory. Exploratory factor analysis is often used in the early stages to research, find information about or explore the interrelationships among a set of variables. Confirmatory factor analysis is the opposite and therefore a more complex and sophisticated set of techniques is used later in the research process to test specific hypotheses or theories concerning the structure underlying a set of variables. Both these methods will be discussed in detail, see 5.6: Step 4 and 5.8: Step 6 of chapter 5.

The term *factor analysis* includes a variety of different and related techniques. One of the main differences is between what is termed principal components analysis and factor analysis. Pallant (2005), states that both methods try to produce a smaller number of linear combinations of the original variables in a way that includes most of the variability in the pattern of correlations. They do, however, differ in a number of ways. In principal components analysis, the original variables are transformed into a smaller set of linear combinations, with all of the variance in the variables being used. In factor analysis, however, factors are estimated using a mathematical model where only the shared variance is analysed (Tabachnick & Fidell, 2001)

4.5.1 Steps involved in factor analysis

There are four main steps involved in conducting factor analysis:

- Step 1: Assessment of the suitability of the data for factor analysis
- Step 2: Factor extraction
- Step 3: Map test and parallel analysis
- Step 4: Methods for fitting the factor analysis model

4.5.1.1 Step 1: Assessment of the suitability of the data for factor analysis

There are two major issues to take into account when determining whether a certain data set is usable for factor analysis, namely sample size and the strength of the relationship between the variables or items (Pallant, 2005). Recommendations with regard to sample sizes are the larger the better. In small samples, the correlation coefficients among the variables are not as reliable as in larger samples. Tabachnick and Fidell (2001) reviewed this issue and suggest that it is ideal to have at least 300 cases for factor analysis. However, they also agree that a smaller sample size, e.g. 150 cases, should be enough if solutions have several high-loading marker variables (above 0.8).

Stevens (1996) suggests that the sample size requirements proposed by researchers have been reduced because of more and more research that has been done on the topic. Some authors suggest that it is not the overall sample size that is of importance but rather the ratio of subjects to items. Nunnally and Bernstein (1987) advised a ten to one ratio: that is 10 cases for each item to be factor analysed. Others suggest that five cases for each item are sufficient in most cases (Tabachnick & Fidell, 2001). The second issue to be dealt with has to do with the strength of the intercorrelations among the items. Tabachnick and Fidell (2001) recommend a search of the correlation matrix for confirmation of coefficients greater than 0.3. If few correlations above this level are found, then factor analysis may not be suitable.

The Kaiser-Meyer-Olkin measure and the Bartlett test of sphericity are used to examine the sampling competence (Kim & Mueller, 1978). The Kaiser-Meyer-Olkin index ranges from 0 to 1 and the guideline advises that values close to 1 point out that factor analysis will be useful with the dat:. 0.6 is recommended as the minimum value for a good factor analysis (Tabachnick & Fidell, 2001). Values lower than 0.5 indicate that factor analysis will not be useful. The Bartlett guideline is that values less than 0.05 indicate that there are most likely noteworthy relationships among the variables. Values higher than 0.10 indicate that the data is not appropriate for factor analysis.

4.5.1.2 Step 2: Factor extraction

There are various techniques that can be used to aid in the decision with regard to the number of factors to keep: Kaiser's criterion, the scree test and parallel analysis. There are a few ways to determine how many factors are appropriate for the data. Usually there should be a much smaller number of factors than variables. Several variables should load on each factor before one can really trust the factor to be significant (Garson, 2008; Habing, 2003).

Mardia, Kent and Bobby (1979) point out that there is a restriction on how many factors one can have and still come up with a model that is easier to interpret than one's raw data. The figure is the difference between the number of unique values in one's data's correlation matrix and the number of parameters in the factor model. According to Habing (2003), the smallest number of variables required for various numbers of factors is as depicted in Table 4.1.

Table 4.1: Minimum number of variables required for different numbers of factors.

Number of Factors	2	3	4	5	6
Variables Required	5	7	8	9	11

Factor extraction includes determining the least number of factors that can be used to best describe the interrelations among the group of variables. There is a selection of methods that can be used to look for the number of underlying factors or dimensions. The most frequently used approach is principal components analysis (Pallant, 2005).

It is up to the researcher to decide on the number of factors that he/she thinks best describes the underlying relationship between the variables. This involves balancing two contradictory needs, firstly the need to uncover a simple resolution with as little factors as possible and secondly, the need to describe as much of the variance in the basic data set as possible. Tabachnick and Fidell (2001) recommend that researchers should assume an exploratory approach, experimenting with various numbers of factors until a suitable solution is found.

Pallant (2005) points out that there are a few techniques that can be used to aid in the decision regarding the number of factors to retain. These are:

- Kaiser's criterion
- Scree test
- Parallel analysis
- Horn

Kaiser's criterion

One of the most frequently used techniques is known as Kaiser's criterion or the eigenvalues rule. When using this regulation, only factors with eigenvalues of 1.0 or more are reserved for further exploration. The eigenvalues of a factor represent the quantity of the total variance described by that factor. But a drawback of this method is that it results in the retention of too many factors in some situations (Pallant, 2005; Garson, 2008).

When using Kaiser's criterion or eigenvalues it is advised to take as many factors as there are eigenvalues larger than 1 for the correlation matrix. Hair, Anderson, Tatham & Black (1998) state that this method is suitable if there are 20 to 50 variables, but it usually takes too little if there are less than 20 and too many if there are more than 50. According to Stevens (2002), it usually takes too many factors if there are more than 40 variables and their communalities are 0.4. Most importantly, it tends to be accurate with 10 – 30 variables and their commonalities are around 0.7.

Scree test

An additional approach that can be used is Cattell's scree test (Cattell, 1966). This method involves plotting all of the eigenvalues and then searching the plot to find a tip at which the contour of the curve changes direction and becomes flat (Garson, 2008). Cattell (1966) recommends that one should keep all factors higher than the break in the plot, as these factors add the most to the explanation of the variance in the data set.

A scree plot it takes the number of factors equivalent to the last eigenvalue before it begins to level off. Hair et al. (1998), report that it sometimes keeps one or two factors more than Kaiser's criterion. Stevens (2002) concludes that both Kaiser and scree are precise if $n > 20$ and communalities are larger than 0.6.

Parallel analysis

This procedure is used when making decisions on the number of components, which involves taking eigenvalues from random data sets that match the actual data set with regard to the number of cases and variables (O'Connor, 2000). This method will be discussed in more detail in the next step of determining factor analysis.

Horn's method

Horn's method ensures the most precise estimation of the amount of true factors in a complex data set. Horn's method includes contrasting the eigenvalues of a correlation matrix of random uncorrelated variables with those of the original data set. The random data set will be based on a similar sample size and number of variables. Factors of the original matrix, which have eigenvalues larger than those of the comparison random matrix, should be kept (Horn, 1965).

4.5.1.3 Step 3: Map test and parallel analysis

O'Connor (2000) states that researchers making use of factor and principal components analyses are urged to make decisions on a variety of technical issues. These issues include deciding on the number of factors to keep, extraction and rotation techniques and the procedure for determining factor scores. It should be noted that it is crucial to make a decision with regard to the number of factors to retain.

Some problems that emerge when non-optimal numbers of factors are extracted are:

- Conflicting findings;
- Under-extraction forcing the variables into a small factor space, which can lead to a loss of vital information;
- Disregarding potentially significant factors;

- A distorted combination of two or more factors;
- Error loadings that are increasing;
- Over-extraction spreading variables across a large factor space, which has the potential to result in factor splitting, in factors with few high loadings and in researchers giving excessive substantive importance to trivial factors (Wood, Tataryn & Gorsuch, 1996; Zwick & Velicer, 1986).

Many users of software packages simply trust the default decision rule, which typically is the eigenvalues greater-than–one rule. Other researchers examine scree plots of eigenvalues, which are also accessible in accepted statistical packages, such as SPSS and SAS. Regrettably, these highly popular judgement rules are showing results that are problematic. The eigenvalues greater-than-one rule usually overestimates, and occasionally underestimates, the number of components (Zwick & Velicer, 1986).

The scree test has been a powerfully promoted optional rule of thumb (Cattell & Vogelmann, 1977). But using this method requires eyeball searches of plots with sharp differences between the eigenvalues for main and insignificant factors. Such differences do not always appear in practice. Therefore, the reliability of scree plot interpretations is low, even with most professionals and experts (Crawford & Koopman, 1979; Streiner, 1998).

There are however increasing agreement among experts that two less well-known procedures namely parallel analysis and Velicer’s minimum average partial (MAP) test, are of a better quality when compared with other procedures and that they provide optimal solutions to the number of components to extract dilemma (Wood et al., 1996; Zwick & Velicer, 1986). The MAP and Parallel methods are based on statistical methods, rather than being subjective rules of thumb.

The MAP test:

Velicer's (1976) MAP test involves a total principal components analysis, which is followed by the assessment of a series of matrices of partial correlations. On the first step, the first principal component is taken out of the correlations between the variables of interest, and the average squared coefficient in the off-diagonals of the consequential partial correlation matrix is determined. The second step consists of steps to ensure that the first two principal components are partialed out of the basic correlation matrix and the average squared partial correlation is determined again. These procedures are repeated for K - 1 (the number of variables minus one step).

According to O'Connor (2000), the average squared partial correlations that are identified will then be lined up, and the number of components is decided on by the step number in the analyses that led to the lowest average squared partial correlation. The average squared coefficient in the first correlation matrix is also determined, and if this coefficient is lesser than the lowest average squared partial correlation, then this would mean that none of the components should be extracted from the correlation matrix. O'Connor (2000) explains that according to the statistics, the components are kept as long as the disparity in the correlation matrix shows systematic differences. Components are not retained when there is proportionately more unsystematic differences than systematic differences.

Parallel analysis

Another technique that is gaining popularity among researchers and particularly in the social science literature is Horn's parallel analysis (Horn, 1965; Choi, Fuqua & Griffin, 2001; Stober, 1998). Garson (2008) also states that Parallel analysis concerns comparing the size of the eigenvalues with those identified from a randomly generated data set of a similar size. In this technique, only those eigenvalues that exceed the corresponding values from the random data set are kept. According to O'Connor (2000), the eigenvalues copied from the actual data are then compared with the eigenvalues identified from the random data.

Garson (2008) also explains how parallel analysis works in more detail. He states that parallel analysis identifies the factors which are larger than random. The actual data is factor analysed, a factor analysis of a matrix of random numbers, which is indicative of the same number of cases and variables is done independently. For both actual and random solutions, the number of factors on the *x*-axis and growing eigenvalues on the *y*-axis is plotted. The number of factors to extract is from the spot where the two lines overlap. This approach to identifying the accurate number of components to retain has been shown to be the most exact. Both Kaiser's criterion and Catell's scree test usually overestimate the number of components to extract (Hubbard & Allen, 1987; Zwick & Velicer, 1986).

Horn (1965) also made his views known about the procedure, by stating that the mean eigenvalues from the random data set were used as the comparison baseline, whereas a more recent suggested practice is to use the eigenvalues that match up to the desired percentile (95^{th}) of the distribution of random data (Cota, Longman, Holden, Fekken, & Xinaris, 1993; Glorfeld, 1995). Factors are kept as long as the *i*'th eigenvalues from the real data is larger than the *i*'th eigenvalues from the random data.

4.5.1.4 Step 4: Methods for fitting the factor analysis model

Principal axis factoring

This is a type of factor analysis which looks for the least amount of factors that can explain the common variance of a set of variables, whereas the more frequent principal components analysis in its full form looks to identify the set of factors that can explain all the general and unique variance in a set of variables, therefore the specific plus error (Garson, 2008). Principal axis factoring uses a principal components analysis approach but applies it to a correlation matrix in which the diagonal elements are not 1's as in principal components analysis, but iteratively-derived estimates of the communalities (Garson, 2008).

Principal axis factoring is ideal for purposes of modelling, especially for structural equation modelling (SEM). This method accounts for the co-variation between variables. It is possible to add variables to a model without upsetting the factor loadings of the original variables in the model and this is only achievable with principal axis factoring. For most datasets, principal components analysis and principal axis factoring will lead to similar substantive conclusions (Wilkinson, Blank & Gruber, 1996). Principal components analysis is mostly preferred for purposes of data reduction, while principal axis factoring is mostly chosen when the research purpose is identifying the data structure or for causal modelling.

Factor loadings

Garson (2008) claims that the factor loadings are the correlation coefficients among the variables and factors. The squared factor loading is the percentage of variance in the indicator variable that is explained by the factor. To determine the percentage of differences in all the variables accounted for by each factor, the total of squared factor loadings for that factor can be divided by the amount of variables.

To interpret the loadings, a general rule in confirmatory factor analysis is that the loadings should ideally be 0.7 or higher to verify that independent variables identified are represented by a factor, on the rationale that the 0.7 level associates with half of the variance in the indicator that is being described by the factor. A high standard that most data will not meet is 0.7, therefore some researchers use a lower level such as 0.4 (Raubenheimer, 2004).

Direct oblique rotation

Once the number of factors has been identified, the step that follows is to try to understand them. To aid in this process, the factors are rotated. The rotation does not modify the underlying solution; it actually gives a visual description of the pattern of loadings in a way that is easier to interpret (Pallant, 2005). There are two major approaches to rotation. This can include orthogonal (uncorrelated) or oblique (correlated) factor solutions. According to Tabachnick and Fidell (2001), orthogonal rotation provides solutions that are easier to understand and to describe. This requires the researcher to presume that the underlying constructs are independent. Oblique approaches permit the factors to be correlated, but they are not so easy to interpret, describe and report (Tabachnick & Fidell, 2001).

In practice, the two different methods often result in solutions that are much the same, especially when the pattern of correlations among the items is clear (Tabachnick & Fidell, 2001). Many researchers perform both orthogonal and oblique rotations and then only report the one that gives the clearest result or they look for the easiest score to interpret. What typically should be looked for is a basic understandable structure (Thurstone, 1947). A basic structure is where each of the variables will only load on one component very strongly, and each component would be represented by a number of powerfully loading variables to obtain the hypothetical factor solutions of the locus of control inventory. In accordance with the rational construct approach, the number of defined theoretical constructs will be used to establish the number of factors for rotation purposes (Owen & Taljaard, 1996; Tabachnick & Fidell, 1989)

Rotation is used to ensure that the output is more understandable and it is frequently necessary to facilitate the understanding of factors. According to Garson (2008), the sum of eigenvalues is not affected by rotation, but rotation will change the eigenvalues of particular factors and will alter the factor loadings. Alternative rotations may clarify and lead to the same variance but have different loadings. The factor loadings are used to perceive the meaning of factors, and this means that different interpretations may be attributed to the factors depending on the rotation. This is a dilemma that is often identified as a drawback for factor analysis (Garson, 2008).

The oblique rotations permit the factors to be correlated, and so a factor correlation matrix is developed when oblique is requested. In oblique rotation, not only a pattern but also a structure matrix is generated (Pallant, 2005). The structure matrix can also be seen as the factor loading matrix, which indicates the variance in a measured variable explained by a factor on both a unique and general contributions basis. But in comparison with this, the pattern matrix contains coefficients which only represent unique offerings. The more factors, the lower the pattern coefficients will be because more general contributions to variance will be provided. When interpreting oblique rotation, the researcher will look at both the structure and pattern coefficients before deciding to attribute a label to a factor (Pallant, 2005).

An oblique rotation also keeps the independence between the factors, but these methods should be interpreted with caution (Garson, 2008). The output from Oblimin rotation will allow the researcher to determine how strongly inter-correlated the factors actually are. Usually, an orthogonal method such as Varimax is chosen and no factor correlation matrix is created because the correlation of any factor with another factor is zero. The Varimax method tries to minimise the number of variables that have high loadings on each factor. It is imperative to check the assumptions before deciding which statistical test is appropriate. The assumptions are:

- Valid attribution of factor labels.

  There is subjectivity involved in imputing factor labels from factor loadings. The researcher may decide to include a panel of subject experts in the imputation process (Garson, 2008).

