EVALUATING A PRE-SCREENING QUESTIONNAIRE'S PREDICTIVE VALIDITY AT A LARGE FINANCIAL INSTITUTION

REYGANA ADAMS

ADMREY003

A dissertation submitted in partial fulfilment of the requirements for the award of the Degree of Master of Commerce in Organisational Psychology

Faculty of Commerce

University of Cape Town

2008

COMPULSORY DECLARATION:

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, and has been cited and referenced.

Signed by candidate

Signature:.....

Date:.....06/04/2009.....

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Prof. Burger van Lill for his guidance and advice on completing this dissertation. I believe I have gained new knowledge by working with you and that this knowledge will assist me in my career in future - thank you.

This dissertation would not have been possible had the financial institution not been so willing to assist me on this project. I want to thank the managers and liaison person at the financial institution for providing me with access to the data and the pre-screening questionnaire. I sincerely hope that you will find the outcomes of this dissertation useful.

To my husband, Zawaad Shade, I wish to thank you for your relentless support and encouragement throughout this year. The completion of my studies is an achievement for us both. I want to thank my parents, Aamien and Fatima Adams for always being there whenever I needed you. It is only through your efforts and prayers that I have managed to succeed in my studies. To the rest of my family, thank you for supporting me in this endeavour.

Also, I wish to say thank my work colleagues from Diagnostic Surveys as well as Organisational Development at the Department of the Premier (Provincial Government of the Western Cape) for providing me with a supportive work environment and for believing in me. I am eternally grateful.

ABSTRACT

This study investigated whether a pre-screening questionnaire was able to predict job performance at a large financial institution. The pre-screening questionnaire comprised a biographical, integrity and interest subscale as predictors of job performance. A sample of 2145 job applicants completed the pre-screening questionnaire and these scores were used to establish the internal consistency of the pre-screening questionnaire. For the validity testing, a smaller sample of 449 job applicants was drawn to compare their test scores of the pre-screening questionnaire to their performance scores obtained, as the criterion. Each subscale was evaluated in terms of its predictive validity. The loglinear analysis revealed that the biographical subscale was a valid predictor of job performance. Despite being reliable i.e. integrity subscale with Kuder-Richardson coefficient of (r =.63) and interest subscale with an overall Cronbach Alpha of (r =.98), both subscales were found to be poor predictors of job performance. The findings suggest that the pre-screening questionnaire demonstrates predictive validity and utility in the selection process to some extent. Recommendations are listed to improve the overall reliability and validity of the pre-screening questionnaire.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

TABLES

FIGURES

CHAPTER 1: INTRODUCTION

The business context in South Africa has undergone major transformation in recent years. The interacting effect of an emerging democracy as well as globalisation has resulted in organisations having to transform in a short period of time (Denton & Vloeberghs, 2003). South African organisations have had to rise to the challenge of remaining viable in the local markets while simultaneously needing to compete in international markets (Denton & Vloeberghs, 2003). Due to the freedom of trade across borders, organisations have been able to make a wider variety of products and services available globally. According to Berner and Van Tonder (2003), this has resulted in a more demanding and consumption-driven society. In order to satisfy rising consumer demands, South African organisations will have to increase their competitiveness and organisational productivity (Sutherland, De Bruin & Crous, 2007).

Job performance is central to achieving organisational productivity and the challenge facing organisations is how to select the most talented and productive candidates more effectively (Sackett & Lievens, 2008). Psychological tests and assessment centres have been used in selection procedures to aid decision-making (Van der Merwe, 2002). The different psychological tests provide an objective manner to determine if job applicants possess the specific competencies needed to drive the business of the organisation (Boyatzis, 2008). A wide range of psychological tests have been used in selection processes such as ability (cognitive) tests, personality inventories, polygraph testing, learning batteries (Aiken, 1994) as well as biographical and integrity questionnaires (Boam & Sparrow, 1992; Ones & Viswesvaran, 2007).

In South Africa, the Employment Equity Act, No.55 of 1998 states that "psychometric testing and other similar assessments of an employee is prohibited unless the test or assessment being used has been scientifically shown to be valid and reliable; can be applied fairly to employees and is not

biased against any employee or group" (p. 7). In order to comply with this legislative requirement, South African organisations have begun to review the psychological tests used in their recruitment and selection processes by establishing its reliability and validity. In doing so, organisations are ensuring that any psychological tests are not unfairly discriminating between candidates in selection, where any differences noted in candidates' scores would be indicative of their level of competence and ability rather than being due to bias in the psychological test or assessment itself.

Given the above, the present study will investigate the psychometric properties and predictive validity of an assessment tool used in the selection process at a large financial institution. The assessment tool is known as a pre-screening questionnaire, which was developed internally by the financial institution. The pre-screening questionnaire is a self-report measure that is divided into three sections, namely a biographical subscale, an integrity subscale and an interest subscale. Job applicants are expected to complete the pre-screening questionnaire on-line prior to being interviewed by the financial institution. Essentially, the pre-screening questionnaire is used to filter through job applicants so as to identify prospective employees who demonstrate the potential and competencies to perform well in jobs offered at the financial institution. As a result, the pre-screening questionnaire is discriminating between job applicants in the selection process and is subject to the legal requirement that it must be fair, valid and reliable as an assessment tool.

Therefore, the aim of this study is to:

i. Investigate the psychometric properties of the pre-screening questionnaire; and
ii. Determine if the pre-screening questionnaire is a suitable measure for predicting job performance in a financial sector.

The introduction outlines the rationale for this study, where South African organisations must ensure that all psychological tests used in the selection

process are fair, valid and reliable. Chapter 2 (Literature Review) of this dissertation provides a definition of job performance and psychological assessment as important constructs in the study. It also provides an overview of the different psychological tests currently used in selection processes and discusses which of these have been found to be good predictors of job performance in the past. In chapter 3 (Method), the statistical techniques, secondary data, criterion variables and predictors used in this study are described. Subsequently, chapter 4 (Results) presents the data analysis, where the reliability and validity statistics are reported for the biographical, integrity and interest subscale respectively. In chapter 5 (Discussion), the findings for each subscale are interrogated and compared to previous studies in order to draw conclusions about the pre-screening questionnaire's reliability and validity. The chapter is concluded with a list of limitations and recommendations for future studies. In addition, the financial institution is advised on how to practically improve the reliability and validity of the pre-screening questionnaire.

CHAPTER 2: LITERATURE REVIEW

This chapter provides a definition for job performance and discusses the role that psychological assessment plays in predicting job performance in organisations. An overview is provided of the different psychological assessments used in selection and entails a debate on each one's reliability and validity. Specifically, the predictive validity of biographical, integrity and interest measures is reported based on previous studies.

Job Performance

Before job performance can be measured, evaluated and improved, it has to be clearly defined and understood within the organisational context. The concept of job performance will be explored in more detail before considering how it can be assessed in organisations.

Organisations have always desired outstanding employee performance (Sutherland et al., 2007) because it enhances overall competitiveness. By definition, job performance describes individual behaviours that are related to the organisational goals (Nelson & Quick, 2006). More importantly, each job in the organisation has certain outputs that contribute to a larger purpose. Employees are expected to perform their jobs to a prescribed standard or level of productivity that would require certain skills and behaviours to deliver excellent performance (Nelson & Quick, 2006). It should be noted that the required skills and behaviours will be different for each job and will vary from industry to industry. Job performance is therefore a multi-dimensional construct (Campbell, McCloy, Oppler & Sager, as cited by Sutherland et al., 2007) that can be predicted and measured in different ways.

Job performance has two distinct components, namely task and contextual performance. Sutherland et al. (2007) states that *task performance* defines job-related activities that an employee is expected to perform e.g. switchboard operators are hired to perform job tasks such as answering and re-routing

telephonic enquiries. *Contextual performance* refers to activities that employees are not expected to perform but are necessary for the achievement of organisational goals (Allworth & Hesketh, 1999). These activities are not outlined explicitly in employees' job descriptions, but seem to enhance quality of job performance and include behaviours such as helpfulness, conscientiousness, commitment and altruism (Allworth & Hesketh, 1999). In other words, a switchboard operator may re-route a telephonic enquiry (job task) and do this activity in a helpful and friendly manner (contextual performance). Sutherland et al. (2007) are of the opinion that organisations are focusing more on contextual performance as a basis for selecting successful and productive employees, which has the potential to provide consumers with a better quality product or service.

Considering the literature on job performance, it appears that task performance is more readily understood in organisations than contextual performance (Allworth & Hesketh, 1999; Sutherland et al., 2007). This is because task performance is easier to measure and predict, as it is quantified in terms of tasks and outputs (Nelson & Quick, 2006). Conversely, contextual performance is more complex to measure because it involves inherent qualities, which employees exhibit on the job (Allworth & Hesketh, 1999). Again, the multi-dimensional nature of job performance is highlighted, posing a challenge in its measurement and prediction. To overcome this challenge, Sutherland et al. (2007) suggest that organisations used different assessment measures to predict job performance i.e. task performance should be predicted using cognitive assessments, while contextual performance is better predicted using dispositional or personality measures.

Given that organisational performance depends largely on employee job performance, gaining the competitive edge as an organisation means that employees will have to provide outstanding performance (Meihem, as cited by Sutherland et al., 2007; Kleynhans, 2006). Organisations must be able to predict

future job performance of candidates in order to select them appropriately (Theron, 2007).

Two aspects become important to consider in the discussion of job performance. The first aspect relates to competencies as a driver to maximise performance in organisations (Boyatzis, 2008). Competencies are sets of behaviour patterns that an individual must display in order to perform their job effectively and efficiently (Saunders, 2005). In other words, it is an ability or capability needed in that job. Competence refers to a blend of knowledge, skills, behaviour and aptitude that a person applies in the work environment in order to perform a job to an acceptable standard (Saunders, 2005). Competence is reflected when a person has acquired a level of proficiency or mastery of a particular competency. Boyatzis (2008) advocates for the use of competencies as important predictors of future job performance, stating that when organisations select competent candidates, it is likely that they will perform well on the job in future. Therefore, it is important that selection processes include assessments tools that can test for company-specific competencies that are linked to job performance in an organisation. Depending on the industry in which the company operates in, certain competencies might be viewed as more important than others in driving organisational performance. Ashton's (1996) case study demonstrated that a hotel group was able to improve organisational performance by using a competency model that focussed on seven core competencies, which specifically related to the hospitality industry. Based on this example, it appears that generic competencies may be less useful in predicting job performance as opposed to using company-specific competencies as predictors of job performance in an organisation.

The second aspect relating to job performance is the concept of person-job fit. The selection process is not a one-sided exercise that only involves the organisation's choice about who to recruit as employees. According to Cascio (2003), people are actively involved in the recruitment process and seek jobs and organisations that will fit their personal belief system and values.

Organisations have the opportunity to articulate its corporate culture and values by streamlining their recruitment and selection process so that it can focus on specific competencies which the organisation values (Armstrong, 2003). When there is congruence between organisational culture and employees' personal values, employees are found to be more productive, more satisfied and tend to act in the interest of the company, irrespective of job duties (Cascio, 2003). Thus, it can be inferred from the above discussion that competencies can play a role in predicting job performance and assessing person-job fit in the selection process.

Job performance remains an important indicator of organisational success and therefore, organisations have to ensure that it selects the right candidates with the necessary competencies to perform well on the job. Psychological tests and assessments can be useful devices in selection processes because it can assess core competencies and the person-job fit for each candidate. Therefore, the kinds of psychological tests or assessments used in selection are of utmost importance when making predictions of future job performance.

Psychological Testing and Assessment

The most prevalent methods of predicting job performance are personality tests, achievement tests, aptitude tests, cognitive (ability) tests, interest questionnaires, biographical questionnaires, role-plays and integrity tests (Armstrong, 2003). If a psychometric instrument is well constructed, then the instrument will be a good predictor of job performance (Lindsay, as cited by Barnard & Schaap, 2005). As per the Employment Equity Act, 1998, a well-constructed psychometric instrument should be valid, reliable and fair. To appreciate this statement more fully, the terms *psychological test* and its related concept of *assessment* is defined below.

