

**THE VALIDATION OF A PSYCHOLOGICAL ASSESSMENT BATTERY FOR THE  
SELECTION OF CUSTOMER SERVICE AGENTS IN A SOUTH AFRICAN  
COMMERCIAL AIRLINE COMPANY**

by

**RYAN JONATHAN DAVIS**

submitted in accordance with the requirements for  
the degree of

**MASTER OF COMMERCE**

in the subject

**INDUSTRIAL AND ORGANISATIONAL PSYCHOLOGY**

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF H.A. BARNARD

JUNE 2013

## ACKNOWLEDGEMENTS

I would like to take this opportunity to express my immense appreciation and gratitude to the following incredible individuals, without whom this would not have been possible:

- Prof Antoni Barnard, for selflessly giving of her time and expertise to ensure the successful completion of my dissertation. Her guidance and encouragement through this process is much appreciated.
- Tina Joubert and Kim Dowdeswell of SHL South Africa, for their assistance with the statistical analysis and for their support and advice throughout the process.
- The airline company within which I carried out the research, for allowing me to use the data from this valuable study for my dissertation.
- All airline company staff that so willingly assisted with data collection.
- Dr Renate Scherrer, Jamie Hey and Leigh Wallace, for their support, direction and brilliant leadership throughout this process, as well as in every aspect of my work at JvR Consulting Psychologists (JvRC).
- The whole team at JvRC for all the encouragement as well as for their inspiring commitment to making a sustainable difference in organisations, the lives of individuals, the profession and the world.
- Dr Nicola Taylor, for always having the right advice to get my mindset back into research mode and for helping me see the value in the knowledge gained.
- My mentor, Esther Venter, for so selflessly sharing her knowledge, thoughts and ideas about psychology, business and life from my first day working with her. Aside from her contribution to this study, her input into my professional growth, development and desired future contribution to the profession is invaluable.
- To all my friends, for their encouragement and understanding, even when I had to miss important events because there were not enough hours in the day.
- The most unbelievable family in the world, for putting up with me through this journey but, much more importantly, for always being my pillars of strength, voices of reason, sources of motivation and encouragement (even when I wanted to give up). I cannot explain in words how much you mean to me and how much I love you. I am the luckiest person in the world to have you all in my life.
- G-d, King of the Universe, without Whom none of this would be possible.

## DECLARATION

Student number: **4581-911-4**

I declare that “The validation of a psychological assessment battery for the selection of customer service agents in a South African commercial airline company” is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

---

**SIGNATURE**  
**(R.J. DAVIS)**

---

**DATE**

## **SUMMARY**

### **THE VALIDATION OF A PSYCHOLOGICAL ASSESSMENT BATTERY FOR THE SELECTION OF CUSTOMER SERVICE AGENTS IN A SOUTH AFRICAN COMMERCIAL AIRLINE COMPANY**

by

**RYAN DAVIS**

**SUPERVISOR** : **H.A. BARNARD**  
**DEPARTMENT** : **INDUSTRIAL AND ORGANISATIONAL  
PSYCHOLOGY**  
**DEGREE** : **MCOM (Industrial and Organisational  
Psychology)**

The purpose of the research was to determine whether measures of ability, personality and behaviour would significantly predict job performance of customer service agents in a South African commercial airline company. The Verbal Interpretation Test (VCC1), Numerical Reasoning Test (NP6.1), Basic Checking Test (CP7.1C), Occupational Personality Questionnaire (OPQ32) and a competency based interview were completed by job applicants. Customer Contact Competency (CCC) scores and a Person Job Match (PJM) score were derived from the OPQ32 and ability measures to ensure job relevance during selection assessment. Job performance statistics in the form of training scores and supervisor ratings (from performance appraisals and criterion questionnaires) were obtained for the sample as criterion data. Correlations revealed statistically significant small to moderate correlations between the predictors and the criterion data.

#### Key terms:

Selection; psychological assessment; psychological test; cognitive ability; numerical ability; checking ability; personality; behavioural interview; competency-based assessment; predictive validity; job analysis; correlations; multi-cultural context; customer service agents.

## TABLE OF CONTENTS

	<b>PAGE</b>
ACKNOWLEDGEMENTS	ii
DECLARATION	iii
SUMMARY	iv
LIST OF TABLES	xi
LIST OF FIGURES	xi
<b>CHAPTER 1: SCIENTIFIC ORIENTATION TO THE RESEARCH</b>	<b>1</b>
1.1 INTRODUCTION	1
1.2 BACKGROUND TO AND MOTIVATION FOR THE RESEARCH	1
1.3 PROBLEM STATEMENT	3
1.3.1 Research questions	4
1.4 AIMS OF THE RESEARCH	5
1.4.1 General aim of the study	5
1.4.2 Specific aims of the study	5
1.4.2.1 Specific theoretical aims of the literature review	5
1.4.2.2 Specific aims of the empirical study	5
1.4.2.3 Hypotheses	6
1.5. THE PARADIGM PERSPECTIVE	6
1.6 RESEARCH DESIGN	7
1.6.1 Research approach	7
1.6.2 Research variables	7

1.6.2.1	The predictor/independent variables	8
(i)	Verbal ability test	8
(ii)	Numerical ability test	8
(iii)	Checking ability test	9
(iv)	CCC scores	9
(v)	PJM score	9
(vi)	Competency based interview	9
1.6.2.2	The criterion/dependent variable	10
(i)	Training course scores	10
(ii)	Performance appraisal ratings	10
(iii)	Criterion questionnaire ratings	10
1.7	RESEARCH METHOD	11
1.7.1	Research participants	11
1.7.2	Measuring instruments	11
1.7.3	Research procedure and ethical considerations	11
1.7.3.1	Literature review	12
1.7.3.2	Empirical study	12
1.7.3.3	Ethical considerations	12
1.7.4	Data analysis	13
1.8	RESULTS	13
1.9	CHAPTER LAYOUT	14
1.10	CHAPTER SUMMARY	14
	<b>CHAPTER 2: LITERATURE REVIEW</b>	<b>15</b>
2.1	INTRODUCTION	15
2.2	SELECTION FROM AN IO PSYCHOLOGY PERSPECTIVE	15
2.2.1	Psychological assessment and selection	16
2.2.2	Selection process	16

2.2.3	Job analysis: the foundation to competency based assessment	18
2.3	CONCEPTUALISATION: PSYCHOLOGICAL ASSESSMENT VS PSYCHOLOGICAL TESTING	21
2.3.1	Psychological assessment	21
2.3.2	Psychological testing	21
2.3.3	Assessment of psychological constructs	23
2.3.3.1	Cognitive ability assessment	24
2.3.3.2	Personality assessment	26
2.3.3.3	Competency based assessment	27
2.4	PSYCHOLOGICAL ASSESSMENT IN THE SOUTH AFRICAN MULTI-CULTURAL CONTEXT	28
2.4.1	A historical perspective on psychological testing in South Africa	28
2.4.2	Psychological assessment in the context of labour legislation	30
2.4.2.1	Health Professions Act (56 of 1974)	30
(i)	Regulatory structure and scope of practice	30
(ii)	Test classification	31
(iii)	Registration of test professionals	33
2.4.2.2	Labour Relations Act (66 of 1995)	33
2.4.2.3	Employment Equity Act (55 of 1998)	34
2.4.3	Ethical principles and conduct	35
2.4.4	Fairness and bias	36
2.4.5	A deeper look at culture-fair assessment in the South African context	38
2.5	COMPETENCY BASED ASSESSMENT	43
2.6	ESSENTIAL PSYCHOMETRIC PROPERTIES OF PSYCHOLOGICAL ASSESSMENT TOOLS	45
2.6.1	Norms and standardisation	45
2.6.2	Moderator variables	45
2.6.3	Basic statistical concepts relevant to validation studies	46
2.6.4	Reliability	46
2.6.4.1	Test-retest reliability	47