- No selection bias/ proper specification.

  The omission of relevant variables and the inclusion of irrelevant variables in the correlation matrix will have an impact on the factors which are uncovered (Kim & Mueller, 1978).

- No outliers.

  Cases which are multivariate outliers should be identified and removed from the analysis prior to factor analysis. The researcher can make use of Mahalanobis distance to construct a dummy variable set, and to then regress this copy of the variable set on all other variables. If the regression is not significant, then the outliers are evaluated to be random and there will be fewer threats associated when deciding to retain them (Garson, 2008).

- Interval data is assumed.

  Kim and Mueller (1978) noted that the ordinal data may be used if it is considered that the assignment of ordinal categories to the data do not distort the core metric scaling.

- Linearity.

  Nonlinear transformation of selected variables may be a preprocessing step. It is important to screen data for linearity especially when there is a small sample size (Garson, 2008).

- Multivariate normality of data is necessary for associated significance tests.

  A less used variant of factor analysis, maximum likelihood factor analysis, presumes multivariate normality. It becomes more important to screen the data when the sample size is smaller.

- Homoscedasticity.

  According to Field (2005), homoscedasticity is an assumption in regression analysis that the residuals at each level of the predictor variables have similar variances. To put it another way, at each point along any predictor variable, the spread of residuals should be fairly constant. Since factors are linear functions of measured variables, homoscedasticity of the relationship is supposed. However, homoscedasticity is not considered to be a critical assumption of factor analysis (Garson, 2008).

- Orthogonality.

  This means perpendicular to something. It tends to be equated to independence in statistics because of the connotation that perpendicular linear models in geometric space are completely independent (Field, 2005). The distinctive factors should not be correlated with each other or with the general factors (Garson, 2008).

- Underlying dimensions shared by groups of variables are supposed.

  If this assumption is not met the 'garbage in, garbage out' (GIGO) principle will be applicable. Factor analysis cannot produce valid dimensions if there is not any that exists in the input data (Garson, 2008).

- Moderate to moderate high intercorrelations exclusive of multicolinearity

  Garson (2008) explains that using factor analysis as the application to a correlation matrix with only low intercorrelations will necessitate a solution with almost as many principal components as there are original variables, thereby eliminating the data reduction purposes of factor analysis. But too high intercorrelations may point towards a multicolinearity problem and collinear terms should be joint or otherwise eliminated before conducting factor analysis.

- No perfect multicollinearity.

  Singularity in the input matrix, which can also be called an ill-conditioned matrix, appears when two or more variables are completely redundant. Singularity stops the matrix from being inverted as well as stopping it from providing an appropriate solution.

- Face validity is a necessity for factor interpretations and labels or else it should be entrenched in theory.

  It is not easy to assign valid interpretations to factors. Garson (2008) advises that a suggested practice is to have an expert panel that are not part of the research project, and to request this panel to assign the items to the factor labels. A ground rule is that at least 80% should be accurate.

- Adequate sample size.

  For the minimum to be reached there must at least be more cases than factors (Garson, 2008).

4.6 STEP 4: EXPLORATORY FACTOR ANALYSIS

Tabachnick and Fidell (2001) recommend that researchers assume an exploratory approach, experimenting with diverse numbers of factors until an acceptable solution is found. The goal of factor analysis is to describe the variance in the observed variables in terms of underlying hidden factors. Factor analytic models are sometimes illustrated by path diagrams. For example, each latent variable can be indicated by making use of a circle, and each visible variable is indicated by making use of a square. An arrow would typically indicate causality. Some authors refer to diverse types of factor analysis such as R-factor analysis or Q–factor analysis (Habing, 2003). This refers to what is serving as the variables and what is serving as the observations. This study will only refer to R–factor analysis.

With R–factor analysis, the focus is on searching for latent factors that lie behind the variables (Habing, 2003). This enables the grouping of diverse test questions, which appears to be measuring a similar underlying construct. To establish if the data is appropriate, a considerable number of correlations should be > 0.3. Other methods that will also be used are: Bartlett's test of sphericity and the kaiser–meyer-olkin measure (KMO). For the sample size, a general rule is that there should at least be 50 observations and at least five times as many observations as variables; 150 testees will then be appropriate for this specific research problem (Habing, 2003). In this specific study, three variables are being measured.

Limitations

Exploratory factor analysis has the following limitations. If the predicted factor solution is not found, it is likely that one of two interpretations is possible. Firstly, there are real differences between the actual and predicted factors structure or secondly, the EFA was not able to reveal the hypothesised structure. The most troublesome difficulty with EFA is choosing the most suitable underlying dimensionality (Harvey, Billings, Nilan, 1985).

4.7 STEP 5: INTERPRETATION OF FACTOR ANALYSIS

As soon as the factor loadings matrix is completed, it is necessary to interpret the factors. Significant loadings can be underlined or highlighted to make interpretation easier. Significance can be measured in two ways. Firstly, practical significance is used to determine if the factor loadings are big enough so that the factors actually have a significant effect on the variables. Hair et al. (1998) recommend the following guidelines for practical significance:

- +/- 0.3 Minimal
- +/- 0.4 More important
- +/- 0.5 Practically significant

Secondly, statistical significance should also be determined. This means that the loadings should be statistically significantly different from zero. Stevens (2002) provides guidelines based on sample size, as indicated in Table 4.2

Table 4.2: Sample size

n	50	100	200	300	600	1000
Loading	0.722	0.512	0.384	0.298	0.210	0.162

There should be at least three variables loading on each factor but for more accurate results additional loadings are preferred (Hatcher, 1994).

4.8 STEP 6: CONFIRMATORY FACTOR ANALYSIS

According to Harvey et al. (1985), the use of CFA allows the researcher to prepare and directly test rival hypotheses with regard to the underlying factor structure. A confirmatory factor analysis is done to establish if the theoretical model fits the observed data or, in more simple terms, to determine if the number of factors and the loadings of measured variables react as expected according to pre-established theory (Garson, 2008). Confirmatory factor analysis can be used to verify if the indicators divide themselves into factors to match how well the indicators have been linked to the latent variables. Indicator variables are chosen on the basis of past theory and then factor analysis is used to see if the variables load as expected on the identified number of factors.

Garson (2008) agreed with Kim and Mueller (1978) by mentioning that the total number of factors in the model is mostly hypothesised beforehand; it can therefore be seen as an essential requirement for Confirmatory factor analysis. A model is estimated for the data which is used to obtain unknown parameters, and fit statistics are used to evaluate the adequacy of the model. A few researchers state that a value of 0.90 is considered to be a good quality fit for the following indices: NNFI, CFI, IFI and GFI. One needs to look at the values and determine if they meet the guideline to be acceptable. The RMSEA should have a value at or below 0.05 for a healthy fitting model or else at or below 0.08 for a reasonably fitting model (Garson, 2008; Bentler, 1990; Bentler & Bonnet, 1980; Steiger, 1995).

4.9 STEP 7: STRUCTURAL EQUATION MODELLING

There are two different approaches to confirmatory factor analysis: the traditional method and the structural equation modelling (SEM) approach. The SEM approach will be used in this study. SEM is a very general statistical modelling technique, which is widely used in the behavioural sciences. It can be viewed as an amalgamation of factor analysis and regression analysis. The reason why there is so much interest in SEM could be because of the focus on the theoretical constructs, which are indicated by the latent factors (Hox, 1995).

While SEM is typically used to model underlying relationships between latent variables, it is equally possible to use SEM to explore confirmatory factor analysis measurement models. This is done by deleting from the model all straight arrows linking latent variables, adding curved arrows demonstrating covariance between every couple of latent variables, and leaving in the straight arrows from each latent variable to its indicator variables (Garson, 2008; Pearl, 2000). This measurement model is evaluated similar to other models, using goodness of fit measures produced by the Structural Equation Modelling software EQS.

The SEM process can be used to conduct confirmatory factor analysis because its main aim is to validate the measurement model by getting estimates of the parameters of the model and by determining if the model itself ensures a good fit to the data (Garson, 2008). The SEM method is used because it mainly consists of two steps: validating the measurement model and fitting the structural model. Kline (1998) urges SEM researchers to constantly check the pure measurement model underlying a complete structural equation model first, and only if the fit of the measurement model is found to be adequate, to carry on with the second step of testing the structural model by comparing its fit with that of other dissimilar structural models.

SEM ensures a very broad and convenient framework for numerical examination, which includes several conventional multivariate procedures, for instance, factor analysis and discriminant analysis. Structural equation models are frequently visualised by a graphical path diagram. The model is mostly represented in a set of matrix equations (Hox & Bechger, 1998). Maximum likelihood estimation (MLE) in SEM requires that the assumption of multivariate normality should be met, but Micceri (1989) claims that a lot of social and behavioural data may not be successful by failing to satisfy this assumption. A few studies (Amemiya & Anderson, 1990; Anderson, 1989; Satorra & Bentler, 1994) measured the toughness of the multivariate normality assumption and they came to the conclusion that the parameter estimates stay valid under the assumptions even when non-normal data is used.

4.9.1 Critiques of structural equation modelling (SEM)

Structural equation modelling consists of a variety of analytical techniques. Modern SEM software and computers made it easy to apply SEM models to different kinds of data. These developments can have a positive impact on substantive research in more applied fields, but it also makes it easy to abuse a lot of these techniques (Hox, 1995). Most critiques against the use of SEM are based on two issues.

Hox (1995) and Micceri (1989) claim that one issue is the importance of statistical assumptions and the required sample sizes. There has been much research on the significance of the normality assumption and the sample sizes required so that researchers can have confidence in the results. Since research tends to be technical, applied researchers using SEM are not constantly aware of the technicalities. Fortunately, recent handbooks and articles that were published tried to inform non-statistical users of the errors that should be steered clear of. One of these articles which can successfully be used is an article by McDonald and Ho (2002), which explains the principles and practice in reporting structural equation analysis.

Hox (1995) states that the second more imperative issue with regard to the causal use of SEM is the concern of causal analysis. Most applications of SEM are on data that is not experimental. Many applications of SEM nevertheless understand the final model as a causal model. This may be truthful, but there is of course nothing in SEM that magically changes correlational data into causal conclusions (Hox, 1995). Cliff (1983) explains a few struggles that are the outcome of causal interpretation of correlational data. This evaluation of competing models plays a significant role and could lead to failing to discard a model.

4.10 STEP 8: MODEL EVALUATION AND FIT

Fitting a model to data involves solving a set of equations. On the one hand, there is the model with its parameters, which values one would like to determine. On the other hand, there are the sample statistics that one knows to be high quality estimates of the matching population values. The basic model in statistical modelling is: Data = Model + Error.

SEM software uses difficult algorithms, which take full advantage of the fit of the model and which take all the model limitations into account (Hox, 1995).

4.10.1 Model fit: goodness of fit indices

Statistical tests for model fit have the difficulty that their authority varies with the sample size. If one has a very big sample, the chi-square will nearly always be significant. Thus, with large samples, one will constantly reject one's model, even if the model describes the data very well. Conversely, with a very petite sample the model will always be accepted, even if it fits rather badly (Garson, 2008). Given the sensitivity of the chi-square statistic for sample size, researchers have anticipated a variety of optional fit indices to measure the model fit. Ullman (2001) also believes that fit indices have been designed to solve this problem.

The first step in the research process is to describe the theoretical factor model. This includes choosing the number of factors to be used and explaining the nature of the loadings between the factors and variables. Maximum likelihood estimation (MLE) can be used to assess the coefficients. MLE is the most universal model-fitting method used as it selects estimates that show the most likely chance of reproducing the observed data (Garson, 2008). A limitation of MLE, however, is that it is perceptive to changes from normality (DeCoster, 1998).

The suitability of the model is then determined by making use of goodness of fit measures. Goodness of fit indicators determine whether the model being tested should be accepted or rejected. The EQS programme (Bentler, 1989) can be used for SEM procedures with maximum likelihood (ML) estimation. The EQS is programmed to issue at least 10, different goodness of fit measures.

4.10.2 Chi-square

The model chi-square is used as it is the most frequent goodness of fit test because it is simple to compute and understand. There are two kinds of chi-square tests. The first is called a one-way analysis and the second is called a two-way analysis. With the chi-square, the disparity between the observed data and the proposed model is tested (Garson, 2008). Garson (2008) explains that the reason for both analyses is to decide whether the observed frequencies differ from the frequencies that one would anticipate by chance.

The observed cell frequencies are prearranged in rows and columns, this can characteristically be called a contingency table. In the views of Garson (2008), the chi-square statistic is the total of the contributions from each of the separate cells. Every cell adds something to the overall chi-square statistic. In order for the model to fit the data, the discrepancy between the two should be minimal. This means that the chi-square should not be significant if there is a high quality model fit, while a significant chi-square provides evidence of a lack of suitable model fit (Garson, 2008; Pallant, 2005)

If the sample size is too small, the error terms will be large, which makes it hard to detect a discrepancy between the model and the data (Hox & Bechger, 1998). With large samples, the chi-square will nearly always be noteworthy. When a chi-square is high, one should visually scrutinise the table to determine which cells are to blame.

The optional indices consider not only the fit of the model but also its usability. Even though the goodness of fit indices has been utilised by most researchers, all goodness of fit measures link to some function of the chi-square and degrees of freedom. These alternative indices consider not only the appropriateness of the model but also its ease of use. Even though the goodness of fit indices still rely on sample size and distribution, the reliance is much smaller than that of the routine chi-square statistic (Gravetter & Wallnau, 2000).

The general guideline is that values of 0.05 and smaller indicate a close fit between the theoretical model and observed data. Values of 0.08 and smaller indicate a reasonable fit and values greater than 0.08 or equal to 1 indicate a poor fit and the fact that the model clearly needs work. A non-significant chi-square value indicates a good model fit (Garson, 2005; Pallant, 2005). Carmines and Mclever (1981) maintain that the relative chi-square should be in the 2:1 to 3:1 range for a satisfactory model. Jaccard and Wan (1996) recommend the use of at least three goodness of fit measures. Kline (1998) recommends at least four tests. The following indices of model fit can be used:

- The non-normed fit index (NNFI) (Bentler & Bonnet, 1980)
- The comparative fit index (CFI) (Bentler, 1989,1990)
- The incremental fit index (IFI) (Bollen, 1989)
- The root mean square error of approximation (RMSEA)

An NNFI close to 1 will be an indication of a good fit. According to the standard, the IFI should be equal to or greater than 0.90 for the model, to be accepted, but it can also be greater than 1.0 for the model to be accepted. If the RMSEA is less than or equal to 0.08 a good fit is achieved. Hu and Bentler (1998) recommend that the RMSEA should be smaller than or equal to 0.06 as a cut-off for an acceptable model fit. RMSEA is a trendy measure of fit and works with a better venue of autonomy (Garson, 2008; Steiger, 1995).

The odds of getting a non-significant chi-square become minute with large sample sizes. The ratio of chi-square to degrees of freedom has been projected, although it appears to suffer from arbitrary principles of explanation (Kelloway, 1998). The CFI, NNFI and the IFI are considered to be relatively vigorous in respect of the effects of sample size (Bentler & Bonnet, 1980).

The chi-square has a few disadvantages as a fit index. The first disadvantage involves the fact that it has no clear upper bound and it is not normed so that the values can fall between 0 and 1 (Hoyle & Panter, 1995). It is advised to handle this shortcoming by dividing the chi-square by its degrees of freedom. This method proves to be successful because it is used under a null hypothesis where the chi-square will be approximately equal to its degrees of freedom. A ratio greater than 3 indicates important lack of fit (Kelloway, 1998).

The second disadvantage of the chi-square is that it is particularly sample size dependent, as discussed above. This is the most important problem that led to the development of so many different fit indices. Hoyle and Panter (1995) describe it as follows: "Statistical test of the lack of fit resulting from over-identifying boundaries placed on a model. Contrary to most research the chi-square evaluates the fixed rather than the free parameters in a structural equation model."

