

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.



UNIVERSITY OF CAPE TOWN

**The predictive validity of the Occupational Personality
Questionnaire (OPQ 32i) in assessing competence in the
workplace.**

Adelaide Forbes

FRBADE002

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

2006

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.

Signature:

Signed by candidate

Date:

MAY 2007

DIGITISED

27 JUN 2013

Table of Contents

List of Tables	v
List of Figures	vi
Acknowledgements	vii
Abstract	viii
Chapter 1: Introduction	1
Chapter 2: Literature Review	4
Psychological Assessment	4
Types of measures	5
Personality and Measurement	7
Personality defined	7
Personality Inventories	8
Criticism and challenges of Personality measurement	9
Validity Challenges	10
Respondent concerns	14
Measurement issues	15
Sample size concerns	15
Psychological assessment in South Africa	16
A brief history	16
Psychological assessment in post–Apartheid South Africa	17
Legal framework	17
Cross–cultural concerns	18
Status of test use in South Africa	20
Psychometric properties of assessment tools	21
Validity of Personality assessment	21
Reliability of Personality assessment	26
Conclusion	27

Chapter 3: Method	28
Research design	28
Participants	28
Roles	29
Grade levels	29
Gender distribution	30
Racial Composition	31
Education levels	31
Measuring Instruments	32
Performance Management system	32
OPQ 32i Assessment tool	34
Psychometric Properties of the OPQ 32i	35
Reliability of the OPQ 32i	35
Validity of the OPQ 32i	37
Procedure	38
Chapter 4: Results	41
Job Performance	41
Administration job family	41
OPQ 32i Personality dimensions	43
Administration job family	43
Reliability of the OPQ 32i	45
Validity of the OPQ 32i for the Administration job family	46
Job Performance	47
Finance job family	47
OPQ 32i Personality dimensions	48
Administration job family	48
Validity of the OPQ 32i for the Finance job family	50
Conclusion	51
Chapter 5: Discussion, implications and recommendations	53
Criterion	53
Validity	53

Reliability	54
Predictor	55
Validity	55
Reliability	57
Implications	58
Limitations	58
Recommendations	60
Conclusion	61
Reference List	63

University of Cape Town

LIST OF TABLES

3.1	Grade levels - Administration and Finance job families	30
3.2	Racial composition - Administration and Finance job families	31
3.3	Education levels of employees in the Administration and Finance job families	32
3.4	Descriptor of Performance Rating scale	34
3.5	OPQ 32i Scale Dimensions	35
3.6	Internal consistency of the OPQ 32i Subscales	36
3.7	Validity of the OPQ 32i scales	38
4.1	Descriptive statistics of Performance ratings for the Administration job family	41
4.2	Descriptive statistics of Relationship with People scales in the Administration job family	43
4.3	Descriptive statistics of Thinking style scales for the Administration job family	44
4.4	Descriptive statistics of Feelings and Emotions scales for the Administration job family	44
4.5	Internal consistency of the OPQ 32i Subscales	45
4.6	Correlation between the OPQ 32i dimensions with Performance for the Administration job family	46
4.7	Descriptive statistics of Performance ratings for the Finance job family	47
4.8	Descriptive statistics of Relationship with People scales for the Finance job family	49
4.9	Descriptive statistics of Thinking styles scales for the Finance job family	49
4.10	Descriptive statistics of Feelings and Emotions scales for the Finance job family	50
4.11	Correlation between the OPQ 32i dimensions with performance for the Finance job family	51

LIST OF FIGURES

3.1	Gender Composition- Administration and Finance job families	30
4.1	Performance ratings in the Administration job family	42
4.2	Performance ratings in the Finance job family	48

University of Cape Town

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Chao Mulenga for directing and providing feedback towards the completion of this dissertation. I would also like to thank Associate Professor Burger van Lil, for technical assistance throughout the research process.

I would like to acknowledge Chris Swart and his team from the assessment department at the financial services institution, where this study was conducted, for the support with data requests and additional assistance.

University of Cape Town

ABSTRACT

The dearth of studies available examining the personality–performance relationship have raised criticisms about the predictive validity of personality assessment in recruitment and selection. The Employment Equity Act (Act No.55 of 1998) stipulates the use of scientifically proven, valid and reliable assessment instruments. This study investigated the validity of the Occupational Personality Questionnaire (OPQ 32i) as a valid predictor of job performance at a large financial services institution in South Africa. The extent to which specific personality dimensions could be correlated with job performance was determined. The sample participants comprised 132 employees, performing different roles, across different grade levels in the Administration and Finance job families. The performance rating was the criterion measure against which the predictive validity of the OPQ 32i was measured. This study produced low indices of validity between the criterion and the predictor. The OPQ 32i subscales produced high internal consistency, demonstrating the reliability of the OPQ 32i as an assessment tool. The results do not support previous findings of specific personality dimensions being valid predictors of performance across job categories. The numerous limitations have however highlighted implications for future research, particularly for the human resource performance appraisal process. The need for the identification and measurement of specific personality dimensions during the performance appraisal has been suggested, as well as the use of more than one criterion measure to improve the reliability estimates of the criterion.

CHAPTER 1: INTRODUCTION

Efficiency and economies of scale were two dominant themes in South Africa during the twentieth century. These have been replaced by teamwork, global markets and customer-centric business models (Grobler, Wörnich, Carrell, Elbert, & Hatfield, 2002). South Africa is no longer protected from global production and trade conditions (Cooper & Robertson, 2000); a new European economy, expanding Asian markets and a reformed central Europe, pose competitive challenges to South African organisations. To be successful and sustainable in this global economy, organisations need to be productive and flexible, employing a competently skilled labour force (Grobler et al., 2002). This requires good organisational sourcing, recruitment and selection strategies.

Psychological tests are an inexpensive and time economical means of attaining information which can be used to screen candidates who do not meet the job performance requirements (Bethell-Fox, 1992). Psychometric assessment quantifies three critical individual differences; motivation, ability and personality. By matching individuals with the appropriate aptitudes and skills to the right roles in an organisation, psychometric assessment enables the organisation to achieve its objectives of development, implementation and execution of its strategy (Boerlijst & Meijboom, 1992). The validity of cognitive measures as valid predictors of job success has been confirmed (Guion & Gottier, 1965; Hogan, Hogan & Roberts, 1996; Ree, Earles & Teachout, 1994; Sackett, Gruys & Ellingson, 1998). The use of personality assessment has however vociferously been challenged.

The validity of personality assessment instruments in the employment context is two-fold; they investigate human nature to explain the features which characterise human performance, and they explain individual differences identifying the critical dimensions of human performance (Hogan & Shelton, 1998).

In spite of this understanding of personality, and how personality assessment tools can assist in providing more holistic information about applicants, there are many challenges to their predictive validity in recruitment and selection processes. For instance, poorly defined personality dimensions prevent employers from being cognisant of the specific personality dimensions which influence performance (Guion & Gottier, 1965). A related concern is the existence of low validity coefficients between personality measures and workplace measures (Jackson & Corr, 1998). Another problem is that, when responding to personality assessment tools, test takers are able to fake and distort their responses thereby reducing the accuracy of the responses, and the subsequent validity of the results (Hogan, Hogan & Roberts, 1996; Kluger, Watson, Laidlaw & Fletcher, 2002; Schwab, 1971). The inclusion of some test items which respondents perceive as intrusive and demonstrating no relation to the employment situation also poses a challenge (Harland, Rauzi & Biasotto, 1995; Rosse, Miller & Stecher, 1994). This can affect test takers' motivation levels, which will negatively influence their performance on the measures (Robertson & Kandola, 1982). Finally, there are cross-cultural concerns where assessment measures are used in multi-cultural environments, as the risks of bias and adverse impact of the measures are then heightened (Huysamen, 2002; Schmit, Kihm & Robie, 2000).

This study's aims have included:

1. To assess the ability of the Occupational Personality Questionnaire 32i (OPQ 32i) to measure and predict future job success
2. To determine the extent to which the OPQ 32i's specific personality dimensions can be correlated with job criteria performance success.

The rationale for investigating the predictive validity of the OPQ 32i lies in personality assessment validity studies conducted (Barrick & Mount, 1991; Dakin, Nilakant & Jensen, 1994; Day & Silverman, 1989; Goffin, Rothstein & Johnston, 1996; Hollenbeck & Whitener, 1988; Jackson & Corr, 1998; Robertson & Kinder, 1993), which outlined how, in spite of the challenges to

and concerns about using personality assessment, understanding personality dimensions provides a more holistic understanding of individual behaviour. This in turn informs recruitment and selection decisions which require predictions on future performance.

The Occupational Personality Questionnaire (OPQ) 32i is a 32-scale model of personality, which describes people's preferred styles of behaviour when at work (SHL, 1999). Validation studies conducted by SHL (2004a) in the insurance sector, based on studies across different organisations and industries, attest to it being a valid predictive assessment instrument, with good internal consistency.

This introduction has outlined the changing environmental context within which South African organisations are forced to operate and which necessitates a skilled labour force. The validity of personality assessment has been outlined and five specific challenges to its use summarised. The next chapter of this dissertation (Literature Review) provides an overview of the existing literature on personality assessment, its ability to add predictive value to recruitment and selection decisions in organisations, and a brief discussion of criticisms of its use. The challenges of assessment in a multi-cultural context will also be outlined. The testable hypotheses developed from the review of the literature will be covered. Chapter 3 (Method) will deal with the statistical techniques used to test the hypotheses, detailing the participants, the measuring instrument, and the research procedure. Chapter 4 (Results) presents the analysis of the data, outlining the correlation between the personality dimensions of the OPQ32i and the performance measures in the organisation. In Chapter 5 (Discussion) the results are summarised and incorporated with existing studies. In addition, the limitations and shortcomings of this study are addressed, with recommendations for future studies outlined.

CHAPTER 2: LITERATURE REVIEW

This chapter will review psychological assessment, focusing on the role of understanding personality in the work context, and on criticism of the validity of personality assessment in decision-making processes in the work environment. Concerns regarding the validity and reliability and freedom from bias of assessment measures in a multi-cultural context will be summarised. Finally, the predictive validity of personality measures as reported by different studies will be discussed, and their implications outlined.

Psychological Assessment

Globalisation has changed the world of work, requiring employees to possess different traits and skills and to be more flexible and empowered (Paterson & Uys, 2005). International and local organisations need to ensure that their recruitment and selection policies are effective in attracting and selecting employees with the appropriate skill sets. Bartram (2004) asserts that psychological assessment is able to assist employers in identifying appropriately skilled employees. It is most often used to aid in recruitment and selection decisions. When engaged in the process of selecting employees for an organisation, employers seek to identify individual differences between applicants so as to differentiate and assist in making the correct appointment decisions. Two determinants are considered in the decision-making process: the ability to match the individual to what is required for the job in terms of personal characteristics (as obtained through job analyses processes), or matching the individual according to the specific task requirements and performance outputs of the role (Cascio, 1998; Elkonin, Foxcroft, Roodt & Astbury, 2001).

To enable effective recruitment and selection decisions, a more holistic understanding of how people think and behave is required. A holistic

approach provides a wealth of information on how individuals interact with others at work, as well as how they go about performing their work (Cascio, 1998; Searle, 2003).

Types of measures

Different assessment tools, which follow different formats and can be used for groups or individuals, are available for use by organisations (Gregory, 2000). Psychological assessment is defined by Fernández-Ballesteros (1999, p. 248) as "the discipline of scientific psychology devoted to the study of a given human subject (or group of subjects), in a specified applied field (clinical, educational, work etc.), by means of scientific tools (tests and other measurement instruments), with the purpose of describing, diagnosing, predicting, explaining or changing the behaviour of that subject". The assessment tools include intelligence, achievement, motivation, values, aptitude, creativity, personality assessment and interest inventories (Gregory, 2000). When selecting employees, organisations consider the personal interests and values of applicants. These can be measured using interest and values inventories, where employees select their preferred activities as well as the values relating to moral, spiritual and religious matters which they consider important (Gregory, 2000).

Five main occupational ability-testing categories are identified by Searle (2003). These include verbal, numerical, spatial, dexterity and sensory tests. Many of these ability tests are job specific, with not all the abilities tested in all jobs. Cognitive tests are defined by Cooper and Robertson (2000) as providing insight to the intellectual ability of respondents, where general intelligence as measured by "g" and its sub-components are assessed. The sub-components include spatial, numerical and verbal ability. The validity of cognitive measures to predict performance has been validated by many studies (Hogan, Hogan & Roberts, 1996; Pearlman, Schmidt & Hunter, 1980;

Sackett, Gruys & Ellingson, 1998; Schmidt, Gast-Rosenberg & Hunter, 1980; Schmidt & Hunter, 2004).