Psychological tests are useful because they provide organisations, researchers and psychologists with an objective way of studying human behaviour. A psychological test involves a clear specification about the construct to be

measured, a standardised process, which test takers must follow in order to respond to test materials so that it can be evaluated and scored appropriately (Huysamen, 2002). More importantly, the outcomes of psychological tests allow for inferences and conclusions to be drawn based on a sample of human behaviour (Anastasi & Urbina, 1997). Psychological assessment is defined as a process-orientated activity where a wide range of information is gathered, evaluated and integrated into a conclusion or decision (Foxcroft & Roodt, 2005). Psychological assessment is a broader concept than psychological testing because it evaluates human behaviour using more than one method or tool. According to Foxcroft and Roodt (2005), a psychological test is a single tool used to assess human behaviour. In an assessment, a combination of tests such as cognitive and ability tests, intelligence testing, interviews, personal history inventories, personality tests, integrity tests, aptitude and interest questionnaires are used to observe and analyse human behaviour (Anastasi & Urbina, 1997; Van der Merwe, 1999). Psychological assessments differ from psychological tests in that it provides a more comprehensive evaluation of human behaviour. Psychological assessment may include psychological testing as one method of evaluation but is not necessarily limited to psychological testing only (Paterson & Uys, 2005). Thus, this paper will refer to psychological assessment as an overarching term, which will include psychological testing.

An assessment centre is a very popular method of obtaining information on candidates applying for managerial and supervisory jobs (Gerber, Nel & Van Dyk, 1999). It involves a series of assessment methods grouped together to test certain competencies that relate to the job (Saunders, 2005). Over and above administering psychological tests, assessment centres can also include exercises such as leaderless group discussions, problem solving exercises, in-basket exercises, case studies and role-plays (Cascio, 1991). This is a highly valid method for assessing candidates due to the fact that assessment centre exercises are designed to identify individual human behaviour that relate strongly to job demands, as opposed to merely measuring behavioural attributes (Cascio, 2003; Saunders, 2005).

Psychological assessments have aided organisations with various decision-making processes including recruitment and selection, training, transfers and promotions (Van der Merwe, 2002). The process of selection was originally based on the sink-or-swim strategy used in the industrial era, where all job-seekers who apply for a job were selected and admitted into positions, but only those who performed would be retained (Aiken, 1994). In the beginning of the twentieth century, a more classical approach to selection was developed known as the "classic trio" (Cook, as cited by Taylor, 2008, p. 253). The classic trio follows three steps; first candidates expect to fill in an application form to which they attach their curriculum vitae, secondly they have to attend an interview and thirdly they would receive an offer of employment subject to reference checking. The classic trio has remained a dominant approach to selection for most of the twentieth century, but it provided organisations with little indication of job candidates' potential to perform on the job (Taylor, 2008). Today, it is evident that modern organisations will not benefit from such selection strategies because it is lengthy and inefficient. While machines and speed of production provided the competitive advantage in the past, people or "human capital" have become the strategic asset through which companies gain the competitive advantage in the new knowledge economy (Chowdury, 2000). Modern organisations need recruit and select the right people at the right time in order to add value to the business and the bottom line (Ulrich & Brockbank, 2005). Therefore, the current business context requires a selection process that will identify high performers more quickly and accurately. Psychological tests and assessments can provide this capability in selection processes. In the opinion of Van der Merwe (2002), psychological assessment remains the most time-efficient and inexpensive method of predicting behaviour and job performance in organisations.

Psychometric Properties of Psychological Tests

It is essential that psychological measures are both reliable and valid (Aiken, 1992). Collectively, reliability and validity are known as the psychometric properties of a psychological measure that indicates how trustworthy the

measure is in predicting human behaviour (Spatz & Kardis, 2008). According to the Employment Equity Act, 1998 the use of psychological testing and assessments are prohibited unless they are scientifically proven to be valid and reliable; that they can be applied fairly to all employees and that they are not biased against any employee or race group. This legislation further states that employers may only fairly discriminate between candidates based on the inherent requirements of the job (competencies). Through the Affirmative Action policy and section 15 of the Employment Equity Act, 1998 employers must ensure that qualified designated groups (black people, women and people with disabilities) have equal opportunities to get a job. In order to practice fair discrimination, organisations are encouraged to select candidates from designated groups who demonstrate the necessary competencies required for the job. Therefore, organisations must be able to prove that their selection decisions were based on non-discriminatory criteria and that a fair selection process was observed (Muchinsky, Kriek & Schreuder, 2005). Given that psychological assessments are used to inform hiring decisions in organisations, it is imperative that psychological assessments used in the selection process are valid and reliable in predicting job performance.

Reliability

Reliability refers to how consistently a psychological assessment is measuring the same construct by producing similar scores for the same individuals on repeated attempts (Foxcroft & Roodt, 2005). The degree of reliability of a psychological test is indicated by a specific reliability coefficient that indicates the percentage of the true variance shared between the test scores and the trait or construct being measured (Kerlinger & Lee, 2000). The square root of the reliability coefficient represents the proportion of common variance obtained and is known as the index of reliability (Anastasi & Urbina, 1997).

Statistically, reliability is defined as the proportion of error variance to the total variance yielded by a measuring instrument subtracted from 1.00, where the index of 1.00 indicates perfect reliability (Kerlinger & Lee, 2000). Like any score,

the obtained variance of a test contains two components: a "true" component that the psychological test should measure ideally and an "error" component measured by the psychological test that should be minimised as far as possible (Grobler, Warnich, Carrell, Elbert & Hatfield, 2006). Error variance corresponds to random fluctuations on respondents' performance from one test session to another (Anastasi & Urbina, 1997). The more error variance included in the reliability coefficient, the less reliable the psychological test becomes and vice versa (Kerlinger & Lee, 2000). Thus, the error variance distorts the reliability of the psychological test, making it less precise in its measurement.

There are various methods to establish a psychological test's reliability and each method has a specific reliability coefficient. When using the test-retest method, the psychological test is repeated with the same respondents under similar conditions (Hair, Babin, Money & Samouel, 2003). The scores obtained on each administration of the psychological test are correlated to yield the *coefficient of stability*, which shows the extent to which scores on a test can be generalised over different occasions (Anastasi & Urbina, 1997). The coefficient of stability is not a preferred method of establishing reliability because time sampling can be a main source of error variance (Anastasi & Urbina, 1997). This means that the first time respondents complete the psychological test may influence their responses on the second time they take it or that situational factors may affect the way respondents' motivation and affect the way they answer the test items on the second attempt (Hair et al., 2003).

Another method of establishing reliability is the alternate-form, where the same respondents are measured at two different times using parallel (not identical) forms of the psychological test. The responses obtained from both versions of the psychological test are correlated to determine the coefficient of equivalence (Hair et al., 2003). While time sampling may not be a source of error variance, the coefficient of equivalence is faced with error variance due to content sampling (Anastasi & Urbina, 1997). Alternate forms of the test must be truly parallel i.e. they must contain the same number of items expressed in the same

format and it must cover the same content. The level of difficulty, test instructions, formats, time limits and illustrative examples must be equivalent (Anastasi & Urbina, 1997). Given the above, using the coefficient of equivalence to establish reliability is not a common approach because of the practical challenges of constructing truly equivalent forms of psychological tests (Anastasi & Urbina, 1997).

In the split-half method of reliability, the psychological test is only administered once and is split into two equivalent halves (Kerlinger & Lee, 2000). Each half of the psychological test is scored and the two sets of scores are correlated to obtain the coefficient of internal consistency. This coefficient measures the extent to which certain items contained in the test itself influenced the test scores (Anastasi & Urbina, 1997). In other words, the internal consistency can be improved or reduced depending on which items were included or excluded from the test, altering the content of the psychological test. The challenge with the split-half method is that there are several possible ways to split the items and such a particular split could lead to error variance in content sampling (Kerlinger & Lee, 2000). The Kuder-Richardson formula 20 and Cronbach's coeffient alpha were developed based on the split-half method (Foxcroft & Roodt, 2005), where a psychological test has been split into every possible combination. Essentially, each item is correlated with every other item in the test and the mean of those correlations is found, resulting in either a Kuder-Richardson 20 or coefficient alpha. Where a psychological test has dichotomous items scored as 0 or 1 or true-and-false, the Kuder-Richardson 20 formula will apply as the reliability coefficient (Kerlinger & Lee, 2000). Coefficient alpha can also be used for dichotomous response scales, but is more applicable to assessment measures that have multiple-scored items such as personality or attitude questionnaires (Foxcroft & Roodt, 2005). Thus, these coefficients are more useful in establishing reliability because they minimise the error variance in terms of content sampling and content heterogeneity (Anastasi & Urbina, 1997).

Validity

Validity is the second psychometric property that each assessment measure should demonstrate. A psychological test is considered to be valid if it measures what it claims to measure (Kline, 2000; Lopes, Roodt & Mauer, 2001). Validity is concerned with the relationship between test performance and other independently observable characteristic or criterion (Anastasi & Urbina, 1997). The concept of validity is controversial in the behavioural sciences because it inquires about the nature and meaning of variables or construct being measured (Kerlinger & Lee, 2000). When constructs are directly measureable such as *weight* or *height*, it is simpler to establish validity. However, constructs such as *intelligence* or *personality* is measured indirectly using psychological tests and thus, it becomes more difficult to establish validity for these constructs. Indirect methods to test for psychological constructs are usually so indirect that the validity of the measurement instrument and its products is doubtful (Kerlinger & Lee, 2000).

Validity is also an evolving concept that cannot be reported in general terms (Anastasi & Urbina, 1997). Instead of a test's validity being either high or low, validity is established in relation to the particular use of the test being measured (Anastasi & Urbina, 1997). Therefore, there are different kinds of validity that can be established, but establishing validity must be congruent with the nature of the test being evaluated i.e. verifying the appropriateness or relevance of the content in an aptitude test (content validity) or determining whether a personality test is a valid predictor of future job performance (predictive validity).

Traditionally, validity has been classified into three types namely content, construct and criterion-related validity (Hair et al., 2003). Content validity refers to how well the items in the measure clearly relate to the characteristic or trait that the psychological test or assessment is measuring (Spatz & Kardis, 2008). Content validity helps to identify those items in the psychological test that have no obvious connection to what the test is measuring and is more often used in achievement and occupational measures (Anastasi & Urbina, 1997).

Construct validity indicates how well an assessment is measuring the theoretical concept that the test was designed to measure (Foxcroft & Roodt, 2005). Constructs are broad categories of features that are commonly shared by other variables, but are not themselves directly observable (Anastasi & Urbina, 1997). Examples of theoretical concepts to which construct validity can be applied are emotional intelligence, personality, integrity, and depression to mention a few.

Criterion-related validity refers to how effective an assessment is in predicting an individual's behaviour in a specific situation (Anastasi & Urbina, 1997). In other words, it is an investigation of how well a predictor such as a personality measure can predict an outcome or criterion such as job performance or conscientiousness. There are two kinds of criterion-related validity identified in the literature, namely concurrent and predictive validity. Concurrent validity is concerned with how well a measure can diagnose the existing or current status of a criterion (Foxcroft & Roodt, 2005). Muchinsky et al. (2005) provide an example of concurrent validity, where one wishes to predict a student's average exam marks based on a test score. This can be obtained by collecting data on the average exam marks for several students and then administering a predictor test for correlation. In this case, the concurrent validity of the predictor test would be available immediately, as one would have a data set of average marks as well as the student's scores on the predictor test. Thus, concurrent validity studies focuses more on what an individual is capable of doing at present. Predictive validity differs from concurrent validity in that it determines how well a measure would be able to predict an individual's future behaviour (Foxcroft & Roodt, 2005). In predictor validity studies, predictor information is collected and used to forecast future criterion performance (Muchinsky et al., 2005). An example of a prediction study would be a company administering a particular test to predict whether job applicants will complete a six-month training programme. By implication, the criterion for predictive validity will not be available and therefore, the criterion is usually estimated on the basis of some predictor (Muchinsky et al., 2005). The present study will be an example of a prediction study because a pre-screening questionnaire is used a predict job

applicants' job performance potential and suitability for a position at the financial institution.