2.6.4.2	Alternate-form reliability	47
2.6.4.3	Internal consistency reliability	47
2.6.4.4	Inter-rater reliability	48
2.6.4.5	Factors affecting the reliability of an assessment	48
2.7	VALIDITY	49
2.7.1	Types of validity	49
2.7.1.1	Content validity	50
2.7.1.2	Construct validity	50
2.7.1.3	Criterion-related validity	51
(i)	Concurrent validity	51
(ii)	Predictive validity	51
2.7.2	The evolving application of validity in IO Psychology practice	51
2.7.3	Validation of a psychological assessment battery	53
2.7.3.1	Requirements for a criterion-related validity study	53
2.7.3.2	Major steps in a predictive validation study	54
2.8	CONCLUSION	55
<b>CHAPTER 3: RESEARCH ARTICLE</b>		<b>56</b>
ABSTRACT		56
3.1	INTRODUCTION	57
3.1.1	Background to the study	57
3.1.2	Trends from the research literature	59
3.1.2.1	Selection and psychological testing	59
3.1.2.2	Fair psychological testing in the multi-cultural South African context	61
3.1.2.3	Validity of psychological testing in selection	65
(i)	Cognitive ability assessment	65
(ii)	Personality assessment	66
(iii)	Competency based interviews	66
3.1.3	Research objectives	67
3.1.4	Potential value add of the study	68

3.2	RESEARCH DESIGN	68
3.2.1	Research approach	68
3.2.2	Research method	69
3.2.2.1	Research participants	69
3.2.2.2	Measuring instruments	70
3.2.2.3	Research procedure	78
3.2.2.4	Statistical analysis	80
3.3	RESULTS	82
3.3.1	Correlations between training scores and predictors	82
3.3.2	Correlations between performance appraisal ratings and predictors	85
3.3.3	Correlations between criterion questionnaire ratings and predictors	87
3.4	DISCUSSION	91
3.4.1	Findings	91
3.4.2	Limitations	93
3.4.3	Recommendations	93
3.4.4	Conclusion	94
3.5	REFERENCES	95
	<b>CHAPTER 4: CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS</b>	<b>107</b>
4.1	INTRODUCTION	107
4.2	CONCLUSIONS	107
4.3	LIMITATIONS	110
4.4	RECOMMENDATIONS	111
4.5	CONCLUSION	112
	<b>REFERENCES</b>	<b>113</b>

<b>LIST OF TABLES</b>	<b>PAGE</b>
TABLE 1. PREDICTIVE POWER OF ASSESSMENTS	24
TABLE 2. DEMOGRAPHIC BREAKDOWN OF SAMPLE	70
TABLE 3. CUSTOMER CONTACT COMPETENCIES: DEFINITIONS	72
TABLE 4. CCC SCALE PARAMETERS AND DESCRIPTIONS	73
TABLE 5. PJM MATCH BAND CATEGORIES	73
TABLE 6. COMPETENCY BASED INTERVIEW MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS	74
TABLE 7. PERFORMANCE APPRAISAL MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS	75
TABLE 8. CRITERION QUESTIONNAIRE MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS (KPA, ORGANISATION CULTURE AND KEY BEHAVIOURS ITEMS)	76
TABLE 9. CRITERION QUESTIONNAIRE MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS (TIME KEEPING AND DISCIPLINARY MEASURES ITEMS)	77
TABLE 10. EXAMPLE ITEMS FOR KEY BEHAVIOURS MEASURED IN CRITERION QUESTIONNAIRE	78
TABLE 11. DESCRIPTIVE STATISTICS FOR PREDICTORS AND CRITERIA	81
TABLE 12. CORRELATIONS BETWEEN TRAINING SCORES AND PREDICTORS	83
TABLE 13. CORRELATIONS BETWEEN PERFORMANCE APPRAISAL RATINGS AND PREDICTORS	86
TABLE 14. CORRELATIONS BETWEEN CRITERION QUESTIONNAIRE RATINGS AND PREDICTORS	88

## LIST OF FIGURES

## PAGE

Figure 1.	The selection process	17
Figure 2.	Summary of the major steps undertaken in conducting a predictive validation study	54

# CHAPTER 1

## SCIENTIFIC ORIENTATION TO THE RESEARCH

### 1.1 INTRODUCTION

This study explored how the various components of the selection process related to the job performance of a customer service agent (CSA) in the airline industry. In this chapter, the background to and motivation for the research is discussed. The research problem is formulated, the aims of the research outlined and the paradigm perspective detailed. Following this, the research design and process are described. In conclusion, the layout of the chapters of the dissertation is presented.

### 1.2 BACKGROUND TO AND MOTIVATION FOR THE RESEARCH

Customers are currently more spoilt for choice than they were just a few years ago. They are far more knowledgeable and much more in touch with what is going on around them. They know that if they are dissatisfied with a service or product, another company will be able to provide for their needs. It follows therefore that organisations providing excellent customer service stand a higher chance of retaining their customers. In contrast, poor customer service could negatively impact any business (Machado & Diggines, 2012).

Customer service jobs must meet customer needs, and acceptable behaviour is usually defined by relevant social norms (Huang & Ryan, 2011). A company can differentiate itself from competitors by providing reliable, quality customer service and by ensuring that customers are satisfied by this customer service (Connellan & Zemke, 1993; La Grange & Roodt, 2001). However, one could argue that it is not really the service that is of importance, but rather, the customer's perception of the service. This then implies that the employer should look at the overall service from the customer's point of view (Brink & Berndt, 2008). The performance of their customer service employees should therefore be a priority for managers in the

customer service environment. Furthermore, in a very competitive market, being able to select employees who are likely to perform excellently appears to be essential.

In light of the above, an airline company in South Africa identified the need to explore the effectiveness of its selection process for CSAs. Potential CSAs are required to go through a rigorous selection process in order to meet the airline's standards for quality. The selection process, designed on the basis of thorough job analysis, typically involves curriculum vitae (CV) screening for gross negative disqualifiers; completion of verbal, numerical and checking ability tests; completion of a personality questionnaire and the calculation of the candidate's Person-Job Match (PJM); as well as a competency based interview.

Cronbach and Gleser (1965) describe future job performance as the foundation on which candidates should be evaluated when applying for a job. Therefore, personnel selection can add value to an organisation by maximising employees' performance by controlling the quality of individuals entering, moving around in, and leaving the organisation (Theron, 2007). In addition, recruitment and selection are integral to the majority of organisations' human resources strategies and alternative methods of assessment are utilised in order to evaluate various candidates. While companies need to ensure that their selection methods allow them to select employees who are likely to perform their roles effectively, there are also legal requirements, which they need to comply with when using psychological assessment in their selection processes (Nzama, De Beer & Visser, 2008).

