UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

THE INFLUENCE OF INTROVERSION/EXTRAVERSION BIAS ON LEADERSHIP ASSESSMENT WITH BEHAVIOUR OBSERVATION

by

Esté de Beer

Submitted in partial fulfilment of the requirements for the degree

MAGISTER COMMERCII

in the

FACULTY OF ECONOMIC AND BUSINESS SCIENCES

at the

UNIVERSITY OF PRETORIA

PRETORIA **AUGUST 2011**

Study Leader: Ms. S.M. O'Neil

CONTENTS

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION

I, Esté de Beer, hereby declare that **"The influence of introversion/extraversion bias on leadership assessment with behaviour observation"** is my own original work and that all resources have been accurately acknowledged and referred to in the reference list by means of a comprehensive referencing system. This document has not previously in its entirety or in part, been submitted to any university in order to obtain an academic qualification.

Esté de Beer	29 August 2011 Date

ACKNOWLEDGEMENTS

I wish to give special thanks to the following individuals:

My Creator, Sumari O'Neil, Johnny O'Neil, Monica Myburgh, The Military Psychological Institute, Henriette Rieger, Emily Jarvis, my loving father (Roelf de Beer), my late mother (Sybeth de Beer), my sisters (Elizma and Jana de Beer) as well as my close friends.

Without you the completion of this study would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

THE INFLUENCE OF INTROVERSION/EXTRAVERSION BIAS ON LEADERSHIP ASSESSMENT WITH BEHAVIOUR OBSERVATION

by

ESTÉ DE BEER

STUDY LEADER: S.M. O'Neil

DEPARTMENT : Human Resource Management

DEGREE : M Com (Industrial Psychology)

'Survival of the fittest' aptly describes the work environment. Employees and the organisations for which they work are therefore required to have various skills sets to afford them a competitive advantage in the job market. This is one of the many reasons why private and public organisations make use of the assessment center, and specifically the behaviour observation exercise to evaluate and select future personnel. Although the behaviour observation exercise provides the rater with rich information regarding a candidate's skills, rater errors that are often inadvertent can result from a rater's inherent subjectivity. One such error, central to this study, is Introversion/Extraversion bias. This type of bias plays out when raters rate candidates with personality types similar to their own more favorably than other candidates when the candidate's degree of introversion or extraversion should not be considered relevant to the selection criteria.

This study aims to explore the effect of Introversion/Extraversion bias on the scores of behaviour observation exercises performed during a leadership assessment center in a security environment. The sample consists of 103 participants (14 raters and 89

candidates) all belonging to the same security organisation. The researcher conducts a cross-sectional, non-experimental field study. Candidate as well as rater Introversion/Extraversion preferences are measured by the Jung Personality Questionnaire (JPQ). The scores of two behaviour observation exercises are used to explore the interaction effect between rater Introversion/Extraversion and candidate Introversion/Extraversion. Point-biserial correlations, independent t-tests as well as a one-way ANOVA are used to test the hypotheses.

No interaction effect is identified between rater Introversion/Extraversion and candidate Introversion/Extraversion, indicating that raters did not score candidates with similar personality types to their own more favorably. However, the results indicate that extraverted candidates were rated higher by both introverted and extraverted raters and are consequently perceived to have performed better in both behaviour observation exercises. The study postulates that the nature of the exercises, which require high levels of engagement with fellow team members (a typical strength of extraverts), is one of the main contributors to the perception that extraverted individuals are better performers.

The results of this study not only contribute to the lacuna in research on the topic, but also to the development of an unbiased behaviour observation exercise within this security organisation.

CHAPTER 1: INTRODUCTION AND ORIENTATION TO THE STUDY

1.1 INTRODUCTION AND BACKGROUND

Modern companies operate in a challenging, fast-paced and ever-changing environment. In order to keep up with these changes and still maintain a competitive advantage, managers should engage in various business process strategies (Hetland & Sandal, 2003:147). One of the most effective strategies is to invest in the process of recruiting the right employees into the right positions as employees are the core of any organisation (Hetland & Sandal, 2003:147). The challenge however, is that the organisation needs employees that are not only capable of adhering to a job's requirements but who will also be able to excel in the position. In other words, employees (or applicants) should not only possess the technical skills required to perform their role but they also need to have the necessary 'life skills' (e.g. problem solving skills, motivation, conflict handling skills, interpersonal skills and coping skills) in order to fit into the culture of that specific company. Consequently, managers all over the world have realised the value of assessment centers when it comes to appointing and assessing employees in their organisations (Lievens, 1998:141).

The assessment center is a process used to evaluate the different skills of an individual (Taylor, 2009:12; Schlebusch & Roodt, 2008:2). Joiner (2000:319) defines an assessment center as a standardised method of evaluating behaviour, based on multiple inputs while making use of several techniques and trained observers. The assessment center is not only a common tool used by many organisations, but it has also been a very popular research topic in recent years (Dilchert & Ones, 2009:254; Howard, 1997:13; Lievens, Van Keer & Volckaert, 2010:98; Schlebusch & Roodt, 2008:2).

One activity that is often used in assessment centers is the behaviour observation exercise. It is specially designed to elicit behaviours that are indicative of the job-related performance dimensions of interest such as communication, adaptability, teamwork and leadership (Brummel, Rupp & Spain, 2009:138). Candidates participating in such exercises are observed by trained assessors (otherwise referred to as raters) who take

detailed behavioural observation notes while candidates participate in an activity. They then score the candidate on each dimension using a standardised rating scale, for example the Behavioural Anchored Rating Scale (BARS) or a Likert Scale (Brummel et al., 2009:140).

Although assessment centers and behaviour observation exercises have been proven to be valid and reliable (Lievens, 1998:141; Coleman & Adams, 1999:27; Lievens & Van Keer, 2001:373; Hermelin, Lievens & Robertson, 2007:406), it often happens that raters present a subjectively skewed representation of candidates (Anderson, Lievens & Van Dam, 2006:558; Dean, Roth & Bobko, 2008:685). These subjective representations are not always deliberate mistakes by the rater, they are merely a reality of the process. They can however have a powerful impact on how behaviour is perceived. These subjectively skewed representations are referred to as rater bias (Taylor, 2009:64).

Rater bias may be any conscious or unconscious predisposition that is not performance related (Grobler, Warnich, Carrell, Elbert & Hatfield, 2006:277). It is the most common error that exists in any appraisal method and may stem from either personal characteristics (such as age, gender, disability or race) or organisation-related characteristics (such as seniority, or membership to a certain entity) (Grobler, et al. 2006:277).

The specific rater bias that is explored in this study occurs when raters have certain preferences for individuals with a personality type (specifically referring to [1]Introversion and Extraversion) that is similar to theirs and as a result, rate such candidates higher. For example, an introverted rater would have a preference towards an introverted candidate and would therefore rate him/her higher than they might rate an extraverted candidate. This particular bias is referred to as introversion/extraversion bias. Before the nature of introversion/extraversion bias can be explained, it is necessary to give a brief overview of introversion and extraversion as they are expressed as personality types.

[1] The researcher will from this point onwards refer to Jung's Introversion and Extraversion personality types without capitalising these terms.

The concepts of introversion and extraversion originated from Jung's personality theory (Meyer, Moore & Viljoen, 1997:123) and the differences between them have always been a relevant topic for research (Killgore, Richards, Killgore, Kamimori & Balkin, 2007:355). McDougall (1929[2]:296) states the following regarding introversion and extraversion:

> *"That Jung is attempting by the use of these terms to point to some deep-lying and very important peculiarities of personality I have no doubt. I have found the distinction between introversion and extraversion extremely useful, both in theory and in the practical handling of cases… and also in understanding normal personalities."* (McDougall, 1929:296)

According to Du Toit (1983[3]:4), Bono and Judge (2004:902), Hautala (2005:85) and Myers, McCaulley, Quenk and Hammer (2009:24), extraverts are people who derive their psychic energy (libido) from the outside or objective world and are typically those who would feel a loss of energy when they are alone for too long. They are preoccupied with people and things. They prefer to be active, energetic and to seek excitement. Extraverts are usually also talkative and outgoing, cultivating a positive attitude in life. At times they might appear assertive and dominant but they also tend to add a lot of value to their relationships.

On the other hand, Du Toit (1983:5), Bono and Judge (2004:902), Hautala (2005:85) and Myers, et al. (2009:24) state that introverts derive their libido from being alone. When they are compelled to function in big groups, they tend to lose energy and get tired quickly. People perceive them as being shy, withdrawn and less outgoing. Introverts are often scared of new situations and are hesitant to explore new things. They tend to be critical of themselves and others while preferring to trust their own

[2] The researcher refers to the work of McDougall (1929) to indicate that the differentiation between the concepts of introversion and extraversion has been a relevant research topic since the early 1900's.

[3] Due to a limitation in recent studies on the topic, the researcher refers to the work of Du Toit (1983) who wrote the manual for the Jung Personality Questionnaire. Du Toit explains Introversion/ Extraversion as derived from Jung's theory which forms the basis for this study. The researcher searched for relevant articles on EbscoHost for the following keywords: *Introversion, Extraversion, Jung's Typology and Personality Types*

judgment rather than the judgment of others. Although they are quiet, they have a rich and imaginative inner life and they often have loyal friends. Du Toit (1983:5) gives further insight by stating that introverts tend to hide their best qualities and are therefore often underestimated.

To date, studies have shown that introverts and extraverts differ in more aspects than simply the way in which they prefer to derive their energy. Areas such as physiological functioning, sensory stimulation and deprivation, perceptual ability, learning potential, social interactions and leadership, to name but a few aspects, have all been explored (Eysenck, 1970[4]:30; McCormack & Mellor, 2002:179; Furnham, Taylor & Chamorro-Premuzic, 2008:181).

Another significant area in which the differences between introverts and extraverts have been explored is as they are presented in the working environment (Opt and Loffredo, 2003:566). According to Gilliland (Eysenck, 1970:29), studies on the difference between introverted and extraverted individuals in terms of job performance date back as far as the early 1900's. It is therefore worth investigating whether or not differences between introverted and extraverted behaviour will have an impact on the behaviour observation exercises used in assessment centers.

Although introversion/extroversion bias may impact behavioural observations in all psychological or behavioural dimensions (Hautala, 2005:84), this study will focus specifically on its effects on, and relevance regarding, leadership effectiveness and the assessment thereof in a [5]security environment. The reason for introversion/extraversion bias could be a result of various factors such as the pleasant or unpleasant feelings raters associate with being confronted with a personality type of a candidate that is similar or different than theirs. Another possibility is that extraverted and introverted raters may simply perceive different leadership constructs to be important. A further scenario that may result in rater bias could include raters focusing on individuals in a

[4] The researcher refers to the work of Eysenck (1970) to demonstrate that the concepts of introversion and extraversion have been studied in the field of psychology for some time.

[5] A security environment refers to a group of organisations with the same responsibility: to keep an institution, a geographical area or a country secure and safe from harm.

group who are more outspoken or quiet and therefore assessors could rate them differently to their peers.

Studies conducted on personality and leadership effectiveness usually involve only one party (either the rater's personality or the candidates' personality) (Hautala, 2005:84). In a study done by Opt and Loffredo (2003:566), it was clear that the candidates' personalities impacted the scores obtained during a performance rating (extraverts tended to score higher when they are rated on 'how' they communicate). Another more recent study by McCormack and Mellor (2002:193) also investigated the role of personality types in leadership and found that extraverted candidates are rated higher on transformational leadership (McCormack & Mellor, 2002:192). A study undertaken by Yun, Donahue, Dudley and McFarland (2005:97) on the other hand shows how the rater's personality also had a definite impact on the scores of performance ratings. The researchers prove that raters who scored high on 'agreeableness' gave higher ratings to candidates when they knew they had to give face-to-face feedback afterwards (Yun, et al. 2005:97). It is, however, significant that very few studies have explored the interaction effect between the personality types of both parties (rater introversion/extraversion as well as candidate introversion/extraversion) (Hautala, 2005:84).

A relevant study by Hautala (2005:84) explores how subordinates rate their leaders on transformational leadership and reaches the conclusion that extraverted and feeling-orientated subordinates appraise their leaders more positively than the subordinates who are introverted and thinking-orientated. Hautala's study does not confirm that the subordinates whose personality is similar to their leaders' personality will rate them higher. She suggests that further studies should be directed towards testing whether or not introverted raters would rate introverted candidates more favorably than they would rate extraverted candidates and vice versa (Hautala, 2005:89). The current study aims to explore this possibility.

1.2 PROBLEM STATEMENT

Assessment centers are often used by both public and private organisations for various purposes such as selection of new leaders, promotion of employees and development of managerial talent (Hermelin, et al. 2007:405; Lievens, 1998:141; Love & DeArmond, 2007:21). This is especially true for the security organisation in which this study was conducted.

Unlike other assessment methods, the strength of assessment centers lies in the fact that it combines various techniques to obtain a multi-dimensional view of the candidate (Talor, 2009:33). However, subjectively skewed representations of candidates by raters are still prevalent, especially during leadership selections (Berr, Church & Waclawski, 2000:133; Brutus, Fleenor & McCauley, 1999:417; Eidson & Gurman, 1998:8; Hilliard & Macan, 2009:162; London, Mone & Schott, 2004:319; Schyns & Felfe, 2006:522; Warr & Hoare, 2002:279; Yun, et al. 2005:97).

The problem statement for this study is therefore:

Raters might have a preference for candidates with similar personality types to their own which influences their final scores of these candidates. The more important impact of this bias is that these candidates are ultimately presented in a better light before the selection panel. This study will focus specifically on introversion/extraversion dimensions of personality.

1.3 PURPOSE OF THE STUDY AND RESEARCH OBJECTIVES

The general purpose of this study is to:

> Determine to what extent this introversion/extraversion bias influences the scores of behaviour observation exercises during leadership assessments in a security environment.

In order to achieve the stated purpose, the researcher will take the following steps:

i. Test for introversion or extraversion in candidates and raters by making use of the introversion-extraversion scale of the Jung Personality Questionnaire (JPQ).
ii. Test whether or not there is an interaction effect between the personality types (introversion/extraversion) of the raters and the personality types (introversion/extraversion) of the candidates on the scores that the candidates obtain during the behaviour observation exercise in the assessment center.

1.4 METHODOLOGY

A non-experimental, field study is used in this explorative study. Behaviour observation records were collected during two behaviour observation exercises that were conducted during a leadership assessment of a security company in South Africa. The candidates' as well the raters' introversion/extraversion preferences were measured by means of the JPQ. Statistical analysis was then conducted to test for an interaction effect between candidate and rater introversion/extraversion. Should an interaction effect exist, it would imply that introverted raters rated introverted candidates more favorably than extraverted candidates rated them. Conversely, it would also imply that extraverted raters rated extraverted candidates more favorably than they rated introverted candidates.

1.5 SIGNIFICANCE OF THIS STUDY

The significance of this study is two-fold. It impacts the broader field of leadership and personality studies, as well as the specific organisation in the security environment.

Firstly, there is limited research available on the topic of introversion/extraversion bias and a clear need for such a study is indicated by previous researchers (for example, Hautala, 2005:89). Secondly, the potential impact of this study on the leadership selection practices in this organisation could also be significant. Assessment results impact the future, growth or promotion of an individual and too often the decisions taken on selection panels can label an individual forever. If candidates are represented in a skewed manner, the organisation might lose potential leaders who will proceed and fulfill successful leadership positions outside the organisation. If an interaction effect

exists between a rater's personality type and a candidate's personality type, it can be used as the motivational force to create awareness among raters on their own preferences or bias during behaviour observation exercises. This awareness could better equip raters for the very important task of selecting the right employees for the right positions.

1.6 RESEARCH QUESTION

Based on the theoretical perspectives and the previous studies conducted, the following research question was developed:

> *To what extent does the rater's preference for introversion or extraversion in a candidate influence the scores in leadership assessments?*

The variables are operationalised as follows:

- Independent variable

  Introversion-extraversion scores of i) the rater and ii) the candidate as measured by the JPQ.

- Dependent variable

  Leadership in terms of scores on Communication, Adaptability, Judgment, Decisiveness and Team leading as measured by the raters during two behaviour observation exercises.

1.7 OVERVIEW OF THE CHAPTERS

The chapters of the research will be presented in the following manner:

Chapter 1: Introduction.

A background to introversion/extraversion bias and a statement of the problem, purpose and significance of this study.

Chapter 2: Theoretical perspectives on introversion and extraversion.

A summary of the existing literature related to personality theories, focusing on Jung's theory on personality types and more specifically how introversion and extraversion play out in the workplace.

Chapter 3: Theoretical perspectives on assessment centers and behaviour observation exercises.

A summary of the existing literature regarding assessment centers for leaders focusing specifically on assessment center activities and the (sometimes inadvertent) mistakes that raters make during this process. The topic of similarity bias is especially important in this chapter.

Chapter 4: Research methodology.

An overview of the hypotheses proposed by this research study and the methods used to test them and interpret the results. Some attention will also be given to explaining why the researcher made use of specific instruments.

Chapter 5: Interpretation of research results.

A complete representation and interpretation of the results of this study with specific reference to the description of the sample and an explanation of the hypotheses stated for the research.

Chapter 6: Findings, limitations and recommendations.

A comprehensive discussion on the findings obtained during this study and the implications thereof. Attention is also given to the limitations of this study and suggestions for further research.

CHAPTER 2: THEORETICAL PERSPECTIVES ON INTROVERSION AND EXTRAVERSION

2.1 INTRODUCTION

The introversion/extraversion construct has a long history of research within the field of psychology (Moon, Hollenbeck, Marinova & Humphrey, 2008:143). In order to understand fully the extent to which this construct forms part of the human personality and the way it relates to behaviour in an organisational setting, it is necessary to review the theoretical context from which it has been derived .

This chapter will begin by presenting a definition on personality as well as a broad overview of personality theories. Specific attention will be given to the theory of Carl Gustav Jung from which a discussion on introversion and extraversion will follow. The chapter will conclude with an overview of research relating to how introverts and extraverts are perceived in the workplace.

2.2 DEFINING PERSONALITY

Studies on the human personality have been the cornerstone of psychology since the early 20th century (Mulyanegara, Tsarenko & Anderson, 2009:235). Many attempts to define, rate and interpret human behaviour have been made; however, the complexities inherent in this concept still allow for new findings, definitions and explanations to be revealed (Momberg, 2005:15).

According to Meyer, et al. (1997[6]:11), the personality of an individual can be defined as the "…organisation of all the functions concerning the body, mind and spirit that determine the behaviour of the individual". Momberg (2005:15) adds to this definition and states that an individual's personality refers to the "…distinctive way in which he/she thinks, behaves and adapts to various situations". Both these definitions recognize that personality influences behaviour.

[6] The work by Meyer, Moore and Viljoen (1997) summarizes the different personality theories according to their historical development and the researcher therefore makes use of this reference. No other recent textbook or article provides the same information in the same format. The researcher performed a thorough search on EbscoHost under the following key words: 'personality theories', 'history of personality' and 'historical development of psychology'.

A recent definition by Mulyanegara, et al. (2009:235) highlights a different aspect regarding the organisation of these functions, namely stability. They describe personality as the "…intrinsic organisation of an individual's mental world that is stable over time and consistent over situations". A personality *theory* is therefore an attempt to explain individual differences by using a model of human functioning (Meyer, et al. 1997:20). In other words, personality theories aim "logically and consistently [to] explain, describe, assess and predict human behaviour" (Momberg 2005:16).

2.2.1 Personality theories and psychological approaches

According to Meyer, et al. (1997:8), there are over 30 personality theories worldwide, all of which revolve around the determinants for behaviour. Some of these theories already made their debut in the 19th century and are still being used, studied and verified today. Meyer, et al. (1997;20) states that these theories can be subdivided into *In-depth psychology theories* (including theories from Freud, Jung and Erikson) the *Behaviouristic theories* (theories from Skinner and Dollard), and the *Person orientated theories* (which includes theories from Murray, Maslow, Keelly and Frankl). The following section provides a broad overview of the different schools of thought regarding personality according to the latter categorisation.

2.2.1.1 In-depth psychology approaches

In-depth psychology is the broadest field of thought within psychology and it focuses on the deep, unconscious aspects of personality. This approach is based on the principle that the inner subjective conscious exist in layers that differ in terms of levels and degrees of consciousness. The outer layer is conscious and the deeper one delves the more unconscious the layers become. The inner layers are regulated by their own laws and according to this school of thought these laws are the main determinant of behaviour (Meyer, et al. 1997:45). Freud, Jung, Adler and Erikson have theorised on in-depth personality approaches. Their theories are summarised in Table 2.1.

Table 2.1: Summary of Theories from the In-Depth Psychology Approach

Theory	Description
Psychoanalytical theory of Freud	Freud pioneered in-depth psychology and attributed sex as an important determinant for human behaviour. His theory revolves around the principle that an individual is always in the middle of a power battle between his/her instinctive or subconscious drives and the rules of society. He divides the psyche into three structural concepts: The id, the ego, and the super ego. The id provides energy for human behaviour. The ego has to supply the needs of the id. And the super ego is a representative of the rules of society. (Meyer, et al. 1997:60)
Analytical theory of Jung	Jung performed a deeper analysis of the unconscious and follows a theological drive by placing special emphasis on the spiritual dimension. He believes that although past experiences have an influence on behaviour, an individual is also focused on continuous creative development to complete him/herself. Jung's theory is therefore less deterministic than Freud's theory and places some value on the conscious. (Meyer, et al. 1997:106)
Individualistic psychology of Adler	Adler's theory is based on a theological viewpoint. He believes that individual behaviour is determined by the striving for superiority and rising above his/her potential. However, each person is unique and is free to create his/her own goals and is therefore not entirely given over to the subconscious. (Meyer, et al. 1997:141)

Social-orientated psychoanalytic theory of Horney, Fromm and Sullivan	This theory rejects Freud's viewpoint that biological attributes determine behaviour as well as over emphasis on the sexual drive. It rather emphasises the role of social and cultural factors in the development of personality or behaviour and therefore moves the focus from the unconscious to the conscious. (Meyer, et al. 1997:166)
Ego psychological theory of Erikson	Erikson is famous for his theory about human development. He differentiates between eight development stages that stretch over the entire lifespan of an individual. He emphasises the development and the dynamics of the ego rather than the unconscious. He also espouses that it is possible for an individual to solve any developmental problems that occur at a later stage of his/her life. He therefore defines the individual as a being with a diverse range of needs, potential and opportunities that is accepted by the society. (Meyer, et al. 1997:213)
Post-modernistic approach of Lacan and Hillman	Lacan supports Freud's statement that human behaviour is mainly controlled by the irrational. Hillman places emphasis on the role of the irrational through focusing on the imagination which forms the bridge between the conscious and the unconscious. In this theory there is no place for self-actualisation as these theories postulate that the 'self' can only be found through others. (Meyer, et al. 1997:243)

2.2.1.2 Behaviourist and learning approaches

This school of thought focuses on the determinants of observable behaviour. According to psychologists who follow the behaviorist approach, all knowledge is gained through sensory experience and behaviour is therefore determined by events from the environment rather than internal forces (Meyer, et al. 1997:275). Table 2.2 provides a

summary of Behaviouristic theories which includes work by Skinner, Dollar and Miller as well as the Behaviourists who followed the socio-cognitive learning approach.

Table 2.2: Summary of Theories from the Behaviouristic Approach

Theory	Description
Radical Behaviorism of Skinner	Where abovementioned theorists reason that behaviour is determined by internal factors (for example, needs and drives), BF Skinner adopts a different perspective and defines behaviour as the result of environmental factors. He believes that behaviour will be repeated if it is reinforced and states that an individual is similar to an animal. In other words, the human being produces behaviour and is controlled by the results of that behaviour in the same way that the environment decides which behaviour should and shouldn't be repeated. (Meyer, et al. 1997:294)
Eclectic Behaviorism of Dollard and Miller	Just like radical behaviorism, eclectic behaviorism suggests that an individual's behavior is determined by the environment. However, where radical behaviorists limit themselves in terms of the outer observable factors, mild (or eclectic) behaviorists are willing to take internal factors (for example biological needs urges and fear) into account when explaining behavior and learning methods. (Meyer, et al. 1997:319).

Social-cognitive learning approach	This approach is similar to the previous two approaches because it also views behaviour as something that can be learned. Just as the eclectic and radical behaviourists, it focuses mainly on the observable behaviour of an individual. However, the social cognitive learning approach differs in the way in which it considers unobservable constructs like thoughts, symbolic processes, expectations and convictions when explaining the manifestation of behaviour. Furthermore, psychologists following the social-cognitive learning approach believe that 'reinforcement' does not necessarily play such a significant role in learning behaviour. (Meyer, et al. 1997:339)

2.2.1.3 Person-orientated approaches

The followers of the person-orientated approach do not agree with the psycho-analytic school of thought in which behaviour is only determined by unobservable urges and needs (Meyer, et al. 1997:373). They believe that behaviorists ascribe too much value to external forces when explaining human functioning. In contradiction to both earlier schools of thought, person-orientated followers wish to recover the individual's worth and describe the individual as a being that differs from the animal and that comprises a range of unique human attributes which enables him/her to function on a higher level. The basic principles that form the foundation of this school of thought follow:

i. The individual is a worthy being with unique human attributes like freedom of choice creativity, values, humour, autonomy, growth, actualisation and emotions.
ii. Conscious processes of the individual like the role of conscious decision making.
iii. The person as an active being who determines his/her own behaviour.
iv. Psychological health as a measurement.
v. The individual as an integrated, unique and organised whole.

(Meyer, et al. 1997:373-375)

The theories that evolved from the person-orientated approach are summarised in Table 2.3. This table contains work by Henry and Murray, Allport, Maslow, Rogers, Kelly and Frankl.

Table 2.3: Summary of Theories from the Person-Orientated Approach

Theory	Description
Need theory of Murray	Murray's theory is built on the key concept that the essential purpose of behaviour is to satisfy needs. In contradiction to the Freudians and the Behaviorists, Murray regards 'needs' as a versatile phenomenon consisting of the following: Biological needs, psychological needs, conscious needs and subconscious needs. He also views the process of need actualisation differently and states that satisfaction does not lie in a stress-free state, but in the process of reducing the stress. (Meyer, et al. 1997:391)
Holistic theory of Allport	Allport views the behavioristic and psycho-analytic approaches as incomplete and states that the individual is a complex being whose behaviour is determined by a various factors. He also believes that no two individuals are the same (even if they have the same genetic combinations or were raised in the same environment) and that each individual functions as a whole. He is therefore convinced that individuals should be studied in a unique way. Lastly, he believes that people behave in accordance with their set of values and self-developed goals. (Meyer, et al. 1997:425)

The self-actualisation theory of Maslow

The most important theme in Maslow's work is the human potential and the possibility of a better existence. According to Maslow, a large part of behaviour can be explained by the process of 'need satisfaction'. He arranges the inherent, basic needs of an individual into a hierarchical order of urgency. At the bottom, there are *physiological* needs. It then continues towards *safety* needs, *affiliation* needs; *self esteem* needs and finally ends at *self-actualisation* needs. Maslow states that the higher-level needs can only be satisfied if the lower level needs are fulfilled. The individual will therefore move from one level to another, as each need is fulfilled. In other words, the drive towards self-actualisation is the basic motive that underlies all behaviour. (Meyer, et al. 1997:461)

Self-actualisation theory of Rogers

Rogers bases his theory on three assumptions:

1. The individual has constructive potential;
2. The nature of the individual is driven towards achieving goals; and
3. The individual can change.

He also emphasises the role of the self-concept on personality and behaviour. According to Rogers, the individual is the central figure in reaching his/her own potential and the environment will only either facilitate or inhibit the process. (Meyer, et al. 1997:486)

Personal construct theory of Kelly

Kelly describes the individual as a scientist. Just like a scientist creates hypotheses and conducts experiments to confirm or test these hypotheses, the individual continuously strives to predict and control

	his/her own behaviour. He also focuses more on the cognitive side of behaviour and states that an individual's constructs (interpretations of reality) will determine how he/she will behave. (Meyer, et al. 1997:524)
Existential theory of Frankl	Part of Frankl's theory was developed during his captivity in Auschwitz, a Nazi concentration camp. He realised that a lot of the prisoners survived because they had something to live for. His conclusion is that values play a significant role in a person's life and that the 'search for meaning' is the true driver of human behaviour. He believes that an individual has the freedom to take responsibility for his/her own life. (Meyer, et al. 1997:556)

As can be seen in Tables 2.1–2.3, the field of personality is extremely broad and it is not necessary for the purposes of this study to explain and elaborate on each one of the theories. What is of relevance to this study is the emergence of different measurements of personality. Based on some of these theories, specific personality measurements or tests have evolved that measure either personality traits or personality types.