Personality measurement constitutes another valuable assessment tool for use in the workplace. As defined by Hogan, Hogan and Roberts (1996, p.470) personality measurement is "any procedure that systematically assigns numbers to the characteristic features of a person's interpersonal style according to some explicit rules. These numbers can then be used to make predictions about the person's responses in future settings". This definition highlights two important components of good personality measures; their stability over time and their ability to enable the prediction of behaviour in real world performance situations. Hogan, Hogan and Roberts (1996) discuss how individual reputations are built on their past behavioural patterns, with past behaviour being the best predictor of future performance. This demonstrates how understanding the stable personality traits and behavioural trends of individuals has practical benefits for organisations. Studies confirming the predictive validity of personality measures will be outlined in the section detailing the validity of personality assessment.

The extent of the information which employers receive by utilising these tools within a limited time period as part of their recruitment and selection processes, is acknowledged by Bedell, Van Eeden and Van Staden (1999). Paterson and Uys (2005) cite two main advantages: the enabling of employers to make comparisons between applicants, and the provision of opportunities for applicants to receive feedback about their strengths, which can be enhanced.

This section has summarised psychological assessment, differentiating between the types of tools available. The next section will define personality, and consider how understanding and measuring personality could add value to recruitment and selection by enabling a more holistic understanding of individual behaviour and performance.

Personality and Measurement

Personality defined

Personality is defined as the reasonably stable feelings and behaviours of individuals' resultant from different genetic and environmental influencers and factors (Ivancevich & Matteson, 2002). Personality theorists concern themselves with understanding the influencers on these enduring characteristics. There are six different approaches to understanding personality and its measurement, each with their own context and application (Pervin, 1996; Searle, 2003). The Psychodynamic approach understands behaviour in terms of the unconscious and underlying motives of individuals. The Biological approach emphasises the biological and hereditary influences on individuals' personality development. The Behavioural approach focuses on how learning moulds behaviour, and on the influence of the external environment. Phenomenological and Humanistic approaches focus on the role of an individual's experiences and how their perceptions influence and mould their behaviours, reflecting on how individuals view themselves and depict meaning. The Social-Cognitive approach views the role of cognition as well as understanding the environment and situation as determinants of behaviour (Pervin, 1996; Searle, 2003).

The last of the personality theories is the Trait-based approach, with traits representing an individual's predisposition to behave in a specific manner, with specific consistent personality characteristics regarded as the determinants of individual behaviour (Pervin, 1996; Searle, 2003). These characteristics are believed to be stable across different contexts and environments, which therefore enable individual behaviours to be predicted. While the influence of situational factors is recognised, individuals are seen to have consistent behavioural patterns. Because of this, Robertson and Callinan (1998) argue that the Trait factor-analytic theory constitutes the building blocks of personality. They state that in order to understand personality in the

context of work, the level of interaction with colleagues, the team and the broader organisational environment needs to be considered.

There are several non-cognitive factors which influence job performance. Robertson and Callinan (1998) and Goldberg (1993) identify job proficiency, the actual job performed, work attitudes, individual commitment to roles or career, levels of stress in the work environment and overall sense of life satisfaction as important determinants of job performance. This is how understanding personality can assist. This recognises that performance at work is not solely dependent on cognitive ability, but also on personality. Understanding personality provides insights into not only what work people do, but also how they go about getting their work done, and how they interact with others (Hogan & Shelton, 1998). Criticism of the sole use of cognitive assessment in human resources processes centres on the failure to understand the role of cognitive ability, where for example, strong cognitive ability does not necessarily guarantee successful performance. Goldberg (1993) argues that when individuals lack personality traits that are relevant to job success, poor performance results. Situational factors are also important determinants, these including the role of colleagues, organisational reward strategies, line management, and the design of work, all of which contribute to performance (Robertson & Callinan, 1998).

Personality Inventories

Two types of personality inventories are predominantly used, namely objective and projective tests. Mello (1995) outlines their core purpose as determining the appropriate fit between a job applicant and the requirements of the role, or assessing employee promotion potential. Objective tests are often called "paper and pencil tests" requiring respondents to respond in a true or false, forced choice format, where respondents select their choice with no opportunity for further clarification. They are either normative or ipsative. Normative tools involve measuring a single personality construct along

different dimensions (Mello, 1995). Cattell's 16 Personality factor Questionnaire, the Minnesota Multiphasic Personality Inventory, and the Occupational Personality Questionnaire are examples of normative measures (Mello, 1995). The ipsative assessment tools require respondents to make a choice between two essentially different personality constructs (Van der Maesen de Sombreff & Hofstee, 1992). The more popular ipsative tools include the Gordon Personal Profile, Gordon Personality Inventory, Edwards Personal Preference Schedule (Mello, 1995). Ipsative measures do not enable comparisons between test takers on their relative standing on specific traits. They outline relative positioning of traits relative to other traits (Van der Maesen de Sombreff & Hofstee, 1992). Projective tests require respondents to make sense and interpret ambiguous stimuli, responding in an open-ended format, with the Rorschach test, Thematic Apperception Test the more common projective tests (Gregory, 2000).

This section has provided a summary of how personality assessment can be used in the workplace. The types of personality inventories available, and how they can assist in providing more holistic information than can be inferred from cognitive and other assessment measures has been outlined. In spite of these benefits, personality assessment has been criticised.

Criticism and challenges of Personality measurement

The use of personality assessment has not been free from controversy. Guion and Gottier (1965, p. 160) argue that , "the only acceptable reason for using personality measures as instruments of decision is found only after doing considerable research with the measure in the specific situation and for the specific purpose for which it is to be used. Sometimes, unvalidated personality measures are used as instruments of decision because of 'clinical insight' or of gullibility or superstition or of evidence accumulated in some other setting". The criticism is divided into concerns related to the validity of personality

measurement, cultural challenges, respondent concerns, measurement issues and sample concerns.

Validity Challenges

1. Criterion–validity of personality variables

When assessing job performance, concerns about the criterion validity of personality variables have been raised (Hogan, Hogan & Roberts, 1996; Hogan & Shelton, 1998). The need for clearly identified personality constructs which are regarded as being relevant job predictors have been highlighted. Cascio (1998) and Raymark, Schmit and Guion (1997) note how the job analysis process comprehensively covers the ability and aptitude components in terms of job requirements, but that personality variables tend to be less well-defined, resulting in personality dimensions being regarded as unimportant for selection.

It is for this reason that Barrick and Mount (1991) and Cascio (1998) argue the need for recruiters to know what personality dimensions to look for through utilising job analysis task information. Hogan, Hogan and Roberts (1996) recommend the use of multidimensional inventories, encouraging recruiters to rely on different scales. They caution against, for example, recruiting individuals with high integrity scores on the basis that they will follow rules and instructions, without considering that they could prove inflexible. Similarly, while conscientiousness is generally highly valued, it could be an indicator of an unwillingness to assist others and a lack of team orientation (Hogan, Hogan & Roberts, 1996).

2. Distortion of responses

Employees can play an active role in their destinies. Rosenfeld, Giacalone and Riordan (1995) discuss how employees seek to create favourable images and caution human resource managers against this "impression management". Studies conducted do not dispute respondent distortion and lying (Goldberg,

1993; Hogan, Hogan & Roberts, 1996; Kluger et al., 2002) and Kline (1995) outlines how social desirability pressurises respondents to respond in a manner which enhances how they may be perceived by others. Leary and Kowalski (1990) report distortion to take a number of forms; respondents positively enhancing their self-concept, responding to inventories according to what they understand the role to demand, and in so doing portraying themselves as ideal candidates for roles. This is obviously a matter of concern for organisations, as important decisions are based on information provided by assessment tools.

There is no verification of the information provided by respondents when they complete personality inventories, resulting in a limited likelihood that test takers will get caught (Rosse, Stecher, Miller & Levin, 1998). Personality assessment therefore provides an ideal opportunity for test takers to present themselves more favourably. "Under these circumstances it would be surprising if most job applicants did not fake some of their answers" (Rosse et al., 1998, p. 635). This does not reflect favourably on the integrity of respondents and poses a risk to organisations using such inaccurate or fabricated information. When respondents complete personality assessments which are perceived to demonstrate high face validity, they are tempted to fake their responses to meet the candidate description they believe recruiters are seeking.

These distortion concerns were countered by Furnham and Drakely (2000), by articulating that for candidates to be able to fake their responses, they need to understand how each item is related to the subscales and how to respond favourably to each of them. They show that the ability to predict scale scores does not automatically translate to an ability to fake responses, and point out that where test items demonstrate construct, concurrent and predictive validity it would be more difficult for respondents to predict and fake favourable responses. This therefore constitutes a remedy to the concern

about the ease with which responses on personality measures can be faked and distorted.

3. Failure to understand the role of motivation and ability

Personality inventories do not measure adequately the individual motivation levels that incite employees to perform their jobs well (Mello, 1995). This is important, as all organisations need to focus on understanding what drives their employees to perform at levels which will achieve successful business results over time. People may have the cognitive ability to perform, as well as the predisposition to behave in a certain way (personality), but it is essentially their levels of motivation that will determine the discretionary effort which will be employed to achieve the organisation's goals (Frank, Finnegan & Taylor, 2004; Goldberg, 1993).

4. "Testwiseness"

Testwiseness is defined as the stable skill developed by respondents when taking a test, where test takers respond to the stimuli presented in a test in a positive and desired manner (Hess & Neville, 1997). This also results in test takers distorting their responses by faking and managing the favourable impression they present. Hess and Neville (1997) outlined two dimensions of testwiseness; candidates improving test-taking skills through familiarity with the process, and candidates exploiting their newly acquired skills to alter their responses according to their understanding of the purpose of the test content. As personality assessment information is used in conjunction with recruitment and selection information (interviews), the potential inaccuracy of fabricated and distorted information about candidates owing to their testwiseness, puts the organisation at risk of making incorrect selection choices.

5. Personality–organisational fit model

The interaction of peers, line management, the work environment, and reward structures are a few of the determinants of individual behaviour and performance (Robertson & Callinan, 1998). Edwards (1991) affirms this in his

observation that "the person and the job operate as joint determinants of individual and organisational outcomes". Kristof (1996) explores this in his supplementary versus complementary fit model, and the need-supplies and demands-abilities model dimensions. The supplementary fit dimension involves the observation of similar characteristics across individuals and the group in a specific job environment. Complementary fit occurs when the individual brings their own abilities and specific attributes to a role and the organisation, which positively enhances the organisation. The needs-supplies dimension relates to individuals' needs, goals, interests and values being fulfilled by the organisation. This is different from the abilities-demands dimension, where the individual possesses the skills and abilities, education and experience which are required to be successful in the role.

The needs and abilities dimension provide the "person fit", with the supplies and demands dimension being the "job fit". In the sphere of recruitment and selection decisions, Kristof's model (1996) highlights the importance of the relationship between the individual and the situation to enable the achievement of job outcomes, and demonstrates that personality alone is not the only determinant of job success.

6. Cultural concerns

Industrial psychologists need to ensure that the tests they use are not discriminatory against specific groups, and are cognisant of the issues in assessment in a multi-cultural context (Huysamen, 2002). Test administrators need to be well trained and sensitised to assessment in a multi-cultural society, ensuring appropriate test norms are used when assessing respondents (Bedell, Van Eeden & Van Staden, 1999; Fernández-Ballesteros, 1999; Oakland, 2004; Paterson & Uys, 2005). This section will be outlined in more detail in the psychological assessment in South Africa section.

Respondent concerns

Respondents cite many concerns when they complete personality inventories (Ones & Viswesvaran, 1999). The structure of the questions in personality assessment measures does not openly demonstrate the intent of the question, meaning that the test taker has limited control and awareness in terms of the kind of information they are providing (Harland, Rauzi & Biasotto, 1995; Rosse, Miller & Stecher, 1994). The perceived lack of relevance (face validity) of the test items of the personality inventories contributes to respondents not feeling a sense of control of their responses. Test takers regard personality assessment as defective, demonstrating no clear logic (Smither, Reilly, Millsap, Pearlman & Stoffey, 1993).

Candidates are rarely allowed opportunities to provide feedback on their responses, which may have an influence on how their scores are then interpreted. Because test takers do not always know the intent (face validity) behind the questions, their negative perceptions of personality assessment are exacerbated (Gilliland, 1993; Harland, Rauzi & Biasotto, 1995). Test items are perceived to be subjective and indirect, focusing more on personal information as opposed to work related issues. It is for this reason that their accuracy and relevance is questioned (Harland, Rauzi & Biasotto, 1995; Hogan, Hogan & Roberts, 1996). Respondents have criticised the personal nature of some of the questions, experiencing the sensitivity of the questions intrusive and showing disrespect. Test takers feel uncomfortable answering the questions, regarding some test items as an infringement on their privacy (Harland, Rauzi & Biasotto, 1995; Hogan, Hogan & Roberts, 1996; Mello, 1995; Orpen, 1974; Rafaeli, 1999; Rosse, Miller & Stecher, 1994). This has at times resulted in candidates feeling resentful when required to complete these tests merely for employment or leadership development purposes, at the expense of compromising their dignity and rights.