In South Africa, the Employment Equity Act (Act 55 of 1998) aims to achieve equity in the workplace and has specific relevance to selection and psychological assessment. The Act stipulates that psychological testing and other similar assessments are prohibited, except when they have been scientifically shown to be reliable and valid, when they can be applied fairly to all employees, and when they are unbiased against any employee or group. In addition, the Act states that what is being measured in psychological assessment should be relevant to the particular requirements of the job (Employment Equity Act, 1998). As such, psychological assessments and interviews could be considered effective ways to measure

individuals against objective criteria for specific positions (Nzama et al., 2008). The requirement of selection methods to therefore, be valid in that they measure what they are meant to measure and adequately predict future job performance is essential.

Over the past 20 years or so, the confidence that researchers have displayed in the validity of personnel selection methods has been extremely significant (Robertson & Smith, 2001). In South Africa and abroad, numerous studies have found evidence of predictive validity from ability tests, personality assessments and competency based interviews predicting job performance in customer service and sales environments (Bartram, 2004; La Grange & Roodt, 2001; Nicholls, Viviers & Visser, 2009; Nzama et al., 2008; Swanepoel, 1998).

In addition, validation studies have been carried out in the airline industry for decades (Damitz, Manzey, Kleinmann & Severin, 2003; Murdy, Sells, Gavub & Toole, 1973). In South Africa, pilot selection and trainee pilot selection assessment batteries have also been the focus in some research (see Flotman, 2002; Mnguni, 2011). In a search of the University of South Africa's library catalogue and available e-resources on the Sabinet databases, no validation studies appear to have been conducted regarding the selection of CSAs in the South African airline industry.

### **1.3 PROBLEM STATEMENT**

While high fuel prices and other factors are making conditions difficult for the airline industry, there is also growing competition. Airlines need to look at ways to reduce costs and increase profitability. As a result, increasing and optimising job performance and how smartly employees work in this highly pressurised industry is of the utmost importance. Despite adequate assessment data being available, the company which is the focus for this research had not, to date, carried out a validation study on its CSA selection battery in order to ensure that it was effectively predicting future performance, or that it met the legal requirements of being fair and unbiased and measuring relevant criteria of the job being assessed. The selection battery consists of measures of ability, personality and behaviour. Specifically, ability

measures include the Verbal Interpretation Test (VCC1), the Numerical Reasoning Test (NP6.1) and the Basic Checking Test (CP7.1C). The Occupational Personality Questionnaire (OPQ32) is included as a personality measure and a competency based interview as a measure of behaviour. As part of a competency based selection process, competency scores of the Customer Contact Competencies (CCCs) as reflected in SHL's Customer Contact Competency Inventory (CCCI), and a Person Job Match (PJM) score are derived from the personality and ability measures in the selection battery and thereafter used to select the best applicants.

This study is of benefit to Human Resource Practitioners as well as Industrial and Organisational (IO) Psychologists interested in selection and the validation of selection batteries. It should also add value to professionals and managers in fast-paced customer service organisations, giving them a better understanding of the selection tools that are most effective at predicting future job performance in this industry. In order to ensure best practice, validation studies should be carried out consistently across all assessment practices.

### **1.3.1 Research questions**

In light of the previously discussed problem statement, the following research questions are posed:

- Do the scores of the VCC1 correlate significantly with future job performance?
- Do the scores of the NP6.1 correlate significantly with future job performance?
- Do the scores of the CP7.1C correlate significantly with future job performance?
- Do the scores of the CCCs correlate significantly with future job performance?
- Is a candidate's PJM score a valid predictor of future job performance?
- Do the scores of the competency based interview correlate significantly with future job performance?

## **1.4 AIMS OF THE RESEARCH**

In relation to the above-mentioned research questions, the following general and specific aims were formulated for the study:

### **1.4.1 General aim of the study**

The general aim of the study was to determine the predictive validity of the CSA selection battery by determining whether the ability and competency scores derived from the selection tools correlate with job performance.

### **1.4.2 Specific aims of the study**

The following aims were formulated for the literature review and empirical study respectively:

#### *1.4.2.1 Specific theoretical aims of the literature review*

The aims of the literature review were to:

- Conceptualise ability testing, personality assessment and competency based interviews and how they can be operationalised as measurement constructs.
- Gain an understanding of job performance and how it can be measured.
- Provide clarity on the VCC1, NP6.1, CP7.1C, and the OPQ32 (as a basis for deriving CCC and PJM scores), as well as the competency based interview as assessment tools.

#### *1.4.2.2 Specific aims of the empirical study*

The empirical study aimed to:

- Explore the empirical relationship between ability testing, personality assessment and the competency based interview respectively to job performance.

- Determine whether the scores of the VCC1, NP6.1, CP7.1C, CCCs, PJM and competency based interview predict job performance of CSAs in the South African airline industry.
- Put forward recommendations to the airline company regarding future selection decisions.

#### 1.4.2.3 Hypotheses

Based on the background to the research and the problem identified above, the following statistical hypotheses were formulated:

- H01: The ability scores of the VCC1, NP6.1 and CP7.1C do not significantly predict job performance of a CSA in the airline industry.
- H1: The ability scores of the VCC1, NP6.1 and CP7.1C significantly predict job performance of a CSA in the airline industry.
- H02: The competency scores of the CCCs do not significantly predict job performance of a CSA in the airline industry.
- H2: The competency scores of the CCCs significantly predict job performance of a CSA in the airline industry.
- H03: The PJM score does not significantly predict job performance of a CSA in the airline industry.
- H3: The PJM score significantly predicts job performance of a CSA in the airline industry.
- H04: The competency scores of the competency based interview do not significantly predict job performance of a CSA in the airline industry.
- H4: The competency scores of the competency based interview significantly predict job performance of a CSA in the airline industry.

### 1.5 THE PARADIGM PERSPECTIVE

The research took place in the IO Psychology discipline, which is described as an applied area of psychology concerned with the study of behaviour related to work, organisations and productivity (Cascio, 2001). In addition, the area of selection

practices falls within the sub-discipline of personnel psychology, and psychological assessment can be also regarded as a sub-discipline applicable to the broader Psychology field.

A paradigm can be defined as a collection of rationally linked concepts, assumptions or propositions which orient research and thinking (Bogdan & Biklen, 1998). The research was informed by the positivist paradigm. Positivists tend to test a theory through observation and measurement to forecast and control forces all around us (O'Leary, 2004). In summary, positivism as applied to human behaviour pertains to how objective, measureable, predictive and controllable behaviour can be, as well as the laws and rules which impact on it. Positivism guides research methods that focus on surveys, quantitative analysis and the like (Dash, 1993).

## **1.6 RESEARCH DESIGN**

### **1.6.1 Research approach**

In this study, a quantitative research approach was used in order to be able to predict and describe what was happening (Mouton & Marais, 1992). In order to achieve the specific research goals, a cross-sectional survey was used, assessing relationships between the different variables within a specific population (Struwig & Stead, 2001). Primary data were utilised for the research, which followed a correlational approach when the data were analysed.

### **1.6.2 Research variables**

In the research, the relationships between several independent/predictor variables and a dependent/criterion variable were studied. An independent variable can be manipulated to determine what effect it has on another variable. The variable which is affected by the independent variable is called the dependent variable (Clark-Carter, 2004).

In this study, ability tests, personality assessment and the competency based interview were the predictor measures. Numerical reasoning, verbal interpretation, checking ability, CCC scores, PJM scores and competency based interview scores were defined as independent/predictor variables, while job performance was the dependent/criterion variable.