4.10.3 Modification indices

There are mainly two reasons for changing a SEM model: to test hypotheses and to improve fit (Ullman, 2001). If the fit of a model is not sufficient, it can be modified, by taking out the parameters that are not significant and adding some parameters that improve the fit of the model. To assist in this process, most SEM software can compute modification indices of each fixed parameter. The value of a given modification index is the smallest amount that the chi-square statistic is expected to decrease if the corresponding parameter is freed. Researchers sometimes use information like this to perform model modifications (Hox, 1995). At each step, a parameter is freed, which is responsible for the biggest enhancement in fit, and this process is continued until a satisfactory fit is reached.

Ullman (2001) points out that SEM is a confirmatory technique; this means that when model adjustment is done to optimise fit, the analysis changes from confirmatory to exploratory. A danger that can be linked to sequential model modification is the capitalisation on chance properties of the sample. Advice provided by Hox (1995) is to use modifications only when there is a theoretical validation for them.

Simulation research has shown that model adjustment often fails to identify the correct mode (Spirtes, Richardson, Meek, Scheines & Glymour, 1998) and that models so achieved do not cross-validate very well (MacCallum, 2003; MacCallum, Roznowskei & Necowitz, 1992). Because of this Ullman (2001) advises that when conclusions are drawn from a model that has been through substantial modification, the outcomes should be viewed with extreme caution, and it is essential to ensure that cross-validation should be performed whenever possible.

4.11 SUMMARY

The steps that can be followed when determining the construct validity of the Locus of control inventory were investigated and described in this chapter. All of these methods were used during the course of the study, but only the most significant results, which led to the greatest contribution in answering the research questions and objectives, will be displayed in the results section. In the following chapter, the methodology used as well as the reasons why the specific methods were chosen will be discussed.

CHAPTER 5

METHODOLOGY

5.1 RESEARCH APPROACH

The research approach refers to the broader theoretical perspective of the research process (Creswell, 1994). Creswell (1994) states that it is important to be able to recognise and comprehend the research approach underlying any study because the selection of a research approach has an effect on the questions asked, the methods selected, the statistical analyses used, the conclusions made and the ultimate goal of the research.

For the purpose of this study, the quantitative approach was used. A *quantitative study* can be defined as an investigation into a social or human difficulty, based on testing a theory consisting of variables, measured with numbers and analysed with statistical measures to be able to decide whether the predictive generalisations of the theory is realistic (Creswell, 1994).

Quantitative research includes experiments, surveys and content analysis (Clarke & Dawson, 1999). Quantitative researchers also make use of deductive reasoning to be able to analyse and interpret their results. The researcher usually observes objectively and does not give his/her own opinion and interpretation. The data collection focuses on scales and frequency tables (Creswell, 1994).

There are some assumptions, put together by Creswell (1994), which should be met for the study to use quantitative methods. These are:

- Reality is objective, “out there” and separate from the researcher, therefore reality is something that can be researched objectively.
- The researcher should remain distant and autonomous of what is being researched.

- The values of the researcher do not obstruct, or become part of, the research; therefore the research is value free.

- Research is based mainly on deductive forms of logic and theories and hypotheses are tested in a cause and effect order.

- The goal is to expand generalisations which add to the theory allowing the researcher to predict, explain and understand some phenomenon.

The study made use of the quantitative research method, with data being gathered through the distribution of questionnaires. The collected data was analysed with the aim of discovering the construct validity of the locus of control inventory by Udai Pareek (1998) and a final report written thereafter to make the findings accessible to the public. The rationale for choosing a quantitative approach is that there is a need for valid and reliable questionnaires that can be used in the South African context.

The reasons for choosing a quantitative approach are that it is based on the positivist social science paradigm, which mainly reflects the scientific method of natural sciences (Kitchen & Tate, 2000). A deductive approach to the research process was used by this paradigm. The way in which it commenced in social science was with theories, hypotheses or research questions about a certain phenomenon. It typically requires the gathering of data from a real-world setting and the researcher then has to carry out and analyse the statistics to be able to either support or reject the hypothesis (Veal, 1997). Because of the reasons mentioned above, the quantitative approach was identified as the method that would be able to assist in reaching the objective of the research study, which is to determine the construct validity of the locus of control questionnaire.

A questionnaire was used in the quantitative methodology to be able to abstract data from the participants and eventually to do the statistical analysis of the data. Precision and control are the main strengths of the qualitative approach. Control was achieved through the sampling, design and precise and reliable quantitative measurement. A quantitative methodology and the use of quantitative data allow for statistical analysis (Welman & Kruger, 2001). The method used therefore provided answers, which have a much more stable and accurate basis than a lay person's common sense, intuition or opinion.

Quantitative research can be described as a process used where evidence is validated, hypotheses and theories are polished and sometimes technical advances are also made. Virtually all research in physics is quantitative whereas research in other scientific disciplines, for example, taxonomy and anatomy, may involve a combination of quantitative and other analytical approaches and methods (Kuhn, 1961)

In the social sciences, quantitative research is often contrasted with qualitative research. Kuhn (1961) states that *qualitative research* is the examination, analysis and interpretation of observations for the reason of identifying underlying meanings and patterns of relationships, including classifications of types of phenomena and entities, in a way that does not make use of mathematical models. According to Heidelberger (2004), approaches to quantitative psychology were first modelled on quantitative approaches in the physical sciences by Gustaf Fechner in his work on psychophysics, which built on the work of Ernst Heinrich Weber.

Some critics report the limitations that can be associated with a quantitative approach; it can denigrate human individuality and ability to think (Walle, 1997; Massey, 2003). Another researcher, Gilbert (1993), states that its mechanistic nature tends to exclude some notions of freedom, choice and moral responsibility. Gilbert (1993) and Massey (2003) state that quantitative research does not succeed in taking into account people's exceptional ability to understand their own experience as well as their ability to develop their own meanings and to act on them. It should be noted that quantitative research approaches cannot be totally objective because subjectivity is part of the process in choosing a research problem that is worthy of examination and analysis.

5.2 RESEARCH DESIGN

Clarke and Dawson (1999) as well as Babbie and Mouton (2006) explain that a research design is actually a research plan, which can be used as the architectural blueprint of the research study. Fouche and De Vos (1998) suggest that research design is a blueprint or detailed plan: "this plan, or blueprint, provides the framework with regard to which data are to be collected to examine the research hypothesis or question in the most effective manner".

The research design could therefore be identified as a checklist, which contains all the research process items needed to carry out a useful research project, for example, the population, sample, data collection method, data analysis and interpretation. Every step that is part of the research will be described in the research design (Creswell, 1994).

There are three general types of quantitative methods:

- experiments
- quasi-experiments
- surveys

A non-experimental cross-sectional survey research design was used for the purposes of this study. Creswell (1994) states that surveys comprise cross-sectional and longitudinal studies, by making use of questionnaires or interviews for data collection with the aim of determining the characteristics of a population group of interest by using a smaller sample from that population.

This type of design, as described by Ruane (2005), takes information from a particular group of people at a single point in time with no effort to follow up. When conducting a cross-sectional study, the researcher will ask a broad range of people a sequence of questions in order to deal with the topic of interest. Therefore, cross-sectional research was identified as a logical strategy for pursuing both descriptive and exploratory research projects.

This was identified as the best method to use because the researcher would not have to think about the effects of subject mortality/attrition or other difficulties linked with longitudinal designs. The disadvantage, however, lies in the fact that the design is not appropriate for researching change over a longer period of time. There were also not enough resources available at the researcher's disposal to carry out either an experimental or longitudinal study, because all expenses had to be incurred personally by the researcher and no funding was made available to the researcher.

The survey was conducted by distributing questionnaires to available and willing students currently studying at the University of Pretoria as well as postgraduate students that are currently working. This will be a once-off survey, without the intent of following up. The rationale for sending the questionnaires via e-mail was because it was identified as a convenient and expedient way, with the possibility of an appropriate response rate. The data obtained was then analysed, interpreted and reported, with the main aim of determining the construct validity of the locus of control questionnaire.

5.3 VALIDITY AND RELIABILITY OF THE RESEARCH DESIGN

To be able to evaluate and establish the suitability of the study's research design, it is necessary to consider issues of validity and reliability pertaining to the research design that was just described. This is consequently discussed in the sections that follow.

5.3.1 Internal and external validity.

A characteristic of the internal validity of a study, according to Welman and Kruger (2001), is construct validity, which has to do with the degree to which procedures planned to generate the independent variable of interest indeed succeed in generating this variable and not anything different. This is the main aim of the study and the results will be communicated at the end of the research report.

Threats to construct validity of the study will be minimised by ensuring that the subsequent statements are all kept in mind during all activities that form part of the process whereby the responses will be collected (Welman & Kruger, 2001; Metha, 2004).

Threats to internal validity for the study could include the following:

- **History:** According to Welman and Kruger (2001), the actions which take place at the same time as the intervention may have an influence on the dependent variables. This means that the participants could have been affected by uncontrollable events.

- **Selection bias:** The internal validity of the study could be endangered by selection bias because individuals will not be selected completely at random. These individuals may have features that are common in their groups but which might not occur across all groups of respondents. An effort was made to send the questionnaire to as many students as possible.

- **Other third variable problems**: It is not possible to exert control over every potential problem or intervening variable that might have a detrimental effect. This means that results of the study might be less internally valid.

- **The subject effect**: Participants might be aware that their perceptions and experiences with regard to internal and external locus of control will be measured and might therefore answer the questions in a way which will make the participants appear in a favourable light. This can be identified as a threat of measurement. While most of the respondents will not really know the researcher, it might happen that some of the respondents may want to put the results at risk or improve the results of the study.

- **The experimenter effect**: This can be identified as a possible threat in that the researcher could have certain expectations as to what he/she would like the final results to be and thus may influence and interpret data and the statistics in order to justify his/her expectations. This is a threat that the researcher should always be aware of and due to the comprehensive literature review, with various perspectives that were considered, every attempt was made to avoid this.

For more information about external validity, please refer to Chapter 3 where it was discussed extensively.

5.3.2 Reliability

Reliability was also discussed in detail in Chapter 3. The reliability of the instrument and the results achieved by the study was determined once all the data had been collected and analysed by means of the SPSS programme. Reliability has to do with the internal consistency of the instrument and therefore the level of generalisability of the items within the instrument (Welman & Kruger, 2001). The reliability of the instrument will be discussed in the results and discussion section of the report.

5.4 UNIT OF ANALYSIS

The unit of analysis refers to the 'what' one studies, what object, phenomenon, entity process or event one is focusing on to investigate (Babbie & Mouton, 2006). When this object is an object in World 1 (real-life 'object'), it is called an empirical research problem. In the view of Babbie and Mouton (2006), when this object is an object in World 2 one talks about conceptual or non-empirical difficulties. Likewise, studies that focus on developing theories and models, which are used to analyse concepts or researching the body of knowledge, are all non-empirical studies. The unit of analysis in each of these cases reside in World 2. The unit of analysis for the study was the pre- or postgraduate students who were registered at the University of Pretoria.

5.5 SAMPLING FRAME

If everyone in the population cannot be tested, then the only other choice is to select a sample or a subset of that population. Characteristics of good sampling techniques include that they maximise the degree to which this selected group will represent the population. A *population* is a group of potential participants to whom one wants to generalise the results of a study (Salkind, 2006).

The population from which the sampling frame was chosen consists of diverse pre- and postgraduate students who were enrolled for a course at the University of Pretoria. From there, a specific university residence was approached and the students were asked for their participation. This did not produce enough respondents and so postgraduate students were also approached. The sampling size of 155 students was sufficient for descriptive and exploratory research.

The biographical information on the sample is reported in Table 5.1. The sample consisted of the following: 17% males and 83% females. There were five different race groups that participated in the study. The distribution among these race groups was: 74% white, 2% coloured, 20% black, 3% caucasian and 2% indian.

Eleven different language groups participated in the study. This also showed the enormous amount of diverse cultures and groups that exist in the South African context, especially among students, recent graduates and postgraduates. The languages that were most prominent in the study were: 63% Afrikaans, 17% English, 5% Zulu and 4% Xhosa. With regard to the age groups that participated 50% lies within the 24 – 26 age group. This is closely followed by the 18 – 20 age groups.

Table 5.1: Biographical information of the respondents

	Frequency	Percentage	Valid percentage	Cumulative percentage
RACE				
White	118	76.1	76.1	76.1
Black	31	20	20	96.1
Indian	3	1.9	1.9	98.1
Coloured	3	1.9	1.9	100
Total	**155**	**100**	**100**	
GENDER				
Female	129	83.2	83.2	83.2
Male	26	16.8	16.8	100
Total	**155**	**100**	**100**	
LANGUAGE				
Afrikaans	98	63.2	63.2	63.2
English	27	17.4	17.4	80.6
South Sotho	2	1.3	1.3	81.9
Sepedi	5	3.2	3.2	85.2
Tsonga	3	1.9	1.9	87.1
Zulu	7	4.5	4.5	91.6
Setswana	4	2.6	2.6	94.2
Xhosa	7	4.5	4.5	98.7
Venda	2	1.3	1.3	100
Total	**155**	**100**	**100**	
AGE GROUP DISTRIBUTIONS				
18 - 20	47	30.3	30.3	30.3
21 - 23	43	27.7	27.7	58.1
24 - 26	50	32.3	32.3	90.3
27 - 29	13	8.4	8.4	98.7
30 - 32	2	1.3	1.3	100.0
Total	**155**	**100**	**100**	

5.6 SAMPLING STRATEGY

According to Salkind (2006), when the results can be generalised from a sample to a population then only do the results have significance beyond the limited setting in which they were originally obtained. When results can be generalised, it can be applied to diverse populations with the same features in different settings. He also states that when results cannot be generalised, the results will only be valid for those who formed part of the original research, and not for any other people.

The sampling used was non-probability sampling – convenience sampling, this means that the probability of selecting a single individual was not known. There was an attempt to include students who are currently registered at the University of Pretoria. For this kind of sampling method, since working with a convenient sample, the number of people questioned is less important than the criteria used to select them (Salkind, 2006). Therefore, the study focused mainly on asking students from the University of Pretoria to complete the locus of control Questionnaire. When using convenience sampling, the audience is a captive one, it is very convenient to generate a sample. This is an easy but not a random method, which is only representative to a limited extent (Pallant, 2005).

After targeting students at the university residences, the sample size of 150 was not yet reached and therefore postgraduate students from the University of Pretoria were also approached. This means that additional mechanisms had to be explored. This followed a form of 'quota sampling' in that those additional respondents were reached by the most accessible means to ensure that the sample size was sufficient for quantitative research. Quota sampling selects people with the characteristics one wants but does not randomly select from the population (Salkind, 2006).

The researcher continued to enlist pre- and postgraduate students until the quota of 150 was reached. This sampling technique ensured that the researcher could use a whole range of opinions or views, without being worried about is generalisability. Non-probability sampling is known for its advantage of convenience and economy

(Welman and Kruger, 2001). An alternative term for this method could be sampling for diversity.

The final sample size amounted to 155 pre- and postgraduate students. This only included those students approached on the University of Pretoria campus. The sample included those students who were available and willing to participate in the study. This method was chosen because it ensured that sufficient respondents existed to ensure the possibility of conducting statistical analysis of the data. The method might not be completely ideal for the study, but when taking into account the time constraints, expertise and resources at the researcher's disposal, this was the best method that could be identified.

A larger sample would be more representative of the population and it would also contain a smaller sampling error (Salkind, 2006). Because this is a large-scale research report a sample size of 150 students was considered sufficient for the purpose of the study. But a larger sample of participants would have produced more accurate results and if more students from the University of Pretoria were chosen, the results could have been even more focused and transferable onto the student population of the University of Pretoria.

5.7 DATA COLLECTION METHOD AND PROCEDURE

A self-report questionnaire was used. This questionnaire incorporated the desired constructs of measurement pertaining to locus of control because it was specifically developed to be used for this purpose. Questionnaires were completed anonymously, therefore honesty of answers was enhanced, and bias due to personal characteristics of the interviewer was eliminated, it is inexpensive, saves time, it is easy to compare and analyse, can be administered to many people and is standardised.