Measurement issues

The structure and design of personality inventories pose certain measurement issues. Kline (1995) discusses test takers completing inventories by agreeing with the statements, even though they may not have understood the content of the questions posed. This response set of acquiescence is problematic as incorrect inferences could be made from how candidates have responded. The scales of some personality inventories have a middle or uncertain category, which test takers can use when they are ambivalent or unsure about specific questions (Kline, 1995). When respondents select this middle category option too often, very little can be inferred about their behaviour. This inhibits meaningful understanding of their responses, and the ability to distinguish between individuals. A related concern to the middle category scoring technique is that of test takers who rate extreme scores for items (Kline, 1995).

It is important that the above concerns are clearly understood when reviewing personality data and assessing applicant personality profiles, as they distort the accuracy of the profile. Such distortion has a negative implication for recruitment and selection decisions.

Sample size concerns

When tests of adequate statistical power are used, more accurate inferences around the validity will be estimated (Aron & Aron, 1994; Cohen, 1988; Hollenbeck & Whitener, 1988). Sample size and effect size are the two determinants of statistical power. Aron and Aron (1994) stipulate how when using a large sample, the standard deviation of the distribution of means is reduced; hence, when comparing the means of two populations, there is less overlap in the distribution of the means with a larger sample.

The more participants introduced for a given effect size, the greater the enhancement of the statistical power. Cohen (1988) discusses how as sample

sizes increase, errors are reduced, thereby improving the precision and reliability of results. Increased sample size enables the probability of identifying the phenomenon being tested (Cohen, 1988). "Large samples are not advocated because large numbers are good in and of themselves. They are advocated in order to give the principle of randomization, or simply randomness, a chance to 'work' (Kerlinger, 1986, p. 119)." Validity studies of personality assessment have made these sample size errors, which have prevented meaningful inferences to be made of the results. These studies will be outlined in more detail later.

This section has provided a summary of the central challenges to the use of personality assessment. These challenges are magnified when assessment is conducted in a multi-cultural context.

Psychological assessment in South Africa

The need to establish the reliability and validity of assessment instruments before they are used in the working context is emphasised by the South African Employment Equity Act, 1998 (Act No. 55 of 1998). Wallis (2004) argues that the spirit of this law focuses on the overall ethical use of assessment measures in any context. She accentuates the importance of ensuring assessment tools measure what they intend to measure, so that accurate inferences can be made. This section will summarise the history of psychological assessment in multi-cultural South Africa, addressing assessment issues in a cross-cultural environment, and identifying measures which can reduce bias.

A brief history

Psychological assessment in pre-1994 South Africa was developed in a racially segregated society (Bedell, Van Eeden & Van Staden, 1999; Foxcroft, Roodt & Abrahams, 2001). The psychological assessment movement in South Africa was characterised by the standardisation of measures for Whites only.

Measures were administered across different groups in the absence of analysing the existence of potentially biased measures. The judgments from test results did not take cognisance of the socio-economic, environmental and cultural dimensions which influence test performance, and international measures were used without making adaptations to ensure they were relevant and suitable for the South African population (Foxcroft, Roodt & Abrahams, 2001). Studies conducted during this period of political, economic and racial segregation used results, from example, intellectual ability tests to further emphasise the superiority of certain groups (Foxcroft, Roodt & Abrahams, 2001; Macleod, 2004). As different jobs were reserved for White and black employees, there was no perceived reason for different tests to be developed for the different cultural and racial groups.

Psychological assessment in post-Apartheid South Africa

In the 1980 and early 1990 periods when changes to the political situation were imminent and job reservation was being removed, there was increased pressure to investigate the use of different measures for different groups. Separate norms were developed to enable the comparison of performance of individuals relative to the appropriate norm groups (Foxcroft, Roodt & Abrahams, 2001). Measures previously designed for White South Africans, including the international measures, would be used with the different racial groups even though there were no relevant norms. The administration of these measures was to be administered "with caution" (Foxcroft, Roodt & Abrahams, 2001, p. 25).

Legal framework

When administering tests to different groups, cross-cultural understanding is important to ensure measures are not discriminatory (Bedell, Van Eeden, Van Staden, 1999; Fernández – Ballesteros, 1999; Oakland, 2004; Paterson & Uys, 2005). The introduction of the Employment Equity Act, 1998 (Act No. 55 of 1998) has legislated the protection of individuals against discriminatory

practices in psychological testing. The act states, "Psychometric testing and other similar assessments of an employee are prohibited unless the test or assessment being used:

- a) has been scientifically shown to be valid and reliable
- b) can be applied fairly to employees, and
- c) is not biased against any employee or group"

This is important because "psychological tests can act as a disabling factor, if the test is inappropriately applied, or used in isolation without verifying the results against other measures" (Paterson & Uys, 2005, p. 18). For this reason the policy on the Classification of Psychometric Measuring Devices, Instruments, Methods and Techniques was introduced, which reflects the stance of the Psychometrics Committee of the Professional Board for Psychology regarding test classification (HPCSA, n.d.). It outlines why tests need to be classified, the control of test use, the process to be followed for new measures or existing measures to be adapted, stipulating the existence of the psychometric properties validity, reliability and no bias, as stipulated in the Health Professions Act, Act 56 of 1974. The International Guidelines for Test Use (2001) developed by the International Test Commission, provides a set of guidelines for international use governing what constitutes good test practice. It focuses on the rights of the test taker, the competencies required by practitioners administering tests, test administration and scoring, and report writing and giving feedback to candidates. These guidelines were introduced because of the different levels of statutory control and requirements for test administration in different countries.

Cross-cultural concerns

The importance of new test development in a multi-cultural context is highlighted by Foxcroft (2004) and Oakland (2004). Test developers are required to be cognisant of cultural relevance and bias of test items, as well as of administration and language issues appropriate across different cultural

groups. Meiring, Van de Vijver, Rothmann and Barrick (2005) point to bias and equivalence as critical considerations in the use of psychological assessment in a multi-cultural context such as South Africa. Bias is defined as nuisance factors such as unwanted systematic factors across cross cultural scores which are not consistent with the underlying trait or ability being measured (Meiring et al., 2005). Equivalence results from the presence of the nuisance factors, which prevent score comparisons across cultures (Meiring et al., 2005). Fernández-Ballesteros (1999) recommends that test administrators not only possess adequate educational qualifications and undergo appropriate training but also that they are sensitised to test use in multi-cultural contexts. This is reaffirmed by Foxcroft (2004, p. 10) who states that "test content is closely aligned to the cultural group for which the test is developed as well as the cultural background of the test developer". Shuttleworth-Jordan (1996) discusses how appropriate norm data need to be used for particular populations, but adds that South Africa should not automatically reject all tests developed internationally just because they have not been designed for a particular group of people. She outlines the acculturation process, where the dynamic nature of South Africa's socio-cultural circumstances is resulting in different positioning along this continuum, dependent on the degree of urbanisation, Westernisation and education levels.

There are consistent cognitive processes existing across all human beings, so the rejection of tests because they are not standardised to the South African multi-cultural society, is considered extreme (Shuttleworth-Jordan, 1996). She distinguishes between racial and socio-cultural differences, with racial differences encompassing ethnic composition, and socio-cultural differences including education, socioeconomic status and language usage, all of which contribute to assessment performance. As educational levels and language usage improve this acculturation will result in changes to achievement on cognitive tests. Her arguments emphasise the significance of education and socio-economic background as determinants of test performance.

When focusing on cross cultural test usage, Bedell, Van Eeden and Van Staden (1999) argue that special focus needs to be geared towards issues of race, language, socio-economic status and educational background, with specific attention being paid to language and cultural differences. Language and culture are interwoven, so when conducting personality assessment, ensuring the accurate meaning of constructs is important (Wallis, 2004). Language and culture are seen as important moderators when assessing test scores. This is supported by Foxcroft (2004) who stresses the importance of understanding whether test performance demonstrates the true score of the test taker, or the result of their competence in the language the test was conducted in.

Status of test use in South Africa

Studies conducted by Paterson and Uys (2005) investigating the status of psychological assessment in South Africa, formed part of a broader Psychological Needs Analysis study conducted by the Human Sciences Research Council (HSRC). Conducted over a 5-month period in 2004, the study focuses on the present use and future trends in psychological assessment. The Paterson and Uys (2005) study identifies a general lack of awareness among test administrators and practitioners of whether or not tests are registered with the Health Professions Council of South Africa (HPCSA). Also noted is the tendency for organisations to select the latest well-advertised international assessment tools, a reason for this being that the competition for skilled resources encourages employers to seek tools that are believed to provide additional information to aid their recruitment and selection decisions. This preference for international over locally developed tests can also be ascribed to their relative accessibility. HPCSA tests are registered, their use is controlled and requires the presence of a registered psychologist; unregistered tests without the same restrictions are deemed easier to use.

These findings are of great importance within the context of a society undergoing radical economic and social transformation. The changes in the labour market has significant implications for recruitment in organisations, that must ensure that workforce profiles are representative of the population. Compounding this is the need for South African organisations to compete globally, which necessitate the employment of skilled specialists. To ensure that persons with the right skills are employed, assessment tools need to be not only valid predictors of success, but also culturally fair.

Psychometric properties of assessment tools

Assessment results are but one of the means of providing additional information which can assist in, and improve the decision making process. Foxcroft and Roodt (2001) highlight the importance of guarding against measurement errors inherent in the assessment process, such as the test administration and the scoring and data interpretation, all of which influence how data can be used and interpreted. For this reason, psychometric assessment needs to be viewed as additional information assisting line managers to make decisions based on predictions of the future performance of employees. The need for results to be both accurate and trustworthy is a given in the workplace. For this reason, the importance of validity and reliability in test administration cannot be overstated (Searle, 2003). The validity of an assessment tool refers to the appropriateness of what is being measured, and reliability refers to the tool's accuracy. The description of these concepts will be expanded further, citing specific studies which have endorsed the use of personality assessment.

Validity of Personality assessment

Validity refers to the extent to which a measure fits the conceptual and operational definitions of the construct and whether the instrument is suitable for the purpose for which it is intended (Terre Blanche & Durrheim, 1999). Validity is determined in four different ways; face, content, construct and

criterion-related validity (Terre Blanche & Durrheim, 1999). A test has face validity when on appearance, it measures what it is supposed to measure (Searle, 2003). A criticism of face validity is that test takers are able to distort and fake their responses, as they are able to understand the intent of the test items (Goldberg, 1993; Hogan, Hogan & Roberts, 1996; Kluger et al., 2002).

Content validity refers to the selection tool encompassing the scope of a conceptual domain. Terre Blanche and Durrheim (1999) and Bethell-Fox (1992) cite ability tests (where test takers are required to indicate their competence in a specific area) as a good example.

Determining construct validity of a measure involves theoretically and empirically assessing how related a specific measure is to other measures, which have similar theoretical foundations (Bethell-Fox, 1992). In discussing the construct validity of a measure, Terre Blanche and Durrheim (1999) state that when measures are not theoretically related, they should not be empirically correlated, and when they are, the validity of the measure should be challenged. In the context of multi-cultural assessment research, construct validity is relevant (Wallis, 2004). Criterion-related validity is the final form. Here the extent to which a measure relates to an external criterion is investigated (Searle, 2003). There are two kinds of criterion-related validity namely; concurrent and predictive validity. Concurrent validity is determined where the performance of the test taker on a specific measure can be determined at the same time as assessment on another measure is completed, and is concerned with the current performance levels (Searle, 2003). Predictive validity is determined by investigating the performance of a test taker but at different points in time. This highlights the extent of the relationship between the test result and an identified criterion at a future point, such as performance in the workplace (Searle, 2003).

The use of personality assessment as a predictor of successful performance has been vociferously challenged by many researchers (Ghiselli, 1973; Guion

& Gottier, 1965; Reilly & Chao, 1982; Schmitt, Gooding, Noe & Kirsch, 1984). Guion and Gottier (1965) argue that personality tests widely used in industry, indicate an interest in predicting "will do" as opposed to "can do" behaviours in individuals. They are critical of the use of inappropriate criterion measures with limited hypothesis development, which contribute to the questionable validity of personality assessment. For personality assessment to add value, the criteria and purpose need to be clearly understood. Barrick and Mount (1991) cite the lack of good taxonomies for personality traits as the reason for the low validities reflected in studies. It is difficult to investigate and determine significant relationships between specific personality constructs and the relevant job performance criteria for specific occupational groups without clearly defined personality traits.

Personality assessment needs to be understood within a context, with assessment focusing on how people prefer to behave in specific contexts, as opposed to how they generally prefer to behave (Dakin, Nilakant & Jensen, 1994). In their study relating to recruitment for managerial jobs they comment on the difficulty of specifying situational factors within which individuals are required to operate. This is because such roles typically comprise much work and task variation and involve working under pressure and dealing with conflict situations. For this reason, it is felt that context is best described in general terms (Dakin, Nilakant & Jensen, 1994). This however, undermines the validity of using personality assessment as a sole predictor of performance.