2.2.2 Type or trait

The difference between trait-theories and type-theories is quite simple. Trait-theories are based on the belief that although an individual's behaviour varies to some extent from one occasion to the next, there are still some constructs that remain stable and which define the individual's true nature (Pittinger, 2005:211; Myers, et al. 2009:6). These characteristics are known as personality traits and they are usually measured by using a continuous variable (for example, conscientiousness) that ranges between two extremes (low conscientiousness to high conscientiousness) (Pittinger, 2005:211).

The scores derived from trait measurement indicate the magnitude of a specific construct. In other words, the scores reflect the extent to which a specific trait is demonstrated or how much of that particular characteristic the individual possesses (Matthews, Deary & Whiteman, 2003:4; Pittinger, 2005:211; Myers, et al. 2009:10). Trait research was developed by researchers like Cattell and Eysenc and forms the nucleus of the scientific study of human personality (Matthews, et al. 2003:5).

Type theories, on the other hand, are based on the belief that personality types are formed whenever these personality traits are clustered together into groups. The underlying assumption is that there are different populations of people who express different personality characteristics; the typology sorts individuals into equally valuable groups to which, according to the theory, they already belong (Matthews, et al. 2003:5; Pittenger, 2005:211; Meyers, et al. 2009:10). The study of types amongst human beings was preceded by oriental astrologers and the physiological typology of Greek medicine. Carl Gustav Jung later developed a typology system for personality (Matthews, et al. 2003:12; Pittenger, 2005:211).

An instrument that measures a personality type assumes that an individual might display various traits from different categories, but that one group of characteristics would still be dominant and therefore most preferred. The instrument therefore measures the strength of the preference and not the strength of the behaviour (Matthews, et al. 2003:12).

This study will focus on personality theories that refer to types and not traits, more specifically the Analytical theory of Jung. Psychometric measurements based on Jung's theory have been quite popular in organisational and leadership studies. It is also one of the most commonly used and applied approaches to understanding individual personality (Berr, et al. 2000:134-135; Roush & Atwater, 1992:19). The next section will provide an overview of Jung's theory and how he used it to develop a typology for human behaviour.

2.2.3 Carl Gustav Jung's personality theory

Carl Gustav Jung was born in a small town in Switzerland and grew up in a very religious home, which had a significant influence on his theoretical viewpoints (Meyer, et al. 1997:105). He studied medicine and began practicing as a doctor at the Burgholzli Psychiatric Hospital. This turned his attention more towards psychiatry. Jung also became interested in Freud's psychoanalytic viewpoints and started to apply some of these in his practice. It was not long before Freud and Jung met and became friends. (Meyer, et al. 1997:106)

Freud saw Jung as his next follower but, due to irreconcilable differences in theoretical viewpoints, the two ended their friendship. Jung gradually developed his own school of thought known as *Analytical Psychology* (Du Toit, 1983:33) which is concerned with the movement of psychic energy and the way in which one habitually or preferentially orientates oneself in the world (Sharp, 1987[7]:10). This system of classifying an individual is relatively contradictory to earlier classifications as it does not divide behaviour into temperamental or emotional patterns (Sharp, 1987:12).

Jung believes that there are three basic differences in the way human beings prefer to apply their minds. Firstly, there are differences in the way in which individuals perceive, or take in information during, an experience. Jung refers to this as the *perceiving* part of personality and it is usually experienced irrationally. Individuals can perceive an experience in two contrasting ways, via *sensing* (S), which is directly through their five senses, or via *intuition* (N), which is through indirect attention to associations and relationships in experience (Myers, et al. 2009:22; Opt & Loffredo, 2003:560).

The second difference is related to how individuals process or reach conclusions about that which they have perceived. Jung labels this aspect of personality the *judging* (J) part which refers specifically to rational processing of information. He claims that individuals evaluate their perceptions either by *thinking* (T), which is a logical process or

[7] The researcher refers to the work of Sharp (1987) as it explains and summarises Jung's theory on introversion and extraversion in a clearly and logically.

by *feeling* (F); that is, applying personal or subjective values (Myers, et al. 2009:22; Opt & Loffredo, 2003:560).

Jung believes that the perceiving and judging processes form the core of all human personalities. He also believes that these processes determine an individual's preference for *sensing* over *intuition* or *thinking* over *feeling* (or vice versa) (Myers, et al. 2009:22; Opt & Loffredo, 2003:560).

The third area of difference refers to attitude or the way in which individuals spend or derive their energy. Attitude merely refers to "a readiness of the psyche to act or react in a certain way…having an attitude is synonymous with an *a priori* orientation to a definite thing" (Myers, et al. 2009: 22). In other words, some individuals might prefer to channel their energy more towards the inner world of experience (which refers to *introversion*) where other individuals might prefer to channel their energy more towards the outer world of experience (which refers to *extraversion*) (Opt & Loffredo, 2003:560).

This way of clustering individual preferences is known as a typology and led to the development of eight typological groups: two personality attitudes (introversion and extraversion) and two modes of orientation (judging and perceiving) consisting of four functions (thinking, feeling, sensing and intuition) (Sharp, 1987:12). A function can be defined as a "particular form of psychic activity that remains the same in principle under varying conditions" (Myers, et al. 2009:22). Jung later specified that each attitude can be habitually used in conjunction with a specific function, which leads to a specific type of personality (Myers, et al. 2009: 22).

From this assumption, four scales or dichotomies were developed on which personality can be measured, namely 'introversion vs. extraversion', 'sensing vs. intuition', 'thinking vs. feeling' and 'judging vs. perceiving' (Myers, et al. 2009: 22). In every dimension an individual will exert a stronger preference towards one point of the scale than the other. For example, a person might prefer to practice extraversion rather than introversion or to sense rather than to perceive an experience. When combining all the stronger preferences, a person's personality type emerges (Hautala, 2005:90; Myers, et al. 2009:22).

Considering the above information, it is important to note a comment by Sharp (1987:89) who states that "no system of typology is ever more than a gross indicator of what people have in common and the differences between them". This simply implies that Jung's model only distinguishes by its parameters and identifies the common preferences or differences between individuals. It therefore cannot reveal the uniqueness of an individual. However, Jung's opinion is that "without a model of some kind, we are simply adrift in a morass of individual opinions – lost in a jungle without a compass" (Sharp, 1987:90). In other words, this model is useful when dealing with methodical investigations and understanding differences amongst individuals (Myers, et al. 2009:26; Sharp, 1987:90).

Considering the above explanation, it can be derived that Jung identifies numerous differences between individuals. However, for the purposes of this study it is crucial to obtain a better understanding of two specific elements of Jung's typology: the attitudes of *introversion* and *extraversion*. The next section will provide an overview on these two concepts.

2.2.4 Two kinds of energy: introversion and extraversion

Jung's initial motivation for developing these two concepts was his need to define ways in which Alfred Adler's theory differed from that of Sigmund Freud (Dolliver, 1994:192; Sharp, 1987:25). Jung classified Adler as having an introverted theory of personality (due to the emphasis that Adler places on a person's desire to seek his own security and supremacy) and Freud as having an extraverted theory (due to the emphasis Freud places on how a person is dependent on significant objects such as the parents) (Dolliver, 1994:193; Sharp, 1987:25). In other words, Adler's theory suggests that human behaviour is conditioned by the subject (introversion), where Freud's theory suggests that the determining factor is in the object (extraversion) (Sharp, 1987:26).

Jung expresses his appreciation for both introversion and extraversion when he states:

> *"The Freudian theory is attractively simple, so much so that it almost pains one if anybody drives in the wedge of a contrary assertion. But the same is true of*

> *Adler's theory. It is of illuminating simplicity and explains as much as the Freudian theory"* (Sharp, 1987:26).

It is therefore evident that Jung does not search for evidence to prove that one theory is superior to the other. Rather, he wishes to determine why an individual might prefer an introverted attitude over an extraverted attitude or vice versa. He concludes that it is not a matter of choice but rather a matter of random distribution. Jung explains this phenomenon as follows:

> *"..These attitude types are found in both sexes and at all levels of society. They are not a matter of conscious choice or inheritance or education. Their occurrence is a general phenomenon having an apparently random distribution"* (Sharp, 1987:26).

Jung believes that human beings have a biological, unconscious and instinctive tendency to adapt to the environment which ensures the continual existence of the living organism (Sharp, 1987:26). The two attitudes therefore enable human beings to adapt to their environment in different ways. The reasons why people adapt to their environments in different ways is unknown and remains a debatable topic (Myers, et al. 2009:26; Sharp, 1987:27).

While researchers and psychologists continue to contemplate the causes for introversion or extraversion, it is important for the purposes of this study to understand the essence of each construct, especially given that the current study is based on the assumption of unconscious bias. It is also necessary to understand how these two different attitudes are displayed through behaviour, as this is the only way in which raters will be able to distinguish between introversion and extraversion. Unfortunately, 'introversion' and 'extraversion' are frequently misused terms and should therefore be clarified. The next section will provide a detailed explanation of both concepts as they are referred to in this study.

2.2.4.1 Extraversion

Everyone is affected by external (objective) data but what differentiates the introvert from the extravert is that the latter's thoughts, decisions and behaviour patterns are determined (and not only influenced) by objective conditions rather than subjective (internal) views (Myers, et al. 2009:26). Sharp (1987:37) states that "…the moment when one's conscious orientation is determined by objective reality, the given facts in the outside world, we can speak of an *Extraverted* attitude. When this is habitual, we have an *Extraverted* type".

Extraverts are people who derive and direct their psychic energy (libido) from and towards the outside or objective world. In other words, external factors (objects) are the predominant motivating force for judgments, perceptions, feelings and actions (Bono & Judge, 2004:901; Du Toit, 1983:4; Hautala, 2005:84; McNiel, Lowman & Fleeson, 2010:19; Oviedo-Garcia, 2007:676; Myers, et al. 2009:26; Sharp, 1987:13).

Jung believes that it is possible to distinguish from very early in an individual's life between an introverted and an extraverted attitude. Extraverted children usually adapt to their environment quite quickly and are recognized by the extraordinary attention they give to objects. External factors have little effect on them and they usually move around with confidence as they become familiar with the unknown. They are also not afraid to expose themselves to risks (Myers, et al. 2009:26; Sharp, 1987:38).

The same is true for the adult. Extraverts have an extraordinary capacity to adapt to their environment, which may be both a strength and a limitation. Although they can readily accept external happenings and put their inner life second to outer reality, they may easily forget to pay attention to the stress they put on their own bodies in order to fulfill objective demands. They trust what they receive from the outside world and are similarly disinclined to submit personal motives to critical examination. Prevailing moral standards can therefore dictate the extravert's personal standpoint. (Hautala, 2005: 85; Myers et al, 2009:26; Sharp, 1987:37-44)

The extravert's personality can further be characterised by an outgoing, candid and accommodating nature. They are typically those who would feel a loss of energy when

they are alone for too long. They prefer to be active, energetic and to seek excitement through engaging with the environment. They like to travel and meet new people. They have an innate desire to influence or be influenced by events. Extraverts are usually talkative and they cultivate a positive attitude in life. At times they might appear assertive and dominant but they do tend to add a lot of value to their relationships. Providing that the extravert is not too pushy or assertive, he/she can be a very useful member of the community and an asset in social situations. Lastly, there is evidence that extraverts perform very well in certain aspects of leadership[8]. (Bono & Judge, 2004:901; Du Toit, 1983:4; Eaton, Collings, D'Agati, Moore & Kellar, 2007:175; Hautala, 2005:84; McNiel, et al. 2010:19; Myers, et al. 2009:26; Sharp, 1987:13; Oviedo-García, 2007:676).

2.2.4.2 Introversion

Where external conditions or objective data are the determining factors for the extravert's behaviour, the introvert's thoughts, decisions and behavioural patterns are determined by internal views or subjective data. In other words, energy moves towards the inner world and the inner reality is of primary importance (Hautala, 2005:86; Myers, et al. 2009:26; Sharp, 1987:13). Although introverts are well aware of external conditions, the subjective determinants from within are habitually the decisive motivating force (Myers, et al. 2009:26; Sharp, 1987:65).

From childhood introverts are reflective, thoughtful and shy while distrusting the unknown. They usually assert themselves over familiar objects and attempt to master them. Introverted children also want their own way and when they ask questions, it is not out of mere curiosity but because they want meanings and explanations to protect them against the object (Killgore, et al. 2007:355; Sharp, 1987:68).

The abovementioned attributes manifest in the adult life of introverts too. Introverts are mainly interested in the "world of concepts, ideas and inner experiences" (Myers et al, 2009:26). When introverts are compelled to function in big groups, they tend to lose energy and they tire quickly. However, they are likely to contribute freely when they are

[8] Extraverts' performance in leadership roles will be discussed at a later stage in this study.

familiar and comfortable with a topic, even with strangers. In general, they prefer to seek in-depth involvement with others rather than socialising in big groups. Introverts usually show enormous respect for the individuality of others and expect the same in return. They are however often perceived as withdrawn, hesitant, reflective and always slightly on the defensive. (Du Toit, 1983:5; Sharp, 1987:69; Myers, et al. 2009:26; Killgore, et al. 2007:355)

Introverts feel the need to prove that everything they do is based on their own decisions or convictions and never because they are influenced by anyone (Sharp, 1987:65). This fear to fall under hostile influences takes a lot of energy and the introvert might, therefore, struggle with inner conflict at times during his/her life (Sharp, 1987:69; Myers et al, 2009:26; Killgore, et al. 2007:355).

Lastly, Jung believes that although introverts prefer to spend time on their own, they are by no means a social loss. He states that "…the introvert's retreat is not a renunciation of the world, but a search for quietude" (Sharp, 1987:69). Whenever an introvert is alone it is possible that he/she could make a contribution to the life of the community. In fact, where the extravert usually avoids introspection, it is a pleasure for the introvert and a process with which he/she is quite comfortable. The introvert's best work is done by his/her own resources and on his/her own initiative and it is usually successful (Sharp, 1987:69). Du Toit (1983:4) agrees and gives further input by stating that the reason why introverts are often underestimated is because they tend to hide their best qualities.

While considering the basic, characteristic traits of introverts and extraverts, one should keep in mind that the descriptions offered above are two extreme cases. One may often find an individual that obtains a medium score during a measurement of introversion or extraversion. This simply means that he/she might possess qualities from both personality types as explained in Section 2.2.2. It might also imply that an individual prefers to adopt a different style in different situations.

Sharp (1987:89) is of the opinion that it is impossible to be totally introverted or totally extraverted, nevertheless he believes that classifying an individual as an introvert or an

extravert does help in understanding behaviour. He states the following: *"...one can never give a description of a type, no matter how complete, that would apply to more than one individual, despite the fact that in some ways it aptly characterizes thousands of others. Conformity is one side of a man, uniqueness is the other"* (Sharp, 1987:89).

This typology portrays the way in which individuals from these two personality types are perceived in general. This is no different to the raters' perception of individuals during an assessment situation and it is therefore critical to understand the effect of subjectively skewed perceptions when raters allocate scores, especially if certain attributes of introverts and extraverts are associated with leadership and job performance. The next section will provide a brief overview of introverted and extraverted behaviours in the work context.

2.3 INTROVERSION/EXTRAVERSION IN THE WORKPLACE

The main question that should be asked in the context of this study is whether or not personality, specifically referring to introversion/extraversion, can predict job performance? Almost 50 years ago Guion and Gottier (1965:159) expressed their skepticism regarding the use of personality measures in selection processes. However, Hogan and Holland (2003:100) state the contrary in their study: *"Since 1990, meta-analytic reviews have shown that personality measures are useful predictors of job performance"* (Hogan & Holland, 2003:100).

It is therefore evident that during the past 20 years this topic has undergone a radical transformation. Judge and Erez (2007:573) confirm this statement when they state that contemporary research does show that personality, or at least some measures of it, can influence work behaviour and performance. The next section will focus on studies by Rust (1999), Beaty, Cleveland and Murphey (2001), Opt and Loffredo (2003), Barrick and Mount (2005), Moon, Hollenbeck, Marinova & Humphrey (2008), Judge and Erez (2009), as well as Minbashian, Bright and Bird (2009) which explore this principle, specifically referring to introversion and extraversion.

A study conducted by Rust (1999:99) tested the ability of the Big Five factors[9] to predict the specific behaviours they were designed to measure. Rust made use of a questionnaire called the *Orpheus Personality Questionnaire* which is based on the five factors of the Big Five Model. The Orpheus' measure of *Fellowship* assesses *Extraversion* as one of the Big Five traits. The results show that individuals who scored high on Fellowship are happier to work in a team where individuals who scored lower generally prefer work that allows them to be independent (Rust, 1999:100). This is unsurprising as the preference to function within a group is true of the extraverted type and the preference to be more independent is true of the introverted type, according to Jung.

A few years later, Opt and Loffredo (2003:560) conducted a study that shows that extraverted individuals tend to have a more positive communicator image. This attribute refers to "…the way in which one verbally, nonverbally, and paraverbally reflects how literal meaning is taken, interpreted, filtered or understood" (Opt & Loffredo, 2003:560). Again, this is unsurprising as extraverts are more likely to express their thinking processes to the external world. On the other hand, it was proven during the same study that introverted individuals have higher apprehension scores when it comes to grasping verbal and non verbal messages. This could be due to the fact that introverts are more likely to focus on the internal meaning of a message.

A finding significant to this study and which supports personality as a predictor of job performance is explained in Barrick and Mount's study during which they investigate the predictive validity of the Big Five personality factors in the workplace (Barrick & Mount, 2005:359). They conclude that extraversion is a valid predictor of job performance; however, they emphasise that this finding is only applicable in certain niches or areas in the workplace. For example, extraverted individuals tend to perform in sales jobs in which a significant requirement of the position involves interacting with others, influencing others and obtaining status or power (Barrick & Mount, 2005:364).

[9] The Big Five model of personality was developed during the 1970's by Costa, McCrae, Goldberg and Norman. They believed that personality can be summarised by five broad factors called *Extraversion, Agreeableness, Conscientiousness, Neuroticism* and *Openness (Rust, 1999:99)*.

A study by Judge and Erez (2007:574) supports this finding and states that the strength of the relationship between introversion/extraversion and job performance might differ from one situation to the next. They attempt to improve predictive validity coefficients by testing extraversion together with other traits like emotional stability. They find that the combination of, or interactions between, certain traits are more likely to predict job performance.

Another study that focuses specifically on introversion and extraversion in the workplace is that of Moon, et al. (2008:143). They believe that an individual's level of extraversion is highly transparent in organisations and therefore they investigate the relationship between extraversion and Organisational Citizenship Behaviour (OCB) (Moon, et al. 2008:143). OCB is defined by Moon, et al. (2008:143) as the "…beneficial behaviors exhibited by employees towards their employing organisation". It may include behaviours such as helping others, compliance and sportsmanship. Theory around OCB states that such behaviour contributes to the wellness of the organisation as a whole (Moon, et al. 2008:144). Some facts important to this study are presented in their research.

Firstly, they find that extraversion is the only factor of the Five Factor model that has more than one potential descriptor. In fact, their research shows that since the beginning of personality studies, there has been a measure of frustration among authors to define extraversion as a broad factor. Cattell also viewed the extraversion factor as a complex combination of different personality traits and therefore an irrelevant construct for predictive purposes (Cattell in Moon, et al. 2008:144).

Moon, et al. (2008:144) supports Cattel's view by referring to a study performed by Smith, Organ and Near (1983:660) who are some of the first authors that were interested in testing the relationship between the Big Five personality factors and OCB. Their results fail to show a significant correlation between extraversion (measured as a broad factor) and any behaviour related to organisational citizenship.

The same is true in a study by Beaty, Cleveland and Murphey (2001:147) who also test for extraversion (as a broad factor) and OCB. Although they find some support for an

extraversion/OCB link in a laboratory, the link is considered minimal and not applicable in a field study.

They conclude:

> *"…to date, there has been no substantial link, either empirically or theoretically, establishing a relationship between the broad factor extraversion and an individual's tendency to demonstrate citizenship behaviours in a work environment"* (Beaty, et al. 2001:147).

Moon, et al. (2008:144) therefore propose that extraversion should be broken down into two traits called *Sociability* and *Surgency* (which add to the broader construct of *extraversion*). These traits should then be measured individually when testing for relationships between extraversion and OCB. When using the factor Sociability (interpersonal warmness) to predict organisational related behaviours, Moon, et al. (2008:146) found that such individuals will have the tendency to form and maintain positive work relationships, be drawn to team-orientated organisational cultures, and also to seek higher levels of stimulation in their work environment.

The factor Surgency indicates the tendency to seek rewards. Moon, et al. (2008:146) also find that individuals who scored high on Surgency, are more likely to be sensitive towards reward strategies and will be drawn to an outcome-orientated organisational culture.

Lastly, a study conducted by Minbashian, Bright and Bird (2009:574) also shows that when extraversion is broken down into traits called *Potency* and *Affiliation,* it provides better measures. Minbashian, et al. (2009:574) establish that Potency is positively related to getting ahead in a working environment.

The abovementioned studies show that certain attributes of introversion and extraversion might lead to job performance. However, it is very useful to keep the comment made by Barrick and Mount (2005:364) in mind. They state that although these studies enhance the theoretical understanding of the validity of specific personality traits predicting job performance, they also prevent us from examining the

predictive validity of the whole personality of an individual (Barrick & Mount, 2005:364). In other words, if the aim is to determine how well personality predicts job performance, it is necessary to study people (meaning, the interaction of attributes and traits) and not traits in isolation.

The same is true when using personality measures to predict leadership effectiveness. Just as a combination of different personality preferences might lead to performance in different work situations, so the same combination of preferences could lead to performance in different aspects of leadership. It is therefore worthwhile to investigate the theoretical viewpoints on how introversion and extraversion relate to leadership effectiveness.

2.4 INTROVERSION/ EXTRAVERSION PREDICTING LEADERSHIP EFFECTIVENESS

Various researchers have debated whether or not analysis of introversion/extraversion in isolation predicts effective leadership behavior; however, results have been inconsistent in this regard (Berr, et al. 2000:148; McCormack & Mellor, 2002:179). As a result of the assertiveness and sociable aspects of extraversion, a common perception might exist that extraverted individuals would perform better in leadership positions (Ployheart, Lim & Chan, 2001:811). However, studies have shown otherwise.

Berr, et al. (2000:148) investigate the impact of personality variables on managerial behaviours and find that it is better for managers or practitioners not to value introverted candidates over extraverted candidates or vice versa. While they acknowledge some research which shows that extraverted individuals tend to be more successful in certain job-related aspects than introverted individuals (McCaully & Roush in Berr, et al. 2000:148), they also add that there may be other dynamics (such as attributes of self-awareness) occurring that could lead to the perception that extraverts are better leaders.

McCormack and Mellor (2002:181) also investigate the role of personality in leadership, initially proposing that effective leaders would be associated with high extraversion.

Although the results of their study indicate that personality is an important contributing factor to leadership, their study does not substantiate a correlation between extraversion and leadership effectiveness. In fact, the individuals who are classified as effective leaders actually tested lower on extraversion (McCormack & Mellor, 2002:181).

Useful to this study which is also based on findings from a security environment is Horey and Fallesen's (2004:29) research regarding personality attributes that predict leadership effectiveness in a military environment as indicated by previous literature. Some attributes listed include autonomy, endurance, responsibility, honesty, integrity and resilience, to name a few. It is interesting to note that not one of these attributes points to extraversion or introversion (Horey & Fallesen, 2004:29).

Bono and Judge (2004:908) extend their current knowledge regarding the association between personality and transformational and transactional leadership. Their results indicate that extraversion could be an important trait to predict leadership; however, they only correlated with 0.24, which indicates a very weak statement.

According to Culp and Smith (2005:46), attributes associated with both introversion and extraversion can present certain challenges and benefits in leadership effectiveness. For example, introverts prefer to think things through before voicing their opinion which often leads them to present a well-thought through plan. This might cause them to perform well in areas associated with leadership such as being perceived to be credible, organised and productive. On the other hand, the same attribute can lead them to underperform in areas like boldness, dynamism and friendliness (Culp & Smith, 2005:47). The same is true for extraverts. Their lively nature might cause them to perform well in areas related to leadership like communication, empowering others and obtaining status. At the same time extraverts might be penalised in areas like being focused, respecting others and commitment to ideas (Culp & Smith, 2005:47).

It is therefore logical to conclude that, for the purposes of the current study, the researcher believes that neither introversion nor extraversion in isolation indicates effective leadership behaviour. Theory clearly states that introversion and extraversion alone cannot predict leadership effectiveness and that one should rather investigate the

work demands of a position and how certain traits relating to these two types can predict an individual's success in a specific role (Culp & Smith, 2005:39).

2.5 CONCLUSION

This chapter offers an overview on the concepts of introversion and extraversion. It begins with a reference to the different schools of thought in psychology and then later focuses specifically on the theoretical foundation from which these concepts are derived (namely the personality theory and typology of Jung). Typical attributes associated with introversion and extraversion have also been discussed. This chapter concludes by presenting the viewpoints of previous studies exploring controversial topics such as the predictive validity of introversion/extraversion in job performance and leadership effectiveness.

Considering all the above information, it may be concluded that personality can be used as an indicator of preference towards certain settings but not necessarily for predicting leadership effectiveness per se. On the other hand, the researcher acknowledges that some studies contradict this statement. In spite of limited evidence proving that introversion/extraversion predicts leadership effectiveness, it is still clear that certain assumptions exist with regard to how these two personality types will behave. These assumptions could result from the vast amount of research that has automatically attributed certain characteristics to introverts and certain characteristics to extraverts, and if not acknowledged it could influence a performance measurement setting.

The next chapter will provide a theoretical framework on a well-known performance measurement method, namely the assessment center. It will also explore how with regards to introverts and extraverts, these assumptions can evolve into bias which can in turn influence the performance ratings of candidates.

CHAPTER 3: THEORETICAL PERSPECTIVES ON ASSESSMENT CENTERSAND BEHAVIOUR OBSERVATION EXERCISES

3.1 INTRODUCTION

This study focuses on leadership assessment. This chapter will therefore start with an overview of leadership assessment in a security environment (that is, police, defense force, emergency personnel, and so forth) and then proceed by introducing the concept of assessment centers. It will present a detailed discussion on the different activities used in assessment centers, specifically on the behaviour observation exercise and some of the elements that may influence the scores that raters allocate to candidates. The chapter will then proceed to discuss rater biases, which reflect the theme of this study. The chapter will conclude with a discussion on introversion/ extraversion bias and highlight some previous studies on this topic.

3.2 AN OVERVIEW OF LEADERSHIP ASSESSMENT TECHNIQUES IN THE SECURITY ENVIRONMENT

Before one can explore the various ways to assess leadership, it is first important to obtain an understanding of the concept. Gary Cohen defines leadership as "…the process whereby an individual influences a group of individuals to achieve a common goal" (Cohen, 2009:16). According to Van Iddekinge, Ferris and Hefner (2009:463), organisational leadership is one of the main contributors to organisational effectiveness and it is a widely-researched topic in the organisational sciences. Bennis supports this statement and adds that the urgency to study leadership in such a setting should not be underestimated as the existence of the organisation depends on it (Bennis in Van Iddekinge, et al. 2009:463).

One of the most relevant questions or issues regarding organisational leadership is the assessment thereof. In addition to the vast number of methods available, the legal and ethical requirements to which professionals must adhere should also be taken into consideration (Taylor, 2009:9). These requirements place a heavy burden on professionals which often causes them to apply quick and immediate options that yield inconclusive results (Taylor, 2009:9). In order to understand the true value of

assessment centers, it is first necessary to gain some insight regarding the most frequently used methods used to select leaders in the environment in which this study takes place: the security environment.

An overview of the available literature establishes that four approaches to leadership assessments are commonly used in a security and business environment (Gal & Mangelsdorff, 1991:39; Grobler, et al. 2006:147; Stamoulis, 2009:68). These include paper screening, psychometric testing, assessment centers and construct-orientated psychological assessments (COPA). Each one of these will now briefly be discussed.

3.2.1 Paper screening

Paper screening is the oldest assessment method and can be described as a process whereby a candidate is assessed or selected according to the sources of information he/she provides to the organisation (Stamoulis, 2009: 69). In a business environment, such sources may include application forms, biographical forms, reference reports, curriculum vitae and educational results (Grobler, et al, 2006:147). On the other hand, managers in the security environment might rely more on sources like service records, course results, medical records and performance evaluations (Gal & Mangelsdorff, 1991: 40).

Apart from the heavy administrative burdens it may cause for organisations, managers still prefer using paper screening methods due to the fact that it is cost-effective and very easy to use. If there is a certain set of criteria involved, candidates that are obviously ill-suited for the position can easily be screened out before even entering the assessment process (Gal & Mangelsdorff, 1991:40; Grobler, et al. 2006:148; Stamoulis, 2009:70).