#### *1.6.2.1 The predictor/independent variables*

Applicants to the position of a CSA in the airline company completed three ability tests (verbal, numerical and checking ability), a personality questionnaire and a competency based interview. The scores from the ability tests and the personality questionnaire were also used to compute competency scores for the CCCs and a PJM score based on the necessary competencies selected for the CSA position. These were defined through the job analysis process.

##### *(i) Verbal ability test*

Verbal ability was assessed using the Verbal Interpretation Test (VCC1). This test measures the ability to understand straightforward written information in order to arrive at reasoned conclusions. This task is relevant to sales and customer service jobs where jobholders receive product information in written form, as well as written communication from customers and/or colleagues. The measure has been found to be both valid and reliable (SHL, 2000;2008b).

##### *(ii) Numerical ability test*

Numerical ability was assessed using the Numerical Reasoning Test (NP6.1). It assesses basic reasoning skills with numbers. Questions involve decimals, fractions and graphs, and calculators are not allowed. The measure has been shown to be both valid and reliable (SHL, 1992; 2008c).

*(iii) Checking ability test*

Checking ability was assessed using the Basic Checking Test (CP7.1C), which measures the speed and accuracy of checking at a basic level. It is principally for use with clerical staff, whose jobs include routine checking. The measure has also been found to be both valid and reliable (SHL, 2004; 2008c).

*(iv) CCC scores*

The CCCI model is directly related to the CSA job profile, with the competencies written to model jobs in Customer Service roles (SHL, 2000). The model consists of 16 competencies, the scores of which are derived from the ability scores as well as scores from the OPQ32 (SHL, 2000).

*(v) PJM score*

If a person has the abilities necessary to perform the inherent requirements of a job, then a good person-job fit exists. The PJM score provides an indication of the candidate's 'degree of fit' to a role, with the higher the score the better the person-job fit. The PJM incorporates the ability tests and the OPQ32, with the key behaviours that influence work performance being provided by the CCC scores (SHL, 2013).

*(vi) Competency based interview*

The selection interview is the most popular method to select applicants for a job (Seijts & Kyei-Poku, 2009). Competency based interviews were conducted with all applicants. They were asked structured situational questions on specific competencies identified for the CSA position, based on the job profile and developed through job analysis process. Their responses were then scored objectively by more than one observer.

### 1.6.2.2 *The criterion/dependent variable*

Criterion data used as a measure of performance for CSAs included:

#### *(i) Training course scores*

Training course results are commonly used as performance criterion data in validity studies (Carretta & Ree, 2000). Up until December 2008, only an Overall Training score was available. From January 2009, Self Study Training and Passenger Handling Training courses were introduced, with scores available for each. These, along with Overall Training scores, were incorporated as performance data. These courses had to be passed before the candidate could assume the role of a CSA.

#### *(ii) Performance appraisal ratings*

The airline company's formal appraisal scores were included in the analysis, with the most recent appraisal scores (Half Year 2011) being correlated with the predictors. These scores are formulated during the performance management process and may inform decisions regarding development, remuneration and further disciplinary measures. The sample sizes for the appraisal scores for other years were too small to include in this study and, at the time of the study, the Full Year 2011 scores were not available.

#### *(iii) Criterion questionnaire ratings*

The direct supervisors of the CSAs completed a criterion questionnaire, unrelated to the performance management process, rating the CSAs on key performance areas, organisational culture factors, absence without leave (AWOL), sick leave, time keeping and disciplinary issues and behaviours representing the CSA job-related competencies identified through a job analysis. These ratings are for research purposes only and have no impact on decisions regarding development, remuneration and further disciplinary measures.

## **1.7 RESEARCH METHOD**

### **1.7.1 Research participants**

In order to get the most out of the size of the samples for each correlational matrix, the sample consisted of any individual for which any predictive data was available, that is any applicant to the CSA position in a South African airline company (N=1223). However, the sample sizes for each of the variables used differ due to predictor and criterion data not being available for each step of the selection process. All predictor and criterion data was available for 192 full-time employees working as CSAs in a South African airline company. All these candidates were appointed over a three-year period. Data was available for their ability tests (verbal, numerical and checking), personality questionnaire and competency based interview. While the whole sample completed numerical tests, they were not all assessed on the NP6.1, which is the reason why the size of the sample for the NP6.1 is smaller than those of the other ability tests. In addition, criterion data, in the form of training scores, performance appraisal ratings and criterion questionnaire ratings were available.

### **1.7.2 Measuring instruments**

In terms of measuring the predictor variables, the following instruments/measures were used: VCC1, NP6.1, CP7.1C, OPQ32 and a competency based interview (see descriptions in item 1.6.2.1). Job performance was measured using training course scores, performance appraisal ratings and criterion questionnaire ratings (see descriptions in item 1.6.2.2).

### **1.7.3 Research procedure and ethical considerations**

The research study was divided into two parts, namely the literature review and the empirical study,

### *1.7.3.1 Literature review*

The literature review explored the theory related to psychological assessment in selection with specific focus in the South African multi-cultural context. It also explored psychometric constructs with specific focus on validity and the different approaches that can be followed in a validation study.

### *1.7.3.2 Empirical study*

The empirical study collated all data already collected from ability tests scores, personality assessments, competency based interview ratings, training result scores, and performance appraisal ratings. It further involved conducting sessions with supervisors to rate behavioural competencies of the CSAs reporting to them and integrating this data into the data set. The data were then backed up to ensure safety of information and finally analysis took place.

### *1.7.3.3 Ethical considerations*

The purpose of the research and plan for conducting the research were explained to management. Following this, the company granted permission to conduct the study. Data were documented with sufficient detail so that another researcher would be able to continue the work if necessary, or to replicate it at a later stage.

Secondary data were used in the study. Data were obtained from the recruitment department (competency based interview scores), OD department (psychological assessment and performance data) and training department (training scores). In order to keep the data confidential, only trained professionals who needed to have knowledge of the data and/or those who were directly involved in the initial collection and analysis of the data had access to it. All others involved in the research, such as those providing either predictive or criterion data, were assured of the confidentiality of their contributions to the study.

Candidates signed a written informed consent form giving permission to use the data for research and to share the information with stakeholders that needed to be involved in the decision making process. In order to ensure anonymity, once data had been linked to a candidate, the candidate's name was deleted and replaced by a code/number which was used for the analysis. In order to ensure informed consent prior to conducting the rating sessions with the customer service supervisors, an electronic outline of the research study and its purpose was sent out to each supervisor with an explanation of what was expected of them. They then were required to respond that they understood their role in the research and were content to contribute, after which a session was set up for this purpose. Furthermore, the front page of each criterion questionnaire stated, "This information will only be used for research purposes, is confidential and will in no way effect the current position and status of the employee."

Finally, once the data had been finalised, feedback was given to the company on the findings of the study, limitations and recommendations.

#### **1.7.4 Data analysis**

This study made use of correlation, the denoting of association between two quantitative variables, where the one variable increases or decreases a fixed amount for a unit increase or decrease in the other (Swinscow & Campbell, 1999). Analysis of the research data was done using Pearson's Product Moment Correlations. Once calculated, the correlation coefficients were used to analyse the relationship between the predictors and criterion data. The full data set derived from all CSAs in the study was analysed with IBM SPSS Statistics 20.

### **1.8 RESULTS**

Chapter 3 presents a discussion of the biographical makeup of the sample and the descriptive statistics, taking correlations into account. This is preceded by a write up of the results, with the data being displayed graphically in tables and graphs.