This questionnaire was electronically distributed to the participants. A questionnaire or survey was a suitable choice according to the intent of the research study, which includes that it had to be both descriptive and exploratory

(Ruane, 2005). It was decided that the items of the questionnaire should be closed-ended because of insufficient resources at the researcher's disposal for the analysis of open-ended questions. Open-ended questions are known to be extremely time-consuming and it does not constantly provide useful and accurate results (Ruane, 2005; Foxcroft & Roodt, 2005). The number of responses was too much to analyse on a qualitative basis and therefore the use of closed-ended questions suited the purpose of the study to the largest extent.

In questionnaires where rating scales are used, there is always a possibility of rating errors. That is errors in the data that makes the data less valid to use in answering a research question. The following factors might have caused rating errors: central tendency bias (this occurs when the respondent does not want to give extreme answers and instead clusters the answers around a central choice), logical errors in rating (when the respondent gives similar ratings for questions that are logically connected to the respondent) and acquiescent response set (the respondent is unwilling to give a negative answer and goes along with everything asked) (Foxcroft & Roodt, 2005).

Consent was obtained and a detailed consent form was signed by all participants. Ethical issues were also explained in the attached letter and information sheet that were sent to the candidates with the locus of control inventory. This was done to ensure that the candidates fully understand the purpose of the study and to explain to them what the researcher will do with the results.

The questionnaire was e-mailed to each participant; and where needed delivered and collected at the appropriate university residences. A specific time frame for each response was given to participants, who then e-mailed their completed questionnaires to the researcher in question. There were no requests to provide any identifying details about the participants themselves on the questionnaire. This could have included: their names, contact details, or anything that might be used to identify them. The omitting of names and other details ensured that anonymity was maintained and that anxiety or concern by the participants were minimised.

Biographical data was obtained by making use of open-ended questions. The participants were able to write down answers for variables pertaining to gender, age group and race and language. The next section of the questionnaire consisted of the questionnaire items that participants had to answer by rating it on a numerical five-point scale. This allowed for easier statistical analysis. The rating scale consisted of a standardised response set, which ensures easy analysis and comparison between different cultural or age groups. The advantages of this data collection procedure include:

- Less time was taken for data collection.
- There was an adequate response rate.
- Anonymity was obtained, which also led to more truthful responses.
- Bias because of the personal characteristics and subjective views of the interviewer was lowered.
- Quality of answers improved.

The questionnaire was the only way of data collection for the purposes of the study due to the prohibiting factors already identified in terms of time and finances and, most importantly, the reason for the study was to determine if the questionnaire measures those constructs that it is supposed to measure. It also provided for the best identified source of information for exploratory purposes of the research. A drawback of this procedure is that the researcher was only able to provide guidance to a very small extent with regard to instructions on the completion of the questionnaire.

5.8 SUMMARY

The reasons for using the quantitative research approach were clearly indicated in this chapter. The researcher discussed the research design, which was indicated as a non-experimental cross-sectional survey research design and the use of the questionnaire was also justified by stating the reasons and explaining the theory behind its use. The sampling method used was identified as non-probability sampling, convenience sampling, and the data collection method and procedure was then discussed in more detail. In the chapter that follows, the results of the statistical analysis will be presented and discussed to determine the construct validity of the locus of control instrument.

CHAPTER 6

RESULTS

6.1 INTRODUCTION

The SPSS and EQS programs were used to do the statistical analysis. All the relevant data will now be presented and linked to the theory with regard to the interpretation of relevant values. A discussion of the relevant results and a conclusion will follow. The chapter closes off with the limitations of the study and recommendations that can be made to improve on future research with regard to this subject.

6.2 DESCRIPTIVE STATISTICS

Descriptive statistics in respect of the locus of control inventory were computed for the respondents. The statistics included a descriptive analysis of the biographical data which is presented as part of the sampling frame in Chapter 7

Pallant (2005), states that *internal consistency* refers to the degree to which the items that make up the scale 'hang together'. Field (2005) also explains that *reliability* means that a scale should consistently reflect the construct that it is measuring. Therefore, it is important to determine if all the items measure the same underlying construct. One of the most commonly used indicators of internal consistency is Cronbach's alpha coefficient. Ideally, the Cronbach alpha coefficient of a scale should be above 0.7 (Pallant, 2005).

Table 6.1: Reliability coefficient

Reliability coefficient	
Cronbach's alpha	0.819

When looking at the instrument as a whole, a Cronbach's alpha of 0.819 was found and this can be interpreted as high. But when looking at the internal and external locus of control scales, respectively, one will find that the Cronbach alpha for items that measure internal locus of control is 0.611 and the items that measure external locus of control amounts to 0.873.

6.3 EXPLORATORY FACTOR ANALYSIS

As discussed in Chapter 5 factor analysis takes a large set of variables and identifies ways that the data may be reduced or summarised using a smaller set of factors or components. This is done by searching for certain values within the inter-correlation table of a set of variables (Pallant, 2005). Factor analysis also assists a researcher to group different test questions that appear to be measuring a similar underlying construct.

6.3.1 Step 1: Assessment of the suitability of the data for factor analysis

As discussed previously, there are two main issues to take into account when analysing a particular data set for its suitability for factor analysis. This includes sample size. In this case, there was a sample of 155 participants, which was sufficient for the purposes of the study although a larger sample would have provided more accurate results as stated in the discussion of the methodology (Pallant, 2005). And secondly, the values for the Bartlett's test as well as the KMO should be significant.

The 30 items of the locus of control inventory of Udai Pareek were subjected to principal components analysis using SPSS. Prior to performing principal components analysis the suitability of the data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of very few coefficients of 0.3 and above. If no scores above 0.3 is found, then factor analysis should be reconsidered. Because only a few values were identified as higher than 0.3, it is advised that factor analysis should be reconsidered (Tabachnick & Fidell, 2005).

Secondly, statistical significance should also be determined. This means that the loadings should be statistically significantly different from zero.

The Kaiser-Meyer-Olkin measure and the Bartlett test of sphericity were used to determine sampling adequacy (Kim & Mueller, 1978).

Table 6.2: KMO and Bartlett's test

KMO and Bartlett's Test		
Kaiser-meyer-olkin measure of sampling adequacy.		0.802
Bartlett's test of sphericity	Approx. Chi-Square	1783.064
	Degrees of freedom	435
	Sig.	0

The Kaiser-Meyer-Olkin value was 0.802 exceeding the recommended value of 0.6 (Kaiser, 1970, 1974). It can therefore be concluded that factor analysis can be used on the sample. To confirm this, it is advised to consider an additional measure that also looks at the appropriateness of using factor analysis. It is therefore advised to look at the Bartlett's test of sphericity (Bartlett, 1954), which reached statistical significance ($p < 0.05$) and therefore supported the factorability of the correlation matrix.

6.3.2 Step 2: Factor extraction and rotation

Factor extraction includes identifying the least number of factors that can be used to best characterise the interrelations among the set of variables. There are a lot of approaches available that can be used to recognise or remove the number of underlying factors of dimensions (Pallant, 2005).

Some of the techniques that can be used to assist in the decision concerning the number of factors to retain are Kaiser's criterion, the scree test and parallel analysis. Several variables should load on each factor before one can actually trust the factor to be meaningful (Garson, 2008; Habing, 2003; Pallant, 2005).

By making use of Kaiser's criterion, components with eigenvalues of 1 or more should be considered for extraction. To determine how many components meet this criterion, it is necessary to look in the total variance explained table. Principal components analysis exposed the presence of three components with eigenvalues exceeding 1, which explains 23.3 per cent, 10.1 per cent and 7.3 per cent of the variance, respectively. It should be noted that often with Kaiser's criterion too many components are extracted, therefore other methods of extraction should be pursued.

6.3.3 Step 3: Direct oblique rotation

After the number of factors to extract has been determined, the following step is to interpret the values. This is done to support the process where the factors are rotated. It will not change the underlying solution. It represents the pattern of loadings in a way that makes it more convenient to interpret (Pallant, 2005).

Oblimin Rotation

There are three main tables to consider when presenting the results of the Oblimin rotation. This includes the pattern matrix, structure matrix and component correlation matrix. The component correlation matrix shows the strength of the relationship between the three factors. In this case the values are quite, low 0.106, 0.048 and 0.051, respectively. The component correlation matrix provides information with regard to the strength of the relationship between the three factors.

Table 6.3: Total variance explained

Component	Initial eigenvalues			Rotation sums of squared loadings(a)
	Total	% of Variance	Cumulative %	Total
1	6.982	23.273	23.273	6.931
2	3.029	10.096	33.368	2.936
3	2.176	7.252	40.62	2.49
4	1.897	6.324	46.944	
5	1.543	5.144	52.088	
6	1.212	4.041	56.129	
7	1.135	3.783	59.912	
8	1.004	3.346	63.258	
9	0.95	3.165	66.423	
10	0.893	2.976	69.399	
11	0.853	2.842	72.242	
12	0.784	2.613	74.854	
13	0.748	2.494	77.348	
14	0.681	2.27	79.618	
15	0.627	2.091	81.709	
16	0.551	1.836	83.545	
17	0.523	1.742	85.287	
18	0.513	1.711	86.998	
19	0.479	1.597	88.594	
20	0.465	1.551	90.145	
21	0.414	1.38	91.525	
22	0.387	1.289	92.814	
23	0.357	1.189	94.003	
24	0.337	1.125	95.128	
25	0.305	1.016	96.144	
26	0.297	0.991	97.135	
27	0.259	0.864	97.999	
28	0.24	0.801	98.8	
29	0.188	0.628	99.428	
30	0.172	0.572	100	
Extraction method: Principal component analysis.				
a When components are correlated, sums of squared loadings cannot be added to obtain a total variance.				

This provides information to decide whether it was reasonable to assume that the three components were not related. Because of the low values it can be interpreted as not related to each other.

From the pattern and structure matrix, the factor loadings of each of the variables are presented. The highest loading items on each component should be identified because this will enable the researcher to label the component (Pallant, 2005).

In the pattern and structure matrix the highest loadings are highlighted in yellow. This means that these items load high on specific constructs and therefore it can be determined whether the item actually measures what it is supposed to measure. This can be seen by reading the question and looking at the construct where the highest loading is indicated.

By reading a few of these while looking at where the highest loading is, it can be determined what construct each item is measuring. The red highlighted items indicate that there is no clear indication or distinction between the items to see on which construct it loads the strongest. It could be that two or more of the constructs load very close to each other or that there is a very small difference between the three construct loadings. It can therefore not be assumed that the item can differentiate effectively between the three different constructs that it is measuring. These items can and should be deleted to ensure that all the items in the questionnaire are effective in differentiating between the constructs that it is measuring.

Table 6.4: Pattern matrix

	Component		
	1	2	3
VAR00030	0.774	0.015	-0.189
VAR00025	0.761	-0.048	0.03
VAR00022	0.743	-0.036	-0.223
VAR00016	0.72	0.106	-0.153
VAR00021	0.703	-0.12	0.142
VAR00017	0.68	0.022	-0.058
VAR00033	0.584	-0.078	-0.115
VAR00019	0.574	-0.035	-0.029
VAR00015	0.551	-0.25	0.211
VAR00018	0.549	0.163	-0.076
VAR00034	0.53	-0.17	0.073
VAR00028	0.528	-0.08	-0.035
VAR00026	0.469	-0.171	0.325
VAR00010	0.468	-0.208	0.442
VAR00012	0.464	-0.025	0.185
VAR00023	0.446	-0.429	0.046
VAR00011	0.405	-0.054	0.086
VAR00008	0.387	0.029	0.195
VAR00031	-0.125	0.729	0.03
VAR00006	-0.013	0.61	0.368
VAR00005	-0.022	0.569	-0.15
VAR00007	0.057	0.509	0.32
VAR00032	0.201	0.498	0.244
VAR00029	0.023	0.495	-0.112
VAR00020	-0.344	0.439	0.331
VAR00024	-0.246	0.111	0.627
VAR00027	-0.388	0.362	0.571
VAR00014	-0.103	0.307	0.548
VAR00009	0.379	-0.318	0.501
VAR00013	0.423	-0.047	0.437

Extraction method: Principal component analysis.

Rotation method: Oblimin with Kaiser normalization

Table 6.5: Structure matrix

	Component		
	1	2	3
VAR00030	0.774	0.015	-0.189
VAR00025	0.761	-0.048	0.03
VAR00022	0.743	-0.036	-0.223
VAR00016	0.72	0.106	-0.153
VAR00021	0.703	-0.12	0.142
VAR00017	0.68	0.022	-0.058
VAR00033	0.584	-0.078	-0.115
VAR00019	0.574	-0.035	-0.029
VAR00015	0.551	-0.25	0.211
VAR00018	0.549	0.163	-0.076
VAR00034	0.53	-0.17	0.073
VAR00028	0.528	-0.08	-0.035
VAR00026	0.469	-0.171	0.325
VAR00010	0.468	-0.208	0.442
VAR00012	0.464	-0.025	0.185
VAR00023	0.446	-0.429	0.046
VAR00011	0.405	-0.054	0.086
VAR00008	0.387	0.029	0.195
VAR00031	-0.125	0.729	0.03
VAR00006	-0.013	0.61	0.368
VAR00005	-0.022	0.569	-0.15
VAR00007	0.057	0.509	0.32
VAR00032	0.201	0.498	0.244
VAR00029	0.023	0.495	-0.112
VAR00020	-0.344	0.439	0.331
VAR00024	-0.246	0.111	0.627
VAR00027	-0.388	0.362	0.571
VAR00014	-0.103	0.307	0.548
VAR00009	0.379	-0.318	0.501
VAR00013	0.423	-0.047	0.437
Extraction method: Principal component analysis. Rotation method: Oblimin with Kaiser normalization.			

Table 6.6: Component correlation matrix

Component	**1**	**2**	**3**
1	1	-0.106	0.048
2	-0.106	1	0.051
3	0.048	0.051	1
Extraction method: Principal Component Analysis. Rotation method: Oblimin with Kaiser Normalization.			

In both of these tables one can see that Factor 1 indicates high loadings on the following items:

- Item 30: The reason I am acceptable to others in my organisation is a matter of luck.
- Item 25: My success or failure in this organisation is a matter of luck.
- Item 22: Getting people in this organisation to listen to me is a matter of luck.
- Item 16: Receiving rewards in the organisation is a matter of luck.
- Item 21: How much I am liked in the organisation depends on my seniors.
- Item 17: The success of my plans is a matter of luck.
- Item 33: My ideas are accepted if I make them fit with the desires of my seniors.
- Item 19: Preferences of seniors determine who will be rewarded in this organisation.
- Item 15: Being liked by seniors or making good impressions on them influences promotion decisions.
- Item 18: Receiving a promotion depends on being in the right place at the right time.
- Item 34: Pressure groups in this organisation are more powerful than individual employees are, and they control more things than individuals do.
- Item 28: Most things in this organisation are beyond the control of the people who work here.
- Item 12: A person’s career is a matter of chance.

- Item 11: The organisation a person joins or the job he or she takes is an accidental occurrence.

These items can be identified as measuring external locus of control because of certain words that could be identified. This includes: luck, chance, accidental occurrence, beyond the control, depending on one’s seniors and so on. It can also be seen that this construct has the highest and most loadings when compared with the other two constructs that the instrument is supposed to measure. When looking closely at the items, both external others and external chance items can be identified. Therefore, there is no definite distinction that can statistically be made between external others and external chance.

The second factor that was measured showed high loadings on the following items:

- Item 31: I determine what happens to me in the organisation.
- Item 6: The course of my career depends on me.
- Item 5: I determine what matters to me in the organisation.
- Item 7: My success or failure depends on the amount of effort I exert.
- Item 29: The quality of my work influences decisions on my suggestions in this organisation.