Although this method provides a quick way to get an overall view of a candidate, paper screening is still an invalid method if used in isolation since it cannot give a current and real indication of leadership effectiveness. There is the added risk that, the sources of the information could be corrupt. Finally, the role of the psychologist/professional is limited in this process (Gal & Mangelsdorff, 1991:41; Stamoulis, 2009:70).

3.2.2 Psychometric tests

This method makes use of psychometric instruments to determine the cognitive ability, personality preferences or aptitude of an individual (Scroggins, Thomas & Morris, 2008:158). Psychometric instruments rest on the principle that certain skills or abilities (for example numerical reasoning and verbal aptitude) might predict success in a certain position and by measuring these attributes it is possible to distinguish between suitable and unsuitable candidates. This approach is therefore useful when identifying the right candidates for positions where certain technical or academic skills are needed or for positions that involve a high level of complexity (Gal & Mangelsdorff, 1991:41; Grobler, et al. 2006:148; Scroggins, et al. 2008;158).

As with paper screening methods, psychometric instruments also have certain limitations. Although studies have shown that psychometric instruments are very good indicators of whether or not a person will be able to function in a certain environment, this is only true if the instruments that are being used are proven to be valid and reliable in the specific environment in which they are used (Scroggins, et al. 2008;159). This method may also be very expensive and it requires a team of professionals to administer the test and interpret the results (Grobler, et al. 2006:148; Scroggins, et al. 2008:193). Lastly, it cannot be used in isolation, as it only predicts cognitive or technical skills and preferences or capabilities, but it does not predict leadership behaviours. Even though researchers have tried to identify certain abilities that might contribute to leadership effectiveness, it is still not always possible to capture a candidate's leadership abilities in a test situation (Gal & Mangelsdorff, 1991:41; Scroggins, et al. 2008:193).

3.2.3 Assessment centers

Unlike the two previous methods, assessment centers make use of a combination of methods, for example paper screening and psychometric instruments, to assess various aspects like career history, mental capacity, personality preferences and the leadership abilities of an individual (Taylor, 2009:10). One aspect that further differentiates assessment centers from the rest of the assessment approaches is that individuals are

placed in a setting in which actual pre-determined leadership behaviours are elicited while they are being rated. It therefore gives candidates the opportunity to show what behaviours they are capable of. (Gal & Mangelsdorff, 1999:43; Schlebush & Roodt, 2008:2; Taylor, 2009:11)

Some limitations of this method are that it requires a lot of pre-work (job profiling, administration, etc.), multiple raters who are experienced and trained, as well as the use of psychometric instruments which may be costly and might require some time. However, when this method is used correctly it has a high predictive power for leadership effectiveness and a very comprehensive view of an individual is obtained (Coleman & Adams, 1999:27; Hermelin, Lievens & Robertson, 2007:406; Lievens, 1998:141; Lievens & Van Keer, 2001:373; Schlebush & Roodt, 2008:3; Taylor, 2009:12).

3.2.4 Construct-orientated psychological assessments (COPA)

The COPA approach is very similar to the assessment center method discussed in the previous paragraph; however, clinical psychologists are much more involved during the COPA approach than they are during the assessment center (Gal & Mangelsdorff, 1991:41). In other words, the focus of the COPA is usually on assessing psychological health and stability within the work environment, while the assessment center evaluates work-related skills and capabilities.

The COPA approach has the same pitfalls as the other assessment methods discussed; however, emphasis should be put on the fact that clinical- or personality-related tests are not the only indicators of leadership performance and should therefore not be considered in isolation. One aspect that will definitely increase the validity and reliability of the COPA method is the systematic use of techniques that are specifically aimed at measuring a specific profile (Gal & Mangelsdorff, 1991:41).

The COPA approach, together with all the previously mentioned methods add value in an assessment setting. The paper screening method and the psychometric instrument method are especially useful when assessing candidates for academic and technical performance. However, when it comes to leadership performance, the assessment

center and COPA method prove to have the most predictive power (Coleman & Adams, 1999:27; Hermelin, Lievens & Robertson, 2007:406; Lievens, 1998:141; Lievens & Van Keer, 2001:373; Schlebush & Roodt, 2008:3; Taylor, 2009:12). These two approaches do not only combine various elements of the methods discussed, they also elicit the required leadership competencies in the participating candidates (Taylor, 2009:13).

For the purposes of this study, the next part of the chapter will focus on assessment centers and the components that make it a valuable tool for assessing leadership performance.

3.3 ASSESSMENT CENTERS

The assessment center is certainly not a new phenomenon and has been used extensively by the military service (especially for officer selection) in the USA and Great Britain since World War II (Coleman & Adams, 1999:28; Gowing, Morris, Adler & Gold, 2008:437; Schlebusch & Roodt, 2008:1). Before that, children were prepared for military service during their school years and officers were chosen by the Officer in Command of each unit, who based his decision mainly on the candidates' social background or the types of schools they attended.

As the war went on the supply of good officers became increasingly scarce (Coleman & Adams, 1999:28). The British Army then decided to follow the same route as the German Army and handed over the officer selection to psychologists and psychiatrists. This group of professionals revolutionised the officer selection system when they developed a system that could not only screen out the individuals who were unsuitable for a job, but they could also predict what a successful candidate would look like. It was called the 'assessment center' (Coleman & Adams, 1999:28; Schlebusch & Roodt, 2008:1).

The military invested a lot of time into the assessment center and it gradually began to gain more credibility. It was not long before the business and educational environment started using it too. It was Douglas Bray who first introduced assessment centers in a nonmilitary context in 1956 when he was asked to conduct research on the development of managers at AT&T (the largest telecommunications company in the US

at that time) (Gowing, Morris, Adler & Gold, 2008:437; Schlebusch & Roodt, 2008:21). Bray co-aliased with Howard and did some ground breaking work for the next 25 years. In 1973, Bray joined Bill Byham and co-founded the International Congress on Assessment Center Methods (Schlebusch & Roodt, 2008:21).

There was only one problem. Owing to its popularity, other assessment practices threatened the reputation of the assessment center by borrowing its name for procedures that incorporated only parts of the method. Psychologists and researchers developed an ethical code at the 3rd International Congress on the Assessment Center Method held in 1975. Thereafter the assessment center process was more regulated and protected by strict guidelines (Howard, 1997:13).

3.3.1 Assessment centers in South Africa

The Edgars group was the first South African organisation that bought into the assessment center idea. They started in 1974 with the development of two pilot assessment centers for their managers. Since then other organisations like Old Mutual, South African Railway Services, Stellenbosch Farmers Winery (SFW), the Department of Post and Telecommunication Services (Telkom), the South African Army and the South African Police all followed suit (Schlebusch & Roodt, 2008:22).

Old Mutual implemented the assessment center in 1973 and sent some of their personnel and line managers to Bill Byham for training in the USA. Later on they started to develop their own competencies and compared these across Old Mutual's major functional divisions (sales, back office, intermediary and property management, specialists and professionals). The assessment center practice continued to grow within this group and was even used to bring about culture change and to imitate strategic direction by assessing managers on future-state competencies (Schlebusch & Roodt, 2008:22).

SFW (now part of Distell) introduced the assessment center in the 1970's as part of a top-down leadership development process starting with directors and senior managers. The results were excellent and it was not long before the Middle Management Assessment Center (MMDC) for aspiring employees functioning on the middle

management level were developed. The adoption of follow-up development strategies is also one of the attributes that makes SFW a unique contributor to the process of assessment centers in South Africa (Schlebusch & Roodt, 2008:22).

Telkom explored a different purpose of the assessment center in 1978 and used it for succession management. They implemented the model on five different management levels in support of their mission to become a world-class organisation. Specific focus was given to the assessment of leadership competencies (Schlebusch & Roodt, 2008:23).

The Military Psychological Institute (MPI – which forms part of the South African National Defence Force) also played a major part in the development of assessment centers in South Africa. They initially started using the assessment center in 1978 to select Antarctica expedition personnel, military attachés and hostage negotiators. Today, their well-established assessment center practices are still used from entry level to more senior levels in all areas of the South African National Defence Force, for both selection and development purposes (Schlebusch & Roodt, 2008:23).

Lastly, the South African Police also made use of assessment centers during the 1980's to develop middle and senior managers. However, it was phased out in 1994 with the amalgamation of 11 police agencies into the new South African Police Service (SAPS). In 2000 the assessment center was reintroduced in the form of the Emerging Leadership Programme aimed at developing previously disadvantaged leaders and later to serve as a selection tool for managers. In 2006 an assessment center was specifically developed for middle-level managers (Schlebusch & Roodt, 2008:23).

It is therefore evident that the assessment center has come a long way since it was first been developed. Today, the assessment center comprises clearly defined elements that form part of the motivation for many organisations to continue to make use of this method, in spite of its complexity and its time-consuming nature. These elements will now be discussed.

3.3.2 Defining assessment centers

Over the years, many definitions and explanations have been developed to describe the assessment center. Joiner (2000:19) describes it as a "standardized method of evaluating behaviour, based on multiple inputs while making use of several trained observers and techniques". Schlebusch and Roodt (2009:2) give a similar definition and explain that an assessment center is a "simulation-based process employing multiple assessment techniques and multiple assessors to produce judgments regarding the extent to which a participant displays selected competencies required to perform a job effectively".

An assessment center is therefore a process used to evaluate the different sets of skills of an individual. The information gathered during such a process is mainly used for selection, development or diagnostic purposes. Although the core elements of the definition remain the same for each of these purposes, the different purposes for which information is used leads to a distinction between different types of assessment centers, namely the traditional assessment center, the development assessment center and diagnostic assessment center:

3.3.2.1 The traditional assessment center

The traditional assessment center is mainly used for selection purposes. Participants attending such an assessment center usually do not receive feedback on their performance during the assessment center but rather after its completion. This lapse is due to the time it takes to process and integrate all the information gathered and to determine the final ratings for each participant (Schlebusch & Roodt, 2009:3).

3.3.2.2 The development assessment center

The purpose for this type of assessment center is to determine specific strengths and areas of development among the employees of an organisation. It is different from the traditional assessment center because the results obtained from the development assessment center may not be used for any decisions regarding selection. In addition, participants usually receive feedback during the assessment center in the form of a

debriefing session and they also have the opportunity to put in place an individual development plan. Participants may also receive an opportunity to learn certain new behaviours through participating in activities (Schlebusch & Roodt, 2009:3).

3.3.2.3 The diagnostic assessment center

The diagnostic assessment center is used to diagnose the specific needs that assessment centers should address in the organisation. It is therefore a test run to determine whether the intended assessment and its activities will be sufficient to provide the required results. The results usually serve as insights for research projects and pilot studies (Schlebusch & Roodt, 2009:3).

The current study is based on an assessment center designed to select employees for leadership positions, therefore the remaining part of the chapter will focus specifically on the traditional assessment center. According to Schlebusch and Roodt (2009:6), there are certain elements or features that need to be present in a selection process before it can be called an assessment center. These features have been determined by the Assessment Centre Study Group (ACSG) of South Africa and will be discussed next (Schlebusch & Roodt, 2009:6).

3.3.3 Elements of an assessment center

3.3.3.1 Job profile/ Selection criteria

Before the assessment center can begin, the purpose for assessment needs to be established (Lievens & Goemaere, 1999:215; Schlebusch & Roodt, 2009:6; Taylor, 2009:10). A thorough job analysis should follow to identify the attributes, characteristics, aptitudes, qualities, skills, abilities, knowledge and tasks necessary for the specific position. These should be linked to behavioural incidents before being organised into job-relevant dimensions which the assessment center should evaluate (Howard, 1997:15). The behavioural incidents and job dimensions will also give an indication of the type of exercises that will elicit the desired behaviour.

3.3.3.2 Combination of valid and reliable assessment techniques

One of the reasons why the assessment center is a popular method is that the predictive power is enhanced through the use of various valid and reliable techniques to measure the specified dimensions or competencies (Gowing, et al. 2008:437; Howard, 1997:15). These include: paper screening, interviews, questionnaires, psychometric instruments and simulation exercises (also referred to as *behaviour observation exercises*). Schlebusch and Roodt (2009:7) emphasize that one competency should be assessed by various exercises in order to make it a valid process.

3.3.3.3 Various assessors

In order to keep the process as objective, valid and reliable as possible, there should be more than one assessor to observe and evaluate each participant (Schlebusch & Roodt, 2009:7; Taylor, 2009:10). These assessors should be equipped to perform such a task and should be familiar with the specific industry, the organisational culture and its service and product offerings (Howard, 1997:17). Training for assessors and raters is therefore vital. Many organisations have chosen to follow other approaches to assess their leaders, because of the financial and time implications of this requirement (Gowing, et al. 2008:437).

3.3.3.4 Data gathering

When conducting assessment centers, the assessors must use a systematic procedure to record specific information and behavior as it occurs during all the different activities. This may include hand-written notes that are transferred onto a rating sheet and then assessed according to certain criteria. Video or audio recordings are also available options. Whichever option is decided on, it is important for observers to use standardized rating sheets in order to keep the information structured and job-related (Howard, 1997:19; Schlebusch & Roodt, 2009:8; Taylor, 2009:10).

3.3.3.5 Data integration

After the exercises, the raters pool their gathered information and compile final ratings for each participant (Schlebusch & Roodt, 2009:8). Consensus should be reached on

the scores and a thorough report or record of observations should be compiled for each candidate and kept after the completion of the assessment center (Howard, 1997:19; Schlebusch & Roodt, 2009:8; Taylor, 2009:10).

3.3.3.6 Feedback

After the assessment center is completed, feedback should be given to the relevant stakeholders namely the organisation, participants and line managers (Schlebusch & Roodt, 2009:9). The content and depth of the feedback should be agreed on by the various parties before embarking on the process. In addition, information should always be treated with the necessary respect as indicated by the ACSG of South Africa (Schlebusch & Roodt, 2009:9).

It is therefore evident that an assessment center is an extensive process that should adhere to a list of requirements before it can be said to give a true representation of an individual's competencies and skills. The next part of this chapter will investigate the quality, suitability and usefulness of an assessment center by considering two technical properties called validity and reliability.

3.3.4 Validity and Reliability of Assessment Centers

In order to understand the principles of validity and reliability in an assessment center context, it will help to describe the assessment center as a very intricate whole consisting of separate parts. In other words, it is a process consisting of various activities. It should therefore be kept in mind that the way in which each activity is executed will contribute to the overall validity and reliability of the process.

Throughout the literature different studies (which will now follow) show that assessment centers are both valid and reliable. Love and DeArmond (2007:21) mention the substantial evidence in their study proving that the assessment center method has been labeled valid, fair and legally defensible, especially in a security environment. Gowing, et al. (2008:437) support this statement but they add that assessment centers only have these two properties if the subcomponents are properly designed and administered.

3.3.4.1 Validity

In an assessment center context, the concept of 'validity' gives meaning to the test scores by indicating links between test performance and job performance. When an assessment is valid, one will be able to make specific conclusions or predictions about individuals based on their performance scores (Lievens, 2009:103). In other words, assessment center validity refers to how well the activities measure what they are intended to measure (Gravetter & Forzano, 2009:75; Schlebusch & Roodt, 2009:267).

There are three types of validity that are relevant when referring to assessment centers: *criterion-related validity, content-related validity* and *construct-related validity*. These types will be explained next.

3.3.4.1(a) Criterion (prediction)- related validity

Criterion-related validity indicates that there is either a positive or a negative correlation between the test performance and job performance of an individual participating in an assessment center. In other words, the higher the candidate scores during an activity of an assessment center, the better he/she will perform in the specific job (and vice versa). If the criterion is obtained at the same time the test is given, it refers to *concurrent validity*. If the criterion is obtained at a later stage in time, it is called *predictive validity.* (Gravetter & Forzano, 2009:76; Schlebusch & Roodt, 2009:268; Taylor, 2009:12).

Early studies around assessment centers indicate minimum evidence for criterion-related validity. Howard (1997:16) states in her research that "…early criterion-related validity data were sparse, came from too few sources, covered too many variations in components, lacked replication, and suffered from criterion contamination". However, in the past two decades considerably more evidence has been found around the criterion validity of assessment centers, especially when validating it against "career progress, overall performance ratings, dimensional performance ratings, potential ratings, wages and training performance" (Howard, 1997:19).

The assessment center practice in general therefore has solid criterion-related validity. In fact, data suggests that criterion-related validity was higher for assessment centers

than for other available selection instruments. The only predictors in the same range as assessment centers are work samples and supervisor/peer evaluations although these still display lower average validity coefficients (Howard, 1997:19). One should however not underestimate the importance of the correct identification of job dimensions prior to the assessment center, as that will dictate the type of exercises used (Schlebusch & Roodt, 2009:6; Taylor, 2009:12).

3.3.4.1(b) Content-related validity

Content-related validity is explained by Taylor (2009:12) as the evidence that the assessment center activities represent or elicit important job-related behaviors. The activities should also test the whole range of characteristics in the domain. Schlebusch and Roodt (2009:266) add that determining the content validity of an assessment center is much more a process than a specific statistical procedure. A leadership assessment center should therefore measure behaviour and gather information on aspects that are relevant to, and important for, leadership. For example, an activity like role play should measure all the defined aspects of a competency one of which would be 'communication'.

It is difficult to determine the overall content validity of assessment centers as it is solely dependent on the specific purpose of the assessment center (and more specifically the activities that form it). In fact, the most compelling evidence on content validity of assessment centers is derived from meta-analysis, which is the combined result of a large number of research studies which make use of large sample sizes (Taylor, 2009:12).

In spite of the subjective nature of evidence regarding this type of validity during assessment centers, there are two guidelines that administrators can follow in order to ensure a higher content validity. Maldé (2006:541) suggests that "the more multi-layered assessment centers should aim to have fewer activities over a longer time frame to ensure a more candidate-friendly assessment setting. According to Lievens and Van Keer (2001:373), administrators should also pay close attention to the

organisational setting and relevant job characteristics when designing the assessment center.

3.3.4.1(c) Construct related validity

Taylor (2009:13) claims that if the activity used in the assessment center measures the construct or characteristic it claims to measure and if this characteristic is important to successful performance on the job, the assessment center consist of construct validity. For example, if an activity is set to measure 'communication' as part of leadership behaviour, the items in that activity should first of all elicit behaviour that is relevant to the dimension of communication (and not something else) and secondly, it should be proven that a construct like communication does indicate leadership ability (Gravetter & Forzano, 2009:78; Schlebusch & Roodt, 2009:267).

Literature has shown that assessment centers do possess high criterion-related validity (Gravetter & Forzano, 2009:78; Schlebusch & Roodt, 2009:267). However, during the 1980s questions were raised as to whether or not assessment center dimensions do indeed represent meaningful constructs (Lievens, 1998:141). After conducting research studies on this concerning issue, the findings are that "[r]atings on the same dimension across exercises correlated lowly, and ratings on different dimensions in a single exercise correlated highly" (Lievens, 1998:141). In other words, a competency measured during different exercises yields lower construct validity where different dimensions of a single competency measured during only one exercise indicate higher construct validity (Lievens, 1998:141). The number of exercises and the competencies measured during assessment center exercises should therefore be carefully considered.

Chen and Naquin (2006:265) still believe that the issue on assessment centers and construct-related validity remains unresolved up until now and suggest that it should be further investigated. This statement is backed by Thornton and Rupp (2006:10) who find conflicting results for the overall construct validity of Assessment centers. Despite these conflicting findings, they still believe that the findings that did yield evidence of construct

validity are enough to give construct status to competencies that are specifically defined and have an operational content (Thornton & Rupp, 2006:10).

3.3.4.2 Reliability

Reliability can be defined as the consistency with which an activity measures a competency (Schlebusch & Roodt, 2009:259). In an assessment center context it therefore implies that a candidate should obtain the same results when tested repeatedly under the same conditions (Gravetter & Forzano, 2009:82). However, when assessing performance it is not that simple to get the same results twice as the process is open to individual interpretation (Taylor, 2009:17). For that reason, it is irrelevant to discuss the reliability coefficient for assessment centers in general as it might differ from one process to the next. Nevertheless, it is important to be informed on the different aspects regarding reliability in order to ensure the most objective or reliable process possible.

As in the case of validity, there are different types of reliability that an assessment center should consist of. The most relevant types are: *test-retest reliability, alternate/ parallel form reliability* and*, inter-rater reliability*. Explanations of each type follow in the next section.

3.3.4.2(a) Test-retest reliability

Test-retest reliability refers to the repeatability of the scores obtained during an assessment center activity by a candidate over the course of time (Schlebusch & Roodt, 2009:259; Gravetter & Forzano, 2009:83). In other words, the ratings that participants obtain for a particular competency should correlate with the ratings they obtain for the same competency on a second occasion.

One should however consider that it is undesirable to put candidates through repeated measuring sessions as learning may occur between the pre- and post-measurements (Schlebusch & Roodt, 2009:259). It is also noteworthy that some competencies are much more stable than others and that the specific requirements should be noted in the job analysis. For example, an individual's reading ability would indicate much higher

test-retest reliability his/her anxiety levels as reading ability is not likely to change from one day to the next (Lowry, 1996:320). It should therefore be decided which skills are needed as well as the desired level of test-retest reliability.

3.3.4.2(b) Alternate/parallel form reliability

This type of reliability indicates the consistency of scores if an individual participates in two or three activities to measure the same construct (Gravetter & Forzano, 2009:85; Schlebusch & Roodt, 2009:258). For example, communication can be measured by a psychometric test as well as an activity where a candidate needs to give a presentation on a topic.

If an assessment center consists of a high parallel form reliability coefficient, it indicates that the different activities measuring a certain construct are very similar which again implies that it makes virtually no difference which activity an individual participates in. In the previous example, the candidate should then score more or less the same in both the psychometric test and the presentation. A low parallel form reliability coefficient, on the other hand, means that the different activities of an assessment center, which are designed to measure the same construct, are not comparable and that they actually measure different constructs (Gravetter & Forzano, 2009:85; Lowry, 1996:320).

3.3.4.2(c) Inter-rater reliability

Inter-rater reliability indicates the concurrence between two raters who simultaneously rate the behaviour of a candidate (Gravetter & Forzano, 2009:85). This type of reliability mostly applies to simulation exercises; that is, a specific activity commonly used during assessment centers which will be discussed later in the chapter.

A high inter-rater reliability coefficient indicates that the assessment process is stable and the raters are clear on what behaviour they are supposed to rate (Gravetter & Forzano, 2009:85). Owing to the fact that assessment centers usually have a panel of assessors, differences of opinions and judgments can easily occur which may lead to variation among activity scores for a single candidate (Gravetter & Forzano, 2009:85).

As in the case of other forms of reliability, inter-rater reliability will differ from one assessment center to the next (Gravetter & Forzano, 2009:85).

There are two principal ways to increase inter-rater reliability. Firstly, raters or assessors should be trained prior to the assessment center. The purpose of rater training is to equip raters to observe, note, classify and evaluate participant behaviour in a reliable way during behaviour observation exercises (Schlebusch & Roodt, 2009:181). Ideally, the training process should comprise the following steps:

- Step 1 – Raters attend the assessment center as a participant. This is to help them experience the assessment center, and specifically the behaviour observation exercise, from the candidate's side (Schlebusch & Roodt, 2009:181).

- Step 2 – After understanding what the candidate experiences during a behaviour observation exercise, the raters obtain an overview of the total observation process. The raters therefore attend the assessment center again, but this time as observers (Schlebusch & Roodt, 2009:182).

- Step 3 – Proper rater training can now be given to the raters, providing a theoretical background to observing behaviour and making them aware of the potential mistakes that can occur during such a process (Schlebusch & Roodt, 2009:182). Special attention should be given to the verbal and nonverbal behaviour of the candidate, the different reactions from the various role players as well as the changes in behaviour that often play out in a group (Schlebusch & Roodt, 2009:184).

- Step 4 – Raters are now given the opportunity to perform all the tasks associated with behaviour observation, but as assistant observers under the supervision of a qualified observer. Only after the senior observer declares an acting observer as competent, can the acting observer take on the full role of a rater in an assessment center (Schlebusch & Roodt, 2009:181).

A study that shows that rater training does make a difference to the inter-rater reliability of assessment centers is that of Jackson, Atkins, Fletcher and Stillman (2005:17). They

examine the effects of 'Frame of Reference (FOR) Training' which includes making raters aware of the constructs they will be measuring, the type of behaviour the constructs entail and the scale on which it will be measured. They find that the agreement among raters increased subsequent to the FOR Training procedure.

Secondly, the assessment center design team should decide on an appropriate rating scale to record behaviour that was observed during the exercises. According to Schlebusch and Roodt (2009:116), this decision primarily relies on the purpose of the specific assessment center. They are of the opinion that the choice of rating scale should be guided by the following principles:

i. The scale should be validated;

ii. The raters, candidates and any other stakeholders should be able to understand the rating scale; and

iii. The scale should be easy to use when evaluating candidates.

Furthermore, Schlebusch and Roodt (2009:117) believe that the rating scale used for a traditional assessment center should be broader than four points, as a scale with a greater range does not limit the observations and is more able to indicate discrepancies. Lastly, they emphasise the importance of using the same rating scale for every exercise used in the assessment center.

All of the abovementioned validity and reliability aspects will contribute to a fair and effective assessment center. It is especially important to consider these aspects during the design phase (Lievens, 2009:107; Schlebusch & Roodt, 2009:104). According to Schlebusch and Roodt (2009:104), assessment centers should be custom designed by professionals who are familiar with the organisation and the reason for assessing. It is therefore crucial that the activities to be included in the assessment center should be chosen very carefully. The next part of the chapter will explore the options.

3.4 ACIVITIES USED IN ASSESSMENT CENTERS

According to Gal and Mangelsdorff (1991:90), Le Roux (2004:48), Schlebusch and Roodt (2009:118), Stamoulis (2009:69) and Taylor (2009:141), assessment center activities are varied and differ from one another in terms of the utility of the techniques as well as the validity and reliability thereof. The next part of this chapter will give an overview of the various activities used in assessment centers. These include the structured interview, in-basket, oral presentations, fact-finding exercises, psychometric evaluations and behaviour observation exercises.

3.4.1 The structured interview

The structured interview is a one-on-one discussion between a professional and a candidate in order to determine whether the candidate has demonstrated certain competencies in the past (Stamoulis, 2009:68). Table 3.1 presents some advantages and disadvantages of the structured interview.

Table 3.1: Advantages and Disadvantages of the Structured Interview

Advantages	Disadvantages
The process is relatively short and easy to administer (Le Roux, 2004:46). The best predictive power results are obtained when combining structured interviews with other activities like psychometric evaluations (Gal & Mangelsdorff, 1991:82)	Due to the short time span of the interview, candidates can easily exhibit different behavior than they usually do in an everyday environment (Le Roux, 2004:46)

3.4.2 The In-basket

This is a written exercise that presents the typical administrative demands a manager has to face (for example memoranda, emails, etc.) to test whether or not a candidate will be able to lead and manage his/ her department (Le Roux, 2004:46). The candidate is required to respond to demands in the most appropriate way possible (which can

include actions like prioritising, asking for information or giving instructions) (Le Roux, 2004:46). Some advantages and disadvantages relating to the in-basket are presented in Table 3.2.

Table 3.2: Advantages and Disadvantages of the In-Basket

Advantages	Disadvantages
This exercise can be administered to more than one participant at a time and can measure more than one competency or skill (Schlebusch & Roodt, 2009:119). Evidence exists that the ratings of an in-basket are closely related to actual managerial performance in tasks similar to the in-basket (Le Roux, 2004:48).	The In-basket can be complex depending on the organisation and the position that is being simulated. It is also time-consuming to design such an exercise (Le Roux, 2004:46; Schlebusch & Roodt, 2009:119)

3.4.3 Oral presentations

During this exercise, the candidate has to prepare and deliver a business presentation to convince the audience of a certain course of action. The background of the case study and the instructions can be given to the candidate before the assessment center or during the assessment center (Le Roux, 2004:46; Schlebusch & Roodt, 2009:123). Some advantages and disadvantages relating to oral presentations are presented in Table 3.3.

Table 3.3: Advantages and Disadvantages of Oral Presentations

Advantages	Disadvantages
This exercise is particularly useful when assessing oral communication skills and the ability to develop formal and informal presentations when functioning under pressure (Gal & Mangelsdorff, 1991:82; Le Roux, 2004:43).	A disadvantage of the oral presentation is that it can be extremely stressful for the candidate which can prevent him/ her from functioning optimally (Gal & Mangelsdorff, 1991:82; Le Roux, 2004:43).