Meta-analytical studies conducted by Barrick and Mount (1991) did not focus on defining the overall validity of personality measures as a valid predictor of performance, but instead, used specific personality dimensions as predictors for specific occupations. They hypothesised that specific "big five" personality dimensions could be correlated with specific occupations. Their study highlights how when aligning specific personality dimensions to specific occupations, the validities are enhanced. They confirmed their hypothesis that

conscientiousness is a valid predictor of job performance across all jobs and all job criteria. Dakin, Nilakant and Jensen (1994) summarising important pre determinants of the validity of personality assessment recommend that employers use job analysis information to outline the specific criteria of the job which they regard as important. In addition, the personality dimensions which are deemed related to the job need to be identified, and when analysing results the relevant personality dimensions need to be focused on, and the irrelevant discarded.

Similar to the Barrick and Mount (1991) study, Day and Silverman (1989) sought to determine the relationship between specific personality dimensions and a sample of accountants. The study covers orientation towards work, degree of ascendancy and degree and quality of interpersonal orientation as being related to job success. Their results confirmed the hypothesis, supporting Dakin, Nilakant and Jensen's (1994) summary of the importance of identifying specific personality scales with the relevant job success measures.

There are distinct requirements in different occupations, in terms of both cognitive ability and personality dimensions (Hollenbeck & Whitener, 1988). For example, they postulate that different traits are likely to be required for successful performance as an accountant versus those required by a fire fighter. In other words, that for different occupations, different personality dimensions are relevant.

Overall job performance is not just about technical competence requirements. Given that not all occupations allow individuals to function on their own (Cascio, 1998), there are also people requirements that need to be considered. Often the ability of people to work with others influences their performance and success, particularly in jobs where cooperation is required. In spite of the poor correlations reported in personality validity studies, Jackson and Corr (1998) investigated the discrepancy between the perceived validity of personality assessment and the actual use of personality

assessment, through understanding the validity coefficients from individual and then aggregate levels. They postulated that personality–performance correlations exist at aggregate levels (such as in organisations), but not at individual levels of analysis (with individual employees), suspecting that the differences in these correlations accounted for this discrepancy gap, where inferences were drawn for aggregate level analyses, and were then merely extended to individuals, and vice versa. They strongly suggest that test users investigate patterns of behaviour which relate to performance specifically at aggregate levels, as opposed to at individual data levels. They argue that, when aggregating, the magnitude of correlations between the different variables would be enhanced.

For this reason, individual and aggregate levels of analyses would draw different inferences, for example by talking about “how the best sales people are extroverts”, or “the best accountants are introverts” (Jackson & Corr, 1998). Moving from individual scores to the average scores across individuals, enabling an analysis of the average level of performance across each of the aggregate personality groups, is suggested by Jackson and Corr (1998). Their findings support the previous studies conducted by Schmitt et al. (1984). These aggregate correlations would then define the correlations between average groups of individuals and across performance. Their study results suggest that for individual levels of analyses there are no critical levels of significance, while at the aggregate level of analyses, most of the aggregate data correlations are large.

Taking cognisance of the criticism and challenges to the use of personality assessment, these studies demonstrate that when used properly, the results do add value in being able to predict success across different jobs.

Reliability of Personality assessment

Reliability is an important component in measurement as it assesses the degree to which results can be repeated and hence the accuracy of the tool enhanced (Searle, 2003; Terre Blanche & Durrheim, 1999). Reliability relates to scores on measures as well as the entire study. Three different approaches are applied to determining reliability; stability, consistency and equivalence (Searle, 2003).

Test-retest reliability is determined by repeating the measurement procedure for the same respondents using the same measurement tool, under identical conditions (Hair, Babin, Money, & Samouel, 2003). The results from the separate test administrations should yield the same outcome. Assessing the consistency of measures is necessary to establishing that the difference in results yielded is not due to external factors. Two approaches to measure consistency are used; split half testing and Coefficient alpha or Kuder Richardson reliability. The equivalence of results investigates the reliability of a measure by developing two equivalent forms of the same construct and measuring test taker responses during the different administrations, using the different forms of the construct (Hair et al., 2003).

This section has outlined the psychometric properties which are required for assessment measures to be valid and reliable. From the different validity studies reviewed, there is evidence to support the use of personality assessment in organisations. The introductory chapter outlined the guiding aims for this study.

Based on the literature reviewed, the hypotheses for this study are:

1. Specific personality constructs are more important than others to enable job success and competence
2. The OPQ 32i is a valid and reliable predictor of job performance

Conclusion

The review of the literature has provided an overview of the development of psychological assessment in South Africa. The validity of psychological assessment in contributing to effective decision-making has been discussed. Valid challenges to the use of personality assessment have been outlined. In spite of these challenges, validity studies conducted have demonstrated how, when used properly, the results are valid indicators of performance. The next chapter will outline the research design and method used in assessing the predictive validity of the Occupational Personality Questionnaire (OPQ 32i) in a large financial services organisation.

University of Cape Town

CHAPTER 3: METHOD

The literature review provides an overview of the role of psychometric assessment in organisations, specifically outlining the significance of understanding personality and the validity of personality assessment to assist employers in making predictions about employees' performance as part of the recruitment and selection processes. This chapter discusses the choice and explanation of the research design and method employed to investigate the predictive validity of the Occupational Personality Questionnaire (OPQ 32i). The sampling process will be outlined with a detailed description of the sample participants. The measuring instruments used to test the hypotheses will be discussed. At the end of the chapter, the research procedure followed to conduct this study will be summarised.

Research design

A relational research design was applied in this study (Haslam & McGarty, 2003). This design approach was appropriate to investigate the existence of a relationship between the predictor, the OPQ 32i scores, and the criterion, the performance ratings.

Participants

The participants in the sample were selected from secondary data available from the organisation. This secondary data included employee data in the form of employee job title, role, gender, race, job family, performance ratings and psychometric data. The employees were selected as participants on the basis of their being employed in the organisation between the periods June 2003 and March 2005. Being assessed using the OPQ 32i assessment tool formed the second criterion for sample selection. Given that this procedure was followed, the sampling technique described is a purposive non-probability sampling approach (Coolican, 2004).

To enable the predictor (the assessment tool information) to be correlated with the criterion (performance ratings), employees required a performance rating to be available. As the database of the psychometric assessment data was only available from the end of 2003, this 21 month period was selected. The organisation's Human Resources Information System provided the employee data for the 132 employees who were selected as sample participants.

Roles

Employees meeting the sample selection criteria were mostly in the Administration and Finance job families. Job families comprise employees who are in the specific family owing to the work they perform, where similar work outputs are performed, requiring similar competencies and qualifications. The employees in the Administration and Finance job families perform different roles in different business units in the organisation. These roles performed were not used as criteria to be selected in the sample, and were not meant to be representative of the respective job families.

Grade levels

The participants were employed in roles ranging from entry level to senior management. The organisation's 19-grade hierarchy ranges from Grade 17, which is the lowest grade, to two additional grades above grade 1 being the highest. Those two additional grades are for the organisation's executive leadership. Table 3.1 summarises the grade levels of the sample participants.

Table 3.1

Grade levels - Administration and Finance job families

Grade	Administration job family (n=60)	Finance job family (n=72)
4 – 7 (Managerial)	6	25
8 – 10 (First line supervisory)	23	35
11 – 12 (Section head)	30	12
13 -16 (Entry level)	1	0

Gender distribution

The Administration sample gender breakdown (n=60) comprised 32 percent male and 68 percent female. The Finance sample gender breakdown comprised 47 percent male and 53 percent male. Figure 3.1 provides a summary of the gender composition of the sample versus the gender breakdown of the respective job family population.

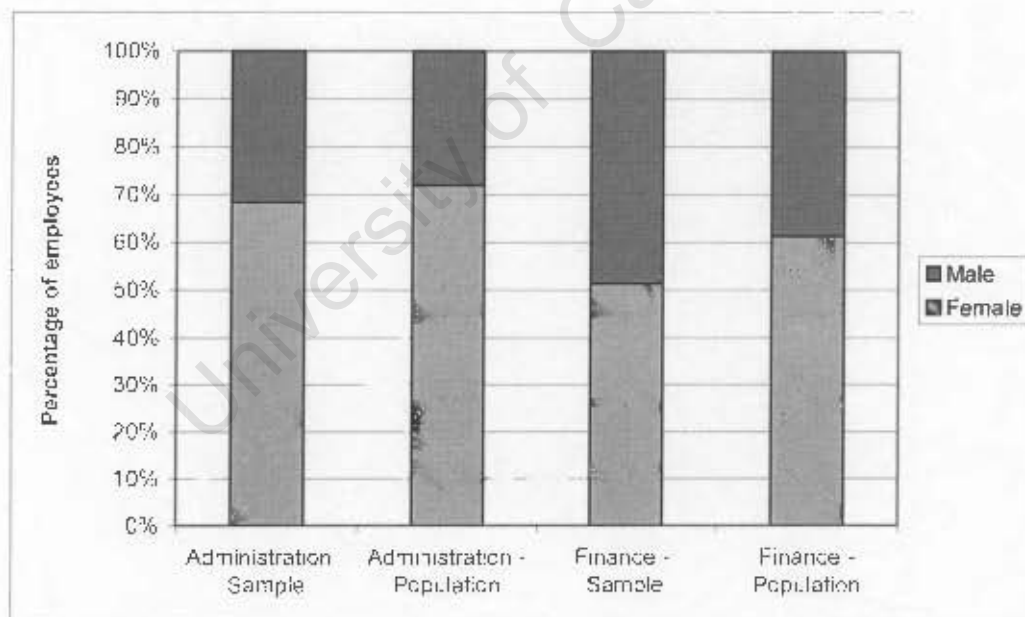


Figure 3.1. Gender composition- Administration and Finance job families.

Figure 3.1 illustrates the Administration job family sample to have a similar gender composition to the Administration job family population in the

organisation. In the Finance job family sample, the gender composition is more evenly distributed compared to the Finance job family population in the organisation.

Racial Composition

The racial composition for the two groups compared to the Administration and Finance job family population is summarised in Table 3.2.

Table 3.2

Racial composition - Administration and Finance job families

Race	Administration job family		Finance job family	
	Sample	Population	Sample	Population
Asian	10%	5%	7%	6%
Black	17%	28%	22%	14%
Coloured	53%	37%	39%	46%
White	20%	30%	32%	34%

Table 3.2 illustrates a relatively higher proportion of Coloured sample participants in the Administration job family relative to the job family population. The Finance job family population has a higher proportion of White employees than the sample participants. With these points noted, the table illustrates how the racial composition of the sample participants is representative of the respective job family populations.

Education levels

Different educational levels are required for the different roles, depending on the technical requirement and seniority of the role. The educational qualification of the sample is summarised in Table 3.3. The qualifications listed reflect the qualification at the time of appointment. Participants who were in the process of studying towards a qualification do not have those qualifications listed as completed.

Table 3.3

Education levels of employees in the Administration and Finance job families

Level of education	Administration	Finance
No school leaving certificate	1	1
Matriculation	31	12
Diploma	14	9
B.Degree	9	19
Postgraduate qualification	4	19
Professional qualification	0	12
Unsure	1	0

Measuring Instruments

Performance Management system

A performance appraisal process refers to the systematic description of the strengths and development areas of an employee, within the context of job specific outcomes (Cascio, 1998). Cascio (1998) affirms that job performance is multidimensional, covering a range of factors constituting performance. In the organisation the performance ratings indicate the performance level of employees, during a specific performance cycle. Two performance management reviews are conducted during the year.

The organisation's performance management system consists of three components; the individual component involves the work performance outcomes measured through key result areas agreed upon by the individual employee and their line manager. These key result areas are based on the cascading of the organisation's strategy down to business unit level, department level and then into individual employee performance contracts. This results in groups of employees in similar roles in different business areas having different key result areas. This individual performance component counts 65 percent towards the final total performance rating.

The second component of the performance management system is the team component, which is the individual rating of the divisional head of the business area. All employees in the same business area have received this rating. This team rating counts 25 percent towards the final total performance rating.

The third component of the performance management system consists of a 360 degree behavioural component. This is in the form of a questionnaire related to how employees' behaviours identify with the organisation's endorsed values. The respondents to the questionnaire are agreed upon by the line manager and employee, as the respondents need to represent a broad range of employees who regularly engage with the work of the employee (under performance review). This component counts 10 percent.

The performance review discussion conducted by the line manager can be concluded manually or electronically. The manual process involves the performance contract, with the line manager manually calculating the final performance rating based on the ratings from the three components. There is a facility for line managers to complete this process on the electronic performance management tool, requiring the performance contract and the result from the performance review to be inputted electronically. The electronic calculator calculates the final performance rating. It is this combined total rating (comprised of all three components) which is used for all other human resources practices, such as the annual salary review process. The performance rating scale applied to assess performance is summarised in Table 3.4.

Table 3.4

Description of Performance Rating scale

4.1 – 5.0	Exceptional performance against agreed standards and outcomes
3.1 – 4.0	Performance exceeds agreed standards and outcomes
2.1 – 3.0	Performance outcomes and standards met
1.1 – 2.0	Performance below agreed standards and outcomes
0.1 – 1.0	Performance well below agreed standards and outcomes

OPQ 32i Assessment tool

The different criterion components have been discussed. The next section will describe the predictor, in terms of the Occupational Personality Questionnaire 32i assessment tool.