These five items loaded high on Construct 2. If the items are read carefully, it can be seen that they all measure internal locus of control because of words like: “I determine, depends on me, my work”.

The third construct measured high loadings on the following items:

- Item 24: The way I work determines whether or not I receive rewards.
- Item 14: Successful completion of my assignments is due to my detailed planning and hard work.

But these two items closely resembled the previous construct, which was internal locus of control. From these results, it can be seen that when the instrument is used in the South African context, it actually only measures 2 different constructs. Therefore, there is no definite distinction that can be made between external others and external chance.

6.4 ITEM ANALYSIS

Item analysis should be performed to determine the characteristic of the items that are included in the constructs of the locus of control scale. The occurrence of item bias should be tested and if detected, the source of bias should then be eliminated (Van de Vijver & Leung, 1997). One of the purposes of item analysis is to discover which items best measure the construct or content domain that the measure aims to assess. Good items consistently measure the same aspect that the total test is measuring (Foxcroft & Roodt, 2005). Firstly, the results of the internal locus of control items will be presented.

Table 6.7: Reliability statistics – Internal locus of control

Cronbach's Alpha	N of Items
0.611	7

Cronbach's alpha of the internal locus of control construct is 0.611. The reliability of the scale can therefore be seen as low. Ideally, the Cronbach alpha coefficient of a scale should be above 0.7. But according to Pallant (2005), it is common to find low Cronbach alpha coefficients because of the sensitivity of the number of items in the sale. It should therefore be noted that it is common to find low Cronbach values with short scales such as the locus of control inventory of Udai Pareek (1998).

Table 6.8: Item statistics - Internal locus of control

	Mean	Std. Deviation	N
VAR00031	3.8516	0.90305	155
VAR00006	4.4581	0.78326	155
VAR00005	3.9742	0.88968	155
VAR00007	4.5032	0.69669	155
VAR00029	4.1097	0.83408	155
VAR00024	4.1419	1.02209	155
VAR00014	4.5677	0.74746	155

Table 6.9: Item total statistics - Internal locus of control

	Scale mean if item deleted	scale variance if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
VAR00031	25.7548	7.641	0.412	0.542
VAR00006	25.1484	7.79	0.483	0.523
VAR00005	25.6323	8.182	0.302	0.582
VAR00007	25.1032	8.6	0.349	0.569
VAR00029	25.4968	8.693	0.229	0.605
VAR00024	25.4645	8.25	0.207	0.624
VAR00014	25.0387	8.414	0.355	0.566

In the column marked corrected item–total correlation, it can be seen that the figures give an indication of the degree to which each of these items correlates with the total score. If these values are low, less than 0.3, it indicates that the item is measuring something different from what the scale is measuring as a whole (Pallant, 2005). Therefore, when looking at the Cronbach alpha and seeing that it is low and therefore less than 0.7, one may need to consider removing some items that currently have a low item–total correlation. In Table 16, it can be seen that

Item 24: 0.207 and 29: 0.229 has scores that are less than 0.3. The researcher can therefore consider removing Items 24 and 29 from the instrument.

In the last column with the heading of alpha if item deleted, the impact of removing each item from the scale is indicated. When these values are compared with the alpha value, it is crucial to look for those values that are higher than the alpha value and it should be considered to remove this item from the scale. The researcher can consider removing Item 24 with a value of 0.624 because it is higher than the Cronbach alpha value of 0.611. External locus of control will now be discussed in more detail.

Table 6.10: Reliability statistics – External locus of control

Cronbach's alpha	N of Items
0.873	15

The external locus of control items presented a value that was well above the desired value of 0.7 (Pallant, 2005). This meant that the reliability of the external locus of control items is higher than the internal locus of control items. The fact that there was a larger number of items that were found to be valid in measuring the external locus of control construct could have had an impact on the reliability value.

Table 6.11: Item statistics - External locus of control

	Mean	Std. deviation	N
VAR00030	1.7032	0.89133	155
VAR00025	1.7355	0.96751	155
VAR00022	1.671	0.88354	155
VAR00016	1.8452	0.94072	155
VAR00021	2.3484	1.1025	155
VAR00017	1.6516	0.92296	155
VAR00033	3.2645	1.14573	155
VAR00019	2.9742	1.05651	155
VAR00015	3.5613	1.2011	155
VAR00018	2.5097	1.14737	155
VAR00034	3.1935	1.15712	155
VAR00028	2.6452	0.99832	155
VAR00012	1.8065	0.93348	155
VAR00011	1.6516	0.95069	155
VAR00008	3.6258	1.08202	155

Table 6.12: Item total statistics – External locus of control

	Scale mean if item deleted	Scale variance if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
VAR00030	34.4839	74.498	0.699	0.858
VAR00025	34.4516	73.782	0.682	0.858
VAR00022	34.5161	74.901	0.678	0.859
VAR00016	34.3419	74.369	0.665	0.859
VAR00021	33.8387	73.097	0.623	0.86
VAR00017	34.5355	75.614	0.597	0.862
VAR00033	32.9226	74.033	0.543	0.864
VAR00019	33.2129	75.35	0.522	0.865
VAR00015	32.6258	74.625	0.481	0.868
VAR00018	33.6774	74.999	0.49	0.867
VAR00034	32.9935	75.552	0.455	0.869
VAR00028	33.5419	76.964	0.462	0.868
VAR00012	34.3806	78.484	0.405	0.87
VAR00011	34.5355	79.497	0.333	0.873
VAR00008	32.5613	78.806	0.316	0.875

When looking at the Cronbach alpha in Table 6.10 and seeing that it is a high and therefore a reliable value, which is more than the required 0.7, one might not need to consider removing some items that currently have a low item–total correlation. In the Item total statistics table, there are no scores that are less than 0.3. Therefore, the researcher does not have to consider removing any additional items, which measure external locus of control from the instrument.

In the last column of Table 6.12, the impact of removing each item from the scale is indicated. The researcher compared the values in the column with the alpha value. There was only one item that revealed a value that was larger than the alpha value. This was Item 8, which could also be deleted so that the validity of the instrument could be increased. The deletion of this item will have an impact on the alpha value, which will now increase to 0.875.

6.5 STRUCTURAL EQUATION MODELLING

The structural equation modelling (SEM) process was used in conjunction with the confirmatory factor analysis (CFA) because it focused on validating the model by identifying estimates of the parameters of the model and by determining if the model itself ensures a good quality fit to the data (Garson, 2008).

SEM Model

The locus of control inventory included 30 items and as discussed above only two factors were extracted and therefore the study will only focus on these two factors as identified in the exploratory factor analysis. The items that measure each factor were also specified while running the EQS programme.

The LM test was used in the statistical analysis of the model. The LM test compares models but only needs the estimation of a single model. It asks if the model would be enhanced if one or more of the parameters in the model, which are fixed at the moment, were estimated or what parameters need to be added to the model to enhance the fit of the model (Ullman, 2001).

6.6 ESTIMATION AND MODEL FIT

Model fit: goodness of fit indices

The model adequacy was evaluated by means of goodness of fit measures. Goodness of fit tests are used to decide whether the model being tested should be kept or rejected. The EQS programme was used for SEM procedures with maximum likelihood (ML) estimation.

Chi-square

For the chi-square to be significant, the value needs to be 0.05 or smaller. In this case, the chi-square should not be significant to indicate a good model fit while a significant chi-square will be interpreted as indicating a lack of reasonable model fit (Garson, 2008; Pallant, 2005).

As discussed above, the general guideline is that values of 0.05 and smaller are a sign of a noteworthy and significant value but it also shows that there is no indication of an acceptable model fit between the theoretical model and observed data. When considering the sample size and the poor model fit that is based on the findings of the significance of the chi-square index. The chi-square/degrees of freedom ratio should also be taken into account. Ratios between 2 and 5 can be interpreted as representing a good fit (Kelloway, 1998).

Table 6.13: Chi-square statistic

Output	**Value**
Probability value for the chi-square statistic	0.00000

In this case the chi-square is 0.0000, which means that the chi-square is significant. The significant chi-square indicates a lack of satisfactory model fit. Because of these limitations, with regard to sample size and the chi-square, it is better to also look at other model fit indices to be able to interpret the data accurately.

The following indices of model fit can be used:

- The non-normed fit index (NNFI) (Bentler & Bonnet, 1980)
- The comparative fit index (CFI) (Bentler, 1989,1990)
- The incremental fit index (IFI) (Bollen, 1989)
- The root mean square error of approximation (RMSEA)

An NNFI close to 1 indicates a good fit. According to the set standard, the IFI should be equal to or more than 0.90 for the model not to be rejected, but it can even be more than 1.0 for the model to be accepted.

Table 6.14: Fit indices

FIT INDICES	
BENTLER-BONETT NORMED FIT INDEX	0.687
BENTLER-BONETT NON-NORMED FIT INDEX	0.782
COMPARATIVE FIT INDEX (CFI)	0.806
BOLLEN'S (IFI) FIT INDEX	0.81
MCDONALD'S (MFI) FIT INDEX	0.562
JORESKOG-SORBOM'S GFI FIT INDEX	0.804
JORESKOG-SORBOM'S AGFI FIT INDEX	0.757

None of these fit indices in Table 6.14 are above 0.9 and it can therefore be noted that these indices also indicate a poor-fitting model.

Table 6.15: RMSEA values

RMSEA value	
ROOT MEAN-SQUARE RESIDUAL (RMR)	0.077
STANDARDISED RMR	0.079
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)	0.083
90% CONFIDENCE INTERVAL OF RMSEA	(0.070 ; 0.095)

The closer the RMR to 0 for a model being tested, the better the model fit. Therefore, the smaller the RMR, the better will the model fit be and a value of 0 will indicate a perfect model fit. In the RMSEA value table, it can be seen that the RMR value is 0.077, which is close to 0 indicating a good model fit.

The RMSEA value in Table 6.15 indicates a good fit if it is less than or equal to 0.08. Hu and Bentler (1998) advise that the RMSEA should be smaller than, or equal to, 0.06 as a cut–off for a good quality model fit. Hair, Anderson, Tatham and Black (1998) consider RMSEA values between 0.05 and 0.08 to be interpreted as being an indication of an acceptable fit, while Steiger (1995) considers RMSEA values of less than 0.10 as up to standard. RMSEA is a popular measure of fit and works with a better venue of independence (Garson, 2005).

In the current study, the RMSEA value is 0.083, which is not smaller than 0.05 and which would have indicated a good fit, but it is close to 0.08, which indicates that it is a reasonably fitting model. It is not more than 1.0, which would have indicated that the model needs some work.

As illustrated in Figure 6.1, the correlation between the two factors that were identified was 0.04, which indicated that confirmatory factor analysis succeeded in providing clear results when the correlations between latent factors were determined. The low correlation between the internal and external locus of control constructs suggests that these constructs can be distinguished as separate items, which measure two different things.

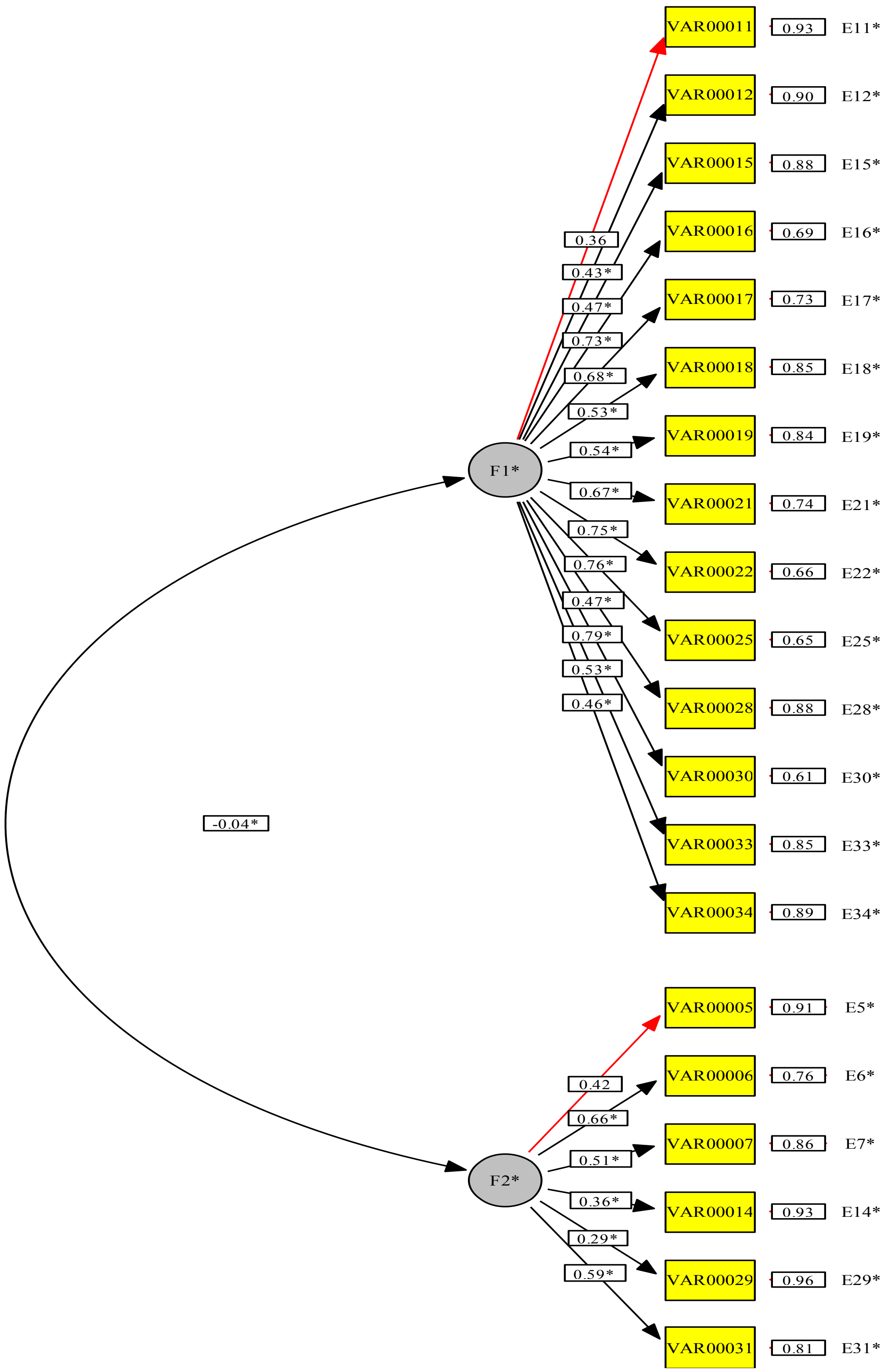


Figure 6.1: Standardised estimated parameters of the locus of control inventory - Two-factor model