3.4.4 Fact Finding Exercises

Fact finding exercises present the candidate with a standardised business problem that has occurred in the organisation. The candidate is then requested to gather as much information as possible in order to make recommendations on how to resolve the problem (Le Roux, 2004:43; Schlebusch & Roodt, 2009:123). Table 3.4 presents some of the advantages and disadvantages relating to the fact finding exercise.

Table 3.4: Advantages and Disadvantages of Fact Finding Exercises

Advantages	Disadvantages
Fact finding exercises can also be used to measure more than one competency (Le Roux, 2004:43). There is no standardised answer which elicits creativity from the candidate.	According to Le Roux (2004:43), fact finding exercises are time consuming to develop as the resource material should be carefully planned. The administrator should also be familiar with the material to provide the responses to the candidate's questions.

3.4.5 Psychometric Evaluations

There are mainly three types of psychometric instruments used in assessment centers. They are *aptitude* or *ability tests*, *personality inventories* and *interest inventories* (Taylor, 2009:90). Aptitude tests measure a candidate's mental ability (for example, verbal

reasoning, numerical reasoning, mechanical skills or problem solving skills) (Taylor, 2009:90). Personality inventories measure a candidate's personality dimensions based on a specific theory (Taylor, 2009:90). Typical examples of these tests include: the Jung Personality Questionnaire (JPQ), the Myers-Briggs Type Indicator (MBTI), the 16 Personality Factor Questionnaire (16PF) and the Big Five Personality Questionnaire (Taylor, 2009:90). The interest inventory is, on the other hand, designed to measure a candidate's professional or occupational preferences (Taylor, 2009:90). Some advantages and disadvantages relating to the use of psychometric evaluations are discussed in Table 3.5.

Table 3.5: Advantages and Disadvantages of Psychometric Instruments

Advantages	Disadvantages
Psychometric instruments usually consist of good validity and reliability coefficients (Gal & Mangelsdorff, 1991:89; Taylor, 2009:90). If the instrument is matched with the appropriate competencies, it has high predictive power to indicate the type of environment an individual would feel comfortable working in. It is especially useful for developmental purposes as it not only determines the specific interest of an individual, but also the underlying motives accompanying the preference (Gal & Mangelsdorff, 1991:89; Taylor, 2009:90).	Although psychometric instruments can predict certain aspects of job performance, it cannot be used in isolation nor can it predict leadership abilities (Gal & Mangelsdorff, 1991:89). Only qualified individuals are allowed to administer psychometric instruments and the instruments need to be standardised. This is overall a very expensive activity (Gal & Mangelsdorff, 1991:89; Taylor, 2009:90)

3.4.6 Behaviour observation exercises

Behaviour observation exercises can take on numerous forms. Two popular practices are the *leaderless group discussion* and the *physical exercise.* The leaderless group

discussion assesses multiple candidates in group format at the same time and in the same location (Stamoulis, 2009:89). The candidates are requested to solve an open-ended problem as a group, while trained raters observe their behaviour. During this exercise specific roles can be allocated to candidates, or assessors can choose not to allocate roles. In both instances, leadership attributes are elicited within the candidates (Stamoulis, 2009:90). The physical task group activity is very similar to the leaderless group discussion. However, with this exercise the group is required to complete a physical task as a team while their behaviour (competencies) is observed by trained raters (Schlebusch & Roodt, 2009:120; Taylor, 2009:215). The advantages and disadvantages to these exercises are discussed in Table 3.6.

Table 3.6: Advantages and Disadvantages of Behaviour Observation Exercises

Advantages	Disadvantages
Behaviour observation exercises are particularly useful when assessing leadership competencies like problem solving, social adaptability, planning and organising as well as influencing skills (Le Roux, 2004:45). The problem that the candidates should solve can be applicable to the specific organisation which creates an opportunity for learning and to interact across different cultures (Schlebusch & Roodt, 2009:120; Taylor, 2009:215)	There are three main challenges associated with behaviour observation exercises. Firstly, every group will differ from the next. Consequently the interaction or discussion that occurs will also differ by group, which makes it difficult for raters to provide objective observations (Le Roux, 2004:46). Secondly, raters need to be trained in order to observe behaviour which is not only time consuming, but also costly. Thirdly, there needs to be at least one rater for each candidate participating in the exercise which can be costly. Lastly, candidates with physical disabilities will not always be able to participate in physical activities.

Section 3.4.1-3.4.6 gives a clear indication of the various activities available to make use of in an assessment center. It is also evident that in order to optimise the predictive power of the assessment center, they should not be used in isolation as the real value lies in how these activities are combined to measure the broad concept of job performance and especially leadership in the end. According to Gal and Mangelsdorff (1991:110), "...the correct combined use of assessment tools results in a success rate of 77% as opposed to a random selection that have resulted in a success rate of 63%". They further state that the combination of activities should be linked to a specific competency profile. The assessment center referred to in the current study uses a selection (otherwise referred to as a 'battery') of carefully combined activities namely paper screening, structured interviews, presentations, psychometric tests and behaviour observation exercises. The behaviour observation exercises consist of two activities: a leaderless group exercise called Winter Survival during which candidates have to solve a problem among themselves as well as a physical task where candidates are required to overcome a range of physical obstacles as a group. The next part of the chapter will provide an in-depth discussion on the nature of behavior observation exercises.

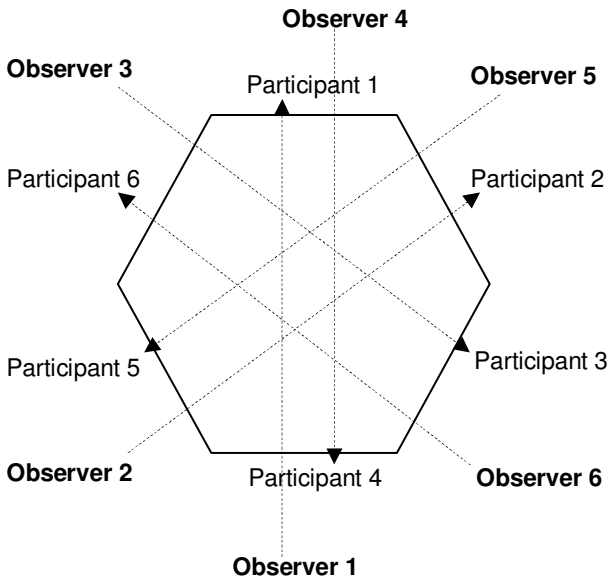
3.5 BEHAVIOUR OBSERVATION EXERCISES

Behaviour observation exercises have a long history as a method for measuring complex human abilities and remain a core element of assessment centers (Joiner, 2000:315; Brummel, et al. 2009:137). Joiner (2000:315) describes this activity as the following: "…an exercise or technique designed to elicit behaviors related to dimensions of performance on the job, requiring participants to respond behaviorally to situational stimuli". These situational stimuli can either be written instructions, information given by an administrator, or reactions from fellow candidates participating in the assessment center (Schlebusch & Roodt, 2009:117). In other words, it involves putting candidates in a situation that is specifically designed to elicit a set of job-relevant competencies and letting a group of assessors (also known as raters) observe and rate the behaviors elicited while scoring these on a rating sheet (Gowing, et al. 2008:437).

According to Howard (1997:15) and Gowing, et al. (2008:438), behaviour observation exercises give the raters rich behavioural information on each candidate through face-

to-face interaction. Figure 3.1 shows the typical set-up of a leaderless group discussion (one of the most commonly-used behaviour observation exercises) indicating face-to-face interaction between candidates and raters (Woodruffe, 1995:155). Brummel, et al. (2009:137) also point out in their study that assessment centers rely on such face-to-face opportunities in order to measure the candidates' competencies and skills that are related to the dimensions of the position. This type of information is difficult to obtain through pen and paper tests (Brummel, et al. 2009:138).

Figure 3.1: Typical Set-up of a Leaderless Group Discussion (Source: Woodruffe, 1995:155)



Even though face-to-face interactions could improve the procedural fairness of an assessment center (Gowing, et al. 2008:438), behaviour observation exercises are very susceptible to internal and external forces that may harm the validity and reliability of the selection process. In recent years, professionals are put under extreme pressure to follow a set of ethical guidelines which ensures that the process is substantially fair and does not put any candidate in a harmful position during the assessment (Foxcroft & Roodt, 2005:220). Disregarding these internal and external factors cannot only

jeopardise a candidate's career, but can also lead to lawsuits which can be costly to the reputation of the organisation (Taylor, 2009:7).

Before the internal and external forces influencing behaviour observation exercises can be discussed, it is necessary to investigate the theories around observing behaviour. Werth, Markel and Förster (2006:103) state that although many organisational factors can be measured using objective standards, some aspects (especially leadership) are difficult to quantify and rely heavily on personal opinion. It is therefore logical to assume that these opinions, whether conscious or subconscious, might influence the way in which behaviour is perceived. Literature on this phenomenon falls under the heading of theories of social judgment.

According to Thornton (1992:96), social judgment refers to the "process of perceiving and evaluating the behaviour of others and forming impressions about their strengths and weaknesses". There are two opposing theories around this concept. The *behaviour-driven theory* explains that behaviour observation is an objective process whereby the observer accumulates various pieces of factual information (Thornton, 1992:96). The observer then integrates this information in a logical and systematic way in order to form judgments (Thornton, 1992:96). It therefore implies that the observer is empowered to take in detailed information and form objective opinions based on actual facts. On the other hand, *the schema-driven theory* states that behaviour observation is mainly influenced by subjective bias and that the observer has limited capacity to remember prior events exactly as they occurred (Thornton, 1992:96). In other words, the observer only stores certain pieces of information (which form the basis of his/ her perceptions). This information is then used to make judgments regarding the behaviour of a candidate.

Ideally the assessment center should be based on a combination of these two theories (Thornton, 1992:96). Raters should be able to record systematically all visible behaviour with the help of specific techniques like rating scales and observation sheets. In addition, the raters should use previous knowledge and experience as a model of managerial performance in order to make professional judgments regarding behaviour (Thornton, 1992:96). A thorough understanding of these theories will lead to an

increased awareness of how judgments regarding observed performance are made. It will also help the rater to identify other factors that might hamper the process. These factors will be discussed in the next section.

3.5.1 Factors influencing behaviour observation exercises

Lowry (1996:308), Maldé (2006:545) as well as Schlebusch and Roodt (2009:115) are of the opinion that the following should be kept in mind when administering behaviour observation exercises: Test-taker's psychological or physical state, environmental factors, instrument format and providing too much information to the raters (information overload). A brief discussion of these factors will now follow.

3.5.1.1 Test-taker's temporary psychological or physical state

Due to the stressful nature of the process, behaviour observation exercises are often associated with differing levels of anxiety, fatigue or motivation amongst candidates (Maldé, 2006:545). It is therefore crucial to be aware of the major influence that extreme assessment situations can have on the applicant's ability to function and to obtain good results (Lowry, 1996:308; Maldé, 2006:545). One way of putting candidates at ease during assessments is to give them a proper introduction in a standard format and in a language that they understand. This introduction decreases the fear of the unknown as the candidates know exactly what is expected of them (Schlebusch & Roodt, 2009:115). The same applies for the tester. According to Lowry (1996:308) and Malde (2006:545), hectic assessment schedules can cause testers to feel tired which can lead to mistakes. Rater errors will be discussed later in this chapter.

3.5.1.2 Environmental factors

Differences in the assessment environment can cause variances in candidates and in turn, influence validity and reliability coefficients. Room temperature, lighting, noise, air supply, movement and access to resources are all examples of environmental factors that should be taken into consideration when conducting assessment centers (Lowry, 1996:309; Schlebusch & Roodt, 2009:114; Westgaard, 1999:89; Woodruffe, 1995:152).

3.5.1.3 Instrument format

Although the problem of different test formats mainly applies to psychometric instruments, it does not exclude simulation exercises. Due to the fact that scores on the dimensions are not provided directly by the candidates, as they are in other testing situations, raters might have their own interpretation of candidates' behaviour (Brummel, et al. 2009:145). For this reason a standardised rating sheet is of utmost importance.

Earlier in Section 3.3.4.2 of this chapter, some guidelines in choosing the correct rating scale to increase inter-rater reliability were mentioned. There are two main types of rating scales used during assessment centers. They are the BARS (Behavioural Anchored Rating Scale) and the Likert-scale.

According to Melchers, Lienhardt, Von Aarburg and Kleinmann (2011:57), the BARS system provides critical behavioural incidents or examples of high, mid-level, and low performance for each element or competency that is measured. It therefore helps the rater to focus on desirable and undesirable incidents of work behaviour (Melchers, et al. 2011:57). Maurer (2002:308) conducted an interesting study on the BARS which finds that ratings made by a sample of inexperienced individuals yield the same accuracy and validity as ratings made by experts who are well-acquainted with the job content and dimensions being rated. This is explained by the fact that both groups use an anchored rating scale which gives them direction in the behaviour observation process. This certainly does not imply that inexperienced raters are allowed to participate in the assessment center, but it does emphasize the importance of having some guidelines while observing the behaviour of others. In addition, Melchers, et al. (2011:53) finds that the anchored rating scale leads to a substantial improvement in both rating accuracy and inter-rater reliability. The only limitation of the BARS is that it takes a lot of time and precision to develop.

Another rating system that can be used is the Likert 5-point scale. This rating scale is based on the same principle as the BARS in the way in which it distinguishes between different levels of competence (Westgaard, 1999:64). An important difference between the BARS and the Likert scale is that the latter is not as descriptive. However, some

behavioural indicators to guide the raters when scoring each dimension may be included, depending on the preference of the assessment center developers.

The Likert scale can either consist of an even number of choices (4-point scale) or an odd number (5-point scale) (Westgaard, 1999:64). In case of an even number, the rater is forced to take a position favouring one extreme over the other. In the case of an odd number of choices, raters can remain slightly impartial and do not have to favour one extreme or the other (Westgaard, 1999:64). The latter is consequently much more sensitive towards the candidate than the first, but the problem arises when raters only rely on the middle option or response. This tendency is known as the *leniency* effect and will be discussed later on in this chapter. (Due to the time constraints that were faced in the current study, the Likert-scale was used. The description of the scale will be discussed in Chapter 4).

3.5.1.4 Information overload

Many questions arise when one considers that the scores candidates obtain during behaviour observation exercises depend on the judgment of raters. This study will consider two questions specifically. Firstly, it is necessary to ask how many behaviour observation exercises are necessary for a rater to make a valid judgment regarding behaviour. Joiner (2000:321) gives a clear answer to this question when he states that it depends on the complexity of the position being assessed and that any reasonable amount will suffice. He does, however, state that the job analysis should clearly indicate that the number of exercises sufficiently simulates a substantial portion of the job being evaluated. He adds that the number of competencies measured in such an activity also plays a big role. He advises developers not to assess every dimension in every exercise. Schlebusch and Roodt (2009:114) share this concern with Joiner and state that there should be no more than six competencies measured during each behaviour observation exercise. Assessing too many competencies may become taxing on the raters and candidates participating in the assessment center, which is also relevant to the next question.

Secondly, it is necessary to ask how many raters should rate one candidate, as the process of observing behaviour is a cognitively demanding task for the rater (Kolk, Van der Flier, Born & Olman, 2002:271). In 80-95% of assessment center cases, raters observe candidates' behaviour during an exercise while taking notes. Raters then proceed to classify the written remarks into behavioural dimensions and allocate a quantitative rating for each dimension. At times, they might evaluate their ratings with co-assessors; however, time and resource constraints may prevent the last step from happening (as in the case of this study) (Kolk, et al. 2002: 271).

Although the process of note-taking is a very important tool to help the rater remember which behaviours were observed, it can be problematic when raters do not have the cognitive processing capacity to do both observing and taking notes (Kolk, et al. 2002:271). Some behaviour may be overlooked while the rater exchanges his/her attention between the scene and the rating sheet or a rater might allocate the wrong behaviour to a specific dimension, especially if the rater is inexperienced in the process (Kolk, et al. 2002:271). The outcome could also be altered if raters do not have anyone with which to validate their observations.

Kolk, et al. (2002:278) investigate the possibility to postpone the note-taking to after the exercise and find that it does not make any difference to results. However, one aspect that does have a significant effect on the results is the experience level of the rater. Kolk, et al. (2002: 278) are certain that the cognitive load on inexperienced raters may lead to rater errors and they urge professionals to ensure that a systematic procedure is in place to record behaviour accurately at the time of occurrence. It is also possible to deduce from this study that more than one rater should be allocated to each candidate for validation purposes.

The cognitive load on raters is only one of the factors that can cause raters to make mistakes. These mistakes are called rater errors and do not only affect the inter-rater reliability of the assessment center, but also the credibility of the selection process as a whole (Kolk, et al. 2002:271; Lowry 1996:309).

The current study focuses by and large on rater errors, therefore the following part of this chapter will give a brief overview of the different types of rater errors that can occur during a behaviour observation exercise. More specifically, the following section will focus on one type of rater error called rater bias.

3.5.2 Rater errors

This phenomenon often occurs among raters participating in behaviour observation exercises. According to Grobler, et al. (2006:277), there are several types of rater errors: The halo effect, central tendency, leniency and the primary/recency effect. Each error will now be discussed.

3.5.2.1 Halo effect

This error occurs when a rater allows one particular aspect of a candidate's performance to influence the evaluation of other aspects of the candidate's performance. For example, a rater might be aware of a candidate's exceptional ability to communicate ideas (possibly through previous assessment encounters with that candidate). The rater may then allow this specific attribute to influence him/her to allocate high scores on other dimensions like 'adaptability' or 'functioning under pressure' too (Bechger, Maris & Hsiao, 2010:608; Grobler, et al. 2006:277; Schlebusch & Roodt, 2009:185).

Conversely, it is possible that because a candidate scored low on 'communication', the rater may mark him/her low on another competency like 'adaptability'. In environment analysed within this study, this rater bias is relevant as raters often have more than one encounter with a candidate.

3.5.2.2 Central tendency

Central tendency occurs when raters evaluate all their candidates as average. This may either be because they find it difficult to rate some candidates higher or lower than others due to the consequences that their rating might have on the specific individual. For example, a high rating might lead to a promotion for the candidate where a low rating might lead to a retrenchment. Alternatively, they could simply be unfamiliar with

the competency being rated (Grobler, et al. 2006:277; Schlebusch & Roodt, 2009:185). This tendency can be overcome by making use of an even-point rating scale which does not provide a 'midpoint' and therefore forces the rater to choose between a favourable and a less favourable rating (Refer section 3.5.1.3; Taylor, 2009:66).

3.5.2.3 Leniency

Leniency occurs when raters are under the impression that the easiest way to appraise performance is simply to give all the candidates a high evaluation (Grobler, et al. 2006:278). This might be due to laziness or it could be because raters do not want to offend candidates by giving them low scores (Taylor, 2009:66). Candidates who actually performed poorly in the assessment will not complain, however, candidates who did perform well may take further steps which will affect all the selections made on that specific project. Furthermore, lenient ratings may also lead to salary implications for already existing employees in an organisation, which could have an effect on the wealth of the organisation (Grobler, et al. 2006:278).

Strictness is the opposite of leniency and occurs when raters consistently give low ratings even though some candidates may have achieved an average or above average performance level. This phenomenon will not only affect the morale of candidates participating in the assessments, but also the true representation of skills observed during a simulation activity (Grobler, et al. 2006:278).

3.5.2.4 Primacy, recency and heuristics

Primacy and recency refers to the truism that first impressions last the longest (Taylor, 2009:66). Raters can make judgments about behaviour within seconds and it often happens that later evidence about behaviour (that proves the initial impression wrong) is disregarded. Heuristics on the other hand refers to the tendency to simplify complex processes (Taylor, 2009:65). A rater might, for example, only recall unusual or infrequent behaviours from candidates. It is also possible that a rater may begin judging a candidate's performance from the viewpoint that he/ she is a good or bad performer (Taylor, 2009:65). For this reason, it is important to make use of behavioural notes that capture all the essential elements necessary to make an objective judgment.

All the above mentioned errors can have a significant impact on the scores candidates obtain during behaviour observation exercises. However, according to Grobler, et al. (2006:277) the most common rater error that exists in any appraisal method is rater bias. Rater bias is central to this study and therefore requires a clear explanation.

3.5.3 Rater Bias

Grobler, et al. (2006:277) and May (2008:302) define rater bias as any conscious or unconscious predisposition of raters that causes them to make incorrect judgments about a candidate's performance. In other words, a candidate's performance might be over- or under-rated based on a specific attribute or action that is not relevant to the position or profile for which the candidate is assessed.

When referring to scores obtained during leadership assessments, Hautala (2005:85) states that most studies focus only on the personality, skills and behaviour of the candidate and that the impact of the rater's personality type and skills have not yet been considered. Werth, et al. (2006:103) emphasises that what raters perceive and how they process it depends on both the personal characteristics of the ratee and the rater. This statement does not necessarily specify the terms of interaction, but it does point to the fact that both parties play a role when formulating judgments during a behaviour observation exercise.

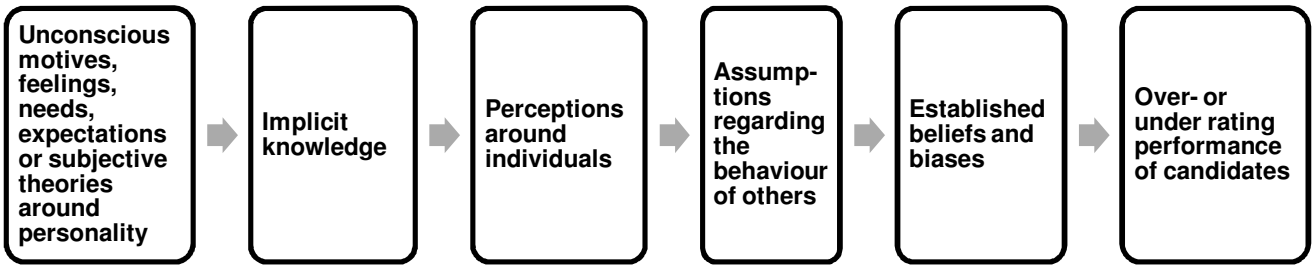
In order to understand the origin and nature of bias, it is valuable to obtain an understanding of the way in which a rater processes information (Werth, et al. 2006:103). As discussed earlier in Section 3.5 of this chapter, the schema-driven theory of social judgment states that bias is mainly caused by subconscious pieces of information which leads to perceptions and assumptions regarding individuals (Thornton, 1992:96). Werth, et al. (2006:103) concur with this theory and refer to these pieces of information as the *implicit knowledge* of raters.

Different findings have been made with regards to the origin of implicit knowledge. Researchers like Strack and Deutsch (2004:220) are of the opinion that implicit knowledge develops through an individual's unconscious motives, needs, expectations and goals. Other researchers like Bretz, Milcovich and Read (1992: 322) believe that

implicit knowledge is derived from situational factors that elicit certain sensations within an individual (for example, nonverbal behaviour from others that causes pleasant or unpleasant feelings). Then there are those who believe that each individual has his/her own subjective theory of personality which forms part of that individual's implicit knowledge. This subjective personality theory of an individual is used to complete missing links in the perceived behaviour of others (Trope & Higgins, 1993:493).

Whichever theory is correct is irrelevant to this study, but what is important is the impact of the implicit knowledge on the rater's judgments. Even though the abovementioned researchers differ slightly in their views on the origin of implicit knowledge, all of them acknowledge the fact that judging behaviour is not merely a process of classifying what is visible to the human eye. There are many more underlying factors that contribute to a rater's decision-making and oftentimes these factors are unknown, even to the rater (Trope & Higgins, 1993:493). In fact, according to Werth, et al. (2006:105), implicit knowledge can drive the process in deciding what information is relevant to a concept like leadership and how to interpret the evidence. Figure 3.2 illustrates how implicit knowledge can lead to bias decisions as stipulated by Werth, et al. (2006:103).

Figure 3.2: The Formulation of Biases and the Effect thereof (Source**:** Werth, et al. 2006:103)



There are two main types of bias in which this implicit knowledge of raters occurs: *stereotyping* and *similarity bias.* Both will be discussed in the section to follow.

3.5.3.1 Stereotyping

Stereotypical behaviour from raters occurs when a single characteristic or attribute of a candidate suggests a group of other qualities supposedly linked to that characteristic (Taylor, 2009:64). For example, the belief that younger candidates will learn faster, or that certain gender groups are better managers, are typical stereotypes echoed by society. These beliefs (also referred to as *prejudices*) may not only cause raters to pre-judge in the absence of substantial evidence, but they may also lead raters to act in a way that eventually elicits the trait they expect a candidate to present (Taylor, 2009:64). Stereotypes can be based on a variety of traits some of which include perceived attractiveness, dress sense, physical features, gender, sexual orientation, culture and ethnic origin (May, 2008:297; Taylor, 2009:64; Werth, et al. 2006:104).

Nobody can claim that they are free of prejudices or stereotypical behaviour; it is a strategy that all individuals develop which helps them to simplify, classify and make sense of the world (Taylor, 2009:65). In some instances stereotyping can contain a measure of truth; however, it is very important for every individual (especially raters) to be fully aware of their personal stereotypes so as to minimize the unjust impact on others around them (Taylor, 2009:65).

3.5.3.2 Similarity bias

Birds of a feather flock together. This truism indicates that in any natural social setting it would not be unusual to find individuals grouped together based on their similar interests, traits or cultures. This is because individuals in general prefer objects, persons or concepts that resemble or are consistent with what they believe to be true about themselves (Murnieks, Haynie, Wiltbank & Harting, 2008:1). A behaviour observation setting is no different: raters may also have a preference for candidates who they perceive as similar to themselves (whether similarities are attitudinal, personality-based, racial or cultural). When raters give more favorable ratings to candidates similar to themselves, the phenomenon is referred to as similarity bias (May, 2008:298).

Merrill and Reid (1999:45) explain similarity bias in terms of the amount of discomfort the rater experiences while interacting with candidates. They state that if there is an attribute in a candidate that closely resembles an attribute of the rater, it creates the least discomfort and may even result in the rater experiencing positive feelings or a sense of familiarity with the candidate. On the other hand, an attribute in a candidate that differs from the rater may cause discomfort or a need to avoid the unknown, which in turn may create tension within a rater. The feelings of discomfort or happiness therefore form part of the implicit knowledge described in Section 3.5.3. A comment from May (2008:312) puts this explanation into perspective:

> *"It is a psychological reality that we like those most who are most like us, and it is a natural tendency to rate those we like in a more favorable manner, regardless of performance."*

A study relevant to this discussion by May and Gueldenzoph (2006:3) documents this bias among employees with similar social styles during performance appraisals. A social style may be described as a series of actions that describes an individual's behaviour (May & Gueldenzoph, 2006:3). It is measured by two reliable scales called *Assertiveness* and *Responsiveness* which indicate the extent to which an individual tries to influence the thinking and actions of others and responds to environmental stimuli (May & Gueldenzoph, 2006:3). When combining these two scales, four styles of observable behaviour emerge namely an *Analytical style* (low on assertiveness and low on responsiveness), a *Driver style* (high on assertiveness and low on responsiveness), an *Expressive style* (high on assertiveness and high on responsiveness) and an *Amiable style* (low on assertiveness and high on responsiveness) (May & Gueldenzoph, 2006:4). The results of the study not only indicate higher scores when employees with the same styles are paired together; it also proves that employees with opposite styles rate one another significantly lower, indicating similarity biases.

If it is possible to base similarity bias on social styles, other areas of personality should be considered. A relevant study in this regard is that of Antonioni and Park (2001:331). They find that candidates who obtain similar scores to their raters on conscientiousness

(a Big Five personality trait) obtain higher ratings in performance appraisals than those who obtain different scores to their raters (Antonioni & Park, 2001:331).

Antonioni and Park (2001:331) do not find any evidence of similarity biases based on the other factors listed on the Big Five Model of Personality; however, one should not exclude the possibility of raters having preferences for other personality types like introversion or extraversion. Rater preferences for introversion or extraversion in candidates are referred to as introversion/extraversion bias from this point in the study and the assumption that this may have a major impact on behaviour observation scores is the main theme of this study.

3.6 INTROVERSION/EXTRAVERSION BIAS

Introversion/extraversion bias occurs when raters favour candidates with the same personality type as their own (specifically referring to introversion or extraversion) and rate such candidates higher as a candidate with a different personality type to theirs. Said differently, introverted raters would for example rate introverted candidates more highly than they would rate extraverted candidates (and vice versa) when neither introverted nor extraverted behaviours form part of the requirements of the position. The current study is conducted in a leadership assessment setting and it is important to note that, for the purposes of this study, it is assumed that leadership effectiveness cannot be predicted by introverted or extraverted behaviours (as discussed in Chapter 2). It is also valuable to distinguish between self preferences for introversion or extraversion (as indicated by the JPQ) and preferences for introversion or extraversion in others (such as candidates). Only if a rater prefers a similar personality type in a candidate to his/ her own personality type, will it be classified as an introversion/extraversion bias.