## **1.9 CHAPTER LAYOUT**

The chapter layout of the study is as follows:

### **Chapter 2: Literature Review**

The literature review provides a more in-depth literature study of the independent and dependent research variables.

### **Chapter 3: Research Article**

This chapter comprises a research article which will include all information regarding the actual research study – methodology, sample, data collection and data analysis. In this chapter, statistical information is provided, the hypotheses are tested and results discussed.

### **Chapter 4: Conclusion, limitations and recommendations**

In the final chapter, a concluding discussion of the findings is presented. Limitations are indicated and suggestions/recommendations for future research are noted.

## **1.10 CHAPTER SUMMARY**

In this chapter, the background and motivation for the research were outlined. The problem statement, the aims, the research paradigm, the research design and method as well as a layout of the chapters were provided. In Chapter 2, a literature review regarding the validation of an assessment battery for selection will be presented.

## **CHAPTER 2**

### **LITERATURE REVIEW**

#### **2.1 INTRODUCTION**

Through a review of relevant literature this chapter discusses the subject of employee selection from an IO psychology best practice perspective. A central focus of the discussion is around selection measures and, specifically, psychometric assessment as part of the selection process. Another central discussion is on psychological assessment in the multi-cultural South African context, competency based assessment and the psychometric concepts that are essential in ensuring fair assessment. Specific attention is given to validity as an essential psychometric property of psychological measures used during the selection process. The chapter concludes with a short summary of the main issues.

#### **2.2 SELECTION FROM AN IO PSYCHOLOGY PERSPECTIVE**

It is essential in our current, rapidly changing, competitive world for companies to ensure that the most competent employees of the highest quality are found and selected in order to propel them forward in terms of optimisation and productivity (McLagan, 1997). The Advisory, Conciliation and Arbitration Services (<http://www.acas.org.uk>) is largely funded by the United Kingdom's Department of Business, Innovation and Skills (BIS) and aims to improve organisations and working life through better employment relations. The Advisory Conciliation and Arbitration Services emphasizes the importance of allocating adequate money and time to selection planning and practice as poor selection can lead to poor results and contribute to the possible downfall of companies. Through effective selection, companies will save costs as they are creating an effective workforce and reducing labour turnover. Attracting large numbers of job applicants is not an issue for companies but selecting the best applicants is many employers' greatest concern (Branine, 2008) and remains a priority for the consulting IO Psychologist (Sackett & Lievens, 2008).

### **2.2.1 Psychological assessment and selection**

Through proper planning of the selection process, the most appropriate methods can be found or developed in order to hire the best possible people. In this regard, organisations carry out assessments as part of a selection process, in order to measure the potential and actual performance of those individuals currently in their employ, as well as those who could potentially work for them in the future (Bartram, 2004; Theron, 2009). Selection is, therefore, very much concerned with the tools and methods used to assess these individuals to ensure that the most efficient decisions are made by organisations (Muller & Schepers, 2003). Psychological assessment is furthermore regarded as a psychological act (Health Professions Act, 1974). The IO Psychologist's knowledge of behavioural dynamics and nuances enhances the efficiency of selection through the application of psychological assessment as a key selection mechanism. IO Psychology further adds value to the selection process as the many possible assessment tools and techniques used during selection should be evaluated on the basis of their reliability, validity, interpretability, practicality and legality.

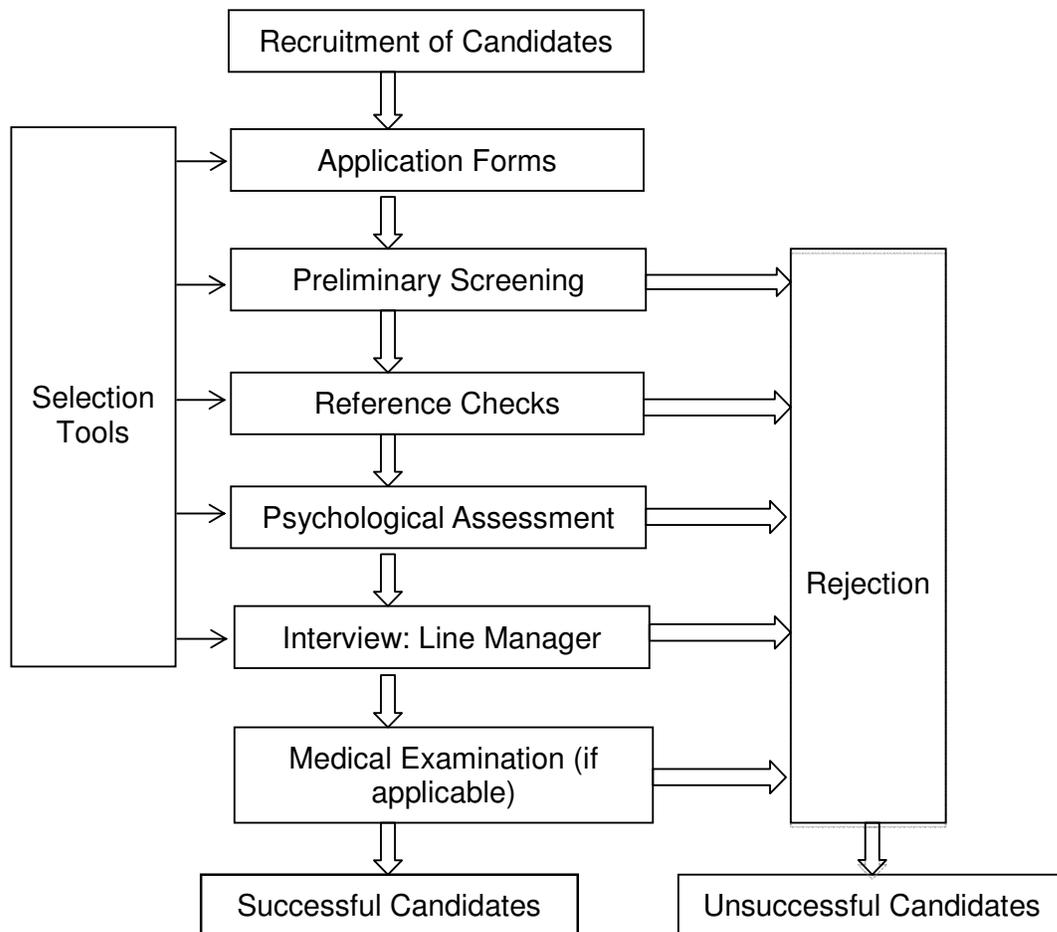
Barrick, Feild and Gatewood (2011) define selection as a process where information about an individual is collected and evaluated so that an offer of employment can potentially be extended, addressing the future interest of both the individual and the organisation. Noe, Hollenbeck, Gerhart and Wright (2003) similarly summarise personnel selection as the process by which organisations decide who will and who will not be given access into their company.

### **2.2.2 Selection Process**

Novit (1979) points out five basic elements in the selection process. The process begins with the definition of organisational goals. From these goals, job design can occur, where duties and responsibilities of the individual are defined. The criteria for job success (and what will contribute to successful performance) are defined, followed by the traits, skills and qualities required in an individual (job specification). Finally, selection instruments are chosen and carried out. These selection

instruments determine if the applicant possesses the desired traits, characteristics and skills.