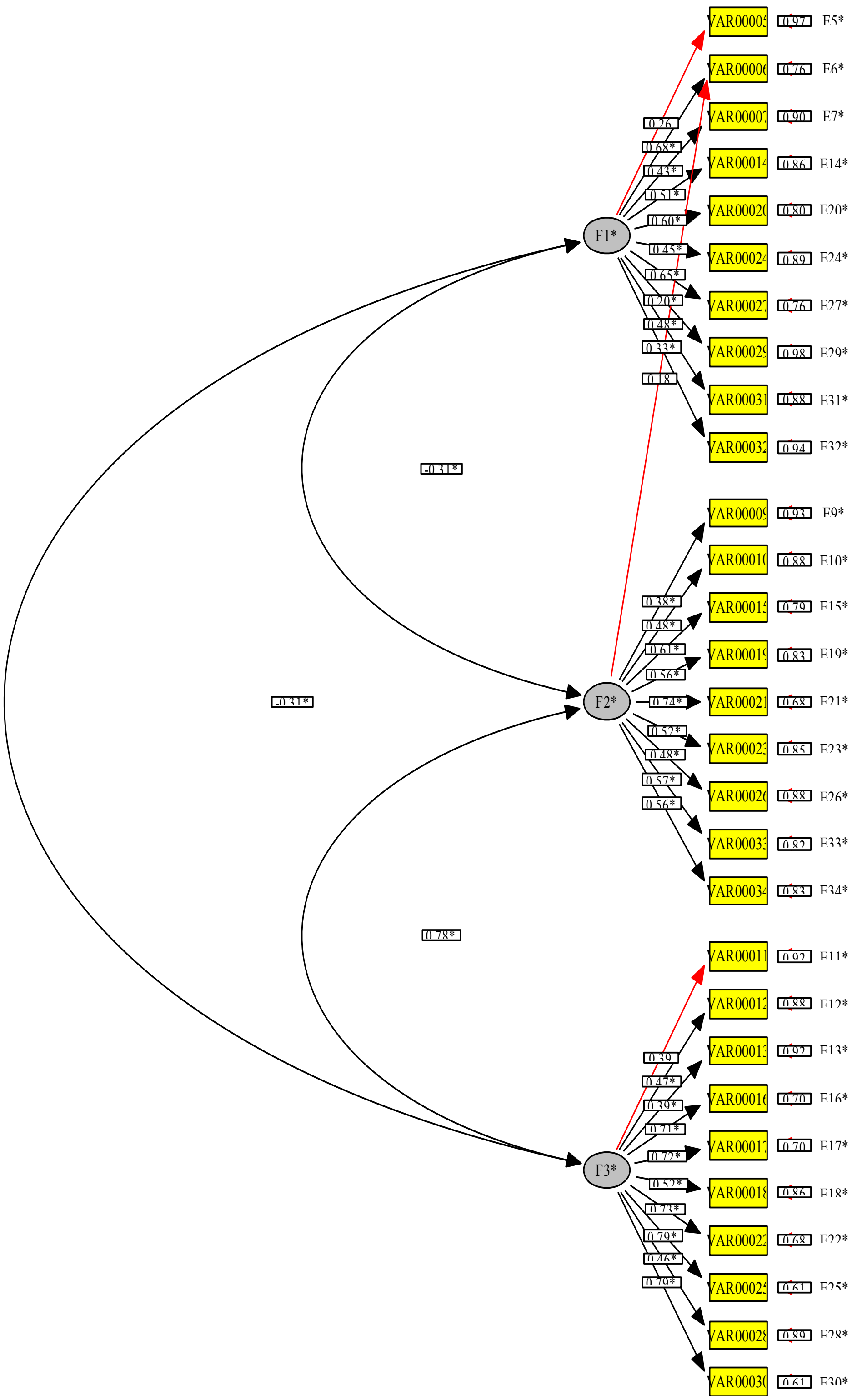


Figure 6.2: Standardised estimated parameters of the locus of control inventory – Three-factor model

With regard to the presentation of the model as Udai Pareek designed it, the following was found by looking closely at the model: the correlation between Factor 1 and 3 was 0.31 and the correlation between 2 and 3 was 0.78. This means that Factors 2 and 3 could actually be measuring the same construct because of the high correlation; and Factors 1 and 3 could be measuring two different constructs.

6.7 SUMMARY

The results of the statistical analysis were presented in table and figure format. This not only enabled the researcher to identify critical patterns that might exist in the results but it also makes further interpretation of the statistics more convenient. The results were also described in more detail so that conclusions can be made from the results. In the following chapter, the findings will be discussed and recommendations will be made for future research studies.

CHAPTER 7

FINDINGS AND RECOMMENDATIONS

7.1 DISCUSSION

To determine the suitability of the data for factor analysis, the following was found. The Kaiser-Meyer-Olkin value was 0.802, which can be identified as more than the suggested value of 0.6 (Kaiser, 1970, 1974) and the Bartlett's Test of sphericity (Bartlett, 1954) was identified as having statistical significance at 0.000, which is supportive of the factorability of the correlation matrix.

Principal component analysis revealed the presence of three components with eigenvalues exceeding 1. Therefore, according to the principal components analysis, three components should be extracted and retained. As explained above, it should be noted that often with Kaiser's criterion too many components are extracted, therefore other methods of extraction should also be pursued to ensure that the correct number of factors are extracted.

The component correlation matrix showed the strength of the relationship between the three factors. The values were identified as 0.106, 0.048 and 0.051, respectively and can be interpreted as quite low. Because of the low values, it can be concluded that the factors do not relate to each other, and therefore measures three different things.

From the pattern and structure matrix the first construct, indicated high loadings on the following items, which all measured external locus of control:

- Item 30
- Item 25
- Item 22
- Item 16
- Item 21
- Item 17
- Item 33
- Item 19
- Item 15
- Item 18
- Item 34
- Item 28

- Item 12

The second construct that was measured showed high loadings on the following items, which measured internal locus of control:

- Item 31
- Item 6
- Item 5
- Item 7
- Item 29

By making use of item analysis, items that were not shown to measure one of the two constructs were detected and eliminated. The purpose of discovering which items best measured the construct that the measure aimed to assess was attained. Good items consistently measured the same aspect that the total test was measuring (Foxcroft & Roodt, 2005).

The Cronbach's alpha of the internal locus of control construct is 0.611. The reliability of the scale was therefore identified as relatively low. In the column marked corrected item – total correlation, items 24 and 29 were deleted because the values were less than 0.3, which means that these two items are measured something different from the scale as a whole. In the alpha if item deleted column, it was also confirmed that item 24, which indicated a value higher than the Cronbach alpha should be deleted. It therefore provided the researcher with enough evidence to delete this item from the instrument.

The external locus of control items were considered to be reliable because of the Cronbach alpha value that was above the value of 0.7. The external locus of control items received a higher value than the internal locus of control items.

The alpha coefficients for the subscales varied from 0.611 to 0.873, which can be interpreted as basically good according to Byrne (2001) for the 0.873, not so good for the 0.611 value, but still close enough to 0.7 for the instrument to be judged reliable. The alpha may be low because of a shortage of homogeneity of variances among the items and it is also lower when there are fewer items per factor.

In the last column of Table 6.12 the impact of removing each item from the scale is indicated. Item 8 was identified as the item to be deleted because it revealed a value that was larger than the alpha value. The deletion of the item had an impact on the alpha value, which increased the reliability to 0.875.

With regard to confirmatory factor analysis the structural equation modelling method was used because it focused on two steps namely, validating the measurement model and fitting the structural model (Garson, 2006).

The Chi-square value was 0.0000, which meant that the chi-square was significant. The significant chi-square therefore indicated a lack of satisfactory model fit. None of the fit indices were above 0.9 and it can therefore be noted that these indices also indicated a poor-fitting model.

The RMR value was 0.077, which was close to 0 and indicated a good model fit. In the current study, the RMSEA value was 0.083, which is not smaller than 0.05 and which would have indicated a good fit, but it is close to 0.08 indicating that it is a reasonably fitting model. It is not more than 1.0, which would have indicated that the model needs some work (Hu & Bentler, 1998; Hair, Anderson, Tatham and Black, 1998; Steiger, 1995; Garson, 2005).

The correlation between the two factors that were identified from the SEM mode was 0.04, which indicated that confirmatory factor analysis succeeded in providing clear results when the correlations between latent factors were determined. The low correlation between the internal and external locus of control constructs suggests that these constructs can be distinguished as separate items which measure two different things.

In the presentation of the SEM model according to the findings of Udai Pareek the following was found by looking closely at the model. The correlation between Factors 1 and 3 was 0.31 and the correlation between 2 and 3 was 0.78. This means that Factors 2 and 3 could actually be measuring the same construct because of the high correlation; and that there was only a significant difference between Factors 1 and 3, which measured two different constructs.

7.2 CONCLUSION

The main purpose of the study was to determine the construct validity of the locus of control inventory of Udai Pareek (1998). With regard to the objectives of the study, the following was achieved. The literature review gave a firm theoretical base to explain the concept of locus of control as well as construct validity and how to achieve it. The steps that were identified in the literature can be put to great use for other researchers who would like to do similar research on other psychometrical instruments. The results from the statistical analysis were displayed and discussed in the previous chapters. Conclusions with regard to the construct validity of the locus of control instrument will now be made.

There is clear evidence that the locus of control inventory measures only two constructs as opposed to three constructs defined by Udai Pareek. The locus of control inventory has been developed and standardised by Udai Pareek using a predominantly Indian sample group. The predominantly white Afrikaans-speaking sample that this study focused on did not provide the same results as those of Udai Pareek, who developed the instrument for use in India. It can therefore also be determined that the instrument only measures two constructs in the South African student population as opposed to three constructs when it was developed and used in India.

From these results, it can be seen that when the instrument is used in the South African context it actually only measures two different constructs, and there is no definite distinction that can be made between external others and external chance. The use of the locus of control inventory in the South African context will therefore not be advised because of the differing results with regard to what the instrument actually measures and how well it discriminates between the factors that it is measuring.

The study emphasised the importance of determining the construct validity of internationally developed tests that are freely available to the South African market. Companies are not advised to make use of instruments that are readily available on the internet but rather to liaise with qualified psychometrists or industrial psychologists to help them determine which assessments to use for the companies' specific needs.

In Chapter 5, specific and detailed steps were mapped out to serve as a guideline to researchers who would like to determine the construct validity of internationally developed tests, which they would like to use in the South African context. It is important to note that the order in which these steps are illustrated in chapter 5 could change depending on the identified need and previous studies done on the instrument. It should also be kept in mind that the approach for the discussion in Chapter 5 was taken to ensure that it is detailed and informative, and therefore covers the most important steps that could be followed during the construct validation process. Next, the limitations as well as the recommendations for the specific study will be discussed.

7.3 LIMITATIONS AND RECOMMENDATIONS

From a survey research standpoint, it should be remembered that the researcher could not be at all the specific locations when the respondents were busy completing the questionnaires. This also means that respondents could not ask for clarification of the questions from the researcher and this could have had an impact on the outcomes.

The demographic arrangement of the sample alerts the researcher that care should be taken when trying to generalise the results to the larger population, especially when taking into account the current situation of South Africa with regard to progress towards employment equity. Most of the respondents were white Afrikaans-speaking females and although a variety of cultures was included in the study only a small percentage of these respondents were included, which is too little to be able to generalise to the population. The sample size of 155 was sufficient for the statistical analysis of the instrument but to be able to generalise to a specific sample, the sample would have to be much larger.

It is recommended that future research should include an exploratory factor analysis (EFA) with data received from a larger, more heterogeneous sample. EFA could offer the chance to explore the experimental data for typical features without forcing a specific model on the data. This could be followed by confirmatory factor analysis, which would be useful in comparing groups based on race, gender, age and so on.

It is suggested that the items of instruments that were identified with low reliabilities should not just be deleted but an effort should be made to rewrite these items to determine the effect thereof on the consequential findings of future validation studies. Therefore, there should be further investigation towards changing those items that have been deleted and it should be reassessed to determine if it could be successfully modified or changed to fit in with a different culture or group of people.

Comparisons between the different cultural groups were not taken into consideration in this study and may be further investigated in another study. If a more representative sample of the South African cultural distribution is obtained, then a comparison can be drawn between the different cultures to see how the SEM model will look for each of the cultures as well as determining which items measure the different constructs the best for each of the cultural groups.

Research could also be conducted to determine the prevalence of selection strategies that are not biased towards any group. This should be a main concern especially in South Africa where attempts are being made to remedy past practices (Snelgar & Potgieter, 2003). Because of the significance of unbiased assessment techniques, the development of bias-free devices should be based on the end-to-end process and not isolated to one feature of measurement.

REFERENCES

Abramowitz, S. (1973). Internal-external control and social–political activism: a test of the dimensionality of Rotter's internal-external scale. *Journal of Consulting and Clinical Psychology*, 40 (2), pp. 196 – 201.

Adams, G.S. (1966). *Measurement and evaluation in education, psychology and guidance.* New York: Holt, Rinehart and Winston.

Aijzen, I. & Madden, T.J. (1986). Prediction of goal–directed behaviour: attitudes, intensions, and perceived behavioural control. *Journal of Experimental Social Psychology,* 22, pp. 453 – 474.

Amemiya, Y. & Anderson, T.W. (1990). Asymptotic chi square tests for a large class of factor analysis models. *Annals of Statistics*, 18, pp. 1453 – 1463.

Anastasi, A. (1990). *Psychological testing*, New York: Macmillan.

Anastasi, A. & Urbina, D. (1997). *Psychological testing* (7th ed.). Upper Saddle River, New Jersey: Prentice-Hall.

Anderson, T.W. (1989). Linear latent variable models and covariance structures. *Journal of Econometrics*, 41, pp. 91 – 119.

Babbie & Mouton (2006). *The practice of social research*. Oxford University Press

Bagozzi, R. P., Yi, Y. & Phillips, L.W. (1991). Assessing construct validity in organisational research. *Administrative Science Quarterly,* 36 (3), pp 421 - 458.

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society*, 16, pp. 296 – 298.

Baumgartel, H.J., Rajan, P.S.S & Newman, J. (1985). Educational environments and attributions of causality: some exploratory research findings. *Quality of work life,* 2 (56), pp. 309 – 328.

Beck , A.T. & Freeman, A.M. (1990). *Cognitive therapy of personality disorders*. New York: Guildford Press.

Beck, A.T., Rush, J., Shaw, B. & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Bedell, B., Van Eeden, R. & Van Staden, F. (1999). Culture as moderator variable in psychological test performance: issues and trends in South Africa. *South African Journal of Industrial Psychology*, 25 (3), pp. 1 – 7.

Bentler, P.M. (1989). *EQS structural equations program: manual*. Los Angeles: BMDD statistical Software.

Bentler, P.M. (1990). Comparative fit indices in structural models*. Psychological Bulletin,* 107, pp. 238-246.

Bentler, P. M. & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, pp. 588 – 606.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bracht, G. H. & Glass, G. V. (1968). The external validity of experiments. *American Education Research Journal*, 5, pp. 437-474.

Byrne, B. (2001). *Structural equation modelling with Amos*. Rahwah, New Jersey: Lawrence Erlbaum.

Cattell, R.B. (1966). The scree test for a number of factors. *Multivariate Behavioural Research*, 1, pp. 245 – 276.

Cattell, R. B. & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioural Research*, 12, pp. 289 - 325

Cascio, W.F. (1998). *Applied psychology in personnel management* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Campbell, D. & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. 56 (2), pp. 81-105.

Carmines, E.G. & Mclever, J. P. (1981). Analysing models with unobserved variables. *Analysis of covariance structures*. Thousand Oaks, CA: Sage.

Chao, G.T, O Leary-Kelly, A.M., Wolf, S., Klein, H.J. & Gardner, P.D. (1994). Organisational socialisation: its content and consequences. *Journal of Applied Psychology*, 79 (5), pp. 730 – 743.

Choi, N., Fuqua, D. & Griffin, B.W. (2001). Exploratory analysis of the structure of scores from the multidimensional scales of perceived self efficacy. *Educational and Psychological Measurement*, 61, pp. 475 – 489.

Clarke, A. & Dawson, R. (1999). *Evaluation research: an introduction to principles, methods and practice.* Thousand Oaks: Sage Publications.

Clemans, W.V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs*, 14, pp. 1 – 54.

Cliff, N. (1983). Some cautions concerning the application of causal modelling mehods. *Multivariate Behavioural Research*, 18, pp. 115-126.

Collins, B.E. (1974). Four components of the Rotter internal external scale: belief in a difficult world, a just world, a predictable world and a politically responsive world. *Journal of Personality and Social Psychology*, 29 (3), pp. 381 – 391.

Cook, T.D. & Campbell, D.T. (1979) Quasi-experimentation: design and analysis issues for field settings. Boston: Houghton Mifflin.

Cota, A.A., Longman, R.S., Holden, R.R., Fekken, G.C. & Xinaris, S. (1993). Comparing different methods for implementing parallel analysis: a practical index of accuracy. *Educational & Psychological Measurement*, 53, pp. 865-875.

Crandall, V.C., Katkovsky, W. & Crandall, W.J. (1965). Children's beliefs in their control of reinforcements in intellectual academic achievement behaviors. *Child Development*, 36, pp. 91 – 109.