The biggest challenge with regards to the topic of introversion/extraversion bias is the lack of research available. Therefore, in order to support the assumption that this study is exploring (even if it is only to a moderate extent), it may be valuable to make note of a study conducted by Hautala (2005:84) who investigates the effect of raters' own personality types on leadership assessments (Hautala, 2005:84). Hautala's study resembles the current study in three ways:

i. There is a strong focus on raters' own preference for introversion or extraversion as a personality *type* (and not a *trait* as with most other studies);

ii. Her study takes place in a similar setting as the current study, that is, a leadership assessment setting; and

iii. She also makes use of a Jungian-based instrument (Myers-Briggs Type Indicator) to measure personality types.

Hautala used the Myers-Briggs Type Indicator (MBTI) and a modified version of the Leadership Practices Inventory (LPI) as tools to test how 167 raters rated their 38 leaders (Hautala, 2005:86). The results that came out are relevant to this study as they indicate that raters with extraverted and/or feeling preferences give significantly higher ratings than those raters with introverted and/or thinking preferences (Hautala, 2005:90). The main conclusion from Hautala's study proves that introverted and extraverted raters rate candidates differently.

Hautala's study does not, however, go further to confirm the assumption that the raters whose personality types are similar to their leaders' personality types will rate their leaders higher. She suggests that further studies should be directed towards hypothesizing on this possibility (Hautala 2005:92). Hautala's suggestion therefore provides the basis and mandate of the current study which aims to address the lacuna in the field by exploring the combined effect (also referred to as the *interaction effect*) of candidate introversion/extraversion and rater introversion/ extraversion during leadership assessments.

3.7 CONCLUSION

It is clear that assessment centers have undergone major developments since they were introduced into a non-military context, especially in South Africa. It is established in this chapter that the key attribute of assessment centers is the way in which they combine various activities in order to offer organisations a versatile and custom-made solution to assessment challenges. Although this method of leadership assessment

proves to be very effective, possibly the most important consideration should be with regards to the validity and reliability of the activities used.

One of the most popular activities used in assessment centers seems to be the behaviour observation exercise. During such an exercise various factors, like the skills and capabilities of the candidates, as well as environmental factors, come into play. However, one aspect that is often neglected when considering factors impacting on scores of observation exercises is the personality and prejudices of the rater. This often plays out in the form of rater bias which results from the implicit knowledge a rater gathers in his/her everyday life experiences.

The current study focuses on one specific rater bias: similarity bias. This type of bias occurs when raters allocate higher scores to candidates with similar personality types (introversion or extraversion) as themselves. Previous research relating to introversion/ extraversion bias is insufficient to substantiate any finding on the topic. For this reason, the researcher proposes to engage in further exploration on the topic. The methodology used to test for similarity bias based on introversion or extraversion will be discussed in the next chapter.

CHAPTER 4: METHODOLOGY

4.1 INTRODUCTION

After viewing the theoretical background to introversion/extraversion and assessment centers, it is necessary to explain the methodology the researcher is using to achieve the purpose of the study, which is:

> To determine to what extent introversion or extraversion bias influences the scores of behaviour observation exercises during leadership assessments in a security environment.

This chapter will begin by explain the research approach and the research design used in this study. The research method will then be discussed, with specific reference to the sample, data collection procedures and analysis as well as the measurement instruments used. It will be concluded with a brief discussion of the ethical regulations considered.

4.2 RESEARCH APPROACH

A quantitative approach is followed during this study. According to Gravetter and Forzano (2006:147), quantitative research is based on measurement of variables which usually consist of numerical values. These values are submitted to statistical analysis for summary and interpretation. This definition applies to this study as the researcher uses the numerical scores for each candidate obtained during the behaviour observation exercises and also the numerical introversion/extraversion scores of both the raters and the candidates to test the interaction effect.

4.3 RESEARCH DESIGN

The researcher makes use of a cross-sectional, non-experimental field study as data is collected within the same group of raters and candidates participating in an assessment center.

The study can further be described in terms of a correlation design. Gravetter and Forzano (2006:333) state that a correlation design is when "…two or more variables are

measured and recorded to obtain a set of scores (usually two scores) for each individual. The measurements are then reviewed to identify whether a relationship exist between the variables and to measure the strength of the relationship". There are two variables present during this study, namely the introversion/extraversion scores of the raters/candidates (independent variable) as well as the scores obtained during the behaviour observation exercises (dependent variable). The researcher investigates whether or not an interaction effect exists between the two.

Also applicable to this study is the fact that, in a correlation design, the variables are not under the control of the researcher and they therefore cannot be manipulated or interfered with (Gravetter & Forzano, 2006:151).

4.4 RESEARCH METHOD

4.4.1 Sample

For the purposes of this study, the researcher makes use of a 'non-probability' sampling technique called *convenience sampling*. Convenience sampling is, according to Ruane (2005:117), one of the oldest sampling strategies. Unlike 'probability sampling', in which each person in the population has the same known probability of being selected, convenience sampling involves building a sample on the basis of finding convenient or available participants (De Vos, Strydom, Fouche & Delport, 2002:202; Ruane, 2005:117). In other words, those who are selected for the sample are those who conveniently just happened to be part of the study.

The sample in this study comprises individuals participating in an assessment center. It consists of 14 raters (who are all from the same security organisation) and a group of 89 candidates applying for leadership positions. Both males and females from different racial groups participated. The sample also represents all 11 official languages of South Africa and the ages range from 19 to 55 years old.

It has to be mentioned that most of the raters consist of industrial psychology interns. Their lack of experience in observing behaviour and allocating scores to candidates

may impact on the study. Further limitations, like the inability to make generalizations to the broader population, will be discussed in the last chapter.

4.4.2 Data collection procedure

In this study, data is collected during three separate stages: prior to the assessment, during the assessment and after the assessment. During the first stage, candidates are required to complete the Jung Personality Questionnaire (JPQ) to determine their introversion/extraversion status. Candidates then enter the second phase and participate in the assessment center as scheduled. This specific assessment center consists of paper screening, psychometric testing, personal interviews, oral presentations and two behaviour observation exercises, namely the 'Letra Course' and 'Winter Survival'. For the purposes of this study, the researcher uses the scores obtained from the two behaviour observation exercises.

During the Winter Survival exercise, candidates participate in a team of approximately six individuals and are given a certain scenario or problem which they are required to solve among one another. There is one rater allocated for each participant during each round who observes the candidates' interaction and contribution to the team during the exercise. This exercise measures five competencies namely *Communication*, *Adaptability*, *Judgment*, *Decisiveness* and *Team leading*. Each rater allocates scores for his/ her specific candidate on a standardized rating scale (which will be discussed later in this chapter). Consequently, each rater rates several candidates during the day.

The Letra Course also takes place in team formation and the candidates are required to overcome a series of physical obstacles (in this case, either to climb over a high wall as a team without falling or to get across a river as a team with only two ropes). Again, each member is required to make use of his/her communication and leadership skills to guide and motivate other team members. This exercise measures the same competencies as Winter Survival and there is also one rater for each candidate in each round. Each rater rates several candidates throughout the day.

Only after the assessment center, the raters in this study enter the third phase and complete the JPQ to determine their personality type (introversion/extraversion). The

raters are therefore not informed of the candidate's personality types while they are observing them. Should raters allocate more favourable ratings to candidates because they have similar personality types as raters, it will be based on the introverted or extraverted behaviour that they observe during the exercise. This situation is typical of subconscious bias.

The raw scores that each candidate obtains in the simulation exercises (Winter Survival and the Letra Course) and the introversion/extraversion scores of the JPQ (which indicates the personality types of candidates and raters) are used in this study.

4.4.3 Data Analysis

The following hypotheses are tested by applying different statistical techniques:

a) **Hypothesis 1a: There is a relationship between candidate introversion/extraversion and scores of behaviour observation exercises**

Point-biserial correlations are run to explore the relationship between candidate introversion/ extraversion and performance on the two simulation exercises. A point-biserial correlation is used when one variable is a scale variable and the other is a discrete dichotomous variable with only two categories. A discrete dichotomy exists when there is no underlying continuum between the categories (Field, 2009:131-132). For the purpose of this hypothesis, a person was categorised as being either introverted or extroverted. By making use of point-biserial correlations, it is possible to determine whether or not the scores obtained by the candidates during the behaviour observation exercises increased/decreased as their scores on the JPQ increased/decreased.

In addition, independent samples t-tests were used to compare the difference in means between introverted and extraverted candidates on the two simulation exercises. Both the point-biserial correlations and the independent sample t-test indicate whether or not one group of candidates (either introverts or extraverts) scored statistically significantly higher than the other group of candidates (Field,

2009:296). Note that the raters' personality types are only considered in the next hypothesis.

b) **Hypothesis 1b: There is a relationship between rater introversion/extraversion and scores of behaviour observation exercises**

This hypothesis explores whether or not introverted raters rated candidates differently overall than extraverted raters. As with Hypothesis 1a, a point-biserial correlation is used to explore the relationship between rater introversion/extraversion and candidates' performance on the two simulation exercises (Field, 2009:132). In other words, it determines if the scores of the candidates obtained during the two exercises increased/ decreased as the JPQ scores of the raters increased/decreased.

c) **Hypothesis 2: Rater introversion/extraversion and candidate introversion/extraversion has an effect on the scores of behaviour observation exercises**

A two-way analysis of variance (ANOVA) statistic is used to test the second hypothesis. According to Field (2009: 734), the two-way ANOVA investigates the combined effect of two or more variables on an outcome. In other words, it tests to establish if there is an interaction effect between rater introversion/extraversion and candidate introversion/extraversion. In the case of this study, an interaction effect would exist when extraverted raters rated extraverted candidates higher than they rated introverted candidates or, conversely, when introverted raters rated introverted candidates higher than extraverted candidates.

Before any valid comparison may be drawn between the scores of introverted raters and extraverted raters to indicate that rater bias plays out during this study, it is essential to show that raters from both introverted and extraverted personality groups rated candidates from both introverted and extraverted personality groups. A Pearson's chi-square test for independence is applied to test whether or not the proportion of introverted/extraverted raters to candidates is equal. Field (2009:682) explains this

technique as a way to compare the frequencies observed in certain categories to the frequencies one expects to arrive at in these categories by chance.

4.4.4 Measurement instruments

Two measurement instruments are used during this study. The participating candidates and raters are asked to complete a questionnaire to determine their own preferences for introversion or extraversion. In addition, the performance of each candidate during the simulation exercises is rated on a standardized six-point rating scale. Both the personality questionnaire and the rating sheet will now be discussed.

4.4.4.1 The Jung Personality Questionnaire (JPQ)

The JPQ is based on Jung's theory. Jung states that individuals differ in terms of the following: attitudes of *introversion* and *extraversion,* psychological functions of *thinking* or *feeling, sensing* or *intuition,* as well as processing information in a rational way (*judging*) or in an irrational way (*perceiving*) (Du Toit, 1996:4). The JPQ was therefore developed to present a straightforward, yet accurate version of an individual's personality structure and it measures these opposing attitudes and functions on four different scales (Du Toit, 1996:4):

i. Extraversion–Introversion (E/I)
ii. Thinking–Feeling (T/F)
iii. Sensation–Intuition (S/N)
iv. Judgment–Perception (J/P)

The first three scales (E/I, T/F and S/N) consist of twenty questions each and the fourth scale (J/P) of fifteen questions. There are three possible responses to each question and the candidate is required to indicate which of two contrasting tasks he/she would prefer to perform. Only when the candidate is not particularly drawn to one of the tasks is he/she allowed to choose the third option which is 'uncertain'. The responses are scored on a three-point scale ranging from 0-2 (Du Toit, 1996:4).

When adding the scores, the first three scales (E/I, T/F and S/N) will each have a raw score ranging between 0 and 40. These figures are each divided by 4 to arrive at a converted score for each scale. The fourth scale (J/P) has a raw score ranging from 0-30 which is then divided by 3. The converted scores are plotted on four different scales with opposing attitudes/functions that ranges from 1 - 10. Lower scores indicate the preference for the attitude/function on the lower end of the scale, where higher scores indicate the preference for the opposing attitude/function on the higher end of the scale. The combination of these scores provides the candidate with his/her preferred combined personality type for example, ESFJ (Extravert, Sensing, Feeling and Judging) (Du Toit, 1996:12).

For the purposes of this study, the researcher only makes use of the first scale which indicates the individual's preference for introversion/extraversion. This scale categorises individuals into three groups: those who have a clear preference for introversion (scores between 1 and 4), those who have a clear preference for extraversion (scores between 6 and 10) and those who are unsure of their preference (called *Ambiverts*, with scores between 4 and 6).

The current study compares the scores of introverted and extraverted candidates during simulation exercises, therefore the inclusion of the ambivert category may obscure the results. The ambivert group is therefore excluded during data analysis: the researcher only makes use of the introversion and extraversion categories for both the candidates and the raters. Although the original sample consists of 89 participants, once the ambivert category is excluded, the final sample size consists of 66 people: 22 introverts (33.3%) and 44 extraverts (66.7%). Among the raters, 6 are introverts (54.5%) and 5 are extraverts (45.5%).

As mentioned in Chapter 2, there is a difference between personality types and personality traits. Jung bases his theory on personality types and the researcher therefore makes use of a questionnaire that measures introversion-extraversion as such. However, some alternative personality questionnaires that measure introversion/extraversion as either a type or a trait are listed in Table 4.1.

Table 4.1: Alternative Measures of Introversion and/or Extraversion

Questionnaire	Types/Traits	Description
Myers-Briggs Type Indicator (MBTI)	Types	The MBTI is based on Jung's theory and measures or identifies the way in which individuals become aware of things and events around them. It categorises preferences of the individual and measures the strength of the individual's preference (Pittenger, 2005:210). The MBTI is a more recent and further developed version of the JPQ
Eysenck's Personality Questionnaire (EPQ)	Traits	The EQP measures the extent to which three personality dimensions called *Neuroticism*, *Extraversion* and *Psychoticism* are displayed and makes use of a scale with opposing ends (Heffernan and Ling, 2001: 322).
Cattell's 16 Personality Factor Inventory (16PF)	Traits	The 16PF measures an individual's self-perception with regards to sixteen personality traits. In addition, these traits are clustered to form five secondary factors namely *Extraversion, Anxiety, Tough-mindedness, Independence* and *Self-control*. (Lange and Rentfrow, 2007: 430).
The Five Factor Model of Personality (the Big 5 Model)	Traits	The Big 5 Model measures five personality factors called *Extraversion, Agreeableness, Conscientiousness, Neuroticism* and *Openness*. (Denissen, Geenen, Van Aken, Gosling and Potter, 2008: 153).

As seen in Table 4.1, a similar instrument to the JPQ is the MBTI. It is also based on Jung's theory and provides a more recent and reliable measure of the two personality types called introversion and extraversion. However, the need for information on only one scale, the unavailability of the MBTI and the financial implications associated with the MBTI all form part of the motivation for this researcher to use the JPQ (refer Appendix A for a overview of the JPQ questions referring to introversion and extraversion).

According to Du Toit (1996:4) the JPQ is reliable, with a total Cronbach's alpha coefficient of 0.87. In the current sample, the reliability for the raters (N=14) who completed the JPQ is sufficient (Cronbach's alpha coefficient of 0.83). However, the JPQ is not as reliable when measuring candidates (Cronbach's alpha coefficient of 0.54).

The researcher therefore explores the lessened reliability of the JPQ when measuring candidates by testing the reliability of the JPQ for each language group involved in the study. The highest reliability scores are obtained for the English-speaking students (Chronbach's alpha of 0.71, which is still acceptable). However, for the Afrikaans-speaking students the reliability is 0.67 and for the African-speaking students, a relatively low 0.39. This is a logical finding as the questions of the JPQ are presented in English (which is not the first language for the Afrikaans and African speaking candidates). The reliability scores of the JPQ during this study are therefore a limitation that will be discussed in the last chapter.

Table 4.2: Reliability Statistics for The JPQ Introversion/Extraversion Scale

	Cronbach's Alpha	**N of Items**
Rater (N=14)	.833	20
Candidate (N=89)	.543	20
African Language (n=62)	*.392*	*20*
Afrikaans (n=20)	*.677*	*20*
English (n=7)	*.714*	*20*

4.4.4.2 The standardised rating sheet

The behaviour observation exercises are designed to elicit the following competencies: *Communication, Adaptability, Judgment, Decisiveness* and *Team Leading*. Each of these competencies comprises different items, for example, a competency like 'Communication' will consist out of 'Audibility', 'Speed', 'Language usage', 'Confidence', Assertiveness' and 'Eye contact'.

During both exercises each of these items are rated on a 6-point Likert scale ranging from 'No evidence' (0) to 'Superior' (5) (refer Appendix B). The first three points ('No Evidence' to 'Less than Acceptable') represent undesirable behaviour while the last three points ('Acceptable' to 'Superior') represent desirable behaviour.

The Likert scale is one of the most long-standing and popular methods used when it comes to rating performance (Grobler, et al. 2006: 270). It is relatively easy to complete and it therefore requires minimum training (Grobler, et al. 2006: 270). A similar method like the Behaviorally Anchored Rating Scale (BARS) could also have been used by the raters. The organisation in the current study prefers to use the Likert scale because the BARS rating scale is time consuming. However, some examples of ideal behaviours on the Likert scale are included to guide the raters.

Each competency has a different number of items, therefore the total score for each competency is converted to a score out of five so that one competency does not contribute more to the overall score of the exercise. This step also makes the interpretation of the score easier. The total score for each exercise is therefore out of 25 points.

In order to test the reliability of the rating sheet during the two behaviour observation exercises of the current study, a reliability coefficient is run. As seen in Table 2, the reliability of the rating scale is also satisfactory for both exercises. The reliability coefficient for the African candidates is 0.971 where the English candidates obtained a reliability coefficient of 0.94 during Winter Survival exercise.

One reason for the satisfactory reliability coefficient may be the training attended by the raters before the assessment center. Raters had been educated on the different rater mistakes and also on the types of leadership behaviour that would be appropriate during the exercises.

Table 4.3: Reliability Statistics for the two Behaviour Observation Exercises

	Cronbach's Alpha	N of Items
WINTER SURVIVAL		
Candidate (N=89)	.973	20
African (n=62)	*.971*	*20*
English/Afrikaans (n=27)	*.940*	*20*
LETRA COURSE		
Candidate (N=89)	.966	20
African (n=62)	*.960*	*20*
English/Afrikaans (n=27)	*.968*	*20*

4.5 ETHICAL CONSIDERATIONS

Conducting research comes with great responsibility. This responsibility is best described by Breakwell (2004:42) who states that when participants agree to be part of a research process, they place themselves in the experimenter's care. She is further of opinion that some research, especially in the field of social psychology, is notorious for using deceptive tactics to manipulate independent variables. Breakwell advises researchers to avoid a cavalier attitude towards their candidates and urges them to make the process a positive experience (Breakwell, 2004:42). In addition, Ruane (2006:16) states that standards of professionalism and honesty should be kept by researchers at all times in order to earn respect from both the participants and the public at large.

In order to guide researchers in acceptable behaviour, certain ethical standards and guidelines have been developed. These guidelines are objective in nature and should

not only be followed when representing the results, but also while conducting the study (Ruane, 2006:17). The following ethical guidelines as indicated by Ruane (2006:17) and Breakwell (2004:120-142) apply to the current study:

4.5.1 Research should not cause harm to respondents

According to Ruane (2006:18) researchers are obliged to safeguard the physical, psychological and emotional well-being of participants.

A full medical team was on stand-by during the physical activities in the assessment center and candidates were fully briefed before they participated.

4.5.2 Researchers should obtain the informed consent of respondents

Participants have the right to decide for themselves whether or not they want to be part of a research project. All candidates are informed about all aspects of the study, including the fact that they are allowed to withdraw their consent from the study at any time (Ruane, 2006:19; Breakwel, 2004:120, 242).

During this study every individual is required to indicate whether or not he/ she wants to participate by signing a consent form (refer Appendix C). This consent form includes a statement of the participant's rights and the option of withdrawing. The briefing session before the assessment center started provides an opportunity for the candidates to ask questions if anything involved in the assessment center requires clarification.

4.5.3 Researchers should respect participants' privacy

Respecting the privacy of respondents mainly refers to three things: to protect sensitive information provided by candidates, to choose a non-invasive location or setting for the research, and to ensure the confidentiality of the results (Ruane, 2006:22-26).

In the current study, neither the content of the measuring instruments or the location of the exercises used in the assessment center poses a threat to the participants. The researcher also respects the participants' privacy by ensuring the following:

i. The names of the participants as well as the name of the security company are not published in this study as requested by management.

ii. The scores obtained during the behaviour observation exercises are linked to candidate names only during the selection, but not for the results of this study.

iii. The introversion/extraversion scores obtained during the JPQ are not for selection purposes.

4.5.4 Researchers should avoid conflicts of interest

According to Ruane (2006:28), researchers are often requested and sponsored by corporate companies to conduct studies that are used as inputs for growth strategies. The risk for bias and manipulation of results therefore cannot be ignored as they impact on the growth of organisations and individuals. As Ruane (2006:28) states, researchers should avoid becoming a “hired gun” for their interest groups.

There are no corporate funders for this study and even though the results may have a significant impact on the future selection procedures of the specific organisation, it is not in the interest of the researcher to manipulate the outcome. A factor which further contributes to the authenticity of this study, is the fact that the raters are not aware of the purpose of the research (which is to test for introversion/extraversion bias) while they allocate scores to candidates. Furthermore, raters are only tested for introversion/ extraversion after the assessment center. Lastly, the results are published in a non-threatening and responsible manner.

According to Breakwell (2004:42), ethical bodies, otherwise known as institutional review boards, exist to govern the ethical reporting of research studies. At the University of Pretoria, the proposed study is reviewed by a departmental research committee, a faculty ethics committee and a postgraduate committee to ensure the quality and ethical adherence of the research is up to standard.

4.6 CONCLUSION

This chapter presents the methodology used to test for introversion/ extraversion bias in behaviour observation exercises. It begins by explaining the research approach and research design that the researcher utilises. Chapter 4 then explores certain aspects relating to the research method namely the sample, data collection procedure, statistical analyses and measurement instruments used in the study. It concludes with the ethical considerations that the researcher acknowledges ensuring a safe, responsible study. The results will be used to guide the future selection processes of this specific organisation and will be discussed in the next chapter.

CHAPTER 5: RESULTS

5.1 INTRODUCTION

Chapter 5 will present the results of this research project. It will begin by describing the sample that was used in terms of gender and race, language, age, introversion and extraversion as well as the scores obtained during the behaviour observation exercises. The researcher will then present and interpret the results which test the hypotheses stated in this study by making use of correlations, independent t-tests and cross-tabulations. Lastly, the study will conclude with the results obtained from the two-way ANOVA's which test whether or not similarity bias is present during the behaviour observation exercises.

5.2 DESCRIPTION OF THE SAMPLE

5.2.1 Sample distribution by gender and race

The sample consists of 89 candidates, with African, Coloured, Indian and White race groups represented to various degrees. Male candidates (n = 72) outnumbered female candidates (n = 17), with the largest group being African males (n = 52) (refer Table 5.1 and Figure 5.1). Coloured and Indian respondents were poorly represented (refer Table 5.1 and Figure 5.2). The rater group consisted of 14 individuals who were predominantly White females (n = 7) (refer Table 5.1).

5.2.2 Sample distribution by language

Although the sample represents all eleven official languages of South Africa, some groups are poorly represented (refer Table 5.2). Sixty two (69.7%) out of the 89 candidates are African-speaking individuals, whereas only 20 (22.5%) are Afrikaans-speaking and 7 (7.9%) are English-speaking. Table 5.2 includes information on the language distribution of South Africa (Statistics South Africa, 2001) and it is clear that the English- and Swati-speaking groups were truly representative of the population while the other languages are not represented in such a way that the sample group of candidates matches the South African population on a much smaller scale.

Table 5.1: Gender and Race Distribution of Raters and Candidates

Candidate/Rater				Gender		Total
				Female	Male	
Candidate	Race	African	Count	12	52	64
			% of Total	13.5%	58.4%	71.9%
		Coloured	Count	2	2	4
			% of Total	2.2%	2.2%	4.5%
		Indian	Count	1	0	1
			% of Total	1.1%	.0%	1.1%
		White	Count	2	18	20
			% of Total	2.2%	20.2%	22.5%
	Total		Count	17	72	89
			% of Total	19.1%	80.9%	100.0%
Rater	Race	African	Count	1	2	3
			% of Total	7.1%	14.3%	21.4%
		Indian	Count	2	0	2
			% of Total	14.3%	.0%	14.3%
		White	Count	7	2	9
			% of Total	50.0%	14.3%	64.3%
	Total		Count	10	4	14
			% of Total	71.4%	28.6%	100.0%

Figure 5.1: Percentage Males and Females among Candidates

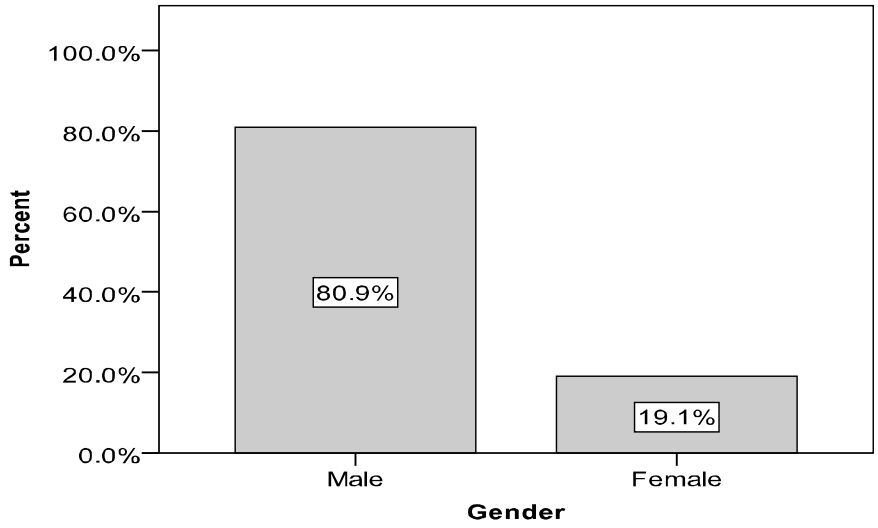


Figure 5.2: Percentage Race Groups among Candidates

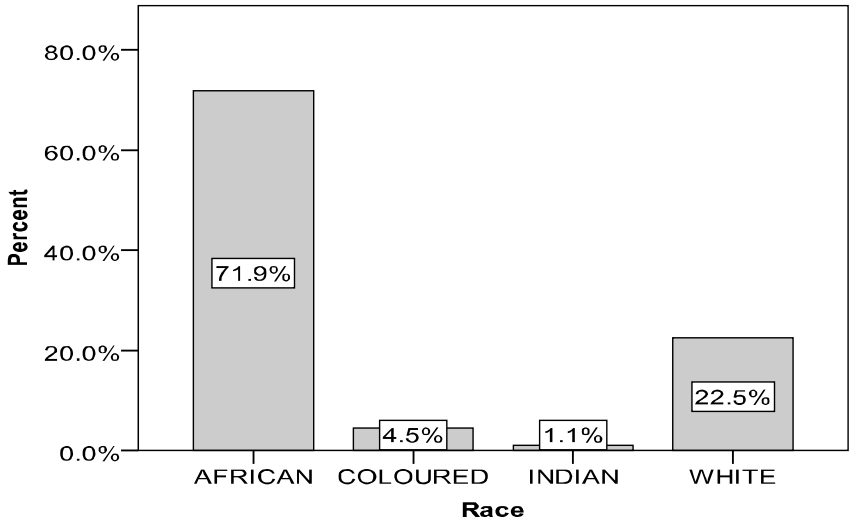


Table 5.2: Language Distribution of Candidates and Raters

Language	Frequency	Percent	Language Distribution in South Africa
Afrikaans	20	22.5	13.35
English	7	7.9	8.2
N Sotho	15	16.9	9.93
Ndebele	5	5.6	1.59
S Sotho	5	5.6	7.93
Swati	3	3.4	2.66
Tsonga	1	1.1	4.44
Tswana	6	6.7	8.2
Venda	5	5.6	2.28
Xhosa	5	5.6	17.64
Zulu	17	19.1	23.82
Other	0	0	0.48
Total	89	100.0	100

5.2.3 Sample distribution by age

From Table 5.3 it is evident that candidates' ages range from 18 to 29 years and that the mean age is 21 years (*SD* = 0.25). From Figure 5.3 it is also clear that the sample is positively skewed (*skewness* = 0.99) which indicates that the candidates are mostly young individuals (between the ages of 18 and 22). This is typical for a selection of this nature in the specific environment. The young sample group may however serve as a limitation when making inferences from the results of this study to a group of older

sample candidates. The ages of the raters range from 24 to 49 and the mean age is 34 years (*SD* = 3.21) (Refer Table 5.3 and Figure 5.4). Age data for seven of the raters is missing.