Contrary to the traditional point of view, measurement experts in the behavioural sciences have advocated for an all-inclusive validity concept. Huysamen (2002) argues that all three validity types are interrelated and states that "all the above forms of validity are subsumed under construct validity" (p. 27). Construct validity is considered to be the most important kind of validity because it explains the properties of the construct being measured and links these properties to a theoretical notion (Kerlinger & Lee, 2000). For instance, an intelligence test would be investigated using construct validity to establish which factors (such as verbal ability, numeracy or abstract reasoning) account for the variance in test performance (Kerlinger & Lee, 2000). With construct validity, it is more important to ensure that a psychological test is actually measuring the underlying psychological construct than evaluating the content or tasks of the test. According to Huysamen (2002), a test should be taken as valid if it demonstrates construct validity because all other forms of validity provide evidence for the existence for construct validity. Content validity and prediction validation procedures ultimately contribute to the definition and understanding of constructs assessed by a test (Anastasi & Urbina, 1997).

While Huysamen (2002) is of the opinion that not all kinds of validity are of equal importance, it does not mean that criterion-related and content validity procedures are not useful. All validity types offer valuable information in their own right in the evaluation of certain tests (Anastasi & Urbina, 1997). This is especially relevant to criterion-related validity or predictive validity, which focuses on predicting a future outcome (criterion) rather than focussing on what is behind test performance (Kerlinger & Lee, 2000). The information supplied by prediction studies are used in the selection and hiring of job applicants, selecting students for admission to university, matching and placing of personnel into training programmes or to screen out applicants who will not fit into the organisational culture (Anastasi & Urbina, 1997). Therefore, predictive validity

has much to offer in improving a psychological test's accuracy in predicting a criterion such as job performance or stress.

Fairness and bias

South African organisations are faced with the on-going challenge of ensuring an efficient and equitable personnel selection in respect of a diverse applicant pool (Theron, 2007). The expectation of eliminating unfair discrimination in selection processes is grounded in the Employment Equity Act, 1998, which states that psychological assessments must be applied fairly to all employees and should not be biased against any employee or group. The concepts of bias and fairness are inter-related, but refer to different issues in psychological assessment.

A psychological test is biased when the items or mode of response in the instrument elicit significantly different responses from various groups being assessed (Foxcroft & Roodt, 2005). Bias relates to the content and construction of the psychological test itself, such as the level of item difficulty, different meanings of words, the language being used, administrative procedures applied, testing environment or instructions given etc. (Muchinsky et al., 2005). The issue of test bias is remedied by using indicators of bias such as differential item functioning (DIF), differential factor structures and differential item difficulties (Foxcroft & Roodt, 2005; Muchinsky et al., 2005).

While test bias is directly related to the statistical or psychometric construction of a psychological measure, fairness is a social concept that is based on a set of values and has different implications for how selection and other personnel decisions are made (Taylor, as cited by Muchinksy et al., 2005). Fairness implies a moral and ethical judgment in selection decisions; "a selection test is considered unfair when it differentially predicts job performance for majority and minority applicant groups" (Van Iddekinge, Eidson, Kudisch & Goldblatt, 2003, p. 150). Fairness in selection procedures using psychological tests refer mainly to how tests are being applied to make a choice of one candidate over another and

comprises all variables involved that could influence the final personnel decision (Muchinsky et al., 2005). Thus, it is possible for a psychological test to be unbiased but can be considered as unfair when it is applied inconsistently to make selection decisions. As far as possible, organisations should investigate the bias inherent in the specific psychological tests used in selection processes as well as establish how fair their selection processes are when hiring from a diverse pool of applicants.

Utility: Usefulness of psychological tests in selection

While research studies attempt to establish the reliability and validity of psychological assessments, this is little indication of how useful any of these are in selection processes (Aamodt, 1999). Therefore, utility of a selection device is defined as the degree to which its use improves the quality of the individuals selected beyond what would have resulted if the selection device were not used (Blum & Naylor, as cited in Cascio, 1991). Ideally, a selection tool that is able to discriminate between successful and unsuccessful job applicants in a cost-effective manner is considered as having utility or usefulness (Bernardin, 2003).

A selection tool's utility is established using the Taylor-Russell tables for which three pieces of information is required, namely the selection ratio, validity coefficient and the current base rate (Cascio, 2000). The selection ratio refers to how many people the organisation has to hire and is determined by dividing the number of openings by the number of applications received for a given position (Aamodt, 1999). The validity coefficient of a test indicates the degree to which a psychological test is measuring what it is meant to measure (Anastasi & Urbina, 1997). The base rate refers to the percentage of current employees who are considered successful (Aamodt, 1999). The concept of utility can be established and applied to different psychological tests using several methods such as the Taylor-Russell model with its specific tables, the Naylor-Shine approach or the Lawshe approach (Aamodt, 1999; Roth, Bobko & Mabon, 2001).

Predictors of Job Performance

Organisations have been concerned with predicting job performance for many years (Nelson & Quick, 2006). Various psychological assessments have been developed and used in selection to predict job performance. In the present study, biographical data; integrity testing and interest subscale of the pre-screening questionnaire will be explored to establish whether they are suitable predictors of job performance in the financial sector.

Biographical data

It is often said that the best predictor of future behaviour is past behaviour (Henry, 1966). Biographical data measures have been developed based on this principle. A biographical data measure is a detailed questionnaire that has a large number of items relating to job applicant's work and personal life (Taylor, 2008). For instance, biographical questions usually refer to age, gender, social status, qualifications, tenure, race group and years of driving experience (Buitendach, Oosthuyzen & van Wyk, 2005). Biographical data intends to elicit responses that provide insight into past behaviour that is associated with successful work performance as a criterion.

There is an on-going debate in the literature regarding the validity of biographical measures. For criterion-related validity, biographical measures to be useful in predicting global job success in insurance and credit companies with a validity coefficient of .4 (Boam & Sparrow, 1992). While Cascio (2003) found a validity coefficient of .35 to be a respectable and sufficient, Schmidt and Hunter's (1998) study concluded that .35 was a low validity coefficient for biographical data measures. Earlier studies have highlighted the fact that biographical data measures tend to lose its validity over time, especially where small samples are used (Hunter & Hunter, 1984). Thus, there appears to be no conclusive opinion about how valid biographical data measures are in practice, where its validity coefficient can range from .2 to .5 (Harold, McFarland & Weekley, 2006). Despite having a low validity coefficient, it should be noted that a psychological test could still be useful if the selection ratio is low (Cascio,

2003). In other words, if the company decided to become more selective in its hiring decision, then biographical measures with low validity coefficients would still be able to discriminate between successful and unsuccessful candidates (Cascio, 2003).

The disparity in validity scores may be associated with the prevalence of social desirability or faking in biographical measures. In the literature, biographical measures have been criticised for being susceptible to faking, where applicants or employees exaggerate and distort their responses so that they appear more favourably (Harold et al., 2006). Despite this, Cascio (2003) argues that biographical measures are less susceptible to faking than personality measures. Other studies have found that people will fake good depending on what was on offer and how much was at stake based on the nature of the items included in the biographical measure (Harold et al., 2006; Schmitt, Oswald, Kim, Gillespie, Ramsay & Yoo, 2003). This meant that individuals were not responding honestly to the items, compromising the validity of the biographical measure. One means of handling the possibility of faking is to check a candidate's previous history or track record with referees or previous employers (Cascio, 2003).

Integrity testing

In recent years, there has been considerable interest the whether or not integrity tests can predict job performance (see Arthur, 2006; Camara & Schneider, 1994; Wanek, 1999). An integrity test is meant to determine a job applicant's attitude towards theft and other counterproductive behaviours in the workplace (Bernardin, 2003). Organisations are particularly concerned with counterproductive behaviours such as theft, vandalism, fraud, corruption, etc (Wanek, 1999). It is believe that if employees engaged in such activities that they would be less focussed on their job performance and organisations have employed integrity testing to reduce the probability of counterproductive behaviours (Schmidt & Hunter, 1998). Integrity tests are usually paper and pencil self-report tests designed to be used on current employees (Wanek, 1999). Organisations can use overt or covert integrity tests. Overt integrity tests

inquire directly about employee's attitude towards theft and their admissions to engaging in counterproductive behaviours e.g. "What is the amount of money or merchandise you have taken from your current employer?" (Wanek, 1999). Covert integrity tests are more personality-related which assess traits such as dependability, social conformity, thrill-seeking, conscientiousness and trouble with authority e.g. "I am more sensible than adventurous" (Wanek, 1999).

There is a concern that integrity tests lacks construct validity as they do not evaluate integrity as an underlying construct *per se*, but seems to cover a broad range of counterproductive behaviours in its evaluation (Carrell et al., 1995). In addition, there is also disagreement about whether integrity is an individual trait or something that is situationally determined (Carrell et al., 1995). According to Kline (2000), integrity tests generally demonstrate very poor predictive validity, stating that honesty of employers are just as important as honesty of employees, yet employers are seldom subjected to integrity tests. Similar to biographical measures, responses in integrity testing can be faked, where an employee can be coached to answer the questions in a desirable way (Lilienfeld, Alliger & Mitchell, 1995). Given the concern regarding the validity of integrity testing, organisations are cautioned not to become over-reliant on its test results in determining job suitability (Arthur, 2006).

Despite these drawbacks, Ones, Viswesvaran and Schmidt's (1995) study established that integrity tests were good predictors of job performance in compared to ability tests, work sample tests and job knowledge tests used in a same selection procedure. Aiken (1994) supports this finding stating that integrity testing demonstrates useful levels of validity when used in combination with other selection procedures. While there is no overwhelming support in the literature for its uses as a predictor of job performance, the fact that few studies have found integrity testing to be a useful predictor cannot be overlooked (see Ones et al., 1995; Ones & Viswesvaran, 2007).

Interest questionnaires

The purpose of interest questionnaires is to determine an individual's preference for specific occupation or career interests (Foxcroft & Roodt, 2005). This type of assessment measure has been reported to be beneficial in career and counselling psychology because it is effective in guiding individuals in making career decisions (Anastasi & Urbina, 1997).

Having stated the above, there is a belief that interest questionnaires are good determinants of job performance because it is assumed that when people are interested in their jobs, they will perform better than those whose interests do not match their jobs. Schmidt and Hunter (1998) have contested this point on the premise that interest questionnaires tend to correlate poorly with the construct of job performance, with a validity coefficient of .1. Aiken (1994) mentions that a possible explanation for the poor predictive validity is that job success correlates better with ability tests than with interest questionnaires. This explanation may be true, where interest questionnaires can predict which jobs people will most likely enter (Carrel, Elbert & Hatfield, 1995). However, once individuals are employed, their level of job performance is largely determined by the mental ability and personality characteristics (Schmidt & Hunter, 1998). As a result, interest questionnaires may be useful at indicating how well a person will adjust to a new work environment and how long they will stay with the organisation (Carrel et al., 1995). It is clear from this discussion that using interest questionnaires is not the best means of predicting future job performance, since the literature remains equivocal on the matter.

Personality measures

One of the more popular methods of predicting job performance is to use personality measures. Personality is defined as a unique combination of cognitive and affective characteristics that are displayed in a consistent pattern over time (Aiken, 1992). Various theories exist for studying personality in industry such as trait, psychoanalytic, phenomenological, biological, behavioural and social learning theory (Forbes, 2006).