There is no one specific selection process that is used universally but most organisations use selection processes based on the successive hurdle technique, where, in order to be appointed to a position, applicants are required to be successful in each step of the process (Van der Merwe, 2002). In line with this, there are likely to be candidates who are rejected after each stage, thus introducing the concept of restriction of range. Since employees need to “pass” each stage of the selection process before eventually being appointed, restriction of range sets in (Shavelson, 1988) and as a result, small and moderate (instead of large) correlations may be observed in validation studies of some selection processes. In Figure 1, Van der Merwe’s (2002) perspective on the “general nature of the selection process,” has been slightly adapted:



**FIGURE 1** The selection process (Van der Merwe, 2002)

A fair amount of work needs to take place by staff responsible for selection before the process is actually carried out with those being selected to the company. The major concern through this process is to collect sufficient information about the candidates that will be closely related to future job performance as well as to utilise this information to identify and select the finest candidates (Barrick et al., 2011). Therefore, a number of steps should take place before a selection battery is finalised in order to ensure that the correct information is being gathered.

After adequate planning has taken place, the first step would be for a thorough job analysis to be conducted. This would then lead to the development of a job description and job specification. Relevant talent requirements, which include job performance dimensions, as well as the knowledge, skills and abilities required for the job, would be identified through this process. Once these dimensions of the job have been decided, selection instruments would need to be found or chosen in order to measure these constructs. These could include application forms, assessment centre exercises, interviews and psychological tests (Barrick et al., 2011; Bartram, 2004; Schmidt & Hunter, 1998). The validation of the assessment instruments is vital as a final step before the measures can be used to assess applicants for the job (Barrick et al., 2011). Next, job analysis as the foundation to selecting appropriate assessment tools, is discussed.

### **2.2.3 Job analysis: the foundation to competency based assessment**

According to Bartram (2004), organisations assess candidates to measure both actual and potential performance. This allows the possibility of predicting performance, rather than random decision making, thereby the criterion (what we wish to predict) is the focus rather than the actual predictor itself (Bartram, 2004; Theron, 2007). It is through job analysis that one is able to access what the criterion entails.

Job analysis focuses on the tasks and activities of the job, its responsibilities and duties, the knowledge and skills required and any other factors relevant to the job performance being successful (Grobler, Warnich, Carrell, Elbert & Hatfield, 2002;

Kriek, 2000; Mirabile, 1997; Potgieter & Van der Merwe, 2002). Cascio (1991) describes the goal of job analysis as defining each particular job with regards to the behaviours necessary to perform it. In addition, he discusses the two components of job analysis: job descriptions and job specifications.

The job description is the most common end product of job analysis and identifies the job, gives a brief job summary and indicates tasks and responsibilities. In summary, it defines the job in terms of its requirements and consists of job characteristics, which include standards of performance and procedures (Voskuijl & Evers, 2008). The information incorporated can be used in competency based selection and other important tasks. When drafting the job description, it is important, as much as possible, to avoid generalisations, prioritise responsibilities, use criteria that can be measurable in the real world, and obtain help from others, if necessary. More than all of this, it is essential for the job description to be as accurate as possible so that the right candidate can be found for the job (Grobler et al, 2002; Noe, 2005).

Job specifications refer to what abilities, skills, knowledge and other personal characteristics are required for the job. The job specifications usually provide a starting point for competency frameworks, against which candidates may be assessed (Voskuijl & Evers, 2008). The competency profile identifies the competencies (areas of personal capability that allow employees to effectively perform their jobs by achieving results or accomplishing tasks) needed for every job as well as the knowledge, skills, behaviour and personality characteristics which underlie each competency (Noe, 2005). This competency profile needs to be in the language of the organisation and needs to represent the organisational culture, be comprehensive and contain observable behaviours only (Grobler et al, 2002; Noe, 2005).

The job analysis will assist in outlining performance criteria to be used to evaluate the employee's successful performance in a job, as it is only possible to describe such success if the specifics of a job are defined (Coetzee & Roythorne-Jacobs, 2010).

Conducting an accurate and effective job analysis essentially involves collecting, recording and analysing information (Muchinsky, Kriek & Schreuder, 2005). It is recommended that the following steps be followed to gather the information (Aamodt, 2010; Coetzee & Rawthorne-Jacobs, 2010):

- Identifying the tasks to be performed, the tools to be used to perform these tasks and the conditions that will be present when the tasks will need to be performed.
- Writing task statements to develop a task inventory which will be outlined in the job description.
- Rating the task statements in order to ascertain the frequency and importance of the tasks being performed.
- Determining the essential knowledge, skills, abilities and other characteristics (or competencies).
- Selecting assessment instruments to measure knowledge, skills, abilities and other characteristics.

The job analysis, being an essential part of the selection process, acts as the starting point for all the other steps. This process defines a number of job competencies most important for competency based selection. These competencies, making up the competency profile, are validated to determine that they are measurable and attainable.

Muller and Schepers (2003) point out that selection is most concerned with the tools and methods utilised to assess potential candidates. Multiple sources of data are indicative of a good selection process. These could include interviews, psychometric data and performance data (Hoffman & McPhail, 1998). Psychological assessment tools are used to identify the most appropriate and needed talent in relation to the job and organisational requirements. As a result, psychological assessment is commonly employed when making decisions regarding employees, including selection decisions. At this point a distinction needs to be drawn between psychological assessment and psychological testing, which is but one component of psychological assessment. The two concepts will be discussed in the following section.

## **2.3 CONCEPTUALISATION: PSYCHOLOGICAL ASSESSMENT VS PSYCHOLOGICAL TESTING**

### **2.3.1 Psychological assessment**

According to Foxcroft and Roodt (2001), many overlapping terms are used in the field of psychological assessment. Gregory (2007, p. 2) defines psychological assessment as “a standardised procedure for sampling behaviour and describing it with categories or scores.” He adds that most tests possess standards and norms, to which the results are compared in order to predict other, more important behaviours. These will be elaborated on later in this chapter. Psychological assessment is a process-oriented activity in which a diverse amount of information is collected. This involves using psychological tests and other measures of behaviour. In the Western World, these include tests of knowledge and skill, tests of ability and personality, interviews, work samples and various assessment centre or simulation exercises (Branine, 2008; Foxcroft & Roodt, 2001; Bartram, 2004).

Shum, O’Gorman and Myors (2006) describe psychological assessment as a process with a broader scope, and psychological testing as the administration process, followed by obtaining and interpreting the test scores of a psychological test. Psychological assessment is, therefore, a multi-dimensional process, with information gathered and integrated from a number of measurement sources of personal attributes and behaviour. Therefore, psychological tests are but one component of psychological assessment (Branine, 2008).

### **2.3.2 Psychological testing**

While only one step in the selection process, psychological tests can be extremely valuable as they are validated and provide an objective method to measure particular aspects of certain behaviour. At the same time, by providing an understanding of a candidate’s strengths and development areas, psychological tests add value to other assessment measures, such as interviews, in the selection process (Van der Merwe,

2002; Parkinson, 1999). These tests have many uses, which contribute to interpretation of individuals, groups and other facets of everyday life.

In the past, psychological testing in South Africa was seen as biased and unfair, but this line of thinking has changed somewhat with the focus shifting to the advantages of good, fair assessment measures. If used with care, psychological tests can enhance the efficiency of the selection process (Foxcroft, 1997; Owen & Taljaard, 1996; Paterson & Uys, 2005).