Table 5.3: Descriptive Statistics of Candidates' Age

		Candidate	Rater
N	Valid	89	7
	Missing	0	7
Mean		21.09	34.00
Std. Error of Mean		.25	3.21
Median		21.00	33.00
Mode		22	38
Std. Deviation		2.32	8.49
Variance		5.38	72.00
Skewness		.99	.78
Std. Error of Skewness		.26	.80
Kurtosis		1.29	.38
Std. Error of Kurtosis		.51	1.59
Minimum		18	24
Maximum		29	49

Figure 5.3: Histogram Displaying Candidate Age

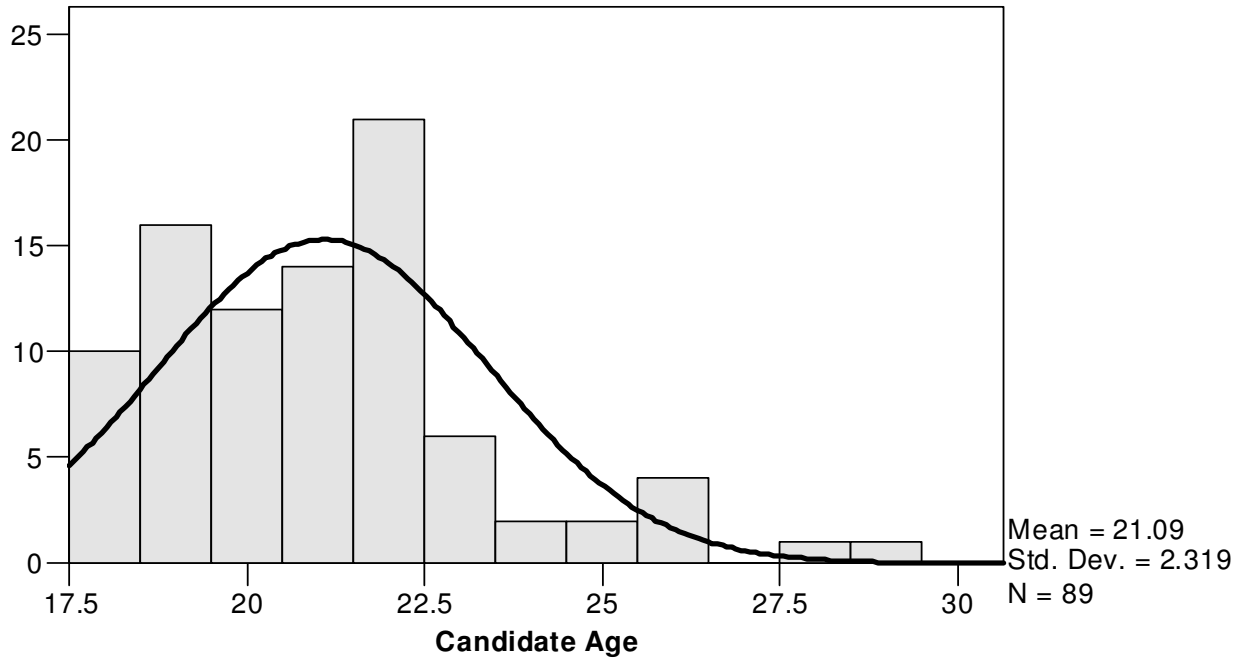
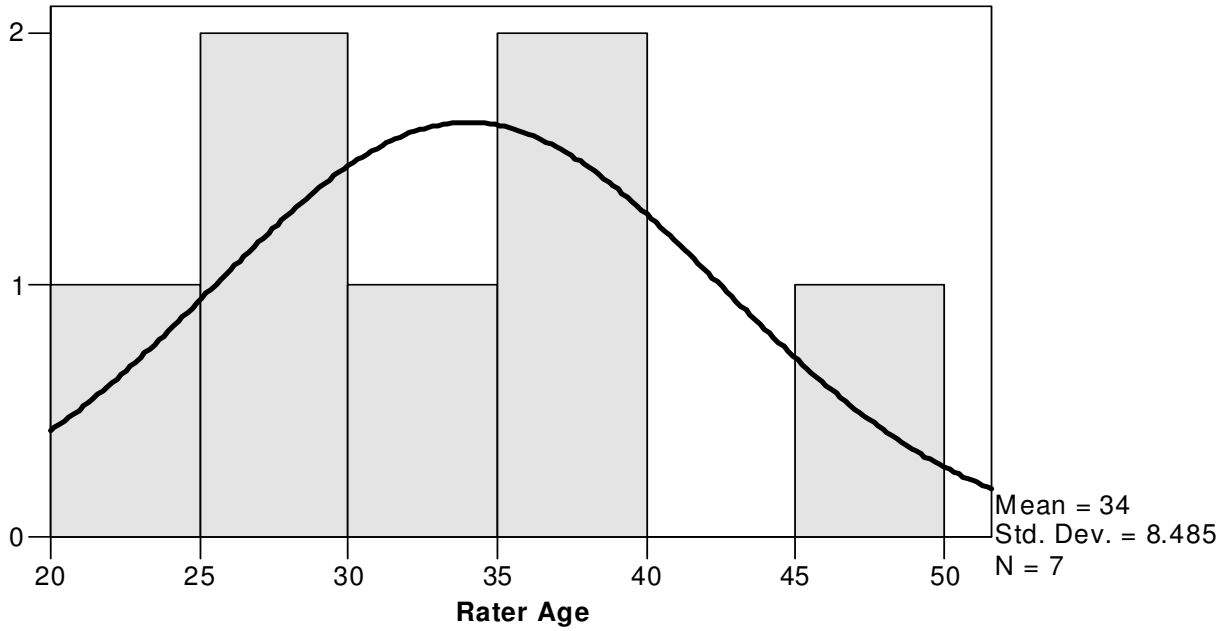


Figure 5.4: Histogram Displaying Rater Age



5.2.4 Sample distribution by introversion/extraversion scores

Possible scores for introversion/extraversion (I/E) as measured on the JPQ range between one and ten. The introversion/extraversion scores for the group of candidates are approximately normally distributed and the scores range from 2.75 and 9.50 (most of the introversion/ extraversion scale is therefore used). The mean score for this group on the scale is 5.55 (*SD* = 1.46) (refer Table 5.4 and Figure 5.5).

The introversion/ extraversion scores obtained by the raters range from 1.75 to 8.50 and the mean score 4.61 (*SD* = 2.26). The distribution of scores is flat (*kurtosis* = -1.07) but that is a logical result as the group of raters is small (refer Table 5.4 and Figure 5.6).

Table 5.4: Descriptive Statistics of Introversion/Extraversion scores for Candidates and Raters

	Candidate	Rater
N (valid)	89	14
Missing	0	0
Mean	5.55	4.61
Std. Error of Mean	.16	.60
Median	5.50	4.75
Mode	6.00	1.75
Std. Deviation	1.46	2.26
Variance	2.15	5.10
Skewness	.45	.26
Std. Error of Skewness	.255	.60
Kurtosis	-.09	-1.07
Std. Error of Kurtosis	.51	1.15
Minimum	2.75	1.75
Maximum	9.50	8.50

Figure 5.5: Histogram Displaying Candidate JPQ Scores

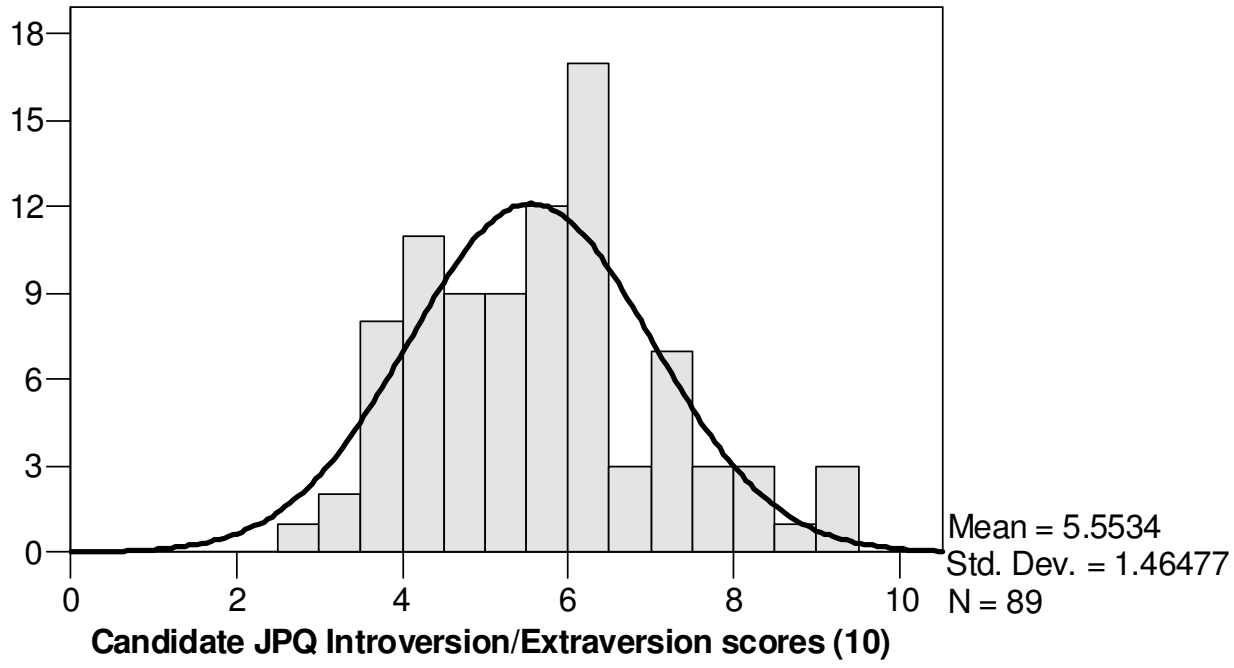
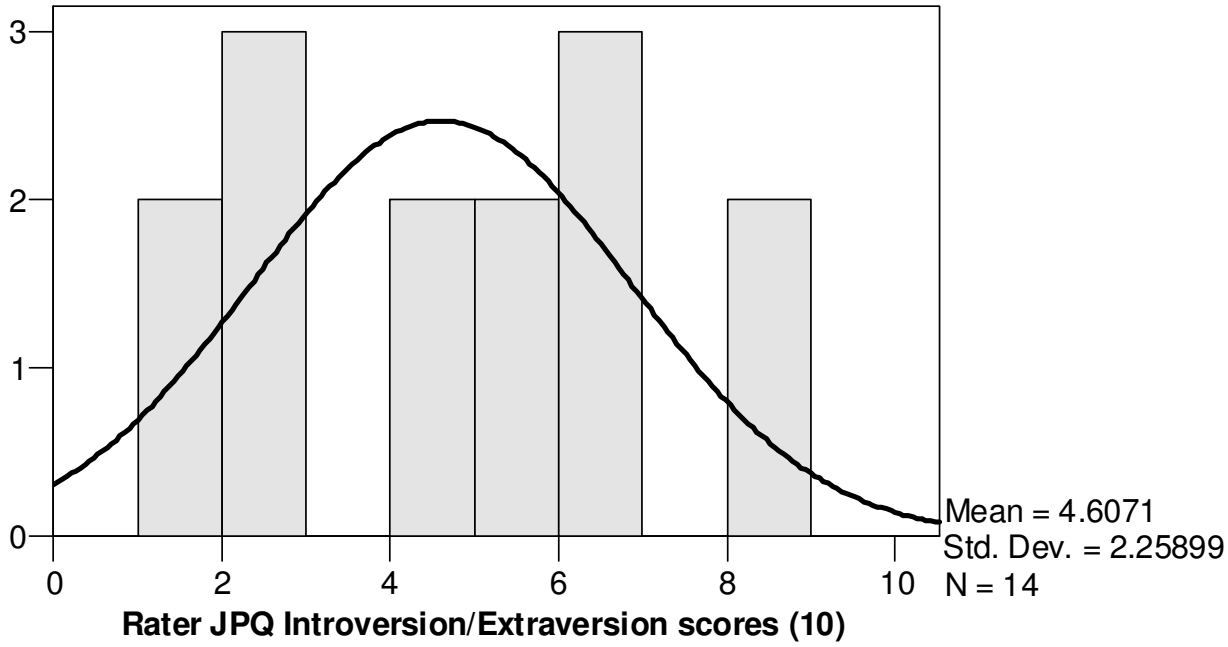


Figure 5.6: Histogram Displaying Rater JPQ Scores



5.2.5 Distribution of scores in Winter Survival and Letra Course

Possible scores for Winter Survival range between 1 and 25; the minimum score received by a candidate is 1.65 and the maximum score is 24.17 *(M* = 14.94, *SD* = 0.58) (refer Table 5.6). The distribution is not normally distributed (*p* = 0.005) (refer Table 5.5) and negatively skewed (*skewness* = -0.80795) which means that more scores lie on the upper end of the distribution (refer Table 5.6, and Figure 5.7). Relatively few candidates therefore performed poorly. However, it is also positively peaked (*kurtosis* = 1.06) which indicates that the majority of candidates scored close to the mean in this exercise (refer Table 5.6).

For Letra Course, possible scores range between 1 and 25; the minimum score received by a candidate is 2.50 and the maximum score was 23.62 (*M* = 14.77, *SD* = 0.46) (refer Table 5.6).The distribution is approximately normally distributed (*p* = 0.367) (refer Table 5.5, and Figure 5.8) with a high kurtosis value (*kurtosis* = 0.95). The distribution shows a slight negative skewness (*skewness* = -0.31) (refer Table 5.6).

Table 5.5: Test of Normality of Scores Obtained During Winter Survival and Letra Course

		Shapiro-Wilk	
Exercise	**Statistic**	**Df**	**Sig.**
Ex 1: Winter Survival	.944	66	.005
Ex 2: Letra Course	.980	66	.367

Table 5.6: Descriptive Statistics of Scores Obtained by Candidates during Winter Survival and Letra Course

		Ex 1: Winter Survival	**Ex 2: Letra Course**
N	Valid	66	66
	Missing	0	0
Mean		14.94	14.77
Std. Error of Mean		.58	.46
Median		15.03	14.32
		17.98	
			12.50, 13.22, 13.27, 14.13
Std. Deviation		4.67	3.76
Variance		21.84	14.16
Skewness		-.80	-.31
Std. Error of Skewness		.30	.30
Kurtosis		1.06	.95
Std. Error of Kurtosis		.58	.58
Range		22.52	21.12
Minimum		1.65	2.50
Maximum		24.17	23.62

Figure 5.7: Histogram Displaying Performance of Candidates during Winter Survival

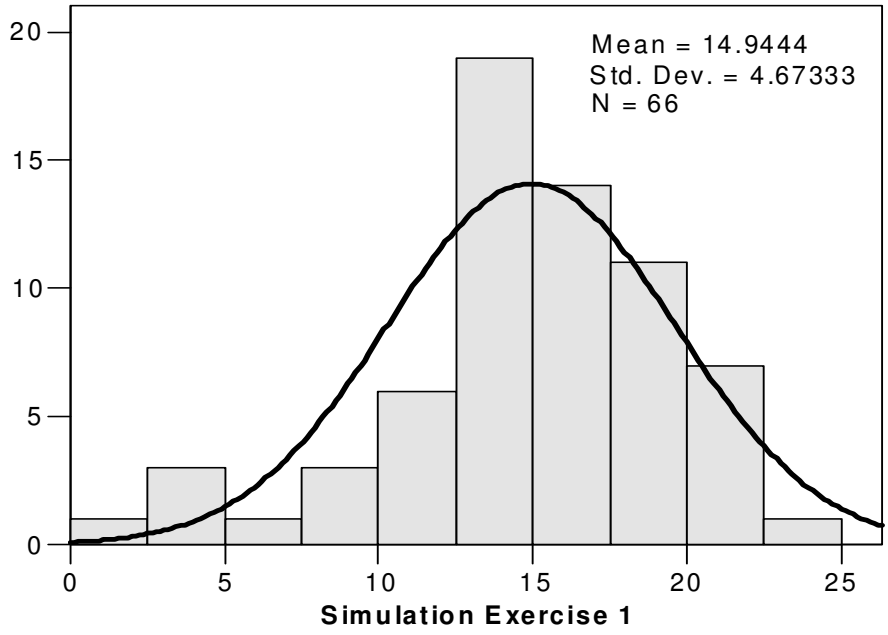
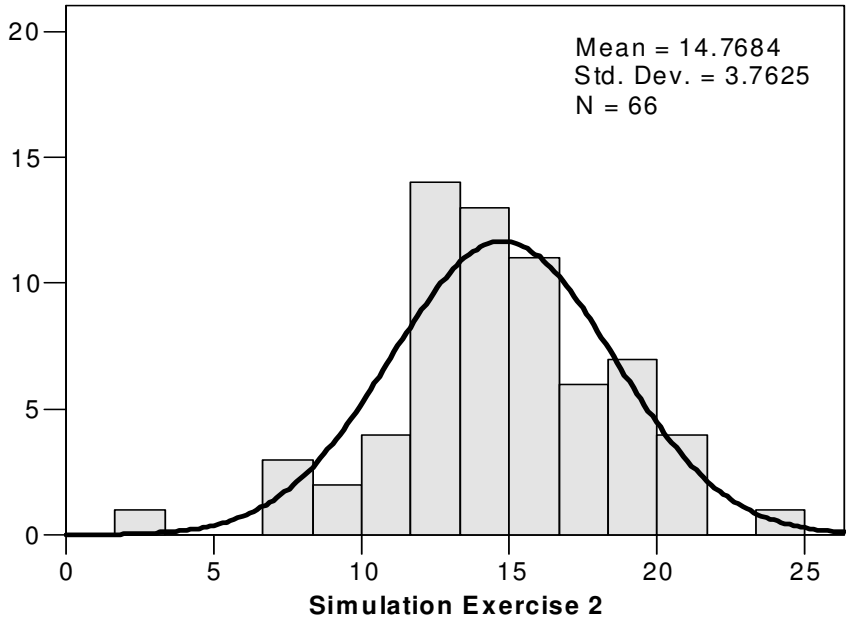


Figure 5.8: Histogram Displaying Perfromance of Candidates during Letra Course



5.3 RELATIONSHIP BETWEEN CANDIDATE INTROVERSION/EXTRAVERSION AND PERFORMANCE ON BEHAVIOUR OBSERVATION EXERCISES

H1a: There is a relationship between candidate introversion/extraversion and scores of behaviour observation exercises

Point-biserial correlations are run to explore the relationship between candidate introversion/ extraversion and performance on the two behaviour observation exercises. The correlation between candidate introversion/ extraversion and Winter Survival (r_{pb} = 0.22, p = 0.075) represents a small effect size (refer Table 5.7). However, there is insufficient power to detect this as a statistically significant result for a sample of 66. This may be a result of the small sample size or the outlier (candidate 6) that was detected in the output for the Winter Survival (refer Figure 5.9). Removal of the outlier increases the correlation only slightly (still a small effect size according to Cohen, 1988), but enough to be considered statistically significant for a sample of 65 (r_{pb} = 0.26, p = 0.033) (refer Table 5.7 and Figure 5.10). Therefore between 4.84% and 6.97% of the variance in the first simulation exercise may be explained by the variance in candidate introversion/extraversion. As a result it can be stated that there is some support for Hypothesis 1 in Winter Survival, but only if it applies to the sample of 65.

Candidate I/E is positively correlated with scores obtained during Letra Course (r_{pb} = 0.32, p = 0.009) (refer to Table 5.7). This indicates that extraverted candidates are more likely to perform better on the exercise than introverted candidates. According to Cohen (1988) this represents a moderate effect, with 10.3% of the variance in the exercise scores explained by candidates being either introverted or extraverted. Even after the scores of candidate 6 are removed, the relationship between candidate introversion/ extraversion and scores obtained during this exercise is still found to be statistically significant (r_{pb} = 0.32, p = 0.008) (refer Table 5.7). Hypothesis 1 can therefore be accepted for this exercise.

Table 5.7: Correlations Between Candidate Introversion/Extraversion and Scores Obtained During Behaviour Observation Exercises

		Candidate I/E (N = 66)	Candidate I/ E (N = 65)
Ex 1: Winter Survival	Pearson Correlation	.22	.26(*)
	Sig. (2-tailed)	.075	.033
Ex 2: Letra Course	Pearson Correlation	.32(**)	.32(**)
	Sig. (2-tailed)	.009	.008

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Figure 5.9: Scatter Plot of the Correlation Between Candidate Introversion/Extraversion and Winter Survival Scores (n = 66)

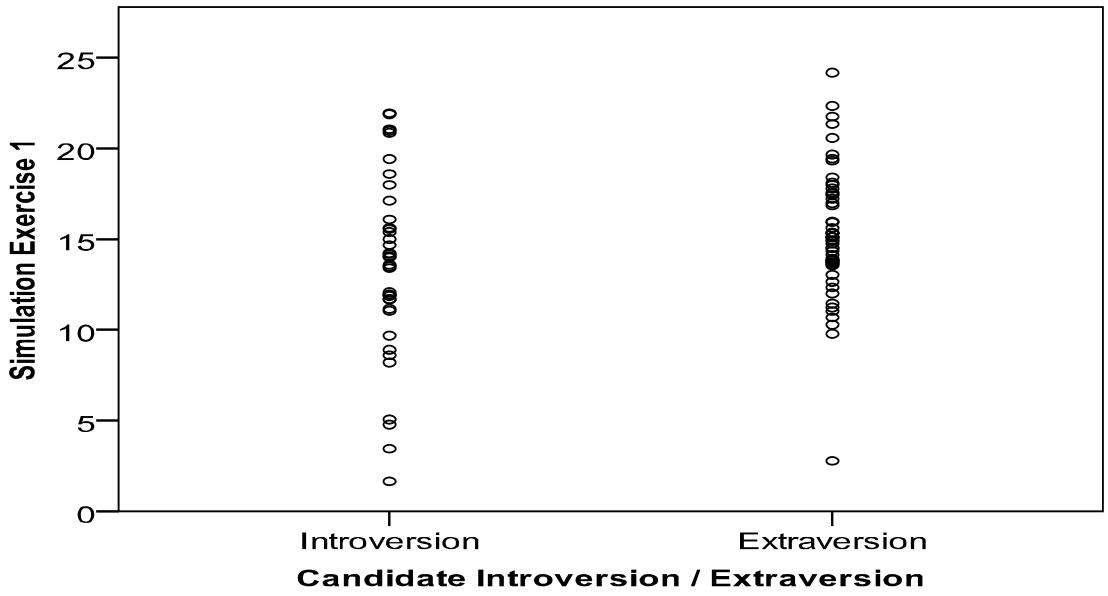
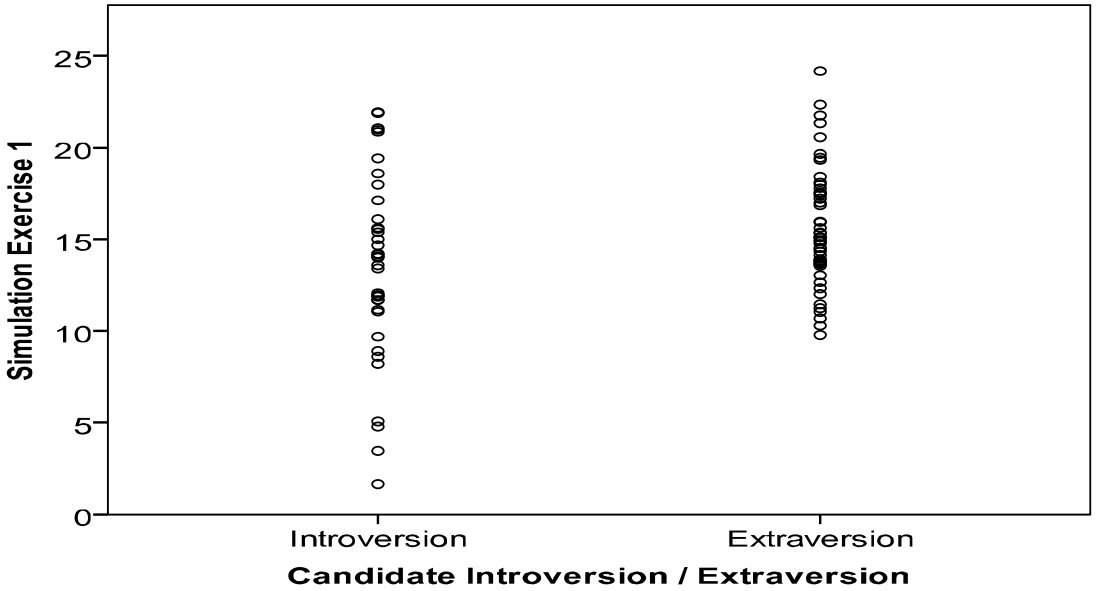


Figure 5.10: Scatter Plot of the Correlation between Candidate Introversion/Extraversion and Winter Survival Scores (n = 65)



5.4 COMPARING MEANS OF INTROVERTED AND EXTRAVERTED CANDIDATES

Independent samples t-tests are run to compare the difference in means between introverted and extraverted candidates on the two behaviour observation exercises. For Winter Survival, introverts achieved a mean score of 13.50 (*SD* = 5.93) compared with extraverts (*M* = 15.67, *SD* = 3.77) (refer Table 5.8). Levene's test is significant (*p* = 0.014) indicating that the assumption of equality of variances is violated. The t-test for unequal variances is therefore utilized. This result fails to reach statistical significance (*p* = 0.129), indicating that the mean scores for introverts and extraverts in Winter Survival do not differ significantly (refer Table 5.8 and Appendix D).

When excluding the outlier (candidate 6) from the results for Winter Survival, the mean scores for introverted candidates (*M* = 13.50, *SD* = 5.93) and extraverted candidates ($\bar{x}$ = 15.97, *SD* = 3.24) still do not differ significantly from each other. However, this result

does come close to reaching statistical significance (*t(27.595)* = -1.82, *p* = 0.080, η^2 = 0.05) (refer Table 5.9 and Appendix D).

For exercise Letra Course, introverts achieve a mean score of 13.07 (*SD*=4.09) compared with extraverts (*M* = 15.63, *SD* = 3.32). The assumption of equality of variances is met (*p* = 0.444) and the t-test for equal variances assumed is therefore utilized. This test is significant (*p* = 0.009), as it indicates that extraverts scored higher than introverts on the second exercise. This translates to a moderate to large effect size (η^2 = 0.103) according to Cohen (1988) (refer Table 5.8 and Appendix D).

When excluding the results for candidate 6 (n = 65) in the results for this exercise, the difference between the mean of introverted candidates (*M* = 13.07, *SD* = 4.09) and extraverted candidates (*M* = 15.65, *SD* = 3.35) remains statistically significant (*p* = 0.008) with a moderate effect size according to Cohen (1998) (η^2 = 0.105) (refer Table 5.9 and Appendix D).

Table 5.8: T-Test for Comparing the Mean Scores of Introverted and Extraverted Candidates on Behaviour Observation Exercises (n = 66)

Candidate Introversion/ Extraversion	Ex 1: Winter Survival			Ex 2: Letra Course		
	n	M	sd	N	M	sd
Introverts	22	13.50	5.93	22	13.07	4.09
Extraverts	44	15.67	3.77	44	15.63	3.32
	t (29.76) = -1.56, p = .129			t (64) = -2.72, p = .009, η^2 = .103		

Table 5.9: T-Test for Comparing the Mean Scores of Introverted and Extraverted Candidates on Behaviour Observation Exercises (n = 65)

Candidate Introversion/ Extraversion	Ex 1: Winter Survival			Ex 2: Letra Course		
	n	M	sd	N	M	sd
Introverts	22	13.50	5.93	22	13.07	4.09
Extraverts	43	15.97	3.24	43	15.65	3.35
	t (27.595) = -1.82, p = .080 η^2 = 0.05			t (63) = -2.72, p = .008, η^2 = .105		

5.5 RELATIONSHIP BETWEEN RATER INTROVERSION/EXTRAVERSION AND RATINGS ALLOCATED DURING BEHAVIOUR OBSERVATION EXERCISES

H1b: There is a relationship between rater introversion/extraversion and scores of behaviour observation exercises.

Point-biserial correlations are run to explore the relationship between rater introversion/ extraversion and candidates' performance on the two behaviour observation exercises. No statistically significant correlations are found for either exercise ($p_{exercise1}$ = 0.425; $p_{exercise2}$ = 0.564), suggesting that raters' evaluation of candidates' performance are not related to raters being either introverted or extraverted (refer Table 5.10). This result is not unexpected as one would hope not to find that the scores allocated by raters are strongly related to their personality as this would render the results of the behavior observation exercises irrelevant based on raters being subjectively influenced by the candidate's personality. Hypothesis 1b is therefore rejected.