Considering the plethora of studies published, there seems to be overwhelming support for the use of personality measures to predict job performance in industry (see La Grange & Roodt, 2001; Rothman & Coetzee, 2003). Ones, Dilchert, Viswesvaran and Judge (2007) conducted a meta-analysis of the Big Five personality constructs and compared its predictability against job performance criteria and found strong correlations ranging from .1 to .45 for personality and job performance, where predictive validity of individual job performance was .23. The results of the meta-analysis revealed that of all five personality facets, *conscientiousness* was the single best predictor of job performance (Ones et al., 2007). Conversely, Rothman and Coetzee (2003) reported in their study that *conscientiousness* had a lower factor loading and did not predict job performance as well as *emotional stability, openness to experience* and *agreeableness.* There have also been arguments that personality is better at predicting task performance (Rothman & Coetzee, 2003), while Sutherland et al. (2007) claims that personality predicts contextual performance better. Irrespective, it is interesting to note that all the facets of personality correlate in some way to job performance in the literature. Based on the above, personality may be a more reliable and valid measure of predicting job performance in industry than other measures.

Ability (cognitive) tests

In ability testing, researchers are interested in determining the individual difference between people (Aiken, 1992). More specifically, ability or cognitive tests are interested in measuring intelligence, which is made up of a general ability (*g)* and special ability (*s*) (Foxcroft & Roodt, 2005). Depending on what is being evaluated, certain ability tests become more useful than others (Kline, 2000). For example, when measuring general occupational and educational success, then verbal and spatial tests are used. However, abilities that are occupation specific will be evaluated using perceptual speed and accuracy, numerical ability, inductive reasoning, ideational fluency and word fluency (Kline, 2000).

In recent studies, general ability has been found to be one of the most valid predictors of job performance across all job sectors (Schmidt & Hunter, 1998; Sutherland et al, 2007). In one study, general ability accounted for 60 percent of the total variance in predicting job performance (Ones, Viswesvaran & Schmidt, 2005). Similarly, other studies have found relatively strong correlations of .51 for general ability and job performance. Authors have suggested that the high correlation may be due to rapid acquisition of job knowledge or high job performance (Barnard & Schaap, 2005; Schmidt & Hunter, 1998). The underlying logic is that when individuals gain more job-related knowledge and expertise, they are able to perform better. General ability (*g*) taps into an individual's accumulated job knowledge and experience to indicate a job applicant's job performance in future. While ability testing is useful at ensuring that job applicants meet the minimum intelligence requirements for a job, high intelligence is itself not a valid predictor of success in the work situation (Gerber et al., 1999).

In terms of fairness, cognitive tests produce large discrepancies between ethnic groups and may not be culturally fair despite being one of the best predictors of job performance. The differences in scores on cognitive ability tests are mainly due to inferior educational experiences of certain race groups (Bernardin, 2003). One means of reducing the adverse impact and making cognitive testing more fair for certain race groups is to set a less stringent cut-off score, but then the utility of the ability test is compromised (Bernardin, 2003). This means that by setting a lower cut-off score, the organisation may hire people who are less minimally qualified for the job.

Summary

The ability to predict job performance will enable organisations to obtain and retain high performing employees, which will ultimately lead to improved organisational performance. However, determining how well an applicant or employee will perform in future may not be immediately evident. It is for this reason that several assessment methods have been developed to assist

organisation to better predict job performance. With a complex concept such as job performance, one psychological test may not be sufficient for prediction. Therefore, organisations have embarked on conducting psychological assessments as a more viable option to measure the competencies that bring about job performance, such as the pre-screening questionnaire being evaluated in the present study.

The pre-screening questionnaire comprises of three subscales. In the biographical subscale, job applicants have to provide details relating to their gender, race, language preference, place of residence, education, previous work experience, qualifications and marital status. Several studies have supported the use of biographical data measures to predict job performance (see Cascio, 2003; Harold, et al., 2006; Schmidt & Hunter, 1998). The most common method of establishing a biographical data measure's validity as a predictor is to use a representative sample of no less than 150 job applicants and correlating their biographical data with their performance scores as a criterion (Bernardin, 2003). As an industry specific consideration, biographical data measures are commonly used in the insurance industry to screen job applicants (Bernardin, 2003; Boam & Sparrow, 1992). Also, biographical data measures are culturally fairer than other psychological tests i.e. cognitive tests because it already takes job applicants' background factors into account (Van Iddekinge et al., 2003). This point is particularly relevant to note when developing a psychological test or assessment in a multi-cultural setting such as South Africa.

The pre-screening questionnaire's integrity subscale is an example of an overt integrity test that requires job applicants to admit and declare their past criminal and credit records, drivers' licence endorsement, history of insolvency and verification of professional registration. Integrity tests are intended to identify a job applicant's propensity to theft (Carrell et al., 1995). According to Wanek (1999), organisations may choose to apply integrity testing in its selection process if they have a serious problem with employee theft, dishonesty,

counterproductive behaviours or filling positions with restricted access to cash or merchandise and require a high degree of honesty and integrity. Thus, integrity tests are most commonly used in jobs where employees have access to money such as insurance companies, banks and retails stores (Bernardin, 2003). In previous studies, regression analysis was conducted to establish the predictive validity of integrity measures (see Ones & Viswesvaran, 2007). In terms of fairness, integrity tests have been found to have no adverse impact as an assessment measure in selection procedures (Bernardin, 2003).

In the interest subscale, job applicants are requested to rate their preference for particular work-related activities such as selling, managing, leading, interacting with people and marketing company products. The intention of interest subscale is to match job applicants' work-related interests to the some of the actual job tasks required by the position being advertised. If there is a match between work-related interests and job tasks, there is an increased likelihood that the job applicant will perform well on the job (Gerber et al., 1999). However, it has been stated that interest questionnaires are not good predictors of job performance (Carrell et al., 1995). In sharp contrast to ability tests, interest questionnaires do not validly predict job performance and seems better suited to occupational counselling and vocational guidance (Foxcroft & Roodt, 2005). Despite lacking utility in predicting job performance, the interest subscale was included in the pre-screening questionnaire as a method of establishing a profile of interests that successful employees would prefer.

Conclusion

This chapter presented the current state of knowledge relating to predicting job performance in industry. The concepts of job performance and psychological assessment have been clearly defined in order to identify the best method of assessing job performance. The literature highlighted several assessment methods that could be applied, where some methods are more reliable and valid than others. For instance, the most powerful predictor of job performance is ability or cognitive tests. Similarly, personality measures such as the 16

Personality Factors and the Occupational Personality Questionnaire have been considered as good predictors for both contextual and task performance. This discussion has built a foundation that will inform the current study on how best to establish the predictive validity of the pre-screening questionnaire.

Previous studies have been insightful regarding the capability of biographical questionnaires, interest questionnaires and integrity tests in predicting job performance. The unique challenge for this research is to generate empirical evidence on the ability of the biographical, integrity and interest subscales to measure competencies linked to job performance at the financial institution. Finally, it is hoped that the findings of this study will add value to the current body of knowledge of using psychological assessments to predict job performance in the South African context.

CHAPTER 3: METHOD

This chapter outlines the research process undertaken. It describes the research design, sample and secondary data chosen for this study. Further, details regarding the criterion measure and predictors are provided in an attempt to establish the pre-screening questionnaire's predictive validity.

This study followed a descriptive research design, which allowed the researcher to describe a situation or phenomena by testing certain hypotheses (Hair et al., 2003).

Participants

With reference to Table 1, a sample of 2145 job applicants applying for sales positions at the financial institution was used in the study. There were 762 female applicants and 1383 male applicants. The majority of the sample was South African citizens, of which most had obtained a Grade 12/ Matric qualification.

In terms of the racial distribution, the sample was fairly well represented. The predominant race group was African/Black with a smaller representation of the other race groups. While all South Africa's official languages were represented in the sample, it appeared that the majority of the sample was Afrikaans speaking, followed by English and Zulu language groups.

Table 1

Gender, Race and Language Distribution of the Sample

		FREQUENCY	%
GENDER DISTRIBUTION			
	Female	762	35.5
	Male	1383	64.5
Total		2145	100
RACE DISTRIBUTION			
	Asian	230	10.7
	Black	934	43.5
	White	778	36.3
	Coloured	203	9.5
Total		2145	100
LANGUAGE DISTRIBUTION			
	Afrikaans	632	29.5
	English	568	26.5
	Afrikaans and	11	0.5
	Xhosa	147	6.9
	Zulu	287	13.4
	Northern Sotho	75	3.5
	Ndebele	17	0.8
	Southern Sotho	43	2.0
	Setswana	18	0.8
	Shona	8	0.4
	Tsonga	55	2.6
	Venda	37	1.7
	Tswana	89	4.1
	Sotho	48	2.2
	Swazi	19	0.9
	Zulu/English	2	0.1
	Pedi	33	1.5
	Portuguese	5	0.2
	Hindi	2	0.1
	German	2	0.1
	French	21	1.0
	Chinese	4	0.2
	Unknown	18	0.8
Total		2145	100

Table 2 reflects the geographical distribution of the sample. A relatively large proportion of the sample resided in the Gauteng area. Compared to Kwazulu-Natal and the Western Cape, a much smaller proportion of the sample lived in the other provinces.

Table 2

Geographical Distribution of the Sample

		FREQUENCY	%
RESIDENCE			
	Eastern Cape	159	7.4
	Free State	99	4.6
	Gauteng	775	36.1
	Kwazulu-Natal	414	19.3
	Limpopo	158	7.4
	Mpumalanga	82	3.8
	North West Province	75	3.5
	Northern Cape	38	1.8
	Western Cape	345	16.1
	Total	2145	100

Measuring instruments

The predictor variables

The pre-screening questionnaire is the measuring instrument used to predict future performance of job applicants at the financial institution. The pre-screening questionnaire was self-developed by the respective financial institution for the purpose of recruitment and selection. In this study, the test scores were obtained by administering the pre-screening questionnaire electronically via the financial institution's internal website as a standard administration procedure. Likewise, standardised procedures were adopted to obtain the performance scores through the financial institution's performance management system. The pre-screening questionnaire comprises three questionnaires or subscales as outlined below.

Biographical data subscale

Biographical questionnaire required respondents to provide personal details such as race, gender, disability and income status and education. A total of 19 items were presented in the biographical questionnaire. Respondents were expected to select the category that best described them on each aspect. Despite being a large subscale, the nature of the items included in the biographical questionnaire was consistent with previous studies that also used variables such as age, gender, race, language, qualifications and years of experience as predictors (see Buitendach, Oosthuyzen & Van Wyk, 2005; Walbrugh & Roodt, 2003; Lopes et al., 2001). It has been reported that biographical data measures have a validity coefficient ranging from .2 to .5 (Harold et al., 2006) and while they demonstrate good validity, it is costly to develop biographical data measures (Taylor, 2008).

Integrity subscale

The integrity questionnaire assessed possible counterproductive behaviours in the workplace. The items in the questionnaire required respondents to honestly declare past crimes or civil cases lodged against them, state of insolvency etc. Job applicants responded to a set of 14 dichotomous items, where a response of 1 indicated a "yes" and a response of 2 meant "no". In previous studies, the integrity questionnaire demonstrated a validity coefficient of .4 when predicting supervisory job performance and students' college performance (Ones et al., 1995; Oswald, 2004). While some authors may disagree, integrity testing has been successful at predicting job performance (Ones et al., 1995; Ones et al., 2007) and seems particularly useful in selection procedures when combined with other measures (Aiken, 1994).

Interest subscale

The interest questionnaire determined if there is was a match between the respondent's preference or interests and the job demands of the sales position (Foxcroft & Roodt, 2005). Responses were measured using a Likert-type scale ranging from 1 to 5, where a score of 1 indicated "strongly disagree" and a score

of 5 indicated "strongly agree". These included having specific interests such working in a team, sales, leadership influencing others. Given the nature of interest questionnaires, it is not commonly used as a predictor of job performance (Carrell et al., 1995; Gerber et al., 1999). In the present study, the information obtained from an interest subscale will be used to determine what kinds of activities tend to interest successful employees at the financial institution. Thus, interest profile will be useful at ascertaining the match between job applicants' career-related interests and nature of jobs on offer at the financial institution.