The OPQ is a 32-scale model of personality which describes people's preferred styles of behaviour when at work. Robertson and Kinder (1993) summarise the OPQ as being similar to trait-based measures of personality with scales deductively developed to assess specific psychological traits for the working population. This makes it suitable for assessing job-relevant behaviour. Using feedback from organisations and personality theories, as well as input from employees in organisations, the general product title of "Occupational Personality Questionnaire" (OPQ) was developed (Barrett, Kline, Paltiel & Eysenck, 1996). An array of roles can be assessed and it is designed as an international model for the study of personality. The OPQ 32i assesses personality along three dimensions namely; Relationships with People, Thinking style and Feelings and Emotions, all within the work context. A fourth dimension, Dynamism, permeates across these three dimensions, focusing on achievement orientation and degree of competitiveness (SHL, 1999).

The OPQ was first designed in the United Kingdom and is used in more than 40 countries, and is available in more than 25 languages (SHL, 1999). Table 3.5 summarises the scales of the OPQ 32i.

Table 3.5

OPQ 32i Scale Dimensions

OPQ 32i Scale Dimensions		
Relationship with People	Thinking style	Feelings and Emotions
Persuasive	Data rational	Relaxed
Controlling	Evaluative	Worrying
Outspoken	Behavioural	Tough-minded
Independent-minded	Conventional	Optimistic
Outgoing	Conceptual	Trusting
Affiliative	Innovative	Emotionally controlled
Socially confident	Variety seeking	Vigorous
Modest	Adaptable	Competitive
Democratic	Forward thinking	Achieving
Caring	Detail conscious	Decisive
	Conscientious	Consistency
	Rule following	

The OPQ 32i is an ipsative objective test, requiring respondents to review sets of four statements, with respondents indicating which statement best and least reflects their typical behaviour or preference. Respondents review 104 sets of four statements from which a model of personality are dissected into the domains namely; Relationship with People; Thinking style and Feelings and Emotions and the fourth domain, Dynamism, which permeates across these three dimensions as noted earlier in the chapter (SHL, 1999).

Psychometric Properties of the OPQ 32i

Reliability of the OPQ 32i

Studies were conducted by SHL (1999) to measure the reliability of the OPQ32i. Different norm groups were established, consisting of a standardisation sample comprising 807 respondents from the United Kingdom. The sample participants included a managerial and professional sample and an undergraduate sample. Two thirds of the standardisation

sample was employed in different roles across seven different organisations and industries. The remaining third consisted of undergraduates from different educational institutions (SHL, 1999). Table 3.6 outlines the internal consistency estimates for this standardisation sample.

Table 3.6

Internal consistency of the OPQ 32i Subscales (N=807)

OPQ 32i Scale Dimensions					
Relationship with People	Cronbach Alpha	Thinking styles	Cronbach Alpha	Feelings and Emotions	Cronbach Alpha
Persuasive	0.81	Data rational	0.88	Relaxed	0.85
Controlling	0.87	Evaluative	0.67	Worrying	0.88
Outspoken	0.76	Behavioural	0.82	Tough-minded	0.82
Independent-minded	0.72	Conventional	0.74	Optimistic	0.80
Outgoing	0.85	Conceptual	0.79	Trusting	0.81
Affiliative	0.82	Innovative	0.88	Emotionally controlled	0.85
Socially confident	0.83	Variety seeking	0.72	Vigorous	0.75
Modest	0.81	Adaptable	0.82	Competitive	0.86
Democratic	0.68	Forward thinking	0.75	Achieving	0.79
Caring	0.78	Detail conscious	0.80	Decisive	0.80
		Conscientious	0.82		
		Rule following	0.84		

The alpha coefficients for the 32 scales were reported between 0.67 - 0.88, with only two scales falling below 0.70, namely Democratic (0.68) and Evaluative (0.67).

Studies were conducted by Baron and Miles (2002) to determine respondent differences between ethnic groups in the United Kingdom, to investigate trends. The results suggested differences between White, Black and Asian groups, but the lower reliabilities may be as a result of lower variability in

other demographic variables, and not differences owing to different respondent patterns. When the personality scales were combined the differences were cancelled out. The results favour the ethnic minority groups. This illustrates how using this personality measure may increase validity as well as reduce adverse impact.

A similar study conducted by Baron and Miles (2001) to investigate the structure of personality across different cultures, comprised 16 different countries and 12 languages. South Africa was included in this sample. The results indicate small score differences for each country for each of the scales. The similarity of the results indicates general individual differences.

Validity of the OPQ 32i

Validation studies were conducted by SHL (2004a) to determine the predictive validity of the OPQ 32i in the insurance sector. The sample participants were middle and senior managers at a large insurance organisation. The validity coefficients are summarised in Table 3.7.

Table 3.7

Validity of the OPQ 32i scales

OPQ 32i Scale Dimensions					
Relationship with People	r_{xy}	Thinking styles	r_{xy}	Feelings and Emotions	r_{xy}
Persuasive	0.11	Data rational	-0.04	Relaxed	0.03
Controlling	0.12	Evaluative	0.18	Worrying	-0.03
Outspoken	0.04	Behavioural	0.10	Tough-minded	0.05
Independent-minded	-0.05	Conventional	-0.14	Optimistic	-0.04
Outgoing	-0.02	Conceptual	0.08	Trusting	-0.03
Affiliative	0.01	Innovative	0.08	Emotionally controlled	-0.03
Socially confident	0.04	Variety seeking	0.04	Vigorous	-0.17
Modest	-0.01	Adaptable	0.09	Competitive	0.12
Democratic	0.06	Forward thinking	0.09	Achieving	0.12
Caring	-0.12	Detail conscious	-0.14	Decisive	0.01
		Conscientious	-0.16		
		Rule following	-0.37		

There were no significant correlations for these scales when correlated with the performance ratings.

Procedure

Ethical clearance was received from the University of Cape Town to conduct the research study. The research proposal with the ethics protocols, including a formal written request was submitted to the financial services institution requesting approval to conduct the study. Written consent to conduct the study was granted on the proviso that the assessment data remained the property of the organisation, and was only to be used for the purposes of this study.

To ensure confidentiality, it was agreed that the names of employees in the sample were removed from the data. Once ethical clearance was approved by the University and permission granted by the institution, the data collection process commenced.

The assessment department identified personality assessment records for test takers in finance roles between the October 2003 and March 2005 period. Finance candidates were initially identified as the sample group as this was a financial services organisation.

The assessment data of the test takers needed to be checked against the Human Resources Information system to establish whether or not those who completed an OPQ 32i assessment were appointed to the organisation. With the list of candidates assessed and recruited, their performance ratings needed to be checked against the records on the system.

Based on these three criteria, there was an insufficient number of Finance employees. To ensure a sample of 120 participants, additional employees in the Administration job family were selected as there was a higher ratio of employees recruited into the organisation in administration roles. When all information related to OPQ 32i assessment data, and performance data was available, a spreadsheet with all data was sent to me for scrutiny.

On receipt of this spreadsheet, further sample selection ensued. The organisation had converted to their new performance management system one and a half years before, and for this reason not all the business areas were using the same rating scale.

The data in the sample representing the old rating scale was excluded from the sample. Once these criteria were met, the data from 132 employees were included in the sample.

This chapter outlined the design of the study, describing the sample and measuring instruments used, and the research procedure. Chapter four will present the results from the study.

University of Cape Town

CHAPTER 4: RESULTS

To achieve the study objective of determining the predictive validity of the OPQ 32i as an instrument to predict performance, this chapter reviews the results of the data collected and analysed. The results are analysed and presented in two sections separately, covering the Administration job family and the Finance job family. Different statistical analyses were conducted, involving the Pearson's product - moment correlation (Coolican, 2004) and the reliability coefficient, using STATISTICA Version 7 and Microsoft Excel 2000.

Job Performance

Administration job family

The Methods chapter outlined the detail of the criterion (the performance ratings), highlighting the components of the organisation's performance management system, the performance review process, and the performance rating scale. The results of the data collected have been analysed for the Administration and Finance job families to enable the relationship between the predictor (the OPQ 32i assessment tool) measures to be correlated with the criterion measure. To determine the link between the OPQ 32i scales and job performance, it was important to analyse the level of performance of the employees as reflected by their performance ratings. These ratings are illustrated in Table 4.1.

Table 4.1

Descriptive statistics of Performance ratings for the Administration job family (n=60)

Administration job family statistics						
M	Median	Min	Max	SD	Skewness	Kurtosis
3.02	3.10	2.30	4.10	0.52	0.39	-1.23

The average and midpoint of the performance ratings were 3.02 and 3.10 respectively. These ratings have illustrated that the performance of the employees in the Administration job family in general exceeded the agreed performance standards and outcomes. Figure 4.1 illustrates the distribution of the performance in the Administration job family.

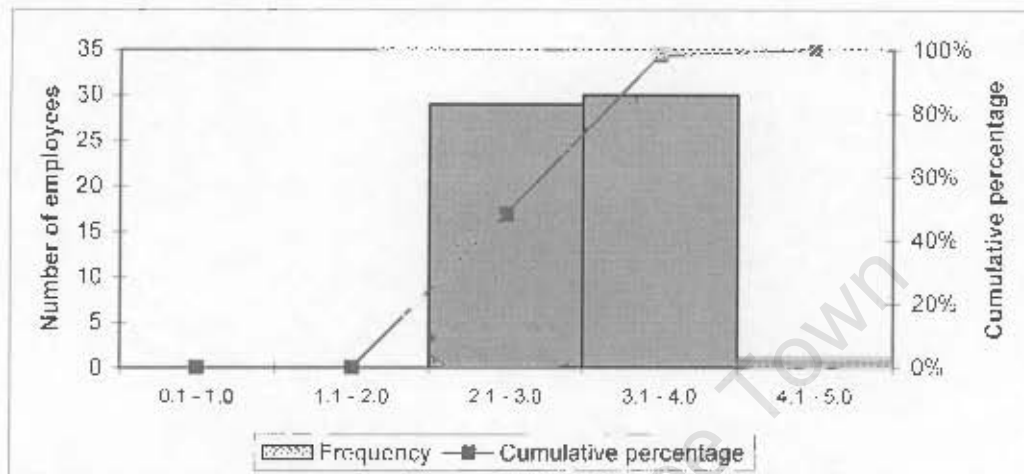


Figure 4.1. Performance ratings in the Administration job family.

Figure 4.1 illustrates how based on the performance ratings, the Administration job family have in general exceeded the performance standards. The scale category "performance outcomes and standards met" accounted for 48 percent of the sample, with 50 percent of participants receiving ratings in the "performance exceeds agreed standards and outcomes" rating category. The highest scale category, "exceptional performance against agreed standards and outcomes" had two percent of the participants with ratings in that category. The implication of these performance ratings, in terms of their being composed of three different components, will be highlighted in the Discussion chapter.

There were no reliability estimates for the performance ratings, as not all business areas conduct two performance reviews. Performance ratings from two different cycles were therefore not available. The intra-rater reliability of the line manager) as well as the consistency of the performance rating could

not be determined. Inter-rater reliability was also not possible. This limitation will be elaborated in the Discussion chapter.

This section has described the criterion information for the Administration family. The next section discusses the data collected for the predictor information.

OPQ 32i Personality dimensions

Administration job family

To illustrate the scores of the OPQ 32i scales for the Administration job family, descriptive statistics of the scales are outlined in Tables 4.2-4.4. The 32 scales were grouped into the three OPQ 32i dimensions namely: Relationship with People, Thinking style and Feelings and Emotions.

Table 4.2

Descriptive statistics of Relationship with People scales for the Administration job family (n=60)

OPQ 32i scales	Mean	Standard Deviation
Persuasive	5.30	2.22
Controlling	5.07	1.88
Outspoken	5.67	1.95
Independent-minded	3.95	1.72
Outgoing	4.35	1.45
Affiliative	3.75	1.51
Socially confident	5.52	1.38
Modest	5.80	1.91
Democratic	5.85	1.71
Caring	5.18	1.59

The highest Mean scores for the Relationship with People dimension were the Democratic, Modest, Outspoken and Socially confident scales.

Table 4.3

Descriptive statistics of Thinking style scales for the Administration job family (n=60)

OPQ 32i scales	Mean	Standard Deviation
Data rational	6.83	1.87
Evaluative	6.13	2.21
Behavioural	4.80	1.58
Conventional	6.65	1.73
Conceptual	5.47	1.66
Innovative	5.07	1.56
Variety seeking	4.27	1.96
Adaptable	3.97	1.96
Forward thinking	5.95	1.99
Detail conscious	6.90	1.54
Conscientious	7.27	1.25
Rule following	7.33	1.69

The Thinking style scales with the highest Mean scores were the Conscientious, Rule following, Detail conscious and Data rational scales.