Table 5.10: Correlations between Rater Introversion/Extraversion and scores on Behaviour Observation Exercises

		Ex 1: Winter Survival	Ex 2: Letra Course
Winter Survival Raters: Introversion Extraversion	Pearson Correlation	.11	
	Sig. (2-tailed)	.425	
	N	55	
Letra Course Raters: Introversion Extraversion	Pearson Correlation		.08
	Sig. (2-tailed)		.564
	N		59

5.6 ALLOCATION OF RATERS TO CANDIDATES

The sample size for the first exercise is 55 (11 missing cases are due to ambivert raters who were excluded). As explained in the Methodology chapter, each rater has an opportunity to rate more than one candidate during each exercise. In order to draw a valid comparison between the scores of introverted raters and extraverted raters (and show that rater bias plays out during this exercise), it is essential that both rater groups

(introverts and extraverts) rate individuals from both candidate groups (introverts and extraverts).

During Winter Survival, introverted raters rate introverted candidates 10 times (18.2% of the time) and extraverted candidates 17 times (30.9% of the time). Extraverted raters rate introverted candidates 7 times (12.7% of the time) and extraverted candidates 21 times (38.2% of the time) (refer Table 5.11).

Pearson's chi-square test for independence is run to test whether or not the proportion of introverted/ extraverted raters to candidates is the same. This test does not yield significant results ($\chi^2 = 0.93$, $df = 1$, $p = 0.334$) as it indicates that introverted/extraverted raters are not disproportionately paired with introverted/extraverted candidates (refer Table 5.11).

Table 5.11: Cross Tabulation of Candidate Introversion/Extraversion and Rater Introversion/Extraversion for Winter Survival

			Rater I/E		
			Introversion	**Extraversion**	**Total**
Candidate I/E	Introversion	Count	10	7	17
		Expected Count	8.3	8.7	17.0
		% of Total	18.2%	12.7%	30.9%
	Extraversion	Count	17	21	38
		Expected Count	18.7	19.3	38.0
		% of Total	30.9%	38.2%	69.1%
Total		Count	27	28	55
		Expected Count	27.0	28.0	55.0
		% of Total	49.1%	50.9%	100.0%

$\chi^2 = 0.93$, df = 1, p = 0.334

0 cells (0%) have an expected count less than 5. The minimum expected count is 8 and the assumption of Chi-square is therefore met.

In Letra Course, the sample size for the second exercise is 59 (7 missing cases are due to ambivert raters who were excluded). Introverted raters rate introverted candidates 13 times (22% of the time) and extraverted candidates 23 times (39% of the time). Extraverted raters rate introverted candidates 6 times (10.2% of the time) and extraverted candidates 17 times (28.8% of the time) (refer Table 5.12).

As in exercise Winter Survival, the Pearson Chi-square test does not yield significant results (χ^2 = 0.65, *df* = 1, *p* = 0.422) as it also indicates that introverted/extraverted raters are not disproportionately paired with introverted/extraverted candidates (refer Table 5.12).

Table 5.12: Cross Tabulation of Candidate Introversion/ Extraversion and Rater Introversion/ Extraversion for Letra Course

			Rater I/E		
			Introversion	**Extraversion**	**Total**
Candidate I/E	Introversion	Count	13	6	19
		Expected Count	11.6	7.4	19.0
		% of Total	22.0%	10.2%	32.2%
	Extraversion	Count	23	17	40
		Expected Count	24.4	15.6	40.0
		% of Total	39.0%	28.8%	67.8%
Total		Count	36	23	59
		Expected Count	36.0	23.0	59.0
		% of Total	61.0%	39.0%	100.0%

χ^2 = 0.65, df = 1, p = 0.422

0 cells (0%) have an expected count less than 5. The minimum expected count is 7 and the assumption of Chi-square is therefore met.

5.7 INTERACTION EFFECT BETWEEN RATER I/E AND CANDIDATE I/E DURING BEHAVIOUR OBSERVATION EXERCISES

H2: Rater introversion/extraversion and candidate introversion/extraversion has an effect on the scores of behaviour observation exercises.

A two-way analysis of variance (ANOVA) statistic is used to test whether or not there is an interaction effect between rater introversion/ extraversion and candidate introversion/extraversion. As explained in Chapter 3 and 4, an interaction effect would exist when introverted raters rate introverted candidates higher than extraverted candidates or when extraverted raters rate extraverted candidates higher than introverted candidates. Two groups of candidates (introverts and extraverts) and raters (introverts and extraverts) in both exercise Winter Survival and Letra Course are used for the two-way ANOVA (as mentioned previously, the ambiverts are excluded).

The mean scores obtained for Winter Survivial by the introverted candidates and extraverted candidates differs to some extent (refer Table 5.13). It would appear that extraverted raters rate extraverted candidates (M = 16.60, SD = 3.10) higher than introverted candidates (M = 15.40, SD = 3.34). Whether or not this interaction effect is significant will now be further explored.

The assumption for Levene's Test of Equality is violated (p = 0.028) which indicates that the variance for the dependent variable is not equal across groups for this exercise (refer Table 5.14). Pallant (2005:234) suggests that a more stringent level of significance should be set if Levene's test is violated. However, even with a more stringent approach (Sig. level at 1%) the results still do not prove to be significant (see explanation below). A post-hoc test is also unnecessary because this study only explores the interaction effect by using a 2-way ANOVA.

Contrary to the hypothesis postulated, no evidence is found to support an interaction effect between candidate and rater introversion/ extraversion [$F_{(1,51)}$ = 0.72, p = 0.401] for exercise Winter Survival, suggesting that extraverted raters do not rate extraverted candidates higher than introverted candidates and vice versa (refer Table 5.15). Indeed, raters' own preference for introversion/ extraversion is not seen to have any effect on simulation exercise scores [$F_{(1,51)}$ = 0.02, p = 0.896]. However, the candidates' preference for introversion/extraversion is seen to have an effect on simulation exercise scores that approach significance at the 0.01 level [$F_{(1,51)}$ = 6.89, p = 0.011] (refer Table 5.15). The magnitude of this effect (η_{p2} = 0.119) is described by Cohen (1988) as moderate (refer Table 5.15). Introverted and extraverted candidates could therefore be seen to differ in terms of their scores for the first simulation exercise, with extraverted candidates obtaining higher scores on average than introverted candidates (see Figure 5.11).

Table 5.13: Descriptive Statistics for Winter Survival

Candidate I/E	Rater I/E	M	Std. Deviation	n
Introversion	Introversion	13.22	6.73	10
	Extraversion	12.34	4.06	7
	Total	12.86	5.64	17
Extraversion	Introversion	15.40	3.34	17
	Extraversion	16.60	3.10	21
	Total	16.06	3.22	38
Total	Introversion	14.59	4.87	27
	Extraversion	15.53	3.78	28
	Total	15.07	4.33	55

Table 5.14: Levene's Test of Homogeneity of Variances for Winter Survival

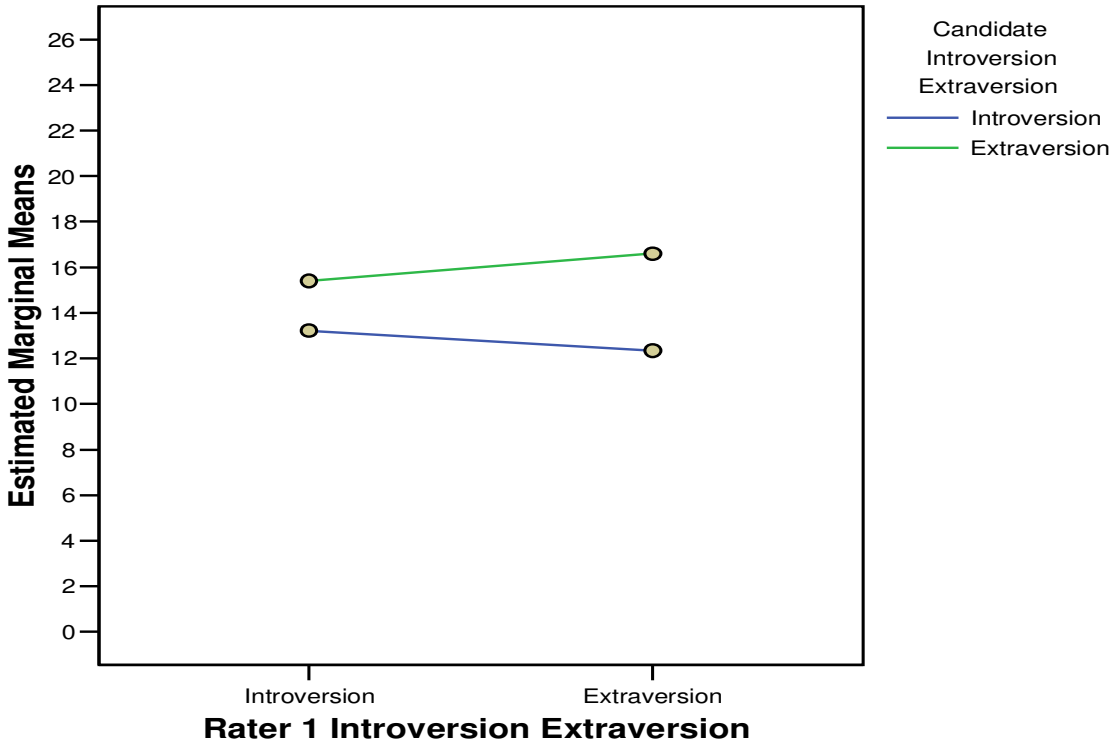
F	df1	df2	Sig.
3.27	3	51	.028

Table 5.15: Tests of Between-Subject Effects for Winter Survival

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	137.07(a)	3	45.69	2.66	.058	.135
Intercept	9484.38	1	9484.38	551.65	.000	.915
Candidate I/E	118.44	1	118.44	6.89	.011	.119
Rater I/E	.30	1	.30	.02	.896	.000
Candidate I/E * Rater I/E	12.34	1	12.34	.72	.401	.014
Error	876.83	51	17.19			
Total	13506.18	55				
Corrected Total	1013.90	54				

(a) R Squared = .135 (Adjusted R Squared = .084)

Figure 5.11: Profile Plots (ANOVA) for Winter Survival



For exercise Letra Course the mean scores of extraverted candidates as being rated by extraverted raters (M = 15.99, SD = 3.39) appear to be higher than those of introverted candidates (M = 15.49, SD = 3.33). (Refer Table 5.16).

In order to test the statistical significance of these differences, a two-way ANOVA is used. Levene's test of equality of error variances is not significant, indicating no violation of this assumption (refer Table 5.17). Between-subjects effects are therefore tested at the 5% significance level. There is a statistically significant main effect for candidate introversion/extraversion [$F_{(1,55)}$ = 9.86, p=0.003]; according to Cohen (1988) this represents a large effect size (η_{p2} = 0.152) (refer Table 5.18). The main effect for rater introversion/extraversion [$F_{(1,55)}$ = 0.01, p = 0.909] and the interaction effect [$F_{(1,55)}$ = 0.13, p = 0.725] does not reach statistical significance (refer Table 5.18).
It can therefore be concluded that introverted and extraverted candidates achieve

different scores during the second exercise as well, with extraverts obtaining higher scores on average than introverts (see Figure 5.12). However, there is no interaction effect between rater preference for introversion/ extraversion and the scores obtained by introverted/extraverted candidates. In other words, raters do not show a bias when rating candidates who share their preference for either introversion or extraversion during this exercise. Hypothesis 2 can therefore be rejected for both Winter Survival and Letra Course.

Table 5.16: Descriptive Statistics for Letra Course

Candidate I/E	Rater I/E	Mean	Std. Deviation	N
Introversion	Introversion	12.54	4.23	13
	Extraversion	12.29	3.79	6
	Total	12.46	3.99	19
Extraversion	Introversion	15.49	3.33	23
	Extraversion	15.99	3.39	17
	Total	15.70	3.32	40
Total	Introversion	14.43	3.89	36
	Extraversion	15.02	3.79	23
	Total	14.66	3.83	59

Table 5.17: Levene's Test of Homogeneity of Variances for Letra Course

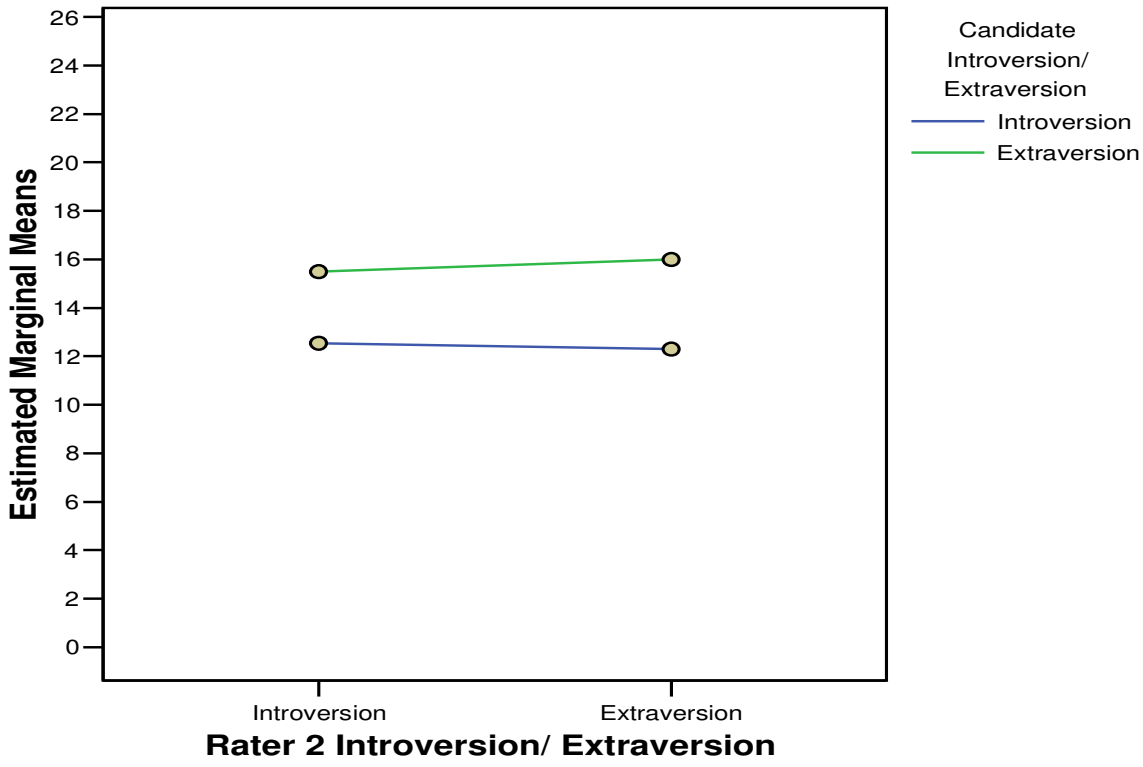
F	df1	df2	Sig.
.17	3	55	.918

Table 5.18: Tests of Between-Subject Effects for Letra Course

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	138.14(a)	3	46.05	3.55	.020	.162
Intercept	9165.58	1	9165.59	705.74	.000	.928
Candidate I/E	127.99	1	127.99	9.86	.003	.152
Rater I/E	.171	1	.171	.01	.909	.000
Candidate I/E*Rater I/E	1.63	1	1.63	.13	.725	.002
Error	714.30	55	12.987			
Total	13529.33	59				
Corrected Total	852.43	58				

(a) R Squared = .162 (Adjusted R Squared = .116)

Figure 5.12: Profile Plots (ANOVA) for Winter Survival



5.8 CONCLUSION

This chapter presents the results obtained during this study. It begins by providing a detailed description of the sample group in terms of gender, race, language, age as well as how many introverted and extraverted participants take part in the study. The distribution of scores obtained during both exercises is also discussed. An explanation is the proffered regarding how the means of the introverted candidates compare with the means of the extraverted candidates. The relationship between rater introversion/extraversion and scores allocated during the two behaviour observation exercises is also explored. The chapter also considers the allocation of raters to candidates. The results on the interaction effect between rater introversion/ extraversion and candidate introversion/ extraversion conclude the chapter.

Two hypotheses (as specified in Chapter 4) are tested during this study in order to explore whether or not introversion/ extraversion bias does exist among raters. Chapter 6 will provide a summary of the findings and a detailed discussion on the implication thereof.

CHAPTER 6: FINDINGS AND RECOMMENDATIONS

6.1 INTRODUCTION

The following chapter will summarise the results obtained during this study, after which a detailed discussion on the main findings will follow. The researcher will also highlight the limitations that could have had an influence on the results. This chapter will conclude with recommendations for future studies.

6.2 SUMMARY OF RESULTS

There is no significant correlation between candidate introversion/extraversion and the scores obtained during Winter Survival. In other words, scores obtained during this exercise do not significantly increase/decrease as the JPQ scores of the candidates increase/decrease. This finding implies that one group of candidates (introverts/extraverts) do not score significantly better than the other group during the exercise. The results of the t-test support this statement as the mean scores that the introverted candidates obtain are not significantly higher than the mean scores that the extraverted candidates obtain and vice versa for Winter Survival.

A noteworthy positive correlation between candidate introversion/extraversion and the scores obtained during Letra Course does exist though. This implies that as the introversion/extraversion scores increase, so do the scores obtained during exercise Letra Course, which indicates that extraverted candidates score better than introverted candidates. Again, this statement is supported with the results from the t-test, which compares the mean scores of the introverted candidates with the mean scores of the extraverted candidates for exercise Letra Course. The mean scores for the extraverted candidates are higher than the mean scores for the introverted candidates.

It must be mentioned that an outlier has a statistically significant effect on the results for both exercises. When a re-run is performed on the correlations (excluding the outlier, n = 65), it proves that extraverted candidates scored higher than the introverted candidates in both Winter Survival and Letra Course. A re-run of the t-tests indicates that the mean scores of the extraverted candidates are not significantly higher than the

mean scores of the introverted candidates for Winter survival. However, it is approaching significance. The re-run of the t-test for exercise Letra Course still indicates that the mean scores for extraverted candidates are higher than the mean scores for introverted candidates, which means that extraverts performed better during this exercise.

It can therefore be stated that some support for Hypothesis 1a exists, since there is a relationship between candidate introversion/extraversion and scores allocated during behaviour observation exercises.

No statistically significant relationship is found between rater introversion/extraversion and scores obtained during both behaviour observation exercises. In other words, raters' evaluations of candidates are not related to raters being either introverted or extraverted. Hypothesis 1b can therefore be accepted.

Cross-tabulations are performed in order to ensure that raters and candidates are paired together proportionally. The results show that for both exercises, introverted/extraverted raters are not disproportionately paired with introverted/extraverted candidates, which allows the researcher to make valid conclusions regarding the interaction effect.

The results from the two-way ANOVA indicate that for both Winter Survival and Letra Course there is no significant interaction effect. In other words, introverted raters did not rate introverted candidates more favourably than they rate extraverted candidates, and extraverted raters did not rate extraverted candidates more favourably than they rate introverted candidates. The outlier is automatically removed and has no impact on the findings, because the ambiverted candidates are excluded from the sample for the ANOVA (as explained in chapter 4). The unequal distribution of variances across the groups for Winter Survival is also considered by making use of Levene's Test for Unequal Variances.

This implies that there is no introversion/extraversion bias present during these two behaviour observation exercises.

The ANOVA does however present interesting additional findings. Results indicate that candidate introversion/extraversion has a significant impact on the scores for both exercises. For exercise Winter Survival, candidate introversion/extraversion has a moderate effect on scores and in exercise Letra Course candidate introversion/extraversion has a significant effect on scores. This again supports the results obtained by the Point-biserial Correlations and the t-tests.

The results can therefore be summarized as follows (refer Table 6.1 and Table 6.2):

- There is some support for Hypothesis 1a as most of the results indicate that extraverts perform better than introverts in Winter Survival and Letra Course.
- There is sufficient support for Hypothesis 1b as all of the results indicate that rater introversion/extraversion does not have any effect on scores for Winter Survival and Letra Course.
- There is no support for Hypothesis 2 as no significant interaction effect is found for either of the exercises.

Table 6.1: High Level Summary of Results With Outlier (n = 66)

Exercise	Correlations	T-test	Two-way ANOVA
Winter Survival	No significant relationship between candidate I/E and scores of Winter survival.	No significant difference between mean scores of introverted and extraverted candidates.	No significant interaction effect. However, candidate I/E did have a moderate effect on the scores.
Letra Course	Significant positive relationship between candidate I/E and scores of Letra Course.	Mean scores of extraverted candidates are significantly higher than introverted candidates.	No significant interaction effect. However, candidate I/E did have a large effect on the scores.

Table 6.2: High Level Summary of Results excluding Outlier (n = 65)

	Correlations	T-test	Two-way ANOVA
Exercise Winter Survival	Significant positive relationship between candidate I/E and scores of Winter Survival.	Approaching Significance in difference between introverted and extraverted mean scores.	Outlier was excluded in original output (the candidate was an ambivert).
Exercise Letra Course	Significant, positive relationship between candidate I/E and behaviour observation scores.	Significant difference between mean scores for extraverted and introverted candidates.	Outlier was excluded in original output (the candidate was an ambivert).

6.3 DISCUSSION

Before the impact of the main findings of this study can be discussed, it is first necessary to review the problem statement again.

The motivation to conduct this study results from the suspicion that assessment center raters in this security environment may have a preference for candidates with similar personality types to their own, which influences the final scores of these candidates during leadership assessments.

Given the lack of research on this phenomenon and the need for evidence, it is appropriate to determine to what extent introversion/extraversion bias influences the scores of behaviour observation exercises. This is done by testing whether or not there is a significant interaction effect between rater introversion/extraversion and candidate introversion/extraversion during two behaviour observation exercises called Winter Survival and Letra Course.

The first significant finding of this study indicates that there is no significant interaction effect between rater introversion/extraversion and candidate introversion/extraversion for both exercises. In other words, it can be stated that introversion/extraversion bias

has no effect on the scores of behaviour observation exercises during this assessment center.

One of the reasons for not observing any introversion/extraversion bias may be the small sample size. Although the statistics presented by the ANOVA prove not to be significant, more participants in the study may have revealed different results. According to Field (2009:360), unequal sample sizes in the ANOVA technique might lead to the F-ratio to be more conservative or lenient. That is, the F-ratio will be either more likely to produce an insignificant result when a real difference does exist, or it will produce a significant result when there is no difference between the groups.

The second finding central to this study indicates that rater introversion/extraversion does not have any significant effect on the scores of both behaviour observation exercises. This finding contradicts the results of previous studies showing that rater personality (especially introversion/ extraversion) does influence leadership assessment scores (Hautala, 2005:84; Yun & McFarland, 2005:97). This finding may result from the training that raters receive prior to the assessment center which is focused on the process of rating behaviour. However, the current study does not imply that rater personality in broad has no impact, as it only considers a small part of personality called introversion/extraversion.

The third and last finding derived from this study is probably the most remarkable. There is sufficient evidence indicating that candidate introversion/extraversion has a significant impact on scores for Winter Survival and Letra Course. More specifically, extraverted candidates tend to score better in both exercises.

This is not an uncommon phenomenon as some previous studies also found that extraverts are rated higher on communication and other attributes related to job performance (McCormack & Mellor, 2002:193; Minbashian, et al. 2009:573; Opt and Loffredo, 2003:566). However, to date, no substantial evidence has been found which indicates that extraverts are better leaders (Beaty, et al. 2001:147; Organ & Ryan, 1995:775). It is then appropriate to explore some reasons for the extraverted candidates performing better in this study.

A logical explanation might be derived from the work of Rust (1999:100) and Moon, et al. (2008:143). They indicate that individuals who prefer extraversion are usually more comfortable with teamwork and usually show a higher level of organisational citizenship. Although most of these attributes can indicate job performance, it is not necessarily a fact that they predict leadership ability. Nevertheless, the perception is present that extraverts may be better leaders which may cause raters to score them higher than introverted candidates.

The nature of the two exercises used in this study therefore contributes largely to extraverted candidates obtaining higher scores. During both Winter Survival and Letra Course, candidates participate in team formation and are rated on competencies such as Communication, Adaptability, Judgment, Decisiveness and Team Leading. All of these competencies require visible behaviour and candidates are therefore expected to interact and communicate with each other throughout the duration of the exercise (especially exercise Letra Course which requires physical participation). If no visible behaviour is observed, the candidate does not get any marks.

According to Bono and Judge (2004), Du Toit (1983:4), Hautala (2005:89) and Myers, et al. (2009:26), extraverts tend to be outgoing and talkative. They also typically enjoy being active and meeting new people. These characteristics would typically lead to increased scores during the exercises used in this study. Introverts, on the other hand, are likely to become tired when having to function in big groups and are usually perceived as reserved or hesitant (Bono & Judge, 2004:902; Du Toit, 1983:12; Hautala, 2005:89; Myers, et al. 2009:26). This in turn can lead to decreased scores because the nature of the exercises favours characteristics typical to extraverts and the opposite of how introverts typically interact.

It is therefore logical to assume that the nature of the exercises can cause raters in general to be positively biased towards extraverted candidates (irrespective of their own personality type). The next section will focus on some of the implications that the findings contain for the security environment in which it is conducted.

6.4 SIGNIFICANCE OF FINDINGS

The results discussed above have practical and theoretical implications. Firstly, they contribute to the relatively small amount of existing knowledge and theory around this topic. Hautala (2005:92) clearly states that there is a need to investigate the effect of rater personality (and especially similarity bias) on performance ratings; a topic central to this study. It also contributes to a relatively new field of study, called introversion/extraversion bias in assessment centers.

Secondly, the results have practical implications for the way in which assessment centers are conducted within this specific security environment. Although no interaction effect is found between rater introversion/extraversion and candidate introversion/extraversion, the tendency of raters in this environment to be positively biased towards extraverted candidates should have enough of an impact to motivate revisitation of the purpose and nature of the behaviour observation exercises.

The decision to include behaviour observation exercises in the selection processes should be carefully considered. If behaviour observation exercises are included, the exercises should be reconstructed to prevent any type of bias towards introverted and extraverted candidates. Over/under representation of candidates on selection boards should also be considered in the context of introversion /extraversion. Effective training on extraversion bias will benefit both candidates and raters.

6.5 LIMITATIONS TO THE STUDY

With any research study there are inherent attributes that can influence the validity of the results obtained. The following limitations were noted during the current study: the small sample size and poor representation of some race and language groups may have impacted the results on the interaction effect between rater introversion/extraversion and candidate introversion/extraversion. This limitation also contributes to the researcher's inability to make generalisations from this study that are representative of the bigger population.

Also, the fact that each candidate is only rated by one rater may have lead to a heavy cognitive load on inexperienced raters (Kolk, et al. 2002:271). This in turn may have caused raters to focus even more on direct and visible behaviour in order to render the rating experience less taxing. It may have been more beneficial to make use of multiple raters, which would also make it possible to test the inter-rater reliability.

As mentioned before, the nature of the competencies measured during both behaviour observation exercises require the candidate to demonstrate visible behaviour. This may have lead to the perception that introverted candidates are performing poorly.

The JPQ is used owing to its accessibility and the financial constraints that would have resulted from using a more updated instrument like the MBTI. Two aspects of the JPQ are regarded as limitations in this study.

Firstly, the JPQ does not measure extraversion as a trait but rather as a type. It does not, therefore, indicate how strong the introverted/extraverted behaviour is demonstrated but only how strong the individual's preference is for introverted/ extraverted behaviour. It would have been useful to compare the individuals who display strong extraverted behaviour with individuals who display strong introverted behaviour.

Secondly, the reliability coefficient for the African language-speaking candidates is very low. This may be due to the fact that the JPQ was developed before African language-speaking candidates were allowed to participate in selection processes of this specific security organisation.

6.6 RECOMMENDATIONS FOR FURTHER RESEARCH

The recommendations for future research are three-fold. Firstly, it is highly recommended that future researchers who are interested in conducting studies on this topic should make use of a bigger sample group.

Secondly, researchers are advised to make use of a more updated personality questionnaire that is standardized for all of the race groups in the population. An example would be a personality test like the MBTI. A personality test that will provide even greater value would be the 16 Personality Factor Questionnaire (16 PF) which

measures introversion/ extraversion as a trait and not a type. Both these instruments will allow the researcher to test for similarity bias based on other areas of personality. The MBTI and the 16 PF are briefly discussed in Chapter 2.

Lastly, it would be highly beneficial to develop a behaviour observation exercise that does not put any personality type in a disadvantaged position.