The criterion measure

The criterion in this study was company-specific performance rating (based on sales performance), which 449 successful job applicants received at the end of the performance cycle. For the purpose of the present research, the performance ratings were converted into scores ranging from 1 - 5. A score of 1 indicating poor performance, 2 was a below average score, a score of 3 meant the performance was adequate, 4 indicated good performance and 5 was considered to be an excellent performance. In order to compute criterion-related validity coefficients, each successful job applicant's raw scores on the pre-screening questionnaire as well as their performance scores were analysed.

Procedure

In this study, secondary data were used. Secondary data are defined as data that the researcher did not collect originally, but it was collected previously for other purposes and processes (Hair et al., 2003). In the present study, the financial institution collected the data at different points in time. The first data set was prospective job applicants' original assessment scores (n=2145) obtained on the pre-screening questionnaire at the time of recruitment. The second was a much smaller data set of performance scores (n=449) for successful job applicants that was collected subsequent to their recruitment. Firstly, the researcher requested and obtained written permission from the financial institution to access the secondary data for this study. Also, the present

research study was approved by an ethical clearance committee prior to the commencement of the study.

Subsequent to obtaining ethical clearance, the secondary data was obtained in electronic format (Microsoft Excel spreadsheet) with all assessment scores and performance scores. To ensure confidentiality, any identifying information such as the names, age and identity numbers of job applicants were omitted from the electronic spreadsheet. Much of the data had to be recoded into a workable format, while other data (such as date of appointment) were excluded from the analysis on the basis of being irrelevant to the study. Thereafter, the data sets were imported into SPSS (Version 16) for analysis to determine if any significant statistical relationships existed between them as evidence of the pre-screening questionnaire's predictive validity.

Summary

This chapter outlined the research design for this validation study. Performance scores were used as the main criterion and assessment scores obtained on the biographical, interest and integrity subscales of the pre-screening questionnaire were the predictors. The research procedure was detailed, where several statistical analyses would be conducted to determine the pre-screening questionnaire's predictive validity.

CHAPTER 4: RESULTS

In this chapter, the results of the data analysis procedures are presented. The chapter is divided into three different sections, where the reliability and validity of biographical, integrity and interest subscales are reported separately. Both inferential and descriptive statistics were used to analyse the data, using SPSS Version 16 and Microsoft Excel 2003.

The three subscales of the pre-screening questionnaire were analysed as follows:

- *Internal consistency (reliability):* Analyses were done on both the integrity and interest subscales to determine their internal consistency. The purpose of measuring internal consistency is to find out how well the items correlate with each other in measuring the underlying construct in each subscale (Tredoux & Durrheim, 2002). The Kuder-Richardson 20 formula was used for the dichotomous items in the integrity subscale, while Cronbach's Alpha reliability coefficient was used for the interest subscale.

- *Zero-order correlation:* A zero-order correlation (Pearson's correlation coefficient) was conducted on both the integrity and integrity subscale to obtain an index of prediction i.e. how well they would be able to predict job performance. "Zero-order correlations are bivariate correlations without controlling for any other variables" (Field, 2005, p. 137). Total scores on both integrity and interest subscales were correlated with employees' performance scores, which were obtained at the end of the performance cycle.

- *Cluster analysis*: The financial institution wished to investigate what kind of work-related activities successful employees were typically interested in. The interest subscale had to be further interrogated using a cluster analysis to obtain a profile of interests for successful job applicants. A

cluster analysis defines a structure or order amongst the interest items by placing the most similar observations into groups (Hair, Black, Babin, Anderson & Tatham, 2006). In this case, job applicants' interests that tend to be similar would be classified into groups.

Biographical Questionnaire

Descriptive statistics was used to analyse the biographical subscale. The raw scores for each item in the subscale were tabulated, using frequency distributions to determine any patterns amongst the frequencies. The results obtained from these statistics provided a description of the sample, as detailed in Chapter 3.

Given the categorical nature of the items in the biographical subscale, a loglinear analysis (extended chi-square) was applied to establish which of the items tended to be closely associated with the criterion. The items were analysed using successful job applicants' performance scores as the criterion. The loglinear analysis created a model of the items in a hierarchical order based on how significantly they predict the criterion in combination.

However, due to the large number of items in this subscale, a decision had to be made regarding which items to select as possible predictors of job performance in the current study. Previous studies used variables such as age, gender, race, language, qualifications and years of experience in prediction studies (see Buitendach et al., 2005; Lopes et al., 2001; Walbrugh & Roodt, 2003). Following from the literature, it was decided that the six items relating to race, gender, language, general qualifications, industry-related qualifications and vehicle ownership would be included in the loglinear analysis. The Pearson correlation was used because there were small numbers in the cells. The results of the loglinear analysis are illustrated in Table 1.

Table 3

Model of the Interaction between Race, Gender and Language Items

			Pearson	
K-way and Higher Order Effects	K	*df*	Chi-Square	Sig.
Race, Gender, Language	1	31	3998.206	.000
Race*Gender, Race*Language, Gender*Language	2	24	1782.643	.000
Race*Gender* Language	3	9	17.248	.045*

The criterion for this analysis was job applicants' performance scores. The model showed that the interaction between the race, gender and language items in predicting the performance scores were significant: χ^2= 17.24 (*df*= 9), p = 0.05. There were several higher order effects and interactions noted. In the first row (K = 1), the one-way interaction was significant between the criterion and the race, gender and language items respectively. This means that removing one of these items will significantly affect how the model fits the data or performance scores. Likewise, the second row (K = 2) indicates that the combined effect of the two-way interaction of race and gender, race and language, gender and language are significant in predicting the criterion. Finally, the third row (K = 3) indicates a main interaction effect between race, gender and language items with the performance scores, meaning that this three-way interaction is a significant predictor of the criterion.

In addition, the race, gender and language items were cross tabulated, using the chi-square to investigate the strength of the relationship between these items

and the criterion. In Table 2, the results indicate that the strength of the relationship between race and criterion was significant with a small effect size. However, the strength of the relationship between the criterion and gender and language items was not significant. While all three biographical items interacted with the criterion in the previous loglinear analysis, the chi-square confirmed that the biographical item of race had the strongest association with the performance scores in this study.

Table 4

Measures of Association between Race, Gender and Language Items with the Performance Scores as Criterion

	Cramer's V	Value of Pearson's Chi-square	*df*	Approx. Sig
Race*Criterion	.147	29.190	12	.004
Gender*Criterion	.063	1.790	4	.774
Language*Criterion	.215	82.280	88	.652

Using the chi-square tests, it was possible to determine the frequency of the different categories in each biographical item in relation to the performance scores. This was obtained by cross-referencing the criterion with the race, gender and language items respectively (see Table 3).

For the biographical item of race, the most frequent category noted in relation to the criterion was the White racial category, followed by the Black racial category. There appears to be an equal count in terms of the Coloured and Asian racial categories. In terms of gender, Table 3 indicates that the male gender had a higher frequency in relation to the performance scores than the female gender.

The highest count reflected in the language category was Afrikaans, followed by English and Zulu languages. While several languages were investigated using in this analysis, the other languages did not yield as high frequencies as Afrikaans and English respectively.

Table 5

Frequency Distribution of Race, Language and Gender Items in relation to Performance Scores

		Performance Scores					
		1	2	3	4	5	
							Total
Race	Asian	22	6	4	7	0	39
	Black	93	34	9	3	0	139
	White	160	47	16	9	2	234
	Coloured	23	14	2	0	0	39
	Total	298	101	31	19	2	451
							Total
Gender	Female	114	35	10	5	1	165
	Male	184	66	21	14	1	286
	Total	298	101	31	19	2	451
							Total
Language	Afrikaans	130	43	11	7	2	193
	English	73	21	11	8	0	113
	Xhosa	14	6	1	1	0	22
	Zulu	25	21	2	1	0	49
	Northern Sotho	7	2	1	0	0	10
	Southern Sotho	5	2	0	0	0	7
	Tsonga	5	1	0	0	0	6
	Tswana	8	3	0	0	0	11
	Sotho	5	4	1	0	0	10
	Swazi	4	1	1	0	0	6
	Other African Languages	10	1	1	1	0	13
	European Languages	5	1	1	1	0	8
	Total (valid)	291	106	30	19	2	448

Integrity Questionnaire

Reliability

The reliability of the integrity subscale was established using a data set of 2145 cases. This subscale comprised of 14 dichotomous questions that was used to evaluate job applicants' integrity in relations to the questions posed. In Table 4, the proportion of yes/no responses to each item is reflected.

Table 6

Frequency Distribution of Positive and Negative Responses to Integrity Items

Item Description		YES	NO
ITG 1	Driver's license endorsed	32 (1.5%)	2102 (98.5%)
ITG 2	Credit agreements	416 (19.7%)	1700 (80.3%)
ITG 3	Previous crimes	34 (1.6%)	2102 (98.4%)
ITG 4	Past civil law suits	65 (3%)	2070 (97%)
ITG 5	Insolvency	27(1.3%)	2099 (98.7%)
ITG 6	Rehabilitation	637 (33.5%)	1262 (66.5%)
ITG 7	Refused contract	11(.5%)	2117(99.5%)
ITG 8	Contracts cancelled	21 (1%)	2107 (99%)
ITG 9	Debarred from practice	6 (.3%)	2122 (99.7%)
ITG 10	Judgments (Professional bodies)	67 (3.1%)	2074 (96.9%)
ITG 11	Membership denied	49 (2.3%)	2084 (97.7%)
ITG 12	Guilty of offense	46 (2.2%)	2091 (97.8%)
ITG 13	Termination of practice	50 (2.3%)	2087 (97.7%)
ITG 14	Disqualified to manage	40 (1.9%)	2094 (98.1%)

The Kuder-Richardson 20 (KR-20) formula was used to establish the integrity subscale's internal consistency. The item variance was determined by calculating the proportion of yes/no responses for each item in the subscale. This resulted in a total item variance of 1.34, which was then applied to the KR-20 formula. A moderate reliability coefficiency of r_{tt} = 0.63 was found for the integrity subscale.

Each item in the integrity subscale was further investigated to ascertain which items were contributing to the reliability of the subscale. First, a total integrity score for each item was calculated. Each item in the subscale was then correlated with its corresponding total integrity score (see Table 5). The analysis revealed integrity items such rehabilitation (ITG 6), cancellation of contracts (ITG 8), guilty of offenses (ITG 12), disqualified to manage (ITG 14) and termination of practice (ITG 13) contributed most significantly to overall reliability of the subscale. However, the item regarding endorsement of driver's licenses (ITG 1) was found to contribute the least the integrity subscale's reliability.

Table 7

Inter-correlation between the Integrity Items and the Total Integrity Subscale

		Correlations (*r*)	
	Item Description		Total Integrity
ITG 1	Driver's license endorsed	Pearson Correlation	.180**
		Sig (2-tail)	.000
ITG 2	Credit agreements	Pearson Correlation	.376**
		Sig (2-tail)	.000
ITG 3	Previous crimes	Pearson Correlation	.246**
		Sig (2-tail)	.000
ITG 4	Past civil law suits	Pearson Correlation	.290**
		Sig (2-tail)	.000
ITG 5	Insolvency	Pearson Correlation	.300**
		Sig (2-tail)	.000
ITG 6	Rehabilitation	Pearson Correlation	.542**
		Sig (2-tail)	.000
ITG 7	Refused remuneration contracts	Pearson Correlation	.401
		Sig (2-tail)	.000
ITG 8	Contracts cancelled	Pearson Correlation	.433**
		Sig (2-tail)	.000
ITG 9	Debarred from practice	Pearson Correlation	.394**
		Sig (2-tail)	.000
ITG 10	Judgments (Professional bodies)	Pearson Correlation	.397**
		Sig (2-tail)	.000
ITG 11	Membership denied	Pearson Correlation	.402**
		Sig (2-tail)	.000
ITG 12	Guilty of offense	Pearson Correlation	.410**
		Sig (2-tail)	.000
ITG 13	Termination of practice	Pearson Correlation	.424**
		Sig (2-tail)	.000
ITG 14	Disqualified to manage	Pearson Correlation	.432**
		Sig (2-tail)	.000

Validity

In terms of the index of prediction for the integrity subscale, a zero-order correlation was conducted between the total integrity scores and the performance scores. The result of the zero-order correlation indicated that the integrity subscale's index of prediction was (r=.017).