Table 4.4

Descriptive statistics of Feelings and Emotions scales for the Administration job family (n=60)

OPQ 32i Scale	Mean	Standard Deviation
Relaxed	5.63	1.56
Worrying	4.55	2.05
Tough-minded	6.15	1.47
Optimistic	6.07	1.51
Trusting	5.07	1.89
Emotionally controlled	5.30	1.71
Vigorous	5.83	1.67
Competitive	4.63	1.83
Achieving	6.50	1.72
Decisive	5.12	1.80
Consistency	4.78	1.83

The Feelings and Emotions scales with the highest Mean scores were the Achieving, Tough-minded and Optimistic scales.

Reliability of the OPQ 32i

For reporting the reliability coefficients for the scales, owing to the small size of the two job families, the reliability coefficient was calculated for the combined group (N=112). Individual item results were only available for 112 of the 132 participants. The results are summarised in Table 4.5.

Table 4.5

Internal consistency of the OPQ 32i Subscales (N=112)

OPQ 32i Scale Dimensions					
Relationship with People	Cronbach Alpha	Thinking styles	Cronbach Alpha	Feelings and Emotions	Cronbach Alpha
Persuasive	0.83	Data rational	0.87	Relaxed	0.79
Controlling	0.86	Evaluative	0.69	Worrying	0.86
Outspoken	0.73	Behavioural	0.72	Tough-minded	0.61
Independent-minded	0.65	Conventional	0.69	Optimistic	0.68
Outgoing	0.76	Conceptual	0.73	Trusting	0.79
Affiliative	0.75	Innovative	0.77	Emotionally controlled	0.75
Socially confident	0.75	Variety seeking	0.70	Vigorous	0.69
Modest	0.75	Adaptable	0.80	Competitive	0.86
Democratic	0.65	Forward thinking	0.84	Achieving	0.69
Caring	0.72	Detail conscious	0.72	Decisive	0.76
		Conscientious	0.75		
		Rule following	0.78		

From the 32 scales, eight demonstrated reliabilities of below 0.70, namely; Independent-minded, Democratic, Evaluative, Conventional, Optimistic, Tough-minded, Vigorous and Achieving. The highest reliability coefficients

were for the Controlling, Worrying and Competitive scales, and the scale with the lowest reliability coefficient was the Tough-minded scale.

Validity of the OPQ 32i for the Administration job family

To determine the predictive validity of the OPQ 32i, correlation analysis was conducted. The results are outlined according to the different OPQ 32i dimensions.

Table 4.6

Correlation between the OPQ 32i dimensions with Performance for the Administration job family (n=60)

OPQ 32i Scale Dimensions					
Relationship with People	r_{xy}	Thinking styles	r_{xy}	Feelings and Emotions	r_{xy}
Persuasive	-.05	Data rational	.02	Relaxed	-.19
Controlling	.18	Evaluative	.13	Worrying	-.16
Outspoken	.08	Behavioural	.01	Tough-minded	-.06
Independent - Minded	.00	Conventional	-.17	Optimistic	.06
Outgoing	-.08	Conceptual	-.11	Trusting	.12
Affiliative	-.17	Innovative	.02	Emotionally controlled	-.03
Socially confident	-.07	Variety seeking	.07	Vigorous	.06
Modest	-.03	Adaptable	.09	Competitive	-.05
Democratic	.14	Forward thinking	.13	Achieving	.20
Caring	-.09	Detail conscious	.00	Decisive	-.00
		Conscientious	-.14		
		Rule following	-.25		

The Relationship with People scale dimensions reflected no significant correlations between the OPQ32i scales and performance. The second domain, Thinking style, there were no significant correlations. The third OPQ 32i domain, Feelings and Emotions, did not reflect significant correlations.

These validity coefficients reflect no significant correlations between the OPQ 32i scales and job performance in the Administration job family.

The first part of the Results chapter focused on analysing the results for the Administration job family. The next section focuses on the results of the Finance job family sample (n=72).

Job Performance

Finance job family

To determine the link between the OPQ32i scales and the actual job success of the finance employees, the performance of the sample participants were analysed, as reflected by their performance ratings.

Table 4.7

Descriptive statistics of Performance ratings for the Finance job family (n=72)

Finance job family statistics						
M	Median	Min	Max	SD	Skewness	Kurtosis
3.50	3.50	2.50	4.50	0.60	-0.2	-1.00

The performance ratings for the Finance job family were higher in terms of the performance rating scale than the Administration job family, (M=3.50) and midpoint (Median = 3.50). These results summarise the Finance participants as having in general performed at levels exceeding the agreed standards and outcomes. To further illustrate the distribution of ratings, Figure 4.2 provides a summary.

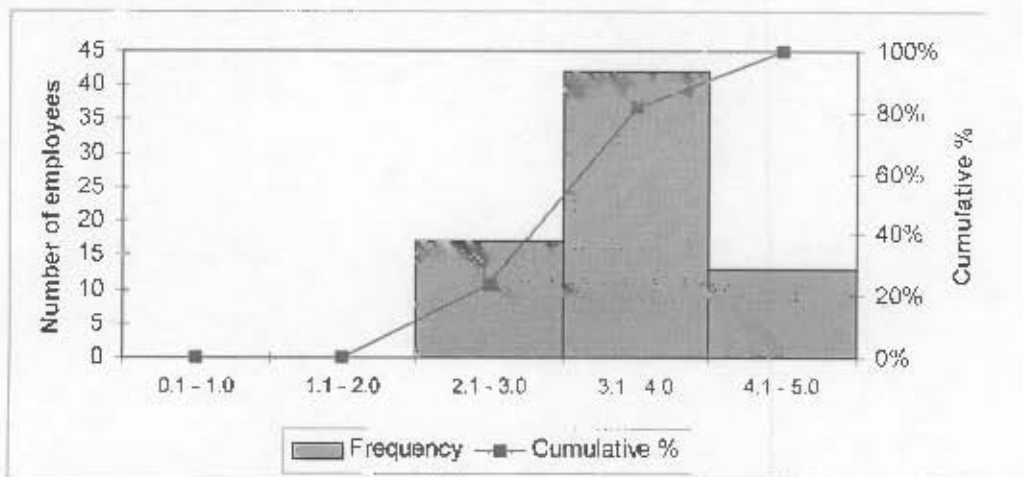


Figure 4.2. Performance ratings in the Finance job family.

No participants in the Finance job family received performance ratings in the “below or ‘well below’ agreed standards and outcomes” performance rating scale categories. In the category “performance outcomes and standards met”, 24 percent of the participants were rated, with 58 percent of participants receiving ratings in the “performance exceeds agreed standards and outcomes” rating category. The highest scale category, “exceptional performance against agreed standards and outcomes”, and 18 percent of the participants received ratings in this category. The implication of these performance ratings, in terms of their being composed of three different components will be highlighted in the Discussion chapter.

OPQ 32i Personality dimensions

Administration job family

To illustrate the scores of the OPQ 32i scales (the predictor) for the Finance job family, descriptive statistics of the scales are outlined in Tables 4.8-4.10. The 32 scales of the OPQ 32i were grouped into their three dimensions namely; Relationship with People, Thinking style and Feelings and Emotions, to aid in the analysis of the findings which will be presented in the discussion chapter.

Table 4.8

Descriptive statistics of Relationship with People scales for the Finance job family (n=72)

OPQ 32i scale	Mean	Standard Deviation
Persuasive	4.90	1.91
Controlling	5.17	1.82
Outspoken	5.92	1.87
Independent-minded	4.81	1.98
Outgoing	4.14	1.70
Affiliative	3.56	1.68
Socially confident	5.28	1.65
Modest	5.78	1.84
Democratic	5.68	1.81
Caring	4.56	1.69

For the Relationship with People domain, the scales with the highest Mean scores were the Outspoken, Modest and Democratic scales.

Table 4.9

Descriptive statistics of Thinking styles scales for the Finance job family (n=72)

OPQ 32i scales	Mean	Standard Deviation
Data rational	8.08	1.73
Evaluative	6.56	1.56
Behavioural	4.28	1.62
Conventional	6.47	1.83
Conceptual	5.58	1.64
Innovative	5.46	1.62
Variety seeking	4.78	2.10
Adaptable	3.04	1.95
Forward thinking	6.19	2.42
Detail conscious	6.54	1.73
Conscientious	6.94	1.74
Rule following	6.97	1.78

For the Thinking style OPQ domain, the scales with the highest Mean scores were the Data rational, Rule following and Conscientious scales.

Table 4.10

Descriptive statistics of Feelings and Emotions scales for the Finance job family (n=72)

OPQ 32i scale	Mean	Standard Deviation
Relaxed	5.56	1.55
Worrying	4.64	1.73
Tough-minded	5.92	1.36
Optimistic	5.68	1.83
Trusting	5.38	2.02
Emotionally controlled	5.18	1.72
Vigorous	5.31	2.03
Competitive	5.35	1.95
Achieving	6.76	1.48
Decisive	5.15	1.90
Consistency	5.10	1.53

For the Feelings and Emotions OPQ domain, the scales with the highest Mean scores were Achieving, Tough-minded and Optimistic.

Validity of the OPQ 32i for the Finance job family

To determine the predictive validity of the OPQ 32i and performance, correlation analysis was conducted. The results are outlined according to the different OPQ 32i dimensions. The results are summarised in table 4.11.

Table 4.11

Correlation between the OPQ 32i dimensions with Performance for the Finance job family (n=72)

OPQ 32i Scale Dimensions					
Relationship with People	r_{xy}	Thinking styles	r_{xy}	Feelings and Emotions	r_{xy}
Persuasive	.17	Data rational	-.17	Relaxed	-.22
Controlling	.13	Evaluative	-.10	Worrying	-.04
Outspoken	.17	Behavioural	.15	Tough-minded	-.17
Independent-minded	-.01	Conventional	-.33	Optimistic	.03
Outgoing	.24	Conceptual	.15	Trusting	-.07
Affiliative	.04	Innovative	.32	Emotionally controlled	-.08
Socially confident	.16	Variety seeking	.14	Vigorous	.11
Modest	-.04	Adaptable	-.11	Competitive	.13
Democratic	.09	Forward thinking	-.07	Achieving	-.16
Caring	-.10	Detail conscious	-.13	Decisive	.16
		Conscientious	-.18		
		Rule following	-.31		

For the Relationship with People domain, the only scale demonstrating a significant correlation was the Outgoing scale. The Thinking style domain reflected only one scale with a significant correlation, namely the Innovative scale. For the Feelings and Emotion domain there were no significant correlations. These results have not produced significant correlations between the OPQ 32i scales and employee performance in the Finance job family.

Conclusion

The OPQ 32i scale scores for the participants in both the Administration and Finance sample did not exhibit significant correlations with performance. The scales with the highest Mean scores on the different OPQ 32i scale dimensions were the same across the two job families. For the Relationship

with People dimension, the Outspoken, Modest and Democratic scales exhibited the highest Mean scores. The scores for the Thinking styles domain reflected Conscientious, Rule Following and Data Rational with the highest Mean scores. In the Feelings and Emotions dimension, the Achieving, Tough-minded and Optimistic scales demonstrated the highest Mean scores. The subscales of the OPQ 32i showed high internal consistency, with eight of the scales scoring below 0.70. Of those eight scales, seven were between 0.65 and 0.69. This demonstrates the reliability of the OPQ 32i as an assessment tool.

The next chapter will outline the implications of these findings as well as the limitations of this study, providing recommendations for future studies.

University of Cape Town

CHAPTER 5: DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

The aims of this study were to investigate whether specific personality constructs are more important than others to enable competence in the job, and whether the OPQ 32i is a reliable and valid predictor of job performance among Administration and Finance staff in a large financial institution. This chapter discusses the outcomes and implications of the results in relation to the literature which was reviewed on personality assessment. The practical implications and limitations of this study are summarised, and the chapter is concluded with a final overview of this study and recommendations for future studies.

Criterion

Validity

In this study the performance ratings were the criterion measure of job success. The performance rating was determined by the line manager, and evaluation was based on the achievement of the agreed performance outputs of the job. A combined global performance rating comprising three components (as outlined in the Method section) was determined, and this rating regarded as the valid indicator of employee performance in the organisation. As the components of this final combined rating are weighted, in instances where the team rating exceeded the individual rating, the 25 percent weighting of the team component, would raise the overall performance rating of the employee. This combined rating in some instances, therefore skews the interpretation of the performance rating.

The challenge of using a single global criterion measure is well articulated by Guion (as cited in Cascio, 1998, p. 53), who states that "the fallacy of the single criterion lies in its assumption that everything that is to be predicted is related to everything else that is to be predicted – that there is a general

factor in all criteria accounting for virtually all of the important variance in behaviour at work and its various consequences of value". In addition, the behavioural components which would be required to perform the key result areas were not identified at the time of performance contracting, and were not measured or reviewed during the performance review. Only the technical, task specific performance outputs were identified and subsequently measured. This therefore challenges what the criterion actually measured and its subsequent validity for the purposes of this study.