Psychological tests generally consist of a standardised series of questions or statements which measure a particular characteristic of an individual. They are usually made up of multiple items which are indicators of the characteristic being measured (Spector, 2012). In order for the test design to be of good quality, the test needs to produce fair, reliable, valid and predictive data. Test publishers also need to ensure that test users are trained and/or have the necessary skills and up-to-date knowledge and information regarding the tests (Parkinson, 1999).

A large number of psychological tests exist and the choice of which tests to use is often aided by the nature of the characteristic that one wishes to measure. Four differentiating characteristics of psychological tests are described (Spector, 2012; Muchinsky et al., 2005) as follows:

- Group versus individual tests: Individual tests are administered to one individual at a time, while group tests are administered to more than one candidate at any given time. It makes sense for group tests to be used when large groups of candidates need to be assessed.
- Paper-and-pencil versus performance tests: Responses to paper-and-pencil tests are in written form, either on a piece of paper or an electronic format. In performance tests, an item of apparatus or an object needs to be manipulated in some way by the candidate.
- Power versus speed tests: Power tests contain difficult items which the candidate may not get all correct. However, there is no time limit, unless this is put in place in order to avoid inconvenience for the administrator. Speed tests, on the other hand, usually have a sizeable number of easier questions that the

candidate is likely to answer correctly. A time limit is imposed. Therefore, the test usually reveals the candidate's speed.

- Closed-ended versus open-ended tests: A closed-ended test usually allows the candidate to select one of several possible choices for responding, whereas an open-ended test involves producing a response, such as an essay, rather than selecting a correct answer.

Different types of psychological tests are developed based on the assessment of different psychological constructs, broadly categorised as cognition, personality and behavioural constructs (Barnard, 2010). As such, psychological tests include cognitive ability/intelligence tests, personality questionnaires, integrity tests, and emotional intelligence tests, amongst others.

### **2.3.3 Assessment of psychological constructs**

In developing an assessment battery, the first thing that must be done is to decide which exercises should be used in order to measure the identified abilities and competencies. This is done by examining the job analysis, the competency profile and the competencies defined as essential or key to the job. Competencies and competency based assessment will be discussed later on in this chapter. Each exercise/instrument must be relevant and related to the position and behavioural, cognitive and personality tests generally need to be represented in order to assess whether the behaviour of the candidate is in line with the competencies required. Any psychological measure must be fair, unbiased and valid (Potgieter & Van der Merwe, 2002). Schmidt and Hunter (1998) explored 85 years of research findings regarding the predictive validity of a number of methods of selection. Common assessment measures which could be relevant to the job profile used in the research study are mentioned in Table 1, below (Schmidt & Hunter, 1998). It is also clear that basing selection decisions on multiple sources of data increases the predictive validity.

**TABLE 1. PREDICTIVE POWER OF ASSESSMENTS**

<b>1</b>	<b>Perfect Prediction (does not exist)</b>
.63	Ability test and structured interviews
.51	Cognitive ability tests
.51	Structured interviews
.40	Personality questionnaires
.38	Unstructured interviews
.37	Assessment centres
.35	Biographical data measures
.26	Reference checks
.18	Job experience (years)
.10	Years of education
<b>0</b>	<b>Random Prediction</b>

Research has shown that personality questionnaires, structured interviews and ability tests all have good validity, which can be generalisable (Barrick & Mount, 1991; Bartram, 2004; Schmidt & Hunter, 1998). Evidence has also shown that the inclusion of more than one assessment (e.g., ability and personality assessment) in a selection process will put the organisation in a more favourable position than if there is only one element included (Mount, Witt & Barrick, 2000; Sackett, Gruys & Ellingson, 1998; Schmidt & Hunter, 1998). Psychological assessment is used to measure different constructs relating to human attributes and thus has a very positive impact on selection and development. The following discussion touches on the potential validity of the specific assessment constructs chosen in this study as predictors in order to assess candidates for selection as CSAs:

#### *2.3.3.1 Cognitive ability assessment*

Spector (2012) defines ability as a person's capacity to learn or carry out a particular task. In certain jobs, psychomotor ability may be relied on. Central to ensuring that an organisation can achieve its goals is the appointing of employees who have the abilities needed to carry out the jobs that they are selected for (La Grange & Roodt,

2001). In customer service jobs, cognitive abilities are relevant to tasks that involve information processing and learning. Many jobs rely on both of these. Cognitive ability assessments have been recognised as very good predictors of job performance and even better predictors of training performance and should, therefore, be included in selection batteries in order for companies to ensure that the companies are performing at their best level (Anderson, Lievens, Van Dam & Ryan, 2004; Barnard & Schaap, 2005; Bartram, 2005; Hunter & Hunter, 1984; Salgado & Anderson, 2002). The types of ability tests that have been used in selection include numerical ability tests, verbal ability tests, deductive reasoning tests, mental ability tests, mechanical ability tests, clerical ability tests, and physical ability tests, amongst others (Bartram, 2005; Taylor, 1994). Companies may look towards verbal and numerical ability tests as these measure skills which may indicate the candidate's ability to cope in the specific role or training. Therefore, it is important that ability tests that correlate with the skill requirements of the job are administered (Taylor, 1994).

There are many cognitive ability tests available in South Africa and elsewhere. These tests measure some form of knowledge which has been acquired up until the time that the candidate takes the test. Most of these tests are classified as crystallised ability measures, the results of which can be strongly affected by schooling or cultural influences (Barrick et al., 2011; Cattell, 1971; Taylor, 1994). Much research is based around using cognitive ability for selection, in particular around whether this has a negative impact on fairness between different cultures (Ployhart, 2006). In addition there appears to be consensus that while cognitive ability assessments are fair and pose some predictive validity, they are not necessary sufficient in predicting future performance and that other measures are also of importance (Bartram, 2004; Murphy, Cronin & Tam, 2003). It is therefore important to combine tests of cognitive ability with other assessments in order to more effectively predict performance. Outtz (2002) found that combining the use of ability and personality assessments added more value than just using ability assessment, as personality assessment adds incremental validity to predicting performance.

### 2.3.3.2 *Personality assessment*

A personality trait is defined by Spector (2012) as the tendency to act in a specific way, across various situations. Recent research has shown support for the fact that personality, or more specifically, a range of characteristics of personality, also predicts future job performance and therefore should be used for selection (Bartram, 2004; 2005; Borman, Penner, Allen & Motowildo, 2001; La Grange & Roodt, 2001).

Specifically, the Big Five personality factors have displayed relationships with job performance (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Ones, Dilchert, Viswesvaran & Judge, 2007; Rothmann & Coetzer, 2003). The five-factor model of personality, made up of Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness, has been developed and elaborated on over the past six decades and represents a structure of traits. It is based on a theory that certain personality traits are stable over time and universal, and is the most commonly used personality model around the world (Nzama et al., 2008; Rothmann & Coetzer, 2003). Research has been conducted around considering traits that are less broad than the five-factor model, and which could provide even greater validity (Ployhart, 2006).

Personality assessments are commonly used in South Africa (Van de Vijver & Rothmann, 2004). They have been shown to be valid predictors of job performance (Goldberg, 1993; Tett & Christiansen, 2007) and, according to certain opinions, unlike cognitive ability tests, do not generally have a differential effect on candidates from different cultural groups (Rothmann & Coetzer, 2003). There are opposing opinions regarding the validity of personality assessment in multi-cultural populations, for example, that certain cultural differences may result in varied reactions to personality questionnaires (Anderson & Witvliet, 2008; Van de Vijver & Rothmann, 2004). However, Bartram (2004) mentions that personality assessments can actually reduce adverse impact. Huang & Ryan (2011) found that personality states were likely to change depending on situational features in customer service work and therefore, in order to understand the impact that situational influences may have on

personality at work, it may be useful to examine personality states in the work environment.