Crawford, C.B. & Koopman, P. (1979). Inter-rater reliability of scree test and mean square ratio test of number of factors. *Perceptual & Motor Skills,* 49, pp. 223 – 226.

Creswell, J.W. (1994). Research design: qualitative and quantitative approaches. Thousand Oaks, CA: Sage Publications.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, pp. 281-302.

DeCharms, R. (1976). *Enhancing motivation: Change in the classroom*. New York: Irvington.

DeCoster, J. (1998). *Overview of factor analysis*. [Online] Available from: http: // www. Stat-help.com/ notes. Html. [Downloaded: 2009-07-20]

Dodgson, P.G. & Wood, J.V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. *Journal of Personality and Social Psychology*. 75, pp. 178 – 197.

Dubinsky, A.J., Skinner, S.J. & Whittler, T.E. (1989). Evaluating sales personnel: an attribution theory perspective. *Journal of Personal Selling and Sales Management*, 9, pp. 9 – 21.

Erasmus, P. & Schaap, P. (2003). *Situation specific work profiling and assessment short course.* JP Expert/PIB SpEEx: Update course presented at the University of Pretoria, Pretoria.

Erwee, R. (1986). Achievement motivation and locus of control of black university students. *Journal of Industrial Psychology*, 12, pp. 1 – 9.

Field, A (2005). *Discovering statistics using SPSS* (2nd ed.).

Fornell, C. & Larcker, F. D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18 (1), pp. 39-50.

Fouche A.S. & De Vos C.B. (1998). Selection of a research design. In Research at grass roots. Pretoria: J.L.Van Schaik.

Foxcroft, C. & Roodt, G. (2005). *An introduction to Psychological Assessment in the South African context.* (2nd ed.). **Place: Publiser**

Franklin, R.D. (1963). *Youth's expectancies about internal versus external control of reinforcement related to N variable*. Dissertation Abstracts, 24, 1684. University Microfilms No.63 - 6493.

Gaa, J & Shores, J. (1979). Domain specific locus of control among Black, Anglo and Chicano undergraduates. *Journal of Social psychology*, 42, pp. 9 – 19

Garson, D.G. (2008) *Statnotes* [Online] Available from: http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm [Downloaded: 2009-05-04]

Gilbert, N. (1993). Research, theory and method. *Researching social life*. London: Sage.

Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational & Psychological Measurement*, 55, pp. 377 – 393.

Gravetter, F.J. & Wallnau, L.B. (2000). *Statistics for the behavioural sciences* (5th E.d.). Belmont, CA: Wadsworth.

Gregory, R.J. (1996). *Psychological testing history, principles and applications* (2nd ed.). Needham Heights, MA: Allyn and Bacon.

Gurin, P., Gurin, G., Lao,R. & Beattie, M. (1969). Internal external control in the motivational dynamics of Negro youth. *Journal of Social Issues*, 25 (3), pp. 29 – 53

Gutkin, J., Robbins, R. & Andrews, L. (1985). The health locus of control scales: Psychometric properties. *Educational and Psychological Measurement*, 45 (2), pp. 407 - 110

Guttentag, M. (1972). *Locus of control and achievement in minority middle school children*. Paper presented at the meeting of the Eastern Psychological Association, Boston, MA.

Habing, B. (2003). *Exploratory factor analysis*. University of South Carolina.

Hair, J.F. Jr., Anderson, R. E., Tatham, R. L. & Black, W.C. (1998). *Multivariate data analysis*, (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hall, C.S. & Lindsey, G. (1985). *Introduction to theories of personality.* New York : John Wiley.

Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: a progress report. *European Journal of Psychological Assessment*, 10, pp. 229 – 244.

Hans, T. (2000). A meta-analysis of the effects of adventure programming on locus of control. *Journal of Contemporary Psychotherapy*, 30 (1), pp. 33 – 60.

Harrison, F.I. (1968). Relationship between home background, school success, and adolescent attitudes. *Merrill Palmer Quarterly of Behaviour and development*, 14, pp. 331 – 344.

Harvey, R.J., Billings, R.S. & Nilan, K.J. (1985). Confirmatory factor analysis of the Job Diagnostic Survey: good news and bad news. *Journal of Applied Psychology*, 70 (3), pp. 461 – 468.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modelling*. Cary, NC: SAS Institute Inc.

Hattie, J.A., Marsh, H.W., Neill, J.T. & Richards, G. E. (1997). Adventure education and outward bound: out-of-class experiences that have a lasting effect. *Review of Educational Research*, 67, pp. 43 – 87.

Health professions Act No 56 of 1974

Heider, F. (1958). *Psychology of interpersonal relations*. New York: John Wiley & Sons.

Holburn, P.T. (1992). *Differential item functioning in the mental alertness test*. Unpublished master's dissertation. Pretoria: University of South Africa.

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30 (2), pp. 179-185.

Heidelberger, M. (2004). *Nature from within: Gustav Theodor Fechner and his Psychophysical Worldview*. Pittsburgh, PA: University of Pittsburgh Press.

Hoyle, R. H. & Panter, A.T. (1995). Writing about structural equation models. *Structural equation modelling* . Thousend Oaks, CA: Sage.

Hox, J.J. (1995). Amos, Eqs and Lisrel for windows: a comparative review. *Structural Equation Modeling*, 2, pp. 79 – 91.

Hox, J.J. & Bechger, T.M. (1998). An introduction to structural equation modelling. *Family Science Review*, 11, pp. 354-373.

Hu, L.T. & Bentler, P.M. (1998). Fit indices in covariance structure modelling: sensitivity to underparameterised model misspecification. *Psychological Methods*, 3, pp. 424 – 453.

Hubbard, R. & Allen, S.J. (1987). An empirical comparison of alternative methods for principal components extraction. *Journal of Business Research*, 15, pp. 173 – 190.

Huysamen, G.K. (1996*). Psychological measurement: an introduction with South African examples.* Pretoria: J.L. van Schaik.

Jaccard, J. & Wan, C.K. (1996). *Lisrel approaches to interaction effects in multiple regression*. Thousand Oaks, CA: Sage.

Kalechstein, A. & Nowicki, S.J.R. (1994). Social learning theory and the prediction of achievement in telemarketers. *Journal of Social Psychology* ,134 (4), pp. 547 – 548.

Kaiser, H. (1970). A second generation little Jiffy. *Psychometrika*, 35, pp. 401 – 415

Kaiser, H (1974). An index of factorial simplicity. *Psychometrika*, 39, pp. 31 – 36.

Kelloway, E.K. (1998). *Using LISREL for structural equation modelling: a researcher's guide.* Thousand Oaks, Califomia: Sage.

Kemp, N. (1999). Psychometric testing is not dead. *People Dynamics*, pp. 16 – 17.

Keshote, K.K. (1989). *Personnel and organizational correlates of conflict management styles*. Unpublished doctoral thesis. University of Gujarat, India.

Kuhn, T.S. (1961). The function of measurement in modern physical science. 52 (168), pp. 161 – 193.

Kim, J & Mueller, C.W. (1978). *Factor analysis: statistical methods and practical issues.* Thousand Oaks, CA: Sage.

Kitchen, R. & Tate, N. (2000). *Conducting research in human geography: theory, practice and methodology*, London: Prentice-Hall.

Klein, H.J. & Weaver, N.A. (2000). The effectiveness of an organizational-level orientation training programme in the socialization of new hires. *Personnel Psychology*, 53, pp. 47-66.

Kline, R. B. (1998). *Principles and practice of structural equation modelling*. New York: Guilford.

Korman, A.K. (1970). Toward a hypothesis of work behaviour. *Journal of Applied Psychology*, 54, pp. 31 – 41.

Krampen, G. & Weiberg, H. (1981). Three aspects of locus of control in German, American and Japanese university students. *Journal of Social Psychology*, 113, pp. 133 – 134.

Kren, A. L. (1992). Budgetary participation and managerial performance: the impact of information and environmental volatility. *The Accounting Review*, 67 (3), 125 – 150.

Lefcourt, H.M. (1966). Internal versus external control of reinforcement. *Psychological Bulletin*, 63(4), pp. 206 – 220.

Lefcourt, H.M. (1976). *Locus of control: current trends in theory and research.* Hillsdale, NJ: Lawrence Earlbaum.

Lefcourt, H.M. & Wine, J. (1969). Internal versus external control of reinforcement and the development of attention in experimental situations. *Canadian Journal of Behavioral Science*, 1, pp. 167 – 181.

Lessing, E.E. (1969). Racial differences in indices of ego functioning relevant to academic achievement. *Journal of Genetic Psychology*. 115, pp. 153 – 167.

Levenson, H. (1972). *Distinctions within the concept of internal external control: Development of a new scale*. Paper presented at the meeting of the American Psychological Association, Hawaii.

Locke, E.A., McClear, K. & Knight, D. (1996). Self-esteem and work. *International review of Industrial Organisational Psychology,* 11, pp. 1 – 32.

Lord, R.G. & Hanges, P.J. (1987). A control systems model of organizational motivation: Theoretical development and applied implications. *Behavioural Science,* 32, pp. 161 – 178.

MacCallum, R.C. (2003). Working with imperfect models. *Multivariate Behavioral Research*, 38, pp. 113 – 139.

MacCallum, R.C., Roznowskei, M. & Necowitz, L.B. (1992). Model modification in covariance structure analysis: the problem of capitalization on chance. *Psychological Bulletin*, 114, pp. 490 – 504.

Mamlin, N. Harris, K.R. & Case, L.P. (2001). A methodological analysis of research on locus of control and learning disabilities: rethinking a common assumption. *Journal of Special Education*.

Mardia, K.V., Kent, J.T. & Bobby, J.M. (1979). Multivariate Analysis. London: Academic Press.

Marquart, J.M. (1989). A pattern matching approach to assess the construct validity of an evaluation instrument. *Evaluation and Program Planning*, 12 (1), pp. 37 – 43.

Maslow, Q.H. (1943). A theory of human motivation. *Psychological Review*, 50, pp. 370 – 396.

Massey, A. (2003). *Methodological triangulation.* [Online] Available from: http://www.freeyourvoice.co.uk/htm/triangulation_3htm [Accessed: 2008-12-09]]

McClelland, D.C., Atkinson, J.W. & Clark, R. A. (1953). *The achievement motive.* East Norwalk, CT, US: Appleton – Century-Crofts.

McClelland, D.E. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

McDonald, R.P. & Ho, M.R. (2002). Principles and practice in reporting structural equation analysis. *Psychological Methods*, 7 (1), pp. 64.

McDonald, A.P. & Tseng, M.S. (1971). *Dimension of internal vs. external control revisited: toward expectancy*. Published paper. West Virginia University.

Messick, S. (1989). Validity. Educational measurement, (3rd ed.).

Mehta, P. (1968). *Increasing achievement motive in high school boys*. New Delhi: National council of Education Research and Training.

Metha, A. (2004). *Threats to construct validity: reasons why inferences about the constructs that characterize study operations may be incorrect.* [Online] Available from: *www.cis.gsu.edu/~dstraub/Courses/MGS%209940/2004/Threats.doc [Accessed 2009-06-25]*

Micceri, T. (1989). The unicorn, the normal curve and other improbable creatures. *Psychological Bulletin*, 105, pp. 156 – 165.

Minton, H. L. (1972). *Internal–external control and the distinction between personal control and system modifiability*. Paper presented at the meeting of the Mid-western Psychological Association, Cleveland, OH, 34, pp. 226-228.

Mirels, H. (1970). Dimensions of internal vs. external control. *Journal of consulting and Clinical Psychology*, 34, pp. 226-228.

Mischel, W. (1966). Theory and research on the antecedents of self-imposed delay of reward. *Progress in experimental personality research,* (3). New York: Academic Press.

Mitchell, T.R., Smyser, C.M. & Wood, S.E. (1975). Locus of control: supervision and work satisfaction. *Academy of Management Journal,* 18 (3), pp. 623-631.

Mueller, S.L & Thomas, A.S. (2001). Culture and entrepreneurial potential: a nine country study of locus of control and innovativeness. *Journal of Business Venturing,* 16 (1), pp. 51 – 74.

Murphy, K.R. & Davidshofer, C.O. (2001). *Psychological Testing Principles and Applications.* (5th ed.). New Jersey: Prentice Hall.

Nunnally, J.C. & Bernstein, I.H., (1987). *Psychometric Theory*. (3rd ed.). Mcgraw-Hill

O'Brian, K.E., (2004). *Self-determination theory and locus of control as antecedents of voluntary workplace behaviours.* University of South Florida.

O'Connor, B.P., (2000). *SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test.*

Owen, K. & Taljaard, J.J. (1996). *Handbook for the use of psychological and scholastic tests of the HSRC*. Pretoria: Human Sciences Research Council.

Pallant, J. (2005). *SPSS survival manual: a step by step guide to data analysis using SPSS version 12.* Chicago: Open university.

Pareek, U. (1982). Internal and external control. *The 1982 annual for facilitators, trainers, and consultants*. San Diego, CA: Pfeiffer & Company.

Pareek, U. (1998). *The Pfeiffer Library.* 8, (2nd ed.). Jossey–Bass/Pfeiffer .

Parkes, K.R. (1991). Locus of control as moderator: an explanation for additive versus interactive findings in the demand-discretion model of work stress. *British Journal of Psychology*, 82, pp. 291 – 312.

Paulhus, D. (1983). Sphere-specific measures of perceived control. *Journal of Personality and Social Psychology, 44, pp.* 1253-1265.

Pearl, J. (2000). *Causality: models, reasoning and inference*. Cambridge, England: Cambridge University Press.

Pennington, D. (2003). *Essential Personality.* Arnold, pp. 37.

Phares, E.J. (1976). *Locus of control in personality*. Morristown, New York: General Learning Press.

Price, J.L. (1997). Handbook of organisational measurement. *International Journal of Manpower*, (18), pp. 305 – 558.

Rasch, G. (1980). *Probabilistic models for some intelligence and achievement tests*. Copenhagen: Danish Institute for Educational Research.

Raubenheimer, J.E. (2004). An item selection procedure to maximize scale reliability and validity. *South African Journal of Industrial Psychology*, 30 (4), pp. 59 - 64.

Reichard, B.D. (1975). *The effect of a management training workshop altering locus of control.* Unpublished doctoral thesis. University of Maryland.

Reynolds, C.R. (1982). The problem of bias in psychological assessment. *The handbook of school psychology.* New York: Wiley.

Rieger, H.S. & Blignaut, D.J.H. (1996). Individualitiet, kollektiwiteit en lokus van beheer as mikroveranderlikes van diversiteit. *Tydskrif vir Bedryfsielkunde*, 22 (3), pp. 35.

Roediger, III, H.L., Capaldi, E.D., Paris, S.G. & Polivy, J. (1991) *Psychology* (3rd ed). New York: Harper Collins.

Rotter, J.B. (1945). *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80 (1), pp. 609

Rotter, J.B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, 35, pp. 651-665.

Rotter, J.B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology,* 43, pp. 56 – 67.

Rotter, J.B., Chance, J.E., & Phares, E.J. (1972). Applications *of a social learning theory of personality.* New York: Holt, Rinehart & Winston.

Ruane, J.M. (2005). *Essentials of Research Methods: A Guide to Social Science Research.* UK: Blackwell.

Rubio, D.M., Berg-Weger, M., Tebb, S.S., Lee, E.S. & Rauch, S. (2004). Objectifying content validity: conducting a content validity study in social work research. *Social Work Research*, 27, pp. 94-104.

Salkind, N.J. (2006). Exploring research. (6th ed.). New Jersey: Pearson Education.

Satorra, A., & Bentler, P.M. (1994). Corrections to standard errors in covariance structure analysis. *Latent variable analysis: applications to developmental research*. Thousand Oaks, CA: Sage.

Schaap, P. (2001). Determining differential item functioning and its affect on the test scores of selected PIB indexes, using item response theory techniques. *Journal of Industrial Psychology*, 27 (2), pp. 32.