In the Administration job family sample, 21 of the 60 sample participants had a performance rating of 2.5. In the organisation, if the time period is too short to engage in a meaningful performance assessment, or if a line manager is still unsure about the performance of a new employee, the line managers tend to provide a 2.5 rating (performance outcomes and standards met). While there are line managers who do feel comfortable to rate performance, others do not engage in a performance review. The prevalence of the 2.5 rating in the sample cannot be investigated, and for this reason, the performance result is skewed, and the validity of the criterion measure further reduced.

Reliability

The participants in the sample were drawn from different departments and business units. Performance management would therefore be experienced differently across the different areas, which would influence the different rating styles of the different line managers. The Guidelines for the Validation and Use of Assessment Procedures for the Workplace (SIOPSA, 2005) stipulates an important precursor for an acceptable criterion measure is that line managers who give the performance ratings are in a position to provide meaningful feedback to candidates on their performance. In this study, different line managers rated the different employees. In the Finance job family, the business areas only rate performance annually, and for this

reason, inter-rater and intra-rater reliability could not be investigated. This would have enhanced the reliability findings of the criterion measure.

This next section will outline the reliability and validity of the OP 32i, with reference to the correlations between the OPQ 32i scales (the predictor) and the performance ratings.

Predictor

Validity

The introduction of the Employment Equity Act (Act No. 55 of 1998) legislated the prohibition of assessment measures which are not valid, reliable and free from bias. To investigate the predictive validity of the OPQ 32i, validity coefficients of the OPQ 32i scales and the performance ratings were calculated. For the Administration job family, there were no significant correlations between any of the OPQ 32i dimensions and the performance ratings. In the Relationship with People domain, the scale with the strongest correlation was Controlling ($r=.18$). In the Thinking styles domain, there were no significant correlations, with the Evaluative ($r=.13$) and Forward thinking scales ($r=.13$) producing the strongest correlations. The Feelings and Emotions domain had one significant correlation, the Achievement scale ($r=.20$). The results from this study produced no evidence of the strength of the correlation between personality and performance.

For the Finance job family, in the Relationship with People dimension, only the Outgoing scale demonstrated a significant correlation ($r=.24$). In the Thinking styles dimension, the Innovative scale was the only scale with a significant correlation ($r=.32$). The Feelings and Emotions dimension showed no significant correlations, with the Decisive scale demonstrating the strongest correlation ($r=.16$). Even though two scales demonstrated a strong correlation and one scale as was the case for the Administration job family results, these results do not provide compelling evidence of the existence of

specific personality constructs which will enable more successful job performance.

Studies conducted by SHL (2004a) investigating the predictive validity of the OPQ 32i, interviews and numerical critical reasoning assessment tools in a large insurance organisation, with middle and senior managers as the sample participants, reflected no correlations between the performance appraisal scores and the personality assessment results. The reason for the low validity was attributed to the fact that the criterion measured cognitive abilities as opposed to personality dimensions, hence the strong correlation between cognitive assessment tools and the performance appraisal (SHL, 2004a). Similar studies conducted in the Food production industry by SHL (2003a), with supervisors being the sample participants, highlight two scales with significant validities; Conscientious ($r=.21$) and Vigorous ($r=.20$). The low correlations between the personality dimensions and the criterion measure were attributed to the small sample size ($N=79$), and the nature of the criterion (SHL, 2003a). It was argued that a single performance score is not a good predictive measure of performance, which is multidimensional in nature (SHL, 2003a). When comparing the results of this study, the low indices of validity between the predictor and criterion information reflect a criterion measure which potentially measures cognitive ability, and not personality. The small sample size in this study for the Administration ($n=60$) and Finance ($n=72$) samples may also have been a contributing factor.

To enhance the correlations between personality variables and job specific criteria, Barrick and Mount (1991) and Dakin, Nilakant and Jensen (1994) advocate aligning specific personality dimensions with specific occupations and job criteria. The Conscientious scale, one of the "Big Five" (Barrick & Mount, 1991) is reported as the one good predictor of job success across many occupations (Barrick & Mount, 1991; Day & Silverman, 1989; Tett, Jackson & Rothstein, 1991). When applying this argument to this study, the OPQ 32i scales represented the behavioural components which were required

for job success, but specific scales were not identified as relevant for job success nor then assessed during the performance review.

Unlike these previous personality validation studies, the Conscientious scale did not show a significant correlation. The findings of this study do not confirm the research findings on personality measures of other studies, where there are significant correlations between personality variables and performance.

Reliability

The internal consistencies of the OPQ 32i subscales scales were calculated. Of the 32 scales, 24 had an alpha coefficient above 0.70. Of the eight scales below the 0.70, only one scale was below 0.65 (Tough-minded). Previous OPQ 32i reliability studies conducted by SHL (2003b) in a large financial services organisations reflected the same scales with low alpha coefficients; Independent-minded, Democratic, Evaluative, Conceptual, Tough-minded and Vigorous scales with alpha coefficients below 0.70. A similar study conducted by SHL (2002) in a large financial services organisation with a sample comprising job applicants from different occupational groups, reported alpha coefficients below the 0.70 for the Independent-minded, Democratic, Caring, Evaluative, Conventional, Variety seeking, Adaptable, Tough-minded, Optimistic, Vigorous and Achieving scales.

When comparing this study's result with prior OPQ 32i studies, the Independent-minded, Democratic, Evaluative and Tough-minded scales seemed to have lower reliability. The other 24 OPQ 32i scales do however demonstrate high internal consistencies. This demonstrates the internal consistency of the OPQ 32i subscales and its reliability as an assessment tool.

The increased use of personality assessment in organisations to aid in recruitment and selection is reported by Baron and Miles (2001). It is

therefore essential that the assessment instruments used in recruitment and selection are valid and reliable.

Implications

The statistical analyses in this study reflect insignificant correlations between the predictor and criterion in the Administration and Finance job family sample. The findings therefore do not support the numerous studies covered in the literature review, which have attested to the relationship between specific personality dimensions and job performance criteria.

The personality factor Conscientious, has been reported in numerous studies as a good predictor of performance across numerous occupations. For both the Administration and Finance job families, the Conscientious scale is negatively correlated. This is a concern as conscientiousness in terms of its definition is what most jobs would require, and based on this study's findings, the relationship is inverted.

The existing literature has stipulated the need for clearly identified and defined personality dimensions to be aligned to specific job criteria. In this study, specific personality scales were not defined as important contributors for job success, and during the performance review these personality dimensions were therefore not measured.

Limitations

When this study commenced, it was envisioned that a homogenous sample comprising approximately 120 employees in financial positions at similar grade levels in the organisation would form part of this sample. Due to constraints related to the availability of data, only data from different employment groups were available for use (from 60 Administration and 72 Finance employees). This sample size constitutes the first limitation for this study, bearing in mind the caution to "use as large samples as possible"

(Kerlinger, 1986, p.117), further highlighting the risk of statistical errors caused by using small samples. The "Guidelines for the Validation and Use of Assessment Procedures for the Workplace" outlined that for a test to be regarded as valid, it needs to demonstrate adequate statistical power (SIOPSA, 2005, p. 12). This prevents the risk of Type II errors being made.

Another difficulty is that the employees in the sample were drawn from different grade levels. When comparing the performance ratings of, for example, an administrator and then an account executive, the different levels of work complexity for the different grade levels must be taken into account. The average performance rating for the administrator would be for work involving less complex outputs as compared to the type of complex output required from the account executive, who also receives an average performance rating (Muchinsky, Kriek & Schreuder, 2005). This results in criterion bias, which SIOPSA (2005, p.14) defines as "systematic error resulting from criterion contamination or deficiency that differentially affects the criterion performance of different sub- groups".

Schmidt and Kaplan (1971) as well as Guion (as cited in Cascio, 1998) challenged the assumption that the criterion encapsulates all behaviour at work. They asserted that combining criterion elements into a composite criterion should not imply a single underlying dimension in job performance. They cautioned against this single underlying dimension being interpreted as being behavioural or psychological in nature. The performance rating does not cover a component which measures personality and behaviours, and this, as reported in previous studies, could account for the low correlations.

A third challenge of using the performance rating as the criterion is that the performance rating is contaminated. In the Administration job family sample, there are 21 employees who have the 2.5 rating which skews the distribution of the performance ratings. In the organisation, when managers are unable to make a good judgment of the performance of the employees, either based on

their being new to the organisation or the role, some choose to have a feedback discussion, allocating a 2.5 performance. However there are other business areas where the line managers give a 2.5 rating based on the employees performing according to what could be expected after such a short period in the role. There is no way of ascertaining whether or not the 2.5 rating is indeed for performance, or owing to the circumstances of the employee being new to the role or the organisation. SIOPSA (2005, p.12) discusses the importance of relevant criterion being able to differentiate the different performance levels between employees, stating further "if such a criterion measure does not exist, use of a criterion-related validation study is not feasible".

The candidates in the sample were pre-selected by way of the recruitment process followed by the organisation. This resulted in restriction of range. Aaron and Aaron (1994) assert that "if a correlation is computed when only a limited range of the possible values on one variable is included in the group studied, the correlation cannot be properly extended to apply to the entire range of values the variable might have among people in general".

Recommendations

Owing to the results and limitations of this study, this section will outline recommendations for future studies.

A follow-up study is recommended where a larger sample needs to be included. The employees in the sample should be in similar grade bands in the organisation, in roles requiring similar complexity.

To avoid the challenge presented in this study of the 2.5 performance ratings, the time period between when employees are recruited and when the performance appraisal is conducted needs to be extended. This could help

ensure a performance rating which is a more valid and true reflection of the employee performance.

To enable the reliability of the performance ratings to be measured, two performance cycles should be included, to determine the inter and intra-rater reliability of the performance ratings.

As organisations follow competency based human resources practices, it is advised that the OPQ 32i scales be used to form a composite competency rating, as competencies are used to measure performance. In addition, there needs to be an additional criterion, such as a line manager competency rating tool, which can then be correlated with the performance rating. This will improve the reliability and validity of the criterion measures.

The specific OPQ 32i scales which are believed to be relevant to job performance need to be identified, and line managers need to rate these behavioural dimensions as part of the performance review process, to ensure that the performance rating encapsulates the cognitive and behavioural performance components.

Conclusion

The aims of this study was to determine the OPQ 32i's ability to measure and predict future job success and to determine the extent to which the OPQ 32i's specific personality dimensions could be correlated with job criteria dimensions regarded as critical for job success. The criticism and challenges to the use of personality assessment in the workplace has been covered in the introduction and literature review chapters. In spite of these challenges, numerous studies conducted (Barrick & Mount, 1991; Dakin, Nilakant, 1994; Day & Silverman, 1989; Goffin, Rothstein & Johnston, 1996) to name a few, cited the validity and use of personality assessment as enabling effective decision making. The findings highlighted the specific conditions which enable

meaningful inferences from personality assessment to be drawn. The results of this study have not supported the findings of a link between personality and performance. There were however, numerous limitations which prevented additional analyses which may have contributed to more resounding evidence.

The limitations in this study have been cited by researchers as reasons perpetuating the negative perception of using personality assessment (Guion & Gottier, 1965; Harland, Rauzi & Biasotto, 1995; Rosse, Miller & Stecher, 1994). The results of this study have not provided convincing validity coefficients that can be generalised. As recruitment and selection are costly exercises in organisations, additional research into the use of personality assessment is recommended, taking cognisance of the limitations of this study. When the appropriate conditions are in place, a more meaningful assessment of their predictive validity will be enabled.