Given that the integrity subscale was validated on a group whose predictor scores do not represent the total range, there is a possibility that the validity coefficient found would be underestimated. This possibility was investigated by calculating using the restriction of range in two-variable situation (Ghiselli, Campbell & Zedeck, 1981). Firstly, the index of prediction was calculated (r=.017). Secondly, the standard deviations were required for the restricted group who were hired by the financial institution (n=449), while the unrestricted group comprised a bigger sample of all job applicants who applied for positions at the financial institution (n= 2145). Using descriptive statistics, the standard deviations were obtained for the restricted group (SD = .834) and unrestricted group (SD = .408). Thereafter, the range restriction was calculated using the following formula:

$$r_{xy} = \frac{r'_{xy}(\sigma_x / \sigma'_x)}{\sqrt{1 - r'^2_{xy} + r'^2_{xy}(\sigma^2 x / \sigma'^2 x)}}$$

$$r_{xy} = \frac{(.017).(.408/.834)}{\sqrt{1-(.017)^2+(.017)^2[(.408)^2/(.834)'^2]}} = \frac{.008}{1} = .008$$

Thus, a range restriction of (r_{xy} = .008) indicates that the validity coefficient found on the restricted group is minimally underestimated in comparison to the unrestricted group in the study.

Interest Questionnaire

The items in the interest subscale relate to general activities at work that some people would be more interested in doing than others. The interest subscale is described in Table 8 below.

Table 8

Descriptive Statistics of the Interest Subscale

Item	Description	Mean	SD	Variance	Skewness	Kurtosis
INT 1	Influence	3.59	1.751	3.066	-1.236	.039
INT 2	Leadership	3.68	1.749	3.058	-1.361	.321
INT 3	Persuasion	3.76	1.788	3.196	-1.382	.303
INT 4	Relational	3.80	1.807	3.267	-1.399	.309
INT 5	Advising	3.77	1.783	3.179	-1.395	.361
INT 6	Outgoing	3.87	1.809	3.273	-1.475	.491
INT 7	Self-confidence	3.87	1.807	3.265	-1.490	.529
INT 8	Management	3.80	1.788	3.197	-1.426	.423
INT 9	Convincing	3.74	1.763	3.109	-1.399	.393
INT10	Presenting	3.61	1.752	3.068	-1.258	.105
INT11	Interacting	3.80	1.807	3.267	-1.402	.320
INT12	Listening	3.88	1.820	3.312	-1.474	.461
INT13	Marketing	3.79	1.804	3.254	-1.395	.310
INT14	Selling	3.66	1.761	3.102	-1.291	.171
INT15	Awards	1.35	.736	.541	-.666	-.885
INT16	Job role	.94	.497	.247	-.128	.960

Reliability

An inter-item correlation matrix of Pearson's Product Moment Coefficients (*r)* was calculated as a basis for determining the internal consistency of the interest subscale (see Table 9). There were significant correlations amongst all 16 interest items, which mean that all the items in the interest subscale seemed to relate well with one another.

More specifically, a strong correlation (r=.903) existed between the items of relational (INT 4) and persuasion (INT 3). Another strong correlation (r=.9) was found between the management (INT 8) and self-confidence (INT 7) items. Given the significant correlations, a reliability analysis was conducted to establish to what extent each item was contributing to the reliability of the interest subscale. The overall Cronbach's alpha for the interest subscale was (r=.985). In Table 10, the results showed that interest subscale's Cronbach's alpha remained consistent (r=.984) irrespective of which items was deleted, meaning that all the items were contributing equally to the reliability of the interest subscale.

Table 9

Matrix of Correlations for Interest Items

Correlations (*r*)

		INT 1	INT2	INT 3	INT 4	INT 5	INT6	INT 7	INT 8	INT9	INT 10	INT 11	INT 12	INT 13	INT 14	INT 15	INT 16
INT 1	Influence		.888*	.867**	.860**	.859*	.853**	.845**	.851**	.869*	.833**	.837**	.853**	.841**	.850**	.885**	.675**
INT 2	Leadership			.883**	.877**	.881*	.878**	.870**	.877**	.871*	.859**	.860**	.872**	.857**	.868**	.703**	.712**
INT 3	Persuasion				.903**	.870*	.874**	.866**	.867**	.869*	.839**	.846**	.870**	.855**	.860**	.713**	.709**
INT 4	Relational					.883*	.888**	.879**	.878**	.865*	.844**	.857**	.880**	.862**	.868**	.714**	.719**
INT 5	Advising						.915**	.898**	.890**	.890*	.871**	.867**	.889**	.897**	.890**	.703**	.685**
INT 6	Outgoing							.910**	.904**	.897*	.883**	.887**	.902**	.893**	.876**	.703**	.712**
INT 7	Self-confidence								.900**	.886*	.877**	.886**	.892**	.891**	.876**	.693**	.698**
INT 8	Management									.893*	.866**	.882**	.893**	.880**	.874**	.708**	.708**
INT 9	Convincing										.868**	.882**	.888**	.882**	.876**	.704**	.699**
INT 10	Presenting											.875**	.874**	.872**	.859**	.638**	.659**
INT 11	Interacting												.897**	.880**	.874**	.681**	.682**
INT 12	Listening													.908**	.892**	.717**	.711**
INT 13	Marketing														.898**	.678**	.681**
INT 14	Selling															.660**	.665**
INT 15	Awards																.682**
IN T16	Job Role																

Table 10

Reliability Analysis on Interest Subscale

Item	Scale Mean	Scale Variance	Corrected Item Total Correlation	Alpha if Item Deleted
INT 1	51.30	520.736	.907	.984
INT 2	51.20	519.250	.928	.984
INT 3	51.13	518.009	.922	.984
INT 4	51.08	516.682	.929	.984
INT 5	51.12	516.889	.940	.984
INT 6	51.02	515.253	.946	.984
INT 7	51.01	515.939	.939	.984
INT 8	51.08	516.840	.938	.984
INT 9	51.15	518.048	.936	.984
INT 10	51.28	520.147	.914	.984
INT 11	51.09	516.913	.926	.984
INT 12	51.01	515.062	.943	.984
INT 13	51.10	516.537	.933	.984
INT 14	51.23	518.760	.927	.984
INT 15	53.53	569.818	.738	.987
INT 16	53.95	578.280	.742	.988

Interest profile

The cluster analysis was used to determine a profile of activities that successful job applicants would be interested in. Although this is more a descriptive statistical technique, the cluster analysis is useful at indicating the items that tend to cluster well together, generating an interest profile of successful job applicants. In figure 1, the dendrogram illustrates the different clusters.

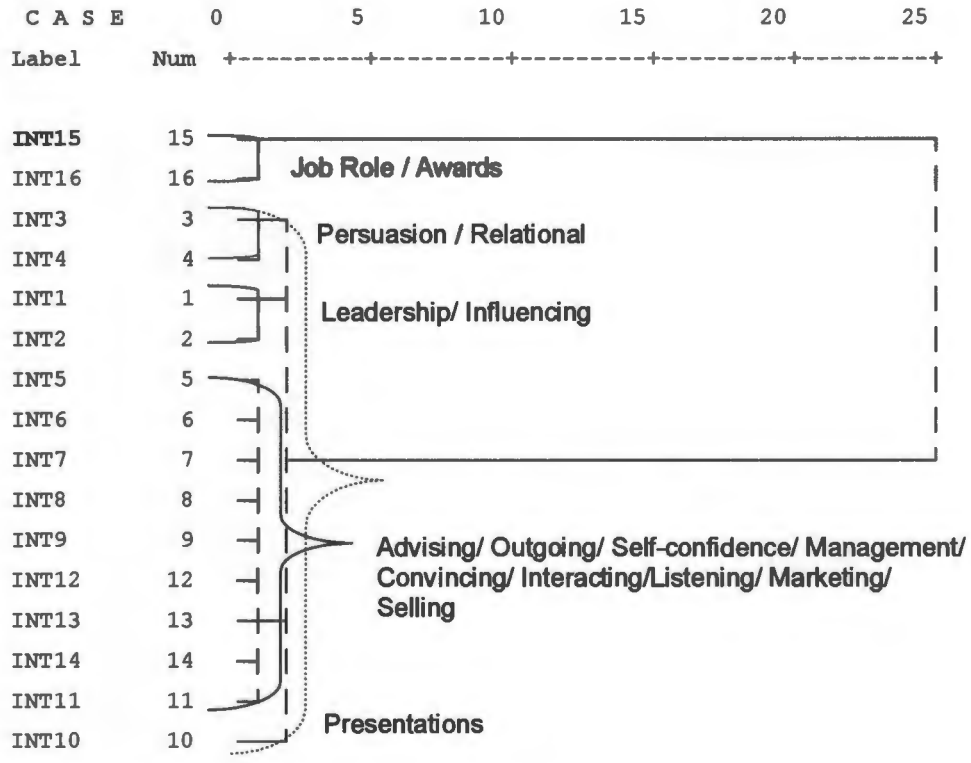

Figure 1. Dendogram: Clustering of items into an interest profile.

It can be seen that the following items have grouped together most significantly, constituting the primary cluster or interest profile: advising (INT 5), outgoing (INT 6), self-confidence (INT 7), management (INT 8), convincing (INT 9), interacting (INT 11), listening (INT 12), marketing (INT 13) and selling (INT 14). The interest profile seems to outline the kinds of activities and work behaviours that successful employees tend to exhibit at the financial institution. Therefore, the

interest profile could provide an indication of how successful a job applicant may be should the financial institution decide to hire him or her.

Three secondary clusters were also noted. The interest items of influencing (INT 1) and leadership (INT 2) have clustered together to a leadership profile, while persuasion (INT 3) and relational (INT 4) form a separate cluster relating to communication style. The final cluster comprised of awards (INT 15) and job role (INT 16), which indicate job knowledge that a successful employee would have. Thus, an interest profile has been generated based on the primary cluster, which will be described in further detail in the discussion section.

It is noted that the item regarding job applicants' interest in conducting presentations (INT 10) is not included in any of the clusters. Instead, this item forms part of a larger tertiary cluster (see Figure 1) that includes all the items in the interest subscale except those of Awards (INT 15) and Job Role (INT 16). Given that it does not group very well with the primary cluster, this item will be omitted from the interest profile.

In addition, a zero-order correlation was conducted between the performance scores and the total interest scores. The result indicated that the index of prediction for the interest subscale was (r=.043).

Summary

The results were presented separately for each subscale. The loglinear analysis revealed that the gender, race and language items in the biographical subscale interacted significantly with the criterion. The integrity subscale was found to have a moderate reliability coefficient and the validity testing revealed that only one item being strongly associated with the criterion. The interest subscale was found to be a highly reliable indicator of work-related interest, yielding high reliability co-efficients. A further analysis on the interest subscale showed that

certain items tended to group together in a primary cluster, forming an interest profile. These results will be interrogated in the forthcoming discussion chapter.

CHAPTER 5: DISCUSSION

The purpose of this study was to determine the pre-screening questionnaire's ability to predict future job performance at the financial institution. The Employment Equity Act (no. 55), 1998 recognizes the utilization of psychological assessments in selection procedures, on condition that such assessments are not biased in terms of culture, gender or race and that they are scientifically valid and reliable. South African businesses have begun to review its assessment instruments in order to comply with this requirement as a means to ensure that more accurate selection decisions are made (Lopes et al., 2001). The pre-screening questionnaire was developed as part of the financial institution's selection process and was subjected to a validation study in terms of the requirements of the Employment Equity Act.