Schepers, J.M. (1995). *Locus of control inventory*. Unpublished Report: Johannesburg Rand Afrikaans University.

Schepers, J.M. (1999). *Die lokus van beheer–vraelys: konstruksie en evaluering van 'n nuwe meetinstrument*. Ongepubliseerde verslag: Johannesburg Rand Afrikaans Universiteit.

Schepers, J.M. (2004*). Overcoming the effects of differential skewness of test items in scale construction*. University of Johannesburg DigiSpace.

Schepers, J.M. (2005).*The construction of a normative scale of locus of control*. UJ DigiSpace at the University of Johannesburg.

Sehlapelo, M. & Terreblanche, M. (1996). Psychometric testing in South Africa: view from above and below. *Psychology in Society*, 21, pp. 49 – 59.

Sen, P.C. (1982). *Personal and organizational correlates of role stress and coping strategies in some public sector banks*. Unpublished doctoral thesis. Univeristy of Gujarat, India.

Shaw, E. & Human, L. (1989). Assessment of managerial potential among black employees in selected South African companies: methods, perceived problems and concerns. *South African Journal of Labour Relations*, 13 (2), pp. 4 – 18.

SIOPSA. (2006). *Code of practice for psychological and other similar assessments in the workplace.*

Siu, O., Spector, P.E., Cooper, C.L., LuYu, S. (2002). Managerial stress in greater China: the direct and moderator effects of coping strategies and work locus of control. *Applied Psychology: an International review*, 51(4), pp. 608 – 632.

Smit, G. J. (1991). *Psigometrika* (2nd ed.) Pretoria: HAUM.

Smith, M. & Robertson, I.T. (1986). *The Theory and Practice of systematic staff selection*. Hong Kong: MacMillan.

Snelgar, R.J. & Potgieter, T. (2003). An exploratory study of group adverse impact in a recruitment and selection strategy. SAJEMS, 6 (1) . Department of Industrial and Organisational Psychology, University of Port Elizabeth. [Online] Available from: http: //www.emeraldinsight.com [Accessed: 2003-10-25]

Spector, P.E. (1982). Behaviour in organisations as a function of employee's locus of control. *Psychological Bulletin*, 91, pp. 482 – 497.

Spector, P.E. (1988). Development of the work locus of control scale. *Journal of Occupational Psychology*, 6, pp. 335 – 350.

Spirtes, P., Richardson, T., Meek, C., Scheines, R. & Glymour, C. (1998). Using path diagrams as a structural equation modelling tool. *Sociological Methods & Resesarch*, 27 (2), pp. 182 – 225.

Steiger, J.H. (1995). *Manual to Statistica-sepath*. Tulsa, Ok: Statsoft Inc.

Stevens, J. (1996). Applied multivariate statistics for the social sciences. Ten Commandments of structural equation modelling. *Reading and understanding* *more multivariate statistics.* Washington, DC: American Psychological Association, pp. 261 – 284.

Stevens, J. (2002) *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, N.J: Lawrence Erlbaum.

Stober, J. (1998). The Frost multidimensional perfectionism scale revisited: more perfect with four dimensions*. Personality and Individual differences*, 24, pp. 481 – 491.

Streiner, E. L. (1998). Factors affecting reliability of interpretations of scree plots. *Psychological Reports*, 83, pp. 687 – 694.

Surti, K. (1982). *Some psychological correlates of role stress and coping styles in working wome.* Unpublished doctoral thesis, University of Gujarat, India.

Tabachnick B.G. & Fidell, L.S. (2001). *Using multivariate statistics* (4th ed.). New York: Harper Collins.

Taylor, J.M. & Radford, E.J. (1986). Psychometric testing as an unfair labour practice. *South African Journal of Psychology*, 16 (3), pp. 79 – 86.

Tharenou. P. (1979). Employee self-esteem: a review of the literature. *Journal of Vocational behaviour*, 15, pp. 316 – 346.

Thurstone, L.L. (1947)*. Multiple factor analysis*. Chicago: University of Chicago Press.

Trochim, W., (1985). Pattern matching, validity, and conceptualization in program evaluation. *Evaluation Review*, 9 (5), 575-604.

Trochim, W.M.K (2006). *Construct validity, research methods knowledge base.*

Ullman, J. B. (2001). Structural equation modelling. Using Multivariate Statistics (4th ed.). Needham Heights, MA: Allyn & Bacon. pp. 653 – 771.

Van de Vijver, F. & Hambleton, R. (1996). Translating tests: some practical guidelines. *European Psychologist*, 1, pp. 89-99.

Van de Vijver, F. & Leung, K. (1997). Methods and data analysis for cross-cultural research. London: Sage.

Van der Merwe, R.P. (1999). Psychological assessment in industry*. South African Journal of Industrial Psychology*, 25 (3), pp. 8 – 1.1

Van Zyl , E & Visser, D. (1998). Differential item functioning in the figure classification test. *Journal of Industrial Psychology*, 24 (2), pp. 25 -33.

Veal, A.J. (1997). *Research methods for leisure and tourism*. A practical guide. London: Pitman.

Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, pp. 321 – 327.

Walle, A.H. (1997) Quantitative versus qualitative tourism research. *Annals of Tourism Research*, 21 (3), pp. 524 – 36.

Welman, C. & Kruger, S.J. (2001) *Research Methodology.* Oxford Southern Africa, Cape Town.

Weiss, H., & Sherman, J. (1973) Internal–external control as a predictor of task effort and satisfaction subsequent to failure. *Journal of Applied Psychology*, 57, pp. 132 – 136.

Wilkinson, L., Blank, G. & Gruber, C. (1996). *Desktop data analysis with SYSTAT*. Upper Saddle River, NJ: Prentice-Hall.

Witt, L.A. (1988). Locus of control and success as a professional money collector. *Journal of Social Psychology*. 128 (5), pp. 703 – 704.

Wolk, S. & DuCette, J. (1984). Intentional performances and incidental learning as a function of personality and task direction. *Journal of Personality and Social Psychology*, 29, pp. 90-101.

Wolman, B.B. (1973) Dictionary of behavioural science. New York: Van Nostrand Reinhold company.

Wood, J.M., Tataryn, D.J., & Gorsuch, R.L. (1996). Effects of under and overextraction on principal axis factor analysis with varimax rotation. *Psychological methods*, 1, pp. 354-365.

Wright B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement* ,14, pp. 97–116.

Zimbardo, P.G. (1985). *Psychology and life.* Glenview, IL: Scott Foreman and Company.

Zwick, W.R. & Velicer W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, pp. 432-442.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Appendix A:

The questionnaire, scoring and interpretation sheet

LOCUS OF CONTROL INVENTORY

By

Udai Pareek

Information: Within psychology, Locus of Control is considered to be an important aspect of personality. The concept was developed originally Julian Rotter in the 1950s (Rotter, 1966). Locus of Control refers to an individual's perception about the underlying main causes of events in his/her life. Or, to put it simpler you can ask the question: Do you believe that your destiny is controlled by yourself or by external forces (such as fate, or powerful others)?

Instructions: The following thirty statements represent employee's attitudes toward their work in an organisation. Read each statement carefully, then indicate the extent to which you agree with it by writing a number in the blank space provided. There are no right or a wrong choice, the one that is right for you is the correct answer. If the responses do not adequately indicate your own opinion, use the number closest to the way you feel

I hereby give my informed consent to participate in the study

Biographical information:

Race:
Gender:
Age:
Language:

Use the following key:

Strongly Agree - 5
Generally Agree - 4
Agree somewhat - 3
Agree only slightly - 2
Seldom or never Agree - 1

1	I determine what matters to me in an organisation.	
2	The course of my career depends on me.	
3	My success or failure depends on the amount of effort I exert.	
4	The people who are important control matters in this organization.	
5	My career depends on my seniors.	
6	My effectiveness in a organization is determined by senior	

	people.	
7	The organisation a person joins or the job he or she takes is an accidental occurrence.	
8	A person's career is a matter of chance.	
9	A person's success depends on the breaks or chances he or she receives.	
10	Successful completion of my assignments is due to my detailed planning and hard work.	
11	Being liked by my seniors or making good impressions on them influences promotion decisions.	
12	Receiving rewards in the organization is a matter of luck.	
13	The success of my plans is a matter of luck.	
14	Receiving a promotion depends on being in the right place at the right time.	
15	Preferences of seniors determine who will be rewarded in a organization.	
16	My success depends on my competence and hard work.	
17	How much I am liked in an organisation depends on my seniors.	
18	Getting people in an organisation to listen to me is a matter of luck.	
19	If my seniors do not like me, I will not succeed in this organization.	
20	The way I work determines whether or not I receive rewards.	
21	My success or failure in an organisation is a matter of luck.	
22	My success or failure depends on those who work with me.	
23	Any promotion I receive will be due to my ability and effort.	
24	Most things in an organization are beyond the control of the people who work there.	
25	The quality of my work influences decisions on my suggestions in this organisation.	
26	The reason I am acceptable to others in a organisation is a matter of luck.	
27	I determine what happens to me in the organization.	
28	The degree to which I am acceptable to others in this organisation depends on my behaviour with them.	
29	My ideas are accepted if I make them fit with the desires of my seniors.	
30	Pressure groups in this organization are more powerful than individual employees are, and they control more things than individuals do.	

LOCUS OF CONTROL INVENTORY

SCORING SHEET

Instructions: The numbers below correspond to the numbers of the items in the locus of Control Inventory. Please transfer the numbers you assigned by writing them n the appropriate blanks below. Then total the numbers you transferred to each column.

Item Number	Number You Assigned	Item Number	Number You Assigned	Item Number	Number You Assigned
1		4		7	
2		5		8	
3		6		9	
10		11		12	
16		15		13	
20		17		14	
23		19		18	
25		22		21	
27		29		24	
28		30		26	
Column Total		**Column Total**		**Column Total**	
	I		**EO**		**EC**

Locus of control inventory

Interpretation sheet

The following information will be helpful in interpreting your scores. These scores represent the way you view what happens in your organization: therefore, no score has to be permanent. If you are happy with the why you have marked the answers, you may create an action plan that will help to change the way you look at things. Select the column with the highest total. Then read the section below that pertains to that column. Next read the section pertaining to your lowest total. Then read the remaining section. The paragraph on rations may also be helpful.

I (Internal)

A person with an internal orientation believes that his or her future is controlled from within. A total I score of 33 or above indicates a very high internality tendency. It represents self confidence in a person's ability to control what happens to him or her n an organization. However, this person may sometimes eb unrealistic in assessing difficulties and may ascribe personal failure to situations over which he or she had no control.

A score from 29 to 32 shows high trust in one's ability and effort and is likely to lead to effective use of these. A score of 18 to 21 indicates that the individual lacks such self trust and needs to examine his or her strengths by using feedback from others. A low score: 17 or less, in this area represents little self confidence and could hinder a person from utilizing his or her potential.

EO (External – others)

A person with an external-others orientation believes that his or her future is controlled by powerful others. Very high EO scores (30 or higher) indicate dysfunctional dependence on significant other people for achieving one's goals. A score of 21 to 29 reflects a realistic dependence on supervisors, peers and subordinates. A score of 17 to 20 shows an independence orientation and a score below 17 indicates counter dependence.

EC (External – Chance)

A person with an external- chance orientation believes that his or her future is controlled primarily by luck or chance. To an extent, the lower the EC score, the better, because a person with a low EC orientation is more likely to utilize another potential in trying to achieve goals. However, a score of 10 or below may reflect problems in coping with frustrations when unforeseen factors prevent achievement of goals.

Ratios of Scores

The ratio of your, I and E scores can also provide information about your orientation. If you're I / total E ration is more than one, this means if you're I score is greater than the total of your E scores, you have an internal orientation. If your EO ration is more than one, you have more internality than externality –other. If you're I/ EC ratio is greater than one, you are more internal and external chance. Rations greater than one is beneficial, and action plans can be created to change ratios that are lower than desired.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Appendix B:

Information Letter

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Information Letter for Research Study on the construct validity of the Locus of control Inventory

This letter serves to provide you with a background of the proposed study. It is important that you read through this letter to familiarise yourself with the purpose of the research so as to better inform your understanding as you answer the questionnaire.

Within psychology, Locus of Control is considered to be an important aspect of personality. The concept was developed originally Julian Rotter in the 1950s (Rotter, 1966) Locus of Control refers to an individual's perception about the underlying main causes of events in his/her life. Or, to put it simpler you can ask the question: Do you believe that your destiny is controlled by *yourself* or by *external forces*?

Rotter's view was that behavior was largely guided by "reinforcements" (rewards and punishments) and that through contingencies such as rewards and punishments, individuals come to hold beliefs about what causes their actions. These beliefs, in turn, guide what kinds of attitudes and behaviors people adopt. A locus of control orientation is a belief about whether the outcomes of our actions are contingent on what we do (internal control orientation) or on events outside our personal control (external control orientation)." (Zimbardo, 1985, p. 275).

Thus, locus of control is conceptualised as referring to a uni-dimensional continuum, ranging from external to internal. An important question is whether an internal or external locus of control is desirable. In general, it seems to be psychologically healthy to perceive that one has control over those things which one is capable of influencing. In simplistic terms, a more internal locus of control is generally seen as desirable

Psychological research has found that people with a more internal locus of control seem to be better off, e.g., they tend to be more achievement oriented and to get better paid jobs. The proposed study will provide the business community with insight in to the importance of the construct validity not only for the locus of control questionnaire but also for ensuring that all psychometrical tests are validated before they are implemented in South Africa.

In the field of Human Resources / Industrial psychology the proposed study aims at expanding the knowledge on the validation of psychometrical instruments for the South African workplace. Human Resource managers have an enormous responsibility with the selection and employment of the most suitable candidates in their companies without discriminating unfairly against certain cultural groups. Thus the ultimate aim of the proposed study for the Human Resources / Industrial psychology perspective is to ensure that the importance of determining the construct validity of psychometrical instruments are known and understood.

Furthermore, the proposed study should stimulate further research to both explore and empirically establish the construct validity of all psychometrical instruments that are currently in use and also those that are being developed in the near future. It is clear that the research on the construct validity of the locus of control questionnaire is few, which represents a substantial need for the proposed research. This represents a substantial need for the proposed study.

The study will be employing a quantitative research method, with data being gathered through the distribution of questionnaires. The accumulated data will then be analysed with the intent of discovering general statistical themes and patterns, and a final report thereafter written to make the findings available to the supervisor at the University of Pretoria for the researchers' completion of her Maters Degree in Industrial Psychology.

The researcher is readily available to address any concerns or attend to any queries you may have, please feel free to contact her.

Researcher: **Corne Engelbrecht**

***Cellular phone number*: 082 336 0590**

***Email address*: corne.engelbrecht@compensation.co.za**

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Appendix C:

Letter of consent

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Letter of Consent

I the undersigned hereby give my informed consent to participate in the study on the construct validity of the locus of control questionnaire. I have read and understand the information letter regarding the study and I realise that the purpose of this study is to state the importance of using validated tests in the South African context for development and employment purposes.

I understand that the study is exploratory and no judgements will be made about me as an individual and that only the researcher will have access to my results. I understand that as a research participant utmost confidentiality and anonymity will be maintained and no results will be linked to me personally in any manner whatsoever. I realise that no records of personal details or identifying factors will be used in the analysis.

I trust that no individual results or profiles will be examined in isolation and therefore the researcher will ensure that no individual will be identified in the final report. I have been assured that the data collected will only be used for the stated purpose of the research and that no personal information related to me will be discussed or shared with anyone without consent.

I furthermore realise that I have the right to withdraw from the study at any time, and I state that I have not been coerced into consenting to participate in this study by the researcher, or anyone else.

I realise that the results of this study will be made available to me by the researcher in question once I have personally contacted her in this regard. Should I have any queries or concerns I am aware that the researcher Corne Engelbrecht, is readily available should I wish to contact her via email (corne.Engebrechtl@compensation.co.za) or telephonically (082 336 0590).

____________________________ ______________________

Signature Date