University of Cape Town

REFERENCE LIST

- Aron, A., & Aron, E. N. (1994). *Statistics for Psychology*. Englewood Cliffs, New Jersey: Prentice Hall
- Barret, P., Kline, P., Paltiel, L., & Eysenck, H. J. (1996). An evaluation of the psychometric properties of the concept 5.2 Occupational Personality Questionnaire. *Journal of Occupational and Organizational Psychology*, *69*, 1–19
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta – analysis. *Personnel Psychology*, *44*, 1 – 26
- Baron, H., & Miles, A. (2001, January, 4-5). Comparability of Personality Profiles around the world. Retrieved August, 12, 2006, from [http://www.shl.com/SHL/en-int/Thought Leadership/opinion/OPO/](http://www.shl.com/SHL/en-int/Thought%20Leadership/opinion/OPO/)
- Baron, H., & Miles, A. (2002, January,3-5). Personality Questionnaire: Ethnic trends and selection. Retrieved August 12, 2006, from [http://www.shl.com/SHL/en-int/Thought Leadership/opinion/OPO/](http://www.shl.com/SHL/en-int/Thought%20Leadership/opinion/OPO/)
- Bartram, D. (2004). Assessment in Organisations. *Applied Psychology: An International Review*, *53*, 2, 237 – 259
- Becker, T. E., & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: An analysis along several dimensions of item type. *Personnel Psychology*, *45*, 2, 389 – 406
- Bedell, B., Van Eeden, R., & Van Staden, F. (1999). Culture as moderator variable in psychological test performance: Issues and trends in South Africa [Electronic version]. *South African Journal of Industrial Psychology*, *25*, 3, 1-7
- Bethell-Fox, C. E. (1992). Psychological testing. In P. Herriot (Ed.), *Assessment and selection in organizations* (pp. 307–329). New York: John Wiley & Sons
- Boerlijst, G., & Meijboom, G. (1992). Matching the individual and the organization. In P. Herriot (Ed.), *Assessment and selection in organizations* (pp. 25 – 44). New York: John Wiley & Sons
- Cascio, W. F. (1998). *Applied psychology in human resource management* (5th ed.). Upper Saddle River, New Jersey: Prentice-Hall

- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc
- Coolican, H. (2004). *Research methods and statistics in psychology* (4th ed.). London: Hodder & Stoughton
- Coolidge, F.L. (2000). *Statistics. A gentle introduction*. London: Sage Publications
- Cooper, D., & Robertson, I.T. (2000). *The psychology of personnel selection: A quality approach*. London: Business Press
- Dakin, S., Nilakant, V., & Jensen, R. (1994). The role of personality testing in managerial selection. *Journal of Managerial Psychology*, 9, 5, 3 – 11
- Day, D.V., & Silverman, S. B. (1989). *Personnel Psychology*, 42, 25 – 36
- Edwards, J. R. (1991). Person–job fit: A conceptual integration, literature review, and methodological critique. In C.L. Cooper, & I.T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 283-357). New York: John Wiley & Sons
- Elkonin, D., Foxcroft, C., Roodt, G., & Astbury, G. (2001). The use of assessment measures in various applied contexts. In C. Foxcroft, & G, Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 265 – 294). Cape Town: Oxford University Press
- Employment Equity Act, Act 55, 1998. Retrieved November 18, 2006, from <http://www.hpcsa.co.za/hpcsa/UserFiles/File/EMPLOYMENT%20EQUITY%20ACT.pdf>
- Fernández–Ballesteros, R. (1999). Psychological assessment: Future challenges and progresses. *European Psychologist*, 4, 4, 248 – 262
- Foxcroft, C., & Roodt, G. (2001). An introduction to psychological assessment in the South African context. Cape Town: Oxford University Press
- Foxcroft, C., Roodt, G., & Abrahams, F. (2001). Psychological assessment: a brief retrospective overview. In C, Foxcroft., & G, Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp.11 – 33). Cape Town: Oxford University Press
- Foxcroft, C. D. (2004). Planning a psychological test in the multicultural South African context. *South African Journal of Industrial Psychology*, 30, 4, 8 – 15

- Frank, F. D., Finnegan, R. P., & Taylor, C. R. (2004). The race for talent: Retaining and engaging workers in the 21st century. *Human Resource Planning, 27*, 3, 12 – 25
- Furnham, A., & Drakeley, R. (2000). Predicting occupational test scores. *The Journal of Psychology, 134*, 1, 103 – 111
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461-477
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review, 18*, 4, 694 – 734
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 1, 26 – 34
- Goffin, R. D., Rothstein, M.G., & Johnston, N.G. (1996). Personality testing and the assessment center: Incremental validity for managerial selection. *Journal of Applied Psychology, 81*, 6, 746 – 756
- Gregory, R. J. (2000). *Psychological testing. History, principles, and applications* (3rd ed.). Needham Heights, Massachusetts: Allyn & Bacon Inc
- Grobler, P.A., Wörnich, S., Carrell, M.R., Elbert, N.F. & Hatfield, R.D. (2002). *Human resource management in South Africa* (2nd ed.). London: Thomson Learning
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135 – 164
- Hair, J. F., Babin, B., Money, A. H., & Samouel, P. (2003). *Essentials of business research methods*. Hoboken, New Jersey: Wiley
- Harland, L. K., Rauzi, T., & Biasotto, M. M . (1995). Perceived fairness of personality tests and the impact of explanations for their use. *Employee Responsibilities and Rights Journal, 8*, 3, 183 – 192
- Haslam, S. A., & McGarty, C . (2003). *Research methods and statistics in psychology*. London : Sage Publications
- Health Professions Act, Act 56, 1974. Retrieved November, 18, 2006 from <http://www.hpcsa.co.za/hpcsa/UserFiles/File/HealthProfessionsAct.doc>
- Health Professions Council of South Africa. (n.d.). Professional Board for Psychology: Policy on the classification of psychometric measuring devices, instruments, methods and techniques (Form 208). Retrieved

November 18, 2006, from
<http://www.hpcsa.co.za/hpcsa/UserFiles/File/PSYCHOLOGY/F208%20update-7%20July%202006-Exco1.pdf>

- Hess, A. K., & Neville, D. (1977). Testwiseness: Some evidence for the effect of personality testing on subsequent test results. *Journal of Personality Assessment*, 41, 2, 170 – 177
- Hogan, J., & Hogan, R. (1989). How to measure employee reliability. *Journal of Applied Psychology*, 74, 2, 273 – 279
- Hogan, R., Hogan, J., & Roberts, B.W. (1996). Personality measurement and employment decisions. *American Psychologist*, 51, 5, 469 – 477
- Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human Performance*, 11, 2/3, 129 – 144
- Hollenbeck, J. R., & Whitener, E. M. (1988). Reclaiming personality traits for personnel selection: Self esteem as an illustrative case. *Journal of Management*, 14, 1, 81 – 91
- 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
- International Test Commission. (2001). International Guidelines for Test Use. *International Journal of Testing*, 1, 2, 93 – 114
- Ivancevich, J.M., & Matteson, M.T. (2002). *Organizational behaviour and management* (6th ed.). New York: McGraw – Hill Inc
- Jackson, C.J., & Corr, P. (1998). Personality–performance correlations at work: Individual and aggregate levels of analyses. *Journal of Personality and Individual Differences*, 24, 6, 815 – 820
- Kerlinger, F. N. (1986). *Foundations of Behavioural research* (3rd ed.). New York: Harcourt Brace Jovanovich College Publishers
- Kline, P. (1995). *Handbook of psychological testing* (2nd ed.). London: Routledge
- Kluger, M. T., Watson, D., Laidlaw, T. M., & Fletcher, T. (2002). Personality testing and profiling for anaesthetic job recruitment: attitudes of anaesthetic specialists/ consultants in New Zealand and Scotland. *Anaesthesia*, 57, 2, 116 – 122

- Kristof, A. L. (1996). Person–organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology, 49*, 1, 1 – 49
- Lanyon, R.I., & Goodstein, L.D. (1982). *Personality assessment* (2nd ed.). New York: John Wiley & Sons
- Leary, M.R., & Kowalski, R.M. (1990). Impression Management: A literature review and two – component model. *Psychological Bulletin, 107*, 1, 34 – 47
- Macleod, C. (2004). South African psychology and 'relevance': Continuing challenges. *South African Journal of Psychology, 34*, 4, 613 – 629
- Meiring, D., Van de Vijver, A. J. R., Rothmann, S., & Barrick, M. R. (2005). Construct, item, and method bias of cognitive and personality tests in South Africa. *South African Journal of Industrial Psychology, 31*, 1, 1 – 8
- Mello, J.A. (1995, October). Personality screening in employment: Balancing information gathering and the law. *Labour Law journal, 46*, 622 – 626
- Moore, D.S. (1995). *The basic practice of Statistics*. New York: WH Freeman and Company
- Muchinsky, P.M., Kriek, H.J., & Schreuder, A.M.G. (2003). *Personnel psychology* (2nd ed.). Cape Town: Oxford University press
- Nikoaou, I. (2003). Fitting the person to the organisation: examining the personality- job performance relationship from a new perspective. *Journal of Managerial Psychology, 18*, 7, 639 – 648
- Oakland, T. (2004). Use of educational and psychological tests internationally. *Applied Psychology: An International Review, 53*, 2, 157 – 172
- Ones, D. S., & Viswesvaran, C. (1999). Relative importance of personality dimensions for expatriate selection: A policy capturing study. *Human Performance, 12*, 3/4, 275 – 294
- Orpen, C. (1974, May). The "correct" use of personality tests: a view from industrial psychology. *Public Personnel Management, 228* – 229
- 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

- Paunonen, S.V. (2003). Big five factors of personality and replicated predictions of behaviour. *Journal of Personality and Social Psychology*, 84, 2, 411 – 424
- Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, 65, 4, 373 – 406
- Pervin, L. A. (1996). *The science of personality*. New York: John Wiley & Sons, Inc
- Rafaeli, A. (1999). Pre-employment screening and applicants' attitudes toward an employment opportunity. *The Journal of Social Psychology*. 139, 6, 700 – 712
- Raymark, P. H., Schmit, M.J., Guion, R.M. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology*, 50, 723 – 736
- Ree, M.J., Earles, J. A., Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology*, 79(4), 518 – 524
- Reilly, R. R. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1 – 62
- Robertson, I., & Callinan, M. (1998). Personality and work behaviour. *European Journal of Work and Organizational Psychology*, 7, 3, 321 – 340
- Robertson, I.T., Kandola, R.S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of Occupational Psychology*, 55, 171 – 183
- Robertson, I.T., Kinder, A. (1993). Personality and job competencies: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology*, 66, 225 – 244
- Rosse, J. G., Miller, J. L., & Stecher, M. D. (1994). A field of job applicants' reactions to personality and cognitive ability testing. *Journal of Applied Psychology*, 79, 6, 987 – 991
- Rosse, J. G., Stecher, M.D., Miller, J. L., & Levin. R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 4, 634 – 644

- Rosenfeld, P., Giacalone, R.A., & Riordan, C.A. (1995). *Impression Management in organizations. Theory, measurement, practice*. London: Routledge
- Sackett, P. R., Gruys, M. L., & Ellingson, J, E. (1998). Ability–personality interactions when predicting job performance. *Journal of Applied Psychology, 83*, 4, 545 – 556
- Searle, R. H. (2003). *Selection & Recruitment. A critical text*. Walton Hall, Milton Keynes: The Open University
- Schmidt, F.L., Gast–Rosenberg, I., & Hunter, J. E. (1980). Validity generalization results for computer programmers. *Journal of Applied Psychology, 65*, 6, 643 – 661
- Schmidt, F.L., & Hunter, J. (2004). General mental ability in the world of work: occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 1, 162 – 173
- Schmidt, F.L., & Kaplan, L.B. (1971). Composite vs. multiple criteria: A review and resolution of the controversy. *Personnel Psychology, 24*, 3, 419-434
- Schmitt, N., Gooding, R. Z., Noe, R, A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37*, 407 – 422
- Schmit, M. J., Kihm, J. A., Robie, C. (2000). Development of a global measure of personality. *Personnel Psychology, 53*, 153 – 193
- Schwab, D, P. (1971). Issues in response distortion studies of personality inventories: A critique and replicated study. *Personnel Psychology, 24*, 637 – 647
- SHL. (2002, October, 15). Reliability studies. Retrieved August 03, 2006, from <http://www.research.shl.co.za/asp/norms.asp?Reliability=1&parentid=85>
- SHL. (2003a, June, 09). Validation studies. Retrieved August 05, 2006, from <http://www.research.shl.co.za/asp/norms.asp?Validation=1&parentid=84>
- SHL. (2003b, December, 15). Reliability studies. Retrieved August 03, 2006, from <http://www.research.shl.co.za/asp/norms.asp?Reliability=1&parentid=85>

- SHL. (2004a, July, 13). Validation studies. Retrieved August 05, 2006, from <http://www.research.shl.co.za/asp/norms.asp?Validation=1&parentid=84>
- SHL. (2004b, December, 21). Validation studies. Retrieved August 05, 2006, from <http://www.research.shl.co.za/asp/norms.asp?Validation=1&parentid=84>
- SHL. (1999). *Opq 32. Manual and User's Guide*. Surrey: SHL group plc.
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R.W. (1993). Applicant reactions to selection procedures. *Personnel Psychology, 46*, 49 – 76
- Shuttleworth–Jordan, A. B. (1996). On not reinventing the wheel: A clinical perspective on culturally relevant test usage in South Africa [Electronic version]. *South African Journal of Psychology, 26*, 2, 1 – 14
- Society for Industrial and Organisational Psychology of South Africa (SIOPSA). (1995). *Guidelines for the validation and use of assessment procedures for the workplace*. South Africa: SIOPSA
- Terre Blanche, M.T., & Durrheim, K. (1999). *Research in practice. Applied method in the social sciences*. Cape Town: University of Cape Town Press
- Tett, R., Jackson, D., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 703-742
- Thurow, R. (1967, July). Personnel testing. *Industrial Management, 3* – 8
- Van der Maesen de Sombreff, P., Hofstee, W. K. B. (1992). Personality questionnaires and inventories. In Herriot, P, (Ed.), *Assessment and selection in organizations* (pp, 353–367). New York: John Wiley & Sons
- Wallis, T. (2004). Psychological tests do not measure what they claim to measure: A re-evaluation of the concept of construct validity. *South African Journal of Psychology, 34*, 1, 101 – 112