6.7 CONCLUSION

This study is based on the dynamic nature of the work environment, which puts organisations under increased pressure to appoint leaders that will lead their employees through highly competitive and ever-changing circumstances. However, the process of appointing leaders who are fit for the job seems easier than it actually is. Various aspects, for example 'rater bias', can influence such a process.

This study aims to determine whether or not introversion/extraversion bias has any effect on the scores of candidates participating in a leadership selection process. Even though the researcher does not find any evidence of such a bias, the results are significant and will impact on future selections.

Considering the general findings of this study, it can therefore be concluded that as long as humans are part of a selection processes, their findings will always be subjectively influenced. It is, however, the responsibility of professionals to try their best to eliminate anything that may influence the under-/over-representation of candidates on the selection board. Not only does the success of the organisation, and the profession of Human Resource Management and Industrial Psychology depend on it, but also the futures of the candidates participating in the process.

LIST OF REFERENCES

Antonioni, D. & Park, J. (2001). The effects of personality similarity on peer ratings of contextual work behaviors. *Personnel Psychology*, 54: 331-360.

Barrick, M.R. & Mount, M.K. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance,* 18: 359–372.

Beaty, J.C. Jr., Cleveland, J.N. & Murphy, K. (2001). The relationship between personality and contextual performance in "strong" versus "weak" situations. *Human Performance*, 14: 125-148.

Bechger, T.M., Maris, G. & Hsiao, Y.P. (2010). Detecting halo effects in performance-based examinations. *Applied Psychology Measurement*, 34: 607-619.

Berr, S.A., Chruch, A.H. & Waclawski, J. (2000). The right relationship is everything: Linking personality preferences to managerial behaviors. *Human Resource Development Quarterly*, 11(2): 133-157.

Bono, J.E. & Judge, T.A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology*, 89(5): 901-910.

Breakwell, G.M. (2004). *Doing social psychology research*. Oxford: Blackwell Publishing Ltd.

Bretz, R.D., Milkovich, G.T. & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management*, 18: 321-352.

Brummel, B.J., Rupp, D. & Spain, S.M. (2009). Constructing parallel simulation exercises for assessment centers and other forms of behavioural assessment. *Personnel Psychology*, 62: 137-170.

Brutus, S., Fleenor, J.W. & McCauley, C.D. (1999). Demograpic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, 18(5): 417-435.

Statistics South Africa. (2001). Census *2001: Primary tables South Africa, Statistics South Africa*. Available from: http://www.statssa.gov.za/census01/html/default.asp (Accessed 15 May 2011).

Chen, H-C. & Naquin, S.S. (2006). An integrative model of competency development, training design, assessment center, and multi-rater assessment. *Advances in Developing Human Resources*, 8(2): 265-282.

Cohen, G. (2009). Defining leadership: questions are now the answer. *Leadership Excellence*, August: 16-17.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates.

Coleman, D.G. & Adams, R.C. (1999). Establishing construct validity and reliability for the NAESP professional development inventory: Simplifying assessment center techniques. *Journal of Personnel Evaluation in Education*, 13(1): 27-45.

Culp, G. & Smith, A. (2005). Leadership effectiveness and behavior. *Leadership and Management in Engineering*, 39(5): 39-48.

De Vos, A.S. (Editor)., Strydom, H., Fouché, C.B. & Delport, C.S.L. (2002). *Research at grass roots: For the social sciences and human service professions*. 2nd edition. Pretoria: Van Schaik Publishers.

Dean, M. A., Roth, P. L. & Bobko, P. (2008). Ethnic and gender subgroup differences in assessment center ratings: A meta-analysis. *Journal of Applied Psychology*, 93(3): 685-691.

Denissen, J.J.A., Geenen, R., Van Aken, M.A.G., Gosling, S.D. & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment*, 90(2): 152-157.

Dilchert, S. & Ones, D. S. (2009). Assessment center dimensions: Individual differences correlates and meta-analytic incremental validity. *International Journal of Selection and Assessment*, 17(3): 254-270.

Dolliver, R.H. (1994). Classifying the personality theories and personalities of Adler, Freud, and Jung with Introversion/Extraversion. *Individual Psychology*, 50(2): 192-202.

Du Toit, L.B.H. (1983). *Manual for the Jung Personality Questionnaire (JPQ)*. Pretoria: Human Sciences Research Council.

Eaton, L.G., Collings, R.D., D'Agati, C.R., Moore, E.M. & Kellar, M.C. (2007). Extraversion and ego-resiliency: Predicting electroencephalograph measured positive emotion. *Individual Differences Research,* 5(3): 175-193.

Eidson, C.E. Jr. & Gurman, E.B. (1998). Similarity bias and ratings of applicants: Fact or artifact? *Journal of Business and Psychology*, 13(1): 81-84.

Eysenck, H.J. (1970). *Readings in Extraversion – Introversion 1, theoretical and methodological issues*. Tinling & Co Ltd: London.

Field, A. (2009). *Discovering statistics using SPSS*. 3rd edition. London: Sage Publications.

Foxcroft, C. & Roodt, G. (2005). *An introduction to psychological assessment in the South African context*. Cape Town: Oxford University Press.

Furnham, A., Taylor, J. & Chamorro-Premuzic, T. (2008). Personality and Intelligence Correlates of Asessment Center Exercises. *Individual Differences Research*, 6(3): 181-192.

Gal, R. & Mangelsdorff, A.D. (1991). *Handbook of military psychology.* Chichester: John Wiley & Sons, Ltd.

Gowing, M.K., Morris, D.M., Adler, S. & Gold, M. (2008). The next generation of leadership assessments: Some case studies. *Public Personnel Management*, 37(4): 435-455.

Gravetter, F.J. & Forzano, L.B. (2009). *Research methods for the behavioral sciences*. 3rd edition. Belmont CA: Wadsworth Cengage Learning.

Grobler, A.P., Wärnich, S., Carrell, M.R., Elbert, N.F. & Hatfield, R.D. (2006). *Human resource management in South Africa*. 3rd edition. London:Thompson Learning.

Guion, R.M. & Gottier, R.F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18(2): 135-164.

Hautala, T. (2005). The effects of subordinates' personality on appraisals of transformational leadership. *Journal of Leadership & Organisational Studies*, 11(4): 84-92.

Heffernan, T.M. & Ling, J. (2001). The impact of Eysenck's extraversion-introversion personality dimension on prospective memory. *Scandinavian Journal of Psychology*, 42: 321-325.

Hermelin, E., Lievens, F. & Robertson, T. (2007). The validity of assessment centers for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment*, 15: 405-411.

Hetland, H. & Sandal, G.M. (2003). Transformational leadership in Norway: Outcomes and personality correlates. *European Journal of Work and Organisational Psychology*, 12(2): 147-170.

Hilliard, T. & Macan, T. (2009). Can mock interviewers' personalities influence their personality ratings of applicants? *The Journal of Psychology*. 143(2): 161-174.

Hogan, J. & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socio-analytic perspective. *Journal of Applied Psychology,* 88: 100–112.

Horey, J.D. & Fallesen, J.J. (2004). Leadership competencies for contemporary army operations: Development, review and validation. Available from: http://www.internationalmta.org/Documents/2004/2004045P.pdf (Accessed 30-03-2011).

Howard, A. (1997). A reassessment of assessment centers: Challenges for the 21st century. *Journal of Social Behavior and Personality*, 12(5): 13-52.

Jackson, D.J., Atkins, S.G., Fletcher, R.B. & Stillman, J.A. (2005). Frame of reference training for assessment centers: Effects on interrater reliability when rating behaviors and traits. *Public Personnel Management*, 34: 17-25.

Joiner, D.A. (2000). Guidelines and ethical considerations for assessment center operations, International task force on assessment center guidelines. *Public Personnel Management*, 29(3):315-331.

Judge, T.A. & Erez, A. (2007). Interaction and intersection: The constellation of emotional stability and extraversion in predicting performance. *Personnel Psychology*, 60: 573-596.

Killgore, W.D.S., Richards, J.M., Killgore, D.B., Kamimori, G.H. & Balkin, T.J. (2007). The trait of Introversion-Extraversion predicts vulnerability to sleep deprivation. *European Sleep Research Society*, 16: 354-363.

Kolk, N.J., Van der Flier, H., Born, M.P. & Olman, J.M. (2002). Assessment center procedures: Cognitive load during the observation phase. *International Journal of Selection and Assessment*, 10(4): 271-278.

Lange, R. & Rentfrow, J. (2007). Color and Personality: Strong's interest inventory and Cattell's 16 PF. *North American Journal of Psychology*, 9(3): 423-438.

Le Roux, E. (2004). *The validity of the assessment centre in predicting managerial performance of business development managers*. Masters dissertation. Pretoria: University of Pretoria.

Lievens, F. (2009). Assessment Centers: A tale about dimensions, exercises and dancing bears. *European Journal of Work & Organisational Psychology*, 18(1): 102-121.

Lievens, F. (1998). Factors which improve the construct validity of assessment centers: A review. *International Journal of Selection and Assessment*, 6(3): 141-152.

Lievens, F. & Goemaere, H. (1999). A different look at assessment centers: Views of assessment center users. *International Journal of Selection and Assessment*, 7(4): 215-219.

Lievens, F. & Van Keer, E. (2001). The construct validity of a Belgian assessment centre: A comparison of different models. *Journal of Occupational and Organisational Psychology*, 74: 373-378.

Lievens, F., Van Keer, E. & Volckaert, E. (2010). Gathering behavioural samples through a computerized and standardised assessment center exercise: Yes, it is possible. *Journal of Personnel Psychology*, 9(22):94-98.

London, M., Mone, E.M. & Scott, J.C. (2004). Performance management and assessment: Methods for improved rater accuracy and employee goal setting. *Human Resource Management*, 43(4): 319-336.

Love, K.G. & DeArmond, S. (2007). The validity of assessment center ratings and 16PF personality trait scores in police sergeant promotions: A case of incremental validity. *Public Personnel Management*, 36(1): 21-32.

Lowry, P.E. (1996). A survey of the assessment center process in the public sector. *Public Personnel Management*, 24(3): 307-322.

Maldé, B. (2006). Do AC's really care about the candidate? *British Journal of Guidance & Counseling*, 34(4): 539-549.

Matthews, G., Deary, I.J. & Whiteman, M.C. (2003). *Personality Traits,* 2nd edition. Cambridge: Cambridge University Press.

Maurer, S.D. (2002). A practitioner-based analysis of interviewer job expertise and scale format as contextual factors in situational interviews. *Personnel Psychology*, 55: 307-327.

May, G.L. & Gueldenzoph, L.E. (2006). The impact of social style on peer evaluation ratings in project teams. *Journal of Business Communication*, 43: 4-20.

May, G.L. (2008). The effect of rater training on reducing social style bias in peer evaluation. *Business Communication Quarterly*, 71(3): 297-313.

McCormack, L. & Mellor, D. (2002). The role of personality in leadership: An application of the Five-Factor Model in the Australian Military. *Military Psychology,* 14(3): 179-197.

McDougall, W. (1929). The chemical theory of temperament applied to introversion and extraversion. *Journal of Abnormal and Social Psychology*, 24: 293-309.

McNiel, J.M., Lowman, J.C. & Fleeson, W. (2010). The affect of state extraversion on four types of affect. *European Journal of Personality*, 24: 18-35.

Melchers, K.G., Lienhardt, N., Von Aarburg, M. & Kleinmann, M. (2011). Is more structure really better? A comparison of frame-of-reference training and descriptively anchored rating scales to improve interviewers' rating quality. *Personnel Psychology*, 64: 53-87.

Merrill, D.W. & Reid, R.H. (1999). *Personal styles and effective performance*. New York: CRC Press.

Meyer, W.F., Moore, C. & Viljoen, H.G. (1997). *Personologie, van individu tot ekosisteem*. Johannesburg: Heinemann.

Myers, I.B., McCaulley, M.H., Quenk, N.L. & Hammer, A.L. (2009). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator Instrument*. 3rd edition. California: CPP.

Minbashian, A., Bright, J. E. H. & Bird, K. D. (2990). Complexity in the relationships among the subdimensions of extraversion and job performance in managerial occupations. *Journal of Occupational and Organizational Psychology*, 82: 537-549.

Momberg, C. (2005). *The relationship between personality traits and vocational interests in a South African context*. Masters dissertation. Pretoria: University of Pretoria.

Moon, H., Hollenbeck, J. R., Marinova, S. & Humphrey, E. (2008). Beneath the surface: Uncovering the relationship between extraversion and organizational citizenship behavior through a facet approach. *International Journal of Selection and Assessment*, 16(2): 143-154.

Mulyanegara, R.C., Tsarenko, Y. & Anderson, A. (2009). The Big Five and brand personality: Investigating the impact of consumer personality on preferences towards particular brand personality. *Brand Management*, 16(4): 234-247.

Murnieks, C.Y., Haynie, M.J., Wiltbank, R.E. & Harting, T. (2007). *I like how you think: The role of cognitive similarity as a decision bias*. Paper presented at the Academy of Management: Boulder, Colorado

Opt, S.K. & Loffredo, D.A. (2003). Communicator image and Myers-Briggs Type Indicator extraversion-introversion. *The Journal of Psychology*, 137(6): 560-568.

Oviedo-García, M.A. (2007). Internal validation of a biodata extraversion scale for salespeople. *Social Behavior and Personality*, 35(5): 675-692.

Pittenger, D.J. (2005). Cautionary comments regarding the Myers-Briggs Type Indicator. *Consulting Psychology Journal: Practice and Research*, 57(3): 210-221.

Ployheart, R.E., Lim, B. & Chan, K. (2001). Exploring relations between typical and maximum performance ratings and the Five Factor Model of personality. *Personnel Psychology*, 54(4): 809-843.

Roush, P.E. & Atwater, L. (1992). Using the MBTI to understand transformational leadership and self-perception accuracy. *Military Psychology*, 4(1): 17-34.

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

Rust, J. (1999). Discriminant validity of the 'Big Five' personality traits in employment settings. *Social Behavior and Personality*, 27(1): 99-108.

Schlebusch, S. & Roodt, G. (2008). *Assessment centres, unlocking potential for growth*. Randburg: Knowres Publishing (Pty) Ltd.

Schyns, B. & Felfe, J. (2006). The personality of followers and its effect on the perception of leadership – an overview, a study, and a research agenda. *Small Group Research*, 37(5): 522-539.

Scroggins, W.A., Thomas, S.L. & Morris, J.A. (2008). Psychological testing in personnel selection, part II: The refinement of methods and standards in employee selection. *Public Personnel Management*, 37(2): 185-198.

Sharp, D. (1987). *Personality Types: Jung's model of Typology*. Toronto: Inner City Books.

Smith, C.A., Organ, D.W. & Near, J.P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68: 653-663.

Stamoulis, D. (2009). *Senior executive assessment: A key to responsible corporate governance*. Chichester: Wiley-Blackwell.

Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8: 220-247.

Taylor, I. 2009. *The Assessment and Selection Handbook. Tools, techniques and exercises for effective recruitment and development*. London: Kogan Page.

Thornton, G.C. III (1992). *Assessment centers in human resource management*. Massachusetts: Addison-Wesley Publishing Company.

Thornton, G.C.III & Rupp, D.E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development*. New Jersey: Lawrence Erlbaum.

Trope, Y. & Higgins, E.T. (1993). The what, when and how of dispositional inference: New answers and new questions. *Personality and Social Psychology Bulletin*, 19: 493-500.

Van Iddekinge, C.H., Ferris, G.R. & Heffner, T.S. (2009). Test of a multistage model of distal and proximal antecedents of leader performance. *Personnel Psychology*, 62: 463-495.

Warr, P. & Hoare, S. (2002). Personality, gender, age and logical overlap in multi-source ratings. *International Journal of Selection and Assessment*, 10(4): 279-291.

Werth, L., Markel, P. & Förster, J. (2006). The role of subjective theories for leadership evaluation. *European Journal of Work and Organisational Psychology*, 15(1): 102-127.

Westgaard, O. (1999). *Test That Work: Designing & Delivering Fair & Practical Measurement Tools in the Workplace*. San Francisco: Jossey-Bass/Pfeiffer.

Woodruffe, C. (1993). *Assessment Centres, identifying and developing competence*. 2nd edition. London: Institute of Personnel and Development.

Yun, G.J., Donahue, L.M., Dudley, N.M. & McFarland, L.A. (2005). Rater personality, rating format, and social context: Implications for performance appraisal ratings. *International Journal of Selection and Assessment*, 13(2): 97-107.

APPENDIX A

JPQ QUESTIONS REFERRING TO INTROVERSION AND EXTRAVERSION

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

JUNG PERSONALITY QUESTIONNAIRE

Please choose the option that applies to you the most and indicate your preference on the given answer sheet.

1. You have to work
   A. in a room in which people often talk
   B. alone in an office where you can concentrate on the work
   C. Uncertain

2. You have to work under conditions where you
   A. will work with a few well-known people and seldom have to talk to strangers.
   B. will constantly meet new people and have to try to remember their names.
   C. Uncertain

3. You often have to
   A. talk to other people and mix with them socially.
   B. listen to what others say about the work.
   C. Uncertain

4. You often have to
   A. investigate methods and make suggestions for making the work proceed more smoothly.
   B. arrange new social functions and be responsible for their smooth running.
   C. Uncertain

5. You have to reply to enquiries
   A. over the telephone.
   B. in writing.

C. Uncertain

6. You have to

A. associate with only a few well-known people.

B. associate with a large variety of people.

C. Uncertain

7. You have to

A. frequently attend parties where many people are present.

B. only rarely attend meetings with a few trusted colleagues.

C. Uncertain

8. You must carry out a task in which you will

A. only occasionally discuss the work with people.

B. often discuss the work with people.

C. Uncertain

9. In your work

A. there is constant variation and change.

B. you have to carry out only one task at a time and concentrate on it.

C. Uncertain

10. You have to work in and office in which the colors of the walls, curtains and carpets are

A. quiet and restful.

B. warm and bright.

C. Uncertain

11. You have to complete a task where you are expected to

    A. be cheerful every day.

    B. be serious at all times.

    C. Uncertain

12. You have to work under conditions where

    A. you rarely hear bits of news and gossip

    B. it is expected that you will always be informed of the latest news and even gossip.

    C. Uncertain

13. You have to work in an environment where other people will

    A. expect you to tell them what you are interested in.

    B. allow you to quietly go your own way.

    C. Uncertain

14. You have to

    A. sometimes attend meetings where you are not required to mix socially with others.

    B. attend a series of social functions where you will always be expected to be sociable.

    C. Uncertain

15. You have to accomplish a task where

    A. enthusiasm is required.

    B. accuracy is required.

    C. Uncertain

16. You have to

A. work for long periods on end with only a few well-known people.

B. meet and work with new people virtually every day.

C. Uncertain

17. A particular work requires that you

A. have to talk with many people on a wide variety of topics

B. consult only a few trusted colleagues.

C. Uncertain

18. You have to

A. attend social functions, but to remain in the background.

B. take the lead in social functions.

C. Uncertain

19. You have to listen to a speaker and

A. give a verbal reply to his arguments.

B. give a considered, written reply to his arguments.

C. Uncertain

20. A particular task requires that you

A. inform others of your views and feelings only occasionally when it is actually necessary.

B. frequently inform others of your views and feelings.

C. Uncertain

APPENDIX B

SIX-POINT LIKERT RATING SCALE USED DURING WINTER SURVIVAL AND LETRA COURSE

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Winter Survival	Letra Course

Surname and Initials or Candidate No: ……

N/E	1	2	3	4	5
No Evidence	**Marginal**	**Less than Acceptable**	**Acceptable**	**More than Acceptable**	**Superior**
	Performed well below expectation as per the behavioural indicators	**Performed below expectation as per the behavioural indicators**	**Meets performance expectation as per the behavioural indicators**	**Performed beyond expectation as per the behavioural indicators**	**Performed well beyond expectation as per the behavioural indicators**

COMMUNICATION The ability to communicate clearly and fluently in English					
Negative behavioural indicators:			**Positive behavioural indicators:**		
Audibility					
• Harsh voice, speaks softly, cannot be heard easily			• Clear understandable tone of voice, speaks loudly enough to be heard easily, does not shout		
N/E	**1**	**2**	**3**	**4**	**5**
Modulation/accent/speed					
• Words are pronounced incorrectly • Speaks too fast/too slowly			• Good articulation/expression, correct pronunciation of words • Comfortable pace		
N/E	**1**	**2**	**3**	**4**	**5**
Language usage					
• Uses slang, jargon or generally inappropriate language specific to the context			• Clear, understandable use of language / speaks fluently in English		
N/E	**1**	**2**	**3**	**4**	**5**
Confident Communication					
• Finds it difficult to express ideas verbally • Verbal communication is difficult to follow • When giving inputs, it is with little confidence			• Communicates persuasively and confidently • Expresses ideas with ease		
N/E	**1**	**2**	**3**	**4**	**5**
Dominance/Assertiveness					
• Tends to over-communicate (talks too much) • Tends to interrupt other speakers • Too submissive • No contribution • Is steamrolled by the group			• Allows time for others to make their point or comments • Able to share point of view • Assertive		
N/E	**1**	**2**	**3**	**4**	**5**
Eye Contact					
• Does not make eye contact or use body language effectively			• Makes continuous eye contact and uses body language effectively		
N/E	**1**	**2**	**3**	**4**	**5**
Total:		**/30**			

ADAPTABILITY The potential to adapt to the situation and accept changes, opposing views and new ideas					
Negative behavioural indicators:			**Positive behavioural indicators:**		
Openness to Ideas					
• Resists adjusting behaviour, ideas and habits • Just agrees with the rest of the group without evaluating inputs			• Open to the ideas of others • Accepts the viewpoints of other people (not just agreeing)		
N/E	**1**	**2**	**3**	**4**	**5**
Acceptance of Change					
• Views change critically without considering merits • Avoids change, attempts to maintain the status quo			• Willing to accept change • Sees change positively and as an opportunity		
N/E	**1**	**2**	**3**	**4**	**5**
Adapts plan					
• Resists adapting plan based on new information, stubborn, wants to stick to previous/original plan			• Adapts plan based on relevant new information		
N/E	**1**	**2**	**3**	**4**	**5**
Emotion Management					
• Gets frustrated by setbacks			• Does not get frustrated by setbacks or problems		
N/E	**1**	**2**	**3**	**4**	**5**
Total:		**/20**			

JUDGEMENT The ability to evaluate information and carefully consider a wide range of possibilities before making a final decision. The ability to arrive at sound conclusions and to display logical reasoning in the solution of problems					
Negative behavioural indicators:			**Positive behavioural indicators:**		
Identifies relevant information					
• Does not contribute to discussion • Ignores relevant details and facts • Introduces irrelevant information			• Contributes ideas and opinions • Asks others for their ideas and opinions		
N/E	**1**	**2**	**3**	**4**	**5**
Evaluates relevant information/possible decisions					
• Bases decisions on too little information • Decision making tends to be impulsive • Fails to consider the impact and consequences of a decision • Unable to motivate decisions			• Evaluates all relevant information before making a decision • Considers more than one course of action • Uses systematic approach to make judgments • Judges other's decisions objectively • Weighs up the options before making a decision • Evaluates impact and consequences of decisions • Makes proposals which impact functioning positively		
N/E	**1**	**2**	**3**	**4**	**5**

Arrives to a sound conclusion					
• Decisions made tend to be inappropriate • Bases decisions on irrelevant facts • Decisions are not well thought through • Tends to jump to conclusions • Bases decisions on rash assumptions			• Bases decisions on relevant facts • Reaches a sound conclusion • Demonstrates logical reasoning • Avoids making rash assumptions • Makes sound decisions which positively impact on area of functioning		
N/E	**1**	**2**	**3**	**4**	**5**
Total:		**/15**			

DECISIVENESS Show readiness and ability to make decisions, render judgment, take action and apply corrective measures whenever the situation demands it					
Negative behavioural indicators:			**Positive behavioural indicators:**		
Convince others of decision/point of view					
• Do not explore alternatives • Unable to convince others of decision			• Explores alternatives and possible consequences • Convince others of decisions		
N/E	**1**	**2**	**3**	**4**	**5**
Decision making process					
• Unable to make a decision within allocated time • When giving inputs, it is with little confidence • Tends to resist making a final decision			• Able to make decisions within allocated time • Stands by and supports decision confidently • Takes action in making a decision (verbally reaching the point where the final decision will be made)		
N/E	**1**	**2**	**3**	**4**	**5**
Emotionally neutral when making a decision					
• May become emotionally involved to the extent that the quality of the decision is compromised			• Remains emotionally neutral when making a decision		
N/E	**1**	**2**	**3**	**4**	**5**
Total:		**/15**			

TEAM LEADING The ability to delegate, give feedback and direction to team members, cooperate with team members, express positive attitude and expectations of team members, encourage others, ask for input from team members and take corrective measures					
Negative behavioural indicators:			**Positive behavioural indicators:**		
Leadership role					
• Is a follower and only agrees with the rest of the group			• Takes the lead in the team/group discussion • Gives direction to the team members		
N/E	**1**	**2**	**3**	**4**	**5**
Value of inputs					
• Inputs do not improve the procedures and techniques of the group • Doesn't add value to the group/team			• Fives inputs to improve methods, procedures and techniques • Adds value to the group/team		
N/E	**1**	**2**	**3**	**4**	**5**

Attitude towards team members					
• Displays a negative attitude towards other team members and/or their inputs			• Expresses a positive attitude and clarifies expectations of team members		
N/E	**1**	**2**	**3**	**4**	**5**
Facilitator role					
• Ignores the others and talks them over			• Encourage others to participate and asks for inputs from team members		
N/E	**1**	**2**	**3**	**4**	**5**
Interpersonal relations					
• Is not part of the team (does not relate well to the other team members) • Is not comfortable working with diverse team members (of different race and gender)			• Is part of the team (relates well with others) • Is comfortable working in a diverse team • Sees everyone as part of the team		
N/E	**1**	**2**	**3**	**4**	**5**
Total:		**/15**			

NOTES

APPENDIX C

CONSENT FORM USED DURING THE STUDY

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

RESEARCH PARTICIPANT CONSENT FORM

The study you are about to participate in is to explore the effect of Introversion/Extraversion biases on scores allocated during behaviour observation exercises in an assessment center.

Please read through the following statements and sign at the bottom of the page if you agree.

- I understand that I will participate in a study conducted on the effect of Introversion/Extraversion biases on leadership assessment which requires that I will have to complete a questionnaire which indicates my preference for Introversion or Extraversion.
- I understand that there will be no costs involved and that my participation in this questionnaire will last more or less 120 minutes.
- I understand that I have the right to withdraw from this study at any point in time.
- I understand that the results of this questionnaire will not have any impact on my assessment scores.
- I understand that the results of this questionnaire will not be used for any selection purposes.
- I understand that there are no risks involved in this study. In the case of potential risks, safeguards will be conducted by the researcher to minimize them as far as possible.
- I understand that my results (output) of the questionnaire will be used for data-processing and report writing which will be published by the researcher at a later point in time. In such case, my identity will remain strictly anonymous.
- I understand that if I have any questions, I may contact Este de Beer (083 7200018) at any time.

If you are sure that you understand this consent form please indicate your participation by signing below

Name of participant: ……

Signature of participant: …… Date: ……

Signature of researcher: …… Date: ……

APPENDIX D

INDEPENDENT SAMPLE T-TESTS FOR WINTER SURVIVAL AND LETRA COURSE (N = 65 & N = 66)

INDEPENDENT SAMPLE T-TESTS FOR WINTER SURVIVAL AND LETRA COURSE (N = 66)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Ex 1: Winter Survival	**Equal variances assumed**	6.323	.014	-1.807	64	.075	-2.16780	1.19956	-4.56419	.22858
	Equal variances not assumed			-1.563	29.757	.129	-2.16780	1.38660	-5.00058	.66498
Ex 2: Letra Course	**Equal variances assumed**	.593	.444	-2.715	64	.009	-2.54583	.93756	-4.41883	-.67284
	Equal variances not assumed			-2.532	35.231	.016	-2.54583	1.00554	-4.58671	-.50496

INDEPENDENT SAMPLE T-TESTS FOR WINTER SURVIVAL AND LETRA COURSE (N = 65)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Ex 1: Winter Survival	**Equal variances assumed**	9.653	.003	-2.174	63	.033	-2.46742	1.13484	-4.73522	-.19963
	Equal variances not assumed			-1.817	27.595	.080	-2.46742	1.35794	-5.25088	.31603
Ex 2: Letra Course	**Equal variances assumed**	.494	.485	-2.718	63	.008	-2.57607	.94763	-4.46977	-.68238
	Equal variances not assumed			-2.548	35.792	.015	-2.57607	1.01098	-4.62684	-.52531