The pre-screening questionnaire comprises three subscales, each intended to provide different kinds of information about job applicants. The first subscale is a biographical questionnaire, which was studied in order to identify which items in this subscale appear to be good predictors of future job performance. The integrity questionnaire was the second subscale, which was evaluated in terms of its reliability in measuring counterproductive, deviant or dishonest behaviour of job applicants as well as determining its predictive validity. The final subscale was the interest questionnaire that was analysed in terms of reliability as well as the formulation of an interest profile of successful job applicants. The findings for each subscale are discussed.

Biographical Questionnaire

Based on the loglinear analysis, it was found that the race, gender and language items in the biographical subscale were significantly associated with the performance scores. In other words, these three items in combination seem to be able to predict how well job applicants will perform in future, with the item of race being the most significant predictor of job performance at the financial institution.

This is consistent with the findings in the literature, which states that biographical measures are useful predictors of global job success in financial companies (Boam & Sparrow, 1992; Taylor, 2008). Further, the chi-square tests revealed that employees who tended to be the most successful in their jobs at the financial institution were white males who spoke either Afrikaans or English. Thus, the biographical subscale in this study is useful in identifying successful job applicants to some extent, but it is considered to be a moderate predictor of future job performance. However, the findings should be interpreted with caution, as the present study did not conclusively evaluate possible item bias in the pre-screening questionnaire or fairness of the financial institution's selection process.

Integrity Questionnaire

The integrity subscale's internal consistency was found to be moderate. There appeared to be a wide variation in responses across the subscale, meaning that some of the items did not consistently measure integrity as the underlying construct. For instance, the item regarding driver's license endorsements reduced the reliability of the subscale, while other items such as rehabilitation, and guilty of offenses increased its reliability.

One explanation for the moderate internal consistency could be the direction of responses in the integrity subscale. The purpose of an integrity questionnaire is to assess a job applicant's dishonesty in the workplace (Camara & Schneider, 1994; Ones & Viswesvaran, 2007). Therefore, integrity questionnaires should be sensitive in detecting counterproductive behaviors. However, it is noted that the majority of job applicants responded by saying "no" to each integrity item. This is to be expected, given that the integrity questions relate to counterproductive behaviours. To put it more clearly, a negative response ("no") would be appropriate when answering an integrity question such as "Have you ever been declared insolvent before?". Based on this logic, job applicants were allocated 1 score for every positive response ("yes") and 2 scores for every negative response ("no"). This has an influence the reliability of the integrity subscale

because it is detecting socially acceptable behaviours rather than counterproductive ones. The direction of the responses may be a limiting factor and the response set should be revisited so that it measures counterproductive behaviours as well as socially appropriate responses more accurately. This will allow the integrity subscale to flag negative behaviours with a higher point score.

The inconsistency of this subscale could mean that it is measuring different aspects of integrity, but may not be measuring integrity as a unitary concept. Due to the nature of the items in the subscale, it appears that the financial institution considers legitimacy of professional registrations, insolvency and misconduct as indicators of integrity. Other interpretations of the meaning of integrity are expressed in the literature. For example, a person's level of integrity is related to positive work behaviours such as impulse restraint, conscience and reduced hostility and is evaluated by measuring counterproductive behaviours such as theft, tardiness, property damage, rule-breaking and violence (Ones & Viswesvaren, 2007). Bartram, (as cited by Paterson & Uys, 2005) views integrity as an extension of the Big Five personality factor of *conscientiousness*, which means that it would best be measured using personality assessments. The challenge with using personality traits to measure integrity is that they do not adequately detect workplace theft and dishonesty, yet this is the main reason for using integrity tests in industry (Camara & Schneider, 1995).

There appears to be different opinions about the meaning of integrity as well as how best to measure it in the workplace. For the financial institution, the integrity items related more to legitimacy of professional registration, status of criminal record, fraud and professional misconduct. This indicates that the financial institution has defined the concept of integrity in terms of the specific business context in which it operates and therefore differs to academic definition of integrity as a personality trait. Despite this, Bartram, (as cited by Paterson & Uys, 2005) maintains that the most effective assessment batteries in personnel

selection are those that include ability tests and the personality traits of both conscientiousness and integrity.

In terms of validity, the prediction index obtained indicated that a weak correlation existed between the performance scores and the integrity subscale. Therefore, the integrity subscale seems to be poor predictor of job performance at the financial institution. This finding is consistent with a large proportion of the research, which states that integrity testing cannot accurately predict job performance (Camara & Schneider, 1994; Kline, 2000; Arthur, 2006). The finding in this study contradicts Ones and Viswesvaran's (2007) recent research which proved the predictive validity of integrity testing. While the integrity test correlated strongly with the performance scores in their study, it was noted that the integrity test was combined with job knowledge, which is a good predictor of maximal job performance (Schmidt & Hunter, 1998). Thus, Ones and Viswesvaran's (2007) conclusion is not unexpected, as integrity tests usually demonstrate more useful levels of validity when they are combined with other selection tools (Aiken, 1992). Also, the present study was predicting typical job performance which refers to how people will perform in a job, while maximal job performance refers to how people can perform (Ones and Viswesvaran (2007). Therefore, maximal and typical job performance is considered to be two distinct criteria, although each of them relate to job performance as a construct.

Utility analysis

Despite having low predictive validity, the integrity questionnaire may still be a useful tool in selection. In order to establish its usefulness in the financial institution's selection process, a utility analysis must be done. In this study, a total of 2145 job applications were received for 449 positions that had to be filled at the financial institution, resulting in a selection ratio of .2. Suppose the integrity subscale has a base rate of .5 and a validity coefficient of .63 and a selection ratio of .2, it would be possible to calculate the utility. According to the Taylor-Russell tables, the integrity test estimates a success ratio of .84, meaning that

84% of job applicants will perform successful at the financial institution. This represents a 34% gain over the base rate (.84 - .5 = .34) if the integrity subscale is used in selection processes at the financial institution. The financial institution could see greater economic gains if it were more selective i.e. only 1 position is available and 10 job applications are received. A selection ratio of .1 would result in a success ratio of .9 or 90%, meaning that the integrity subscale shows a 40% improvement over the current base rate (.9 - .5 = .4). The integrity subscale therefore demonstrates utility and is considered to be useful in the financial institution's selection process.

Interest Questionnaire

The interest subscale demonstrated high internal consistency. This means that the interest items correlated significantly with one another in their ability to consistently measure the work-related interest of job applicants at the financial institution. This finding is consistent with previous studies that found interest questionnaires to be reliable in indicating person-job fit, where the job applicant's interests can matched with activities and behaviours required by the job (Anastasi, 1982; Carrel et al., 1995; Foxcroft & Roodt, 2005). More interestingly, the interest items of outgoing (INT 6) yielded strongly associated with five other items in the subscale, namely listening (INT 12), self-confidence (INT 7), management (INT 8), advising (INT 5) and marketing (INT 13). Thus, the item of outgoing (INT 6) seems to be a common factor that is being measured by the interest subscale. In other words, this item evaluates a job applicant's preference for a job where he/she must like to meet and constantly relate to other people. The interest subscale seems reliable in determining a job applicant's preference for doing work where he or she has to constantly meet and engage with new people

Interest questionnaires are usually used to assist people to make career decisions (Aiken, 1994). Thus, prediction studies have been conducted with interest inventories in terms of long-term career choice (Harrington, as cited by

Chope & Healy, 2006), but have been considered as poor predictors of job performance. Schmidt and Hunter (1998) reported that interest questionnaires that demonstrate a predictive validity of .1 are not considered to be good predictors because their validity is low. In the present study, a low index of prediction was obtained for the interest subscale, meaning that there was a poor correlation between the performance scores and the interest items. Consistent with previous studies' findings, the interest subscale was found to be a poor predictor of job performance.

The financial institution wanted to investigate whether there would be a pattern or profile of work-related activities that successful employees would typically be interested in. The cluster analysis indicated that high performing employees are self confident with strong selling and leadership skills. They enjoy work environments where they constantly meet different people whom they can advise on which company products will suit their needs. The financial institution would be able to utilize this profile to identify company specific competencies as indicators of performance drivers in the organisation (Boyatzis, 2008). The interest profile is also useful in establishing person-job fit (Carell et al., 1995) because it is able to discriminating between job applicants by eliminating those who do not prefer working in sales or who tend to avoid circumstances where they have to engage with different people.

Utility analysis

In order to determine the interest subscale's utility, the same selection ratio of .2 and base rate of .5 will be used as for the integrity subscale. The validity coefficient for the interest subscale was .1, as per the regression analysis in the previous chapter. Thus, the success ratio of .56 is indicated on the Taylor-Russell table, meaning that there is only a 6% improvement over the base rate if the interest subscale is used in the financial institution's selection process. The interest subscale demonstrates a much lower utility than the integrity subscale.

Implications

Given the emphasis on the use of reliable, valid and fair psychological tests in selection, this study investigated whether the pre-screening questionnaire demonstrated satisfactory levels of reliability and validity as required by the Employment Equity Act, 1998. More importantly, the findings of this study provide insight into the pre-screening questionnaire's ability to predict job performance at the financial institution. The pre-screening questionnaire uses three subscales as specific predictors of job performance at the financial institution and each of the subscales were analysed in this study.

The biographical subscale seems to be the strongest predictor of job performance in the pre-screening questionnaire. Although rudimentary in nature, background factors such as the race, gender and language items in the biographical subscale were able to predict which job applicants would perform well on the job in future. It also appears that the biographical subscale is a suitable predictor of job performance in the specific industry within which the financial institution operates.

The integrity subscale was reasonably reliable in measuring integrity as understood by the financial institution. While the integrity subscale was not able to predict job performance adequately, it demonstrates a high utility (or usefulness) as a selection tool.

This study found the interest subscale to be highly reliable in measuring job applicants' work-related interests. With a low index of prediction, the interest subscale was not a good predictor of job performance. The interest subscale also had low utility, meaning that it might not be the most useful device to use in the financial institution's selection process. Despite its lack of utility, a profile of interests was generated based on the interest subscale that can assist in determining person-job fit, i.e. whether job applicants will fit in and adjust to the organizational culture of the financial institution. These findings provides support

for the financial institution's decision to use the interest subscale to obtain a profile of job applicants' work-related interests, rather than using it as a predictor of job performance.

The combination of biographical, integrity and interest subscales makes the pre-screening questionnaire a unique instrument. Instead of using them separately, the pre-screening questionnaire combines the biographical and integrity subscales and uses each of them as specific predictors of job performance to select job applicants, while the interest subscale identifies an interest profile for each job applicant. Generally, personality inventories, cognitive ability tests and job knowledge tests have been the more popular approaches to predicting job performance, but it seems that the pre-screening questionnaire has specific relevance for the financial institution's recruitment and selection process.

Limitations

One of the main limitations of this study was the format in which the secondary data has been collected. There were instances where performance scores were indicated as random numbers that had to be converted into conventional ratings on a scale of 1 to 5. Some of the assessment data had to be recoded so that it would be compatible with the statistical package. This became a time-consuming exercise and it is advised that future researchers take the nature and format of secondary data into account. According to Hair et al. (2003), this challenge arises because secondary data seldom fit the purpose of the study at hand, since the data was collected for different reasons e.g. assessment scores are collected for selection and performance ratings, as opposed to being used in a validation study.

The use of performance scores, as the criterion measure was a second limitation noted in this study. Generic performance scores are actual ratings given by supervisors, which are susceptible to judgmental errors or bias (Anastasi & Urbina, 1997). Subjective performance ratings may not represent the construct of job performance accurately, thereby reducing its accuracy as a criterion. It is for this reason that a criterion is only relevant if it is based on the understanding of the construct (Muchinsky et al., 2005). In this study, the performance ratings were based on each employee's sales volume to which a supervisory rating was ascribed. Therefore, the performance scores may not have been an accurate criterion of the job performance at the financial institution, which could have influenced the outcomes of this study.

Another limitation was that there were rather few validation studies on biographical, integrity and interest questionnaires published in the South African context. It was difficult to compare this study's findings to that of similar research, since previous studies seemed to focus on using well-established instruments such as Ability Processing of Information and Learning Battery (APIL-B), Occupational Personality Questionnaire (OPQ) and Customer Contact Sales

Questionnaire (CCSQ) to predict job performance (cf. La Grange & Roodt, 2001; Lopes et al., 2001; Forbes, 2006). The dearth of research on prediction studies using biographical, integrity and interest questionnaires is indicative of the fact these instruments are not as extensively explored as other devices in the selection procedure. Thus, this validation study can be considered as an addition to the existing body of research using these instruments in predicting job performance in the South African context.

Recommendations

- A follow up study should investigate the reasons as to why white male employees who speak either Afrikaans or English seem to perform well consistently in their jobs at the financial institution. This could be linked to customer service, type of experience or specific skills that enable this category of employees to excel in their jobs. The investigation will identify drivers of job performance, which the financial institution can leverage to improve organisational performance.

- Linked to the first recommendation, future studies should focus on the item bias and fairness of the pre-screening questionnaire to assist the financial institution in eliminating any possible unfair discrimination that ***may*** exist in its selection process.

- In the integrity subscale, it was found that job applicants received low scores each time they admitted to engaging in counterproductive behaviours and had high scores when they did not. However, this scoring method should be reversed to determine job applicants' propensity to engage in counterproductive behaviours as opposed to determining socially acceptable behaviours. An integrity questionnaire is not meant to measure pro-social behaviours in the manner that integrity subscale currently does. By scoring job applicants on their tendency towards counterproductive behaviours, it will improve the integrity subscale's reliability.

- Another means of improving the integrity scale's reliability is to reduce the number of items in the subscale from a total of 19 items to a prioritized 14 items as outlined in this study. It is recommended that the item of driver's license endorsements be removed to improve the internal consistency of the integrity subscale.

- It may be useful to re-evaluate the integrity subscale in a future study by evaluating its construct validity. Given that the integrity subscale is currently measuring different aspects of integrity, it may be influencing the reliability and validity of subscale all together because it may not be measuring a unitary concept of integrity as it should.

- On a more practical level, it is recommended that the financial institution use the interest subscale as a means to indicate of person-job fit in the selection process. While it is not advisable to use the interest subscale as a predictor of job performance, it does provide insight into which job applicants would be most suitably placed in positions at the financial institutions.

REFERENCES

Aamodt, M.G. (1999). *Applied industrial/organizational psychology* (3rd ed.). Belmont, CA: Brooks/Cole Wadsworth.

Aiken, L. (1994). *Psychological testing and assessment* (8th ed.). Massachusetts: Allyn and Bacon.

Allworth, E., & Hesketh, B. (1999). Construct-orientated biodata: Capturing change-related and contextually relevant future performance. *International Journal of Selection & Assessment*, *7*(2), 97-111.

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: MacMillian Publishing Co.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall International.

Armstrong, M. (2003). *A handbook of human resource management practice* (5th ed.). London: Kogan Page.

Arthur, D. (2006). *Recruiting, interviewing, selecting & orienting new employees* (4th ed.). New York: American Management Association.

Ashton, C. (1996). How competencies boost performance. *Management Development Review*, *9*(3), 14-19.

Barnard, M., & Schaap, P. (2005). The concurrent validity of the technical test battery as an indicator of work performance in a telecommunications company. *South African Journal of Industrial Psychology*, *31*(2), 71-77.

Bernardin, H.J. (2003). *Human resource management: An experiential approach* (3rd ed.). Boston: McGraw-Hill Irwin.

Berner, A., & Van Tonder, C.L. (2003). The postmodern consumer: implications of changing customer expectations for organisation development in service organisations. *South African Journal of Industrial Psychology*, *29*(3), 1-10.

Boam, R., & Sparrow, P. (1992). *Designing and achieving competency: A competency based approach to developing people and organisations*. England: McGraw-Hill.

Boyatzis, R.E. (2008). Competencies in the 21st century. *Journal of Management Development*, *27*(1), 5-12.

Buitendach, J.H., Oosthuyzen, A., & Van Wyk, C. De W. (2005). Predicting job insecurity from background variables. *South African Journal of Industrial Psychology*, *31*(3), 70-78.

Camara, W.J., & Schneider, D.L. (1994). Integrity tests: Facts and unresolved issues. *American Psychologist*, *49*(2), 112-119.

Camara, W.J., & Schneider, D.L. (1995). Questions of construct breadth and openness of research in integrity testing. *American Psychologist*, *50*(6) 459-460.

Carrell, M.R., Elbert, N.F., & Hatfield, R.D. (1995). *Human resource management: Global strategies for managing a diverse workforce* (5th ed.). Englewood Cliffs, New Jersey: Prentice Hall.

Cascio, W.F. (1991). *Applied psychology in personnel management* (4th ed.). Englewood Cliffs, New Jersey: Prentice-Hall International.

Cascio, W.F. (2003). *Managing human resources, productivity, quality of worklife, profits* (6th ed.). Boston: McGraw-Hill Irwin.

Chope, R.C., & Healy, C.C. (2006). Introduction to the special issue on interest measurement. *Measurement and Evaluation in Counseling and Development*, *38*, 195-197.

Chowdhury, S. (2000). *Management 21C: Someday we'll all manage this way.* Great Britain: Prentice Hall.

Denton, M., & Vloeberghs, D. (2003). Leadership challenges for organisations in the new South Africa. *Leadership and Organisational Development Journal*, *24*(2), 84-95.

Employment Equity Act, Act 55, 1998. *Republic of South Africa Government Gazette 400*: 19370.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

Forbes, A. (2006). *The predictive validity of the occupational personality questionnaire (OPQ) in assessing competence in the workplace.* Unpublished master's thesis, University of Cape Town, South Africa.

Foxcroft, C., & Roodt, G. (eds.). (2005). *An introduction to psychological assessment in the South African context* (2nd ed.). Oxford: Oxford University Press.

Gerber, P.D., Nel. P.S., & Van Dyk, P.S. (1999). *Human resource management* (4th ed.). Cape Town, South Africa: Oxford University Press.

Ghiselli, EE., Campbell, J.P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences.* San Francisco: W.W. Freeman & Co.

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

Hair, J.F., Babin, B., Money, A.H., & Samouel, P. (2003). *Essentials of business research methods.* Hoboken, New Jersey: Wiley.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., & Tatham, R.L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, New Jersey: Pearson Education Inc.

Harold, C.M., McFarland, L.A., & Weekley, J.A. (2006). The validity of verifiable and non-verifiable biodata items: An examination across applicants and incumbents. *International Journal of Selection and Assessment*, *14*(4), 336-346.

Henry, E.R. (1966). Conference on the use of biographical data in psychology. *American Psychologist*, *21*(3), 247-249.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*(1), 72-98.

Huysamen, G.K. (2002). The relevance of the new APA standards for educational and psychological testing for employment testing in South Africa. *South African Journal of Psychology*, *32*(2), 26-33.

Kerlinger,F.N., & Lee, H.B. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth, TX: Harcourt College Publishers.

Kleynhans, E.P.J. (2006). The role of human capital in the competitive platform of South African industries. *South African Journal of Human Resource Management*, *4*(3), 55-62.

Kline, P. (2000). *Handbook of psychological testing* (2nd ed.). London: Routledge.

La Grange, L., & Roodt, G. (2001). Personality and cognitive ability as predictors of the job performance of insurance sales people. *Journal of Industrial Psychology*, *27*(3), 35-43.

Lilienfeld, S.O., Alliger, G., & Mitchell, K. (1995). Why integrity testing remains controversial. *American Psychologist*, *50*(6), 457-458.

Lopes, A., Roodt, G., & Mauer, R. (2001). The predictive validity of the APIL-B in a financial institution. *Journal of Industrial Psychology*, *27*(1), 61-69.

Muchinsky, P.M., Kriek, H.J., & Schreuder, D. (2005). *Personnel psychology.* (3rd ed.). South Africa: Oxford University Press.

Nelson, D.L., & Quick, J.C. (2006). *Organisational behaviour: Foundations, realities and challenges* (5th ed.). United States: Thomson South-Western Corporation.

Ones, D.S., & Viswesvaran, C. (2007). A research note on the incremental validity of job knowledge and integrity test for predicting maximal job performance. *Human Performance*, *20*(3), 293-303.

Ones, D.S., Dilchert, S., Viswesvaran, C., & Judge, T.A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, *60*, 995-1027.

Ones, D.S., Viswesvaran, C., & Schmidt, F.L. (1995). Integrity tests: Overlooked facts, resolved issues and remaining questions. *American Psychologist*, *50*(6), 456-457.

Ones, D.S., Viswesvaran, C., & Schmidt, F.L. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology*, *90*(1), 108-131.

Oswald, F.L., Schmitt, N., Kim, B.H., Ramsay, L.J., & Gillespie, M.A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology*, *89*(2), 187-207.

Republic of South Africa. (1998). Employment Equity Act (55 of 1998). *Government Gazette*, *400*, No. 19370: Pretoria.

Roth, P.L., Bobko, P. & Mabon, H. (2001). Utility Analysis: A Review and Analysis at the Turn of the Century. In N. Anderson, D.S. Ones, H.K. Sinangil & C. Viswesvaran (Eds.), *Handbook of Industrial, Work and Organizational Psychology: Vol. 1 Personnel Psychology* (pp. 363 – 384). London: Sage.

Rothman, S., & Coetzee, E.P. (2003). Big five personality dimensions and job performance. *South African Journal of Industrial Psychology*, *29*(1), 68-74.

Paterson, H., & Uys, K. (2005). Critical issues in psychological test use in the South African workplace. *South African Journal of Industrial Psychology*, *31*(3), 12-22.

Sackett, P.R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, *59*, 419-450.

Saunders, E. (2005). *Assessing human competence: Practical guidelines for the South African manager*. Randburg: Knowres Publishing (Pty) Ltd.

Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implication of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262-274.

Schmitt, N., Oswald, F.L., Kim, B.H., Gillespie, M.A., Ramsay, L.J., & Yoo, Y.T. (2003). The impact of elaboration on socially desirable responding and the validity of biodata measures. *Journal of Applied Psychology*, *88*(6), 979-988.

Spatz, C., & Kardis, E. (2008). *Research methods in psychology: Ideas, techniques and reports*. New York: McGraw-Hill.

Sutherland, R., De Bruin, G., & Crous, F. (2007). The relationship between conscientiousness, empowerment and performance. *South African Journal of Human Resource Management*, *5*(2), 60-67.

Taylor, S. (2008). *People resourcing* (4th ed.). London: Chartered Institute of Personnel and Development.

Theron, C. (2007). Confessions, scapegoats and flying pigs: Psychometric testing and the law. *South African Journal of Industrial Psychology*, *33*(1), 102-117.

Tredoux, C., & Durrheim, K. (Eds.). (2002). *Numbers, hypotheses and conclusions: A course in statistics for the social sciences*. Cape Town: UCT Press.

Ulrich, D., & Brockbank, W. (2005). *The HR value proposition*. Boston: Harvard Business School Press.

Van der Merwe, R.P. (1999). Psychological assessment in industry. *South African Journal of Industrial Psychology*, *25*(3), 8-11.

Van der Merwe, R.P (2002). Psychometric testing and human resource management. *South African Journal of Industrial Psychology*, *28*(2), 77-86.

Van Iddekinge, C.H., Eidson Jr., C.E., Kudisch, J.D., & Goldblatt, A.M. (2003). A biodata inventory administered via interactive voice response (IVR) technology: Predictive validity, utility and subgroup differences. *Journal of Business and Psychology*, *18*(2), 145-156.

Walbrugh, A., & Roodt, G. (2003). Different age groups' response to employment equity practices. *South African Journal of Human Resource Management*, *1*(2), 28-39.

Wanek, J.E. (1999). Integrity and honesty testing: What do we know? How do we use it? *International Journal of Selection and Assessment*, *7*(4), 183-195.