Personality assessments differ from ability tests in that they do not require a correct or incorrect answer for each question. Rather, responses are subjective and tend to measure underlying characteristics that generally encourage individuals to act in a certain way (Barnard, 2010). While the use of personality in selecting employees was previously met with reservation, certain meta-analytic studies have shown the importance and value that it can add in selection as a predictor of performance (see Schmidt & Hunter, 1998; Mount & Barrick, 1998). Barnard (2010) adds that personality tends to be stable across time and that individuals tend to be distinct from each other with regards to their behaviour.

### *2.3.3.3 Competency based assessment*

One of the criticisms of personality assessment is that it may be difficult to assess because an individual's underlying personality constructs are not necessarily observable by others. In competency based assessment, the candidate is assessed as to whether the required competencies to perform a job successfully are possessed by focusing on an output-based approach (Barnard, 2010). Competency based interviews are considered the most common, and one of the best, selection tools in most of the world (Anderson & Witvliet, 2008; Branine, 2008; Schmidt & Hunter, 1998).

Evidence shows that the structured competency based interview, focusing on job competencies, adds much value in predicting training success and job performance and is an important aspect to consider including in the selection process (Chung-Yan, Hausdorf & Cronshaw, 2005; Nzama et al., 2008; Ployhart, 2006; Salgado & Moscoso, 2002; Schmidt & Rader, 1999; Schmidt & Zimmerman, 2004). The structured interview is also much more effective than an unstructured interview and, therefore, the more structured the interview, the higher the validity (Nzama et al., 2008; Schmidt & Rader, 1999).

Janz (1982) found that structured competency based interviews were made up mainly of questions around descriptions of experience and activity (40%), and behaviour descriptions focusing on past and future behaviour (33%). Lesser components were questions on self-evaluative information (23%) and credential questions (4%). The main reason for the interview appears to be the likelihood of being able to assess social and communication skills because of the face-to-face interaction (Branine, 2008). There could be dangers of subjectivity, which can possibly be resolved by conducting a panel interview with more than one interviewer. In addition, the candidate may not perform optimally due to anxiety (McCarthy & Goffin, 2004) and this should be taken into consideration. As such, it is not advisable to use an interview as the only predictor of job performance.

## **2.4 PSYCHOLOGICAL ASSESSMENT IN THE SOUTH AFRICAN MULTI-CULTURAL CONTEXT**

Psychological assessment measures are frequently utilised for the purpose of development and selection in South Africa, and make a valuable contribution to selection decisions (Van der Merwe, 2002; Van de Vijver & Rothmann, 2004). Van de Vijver & Rothmann (2004, p. 6) describe multi-cultural assessment as “a new branch of the tree of psychological assessment” and confirm that it is likely to be around for some time. To understand the scope and significance of multi-cultural assessment in the South African context, a reflection on historical to current day practices as well as the current legal and ethical context of assessment in the country is required. In discussing the legal and ethical contexts the issues of fairness and bias are also considered.

### **2.4.1 A historical reflection on psychological testing in South Africa**

Early psychological testing in South Africa imported tests to focus on educability and trainability of black South Africans. Other tests developed in South Africa were standardised for whites (Huysamen, 2002). The testing was carried out on blacks against these standards and results were presented without considering the impact of culture and the schooling environment. It was clear at that point that cultural

differences affect test outcomes and that it is not so easy to just create culture-free tests. This strengthened the view that culture may contribute strongly to systematic error in assessment results. According to England (1991, in Foxcroft & Roodt, 2001), it is therefore clear that one of the major aims of psychological testing in South Africa is to facilitate equal opportunity for all, which can easily be compromised through a lack of understanding of other cultures (England, 1991). Even recently, psychological test use in selection has been the subject of much speculation and criticism, largely to do with the issue of fairness (Theron, 2007).

Fairness is a central and essential component of psychological assessment as there is a need to develop and use assessment methods that are fair, relevant, practical and transparent, and which can improve the perception of and support for psychological testing (Potgieter & Van der Merwe, 2002; Skarlicki, 2003). Fairness in assessment practices involves a desire to ensure equal opportunities for every person. This would allow for the candidate identified as most suitable for the role to be selected (Bedell, Van Eeden & Van Staden, 1999; Foxcroft & Roodt, 2001). In addition, there should be appropriate, ethical, fair and professional utilisation of assessment measures and results, bearing in mind the rights and needs of all individuals and groups involved in the assessment process. Moreover, the assessment carried out must be relevant and in line with the purpose of the research. Through all this, the political, cultural and social environment within which the assessment is taking place, and how these elements may alter results of the assessment, must be taken into account (International Test Commission's Guidelines for Test Use Version 2000, 1999; Foxcroft & Roodt, 2001).

The responsible use of psychological assessments is emphasised by psychologists (Van de Vijver & Rothmann, 2004). The selection process takes place within a legal context in that it addresses the future interests of both the individual and the organisation (Barrick et al., 2011). Psychological assessments in the current South African setting are being shaped by a number of factors, including the law and the actions of the government; the need to create fair and unbiased measures that can be used for all cultural groups in the country; the need for assessment practitioners to understand and act on all their responsibilities with regard to carrying out of ethical

testing; the training of professionals as well as practice guidelines for professionals as laid out by both statutory and other bodies (Foxcroft & Roodt, 2001; Mauer, 2000). In this section, the legal and ethical context within which psychological assessment occurs and which regulates psychological assessment practice will be discussed. Then, the crucial constructs of fairness and bias will be further explored. The section will be concluded with a deeper look at culture-fair assessment in the South African context.

## **2.4.2 Psychological assessment in the context of labour legislation**

While companies are required to ensure that the selection process determines which candidates are able to perform the necessary tasks and make a valuable contribution, they also need to comply with legislation (Nzama et al., 2008). Three acts address psychological assessment practice in South Africa, namely the Health Professions Act, the Employment Equity Act and the Labour Relations Act.

### *2.4.2.1 Health Professions Act (56 of 1974)*

While dealing with a number of issues, the Health Professions Act (56 of 1974) focuses on matters central to psychology and deals with issues around the regulation of psychological assessment practice in South Africa. This includes the establishment and mandates of the Health Professions Council of South Africa (HPCSA), the Professional Board for Psychology and the Psychometrics Committee of the Professional Board for Psychology; the requirement for professional registration and scope of practice; and the classification of psychological tests (Mauer, 2000).

#### *(i) Regulatory structure and scope of practice*

The Act makes provision for the establishment of the HPCSA and the Professional Boards beneath it. The Professional Board for Psychology was established in order to regulate psychologists, all of whom need to be registered under the provisions of the Act. The Act outlines the requirements for registration of a psychologist and intern

psychologist. It further states that psychological assessments, which are considered psychological acts, have to be performed by a registered psychologist (Health Professions Act, 1974).

The Board of Psychology used its mandate to establish the Psychometrics Committee, which deals with the issues around non-psychologists, such as psychometrists and psychotechnicians being able to conduct and administer a certain number of tests. This includes being responsible for all theoretical and practical training of these professionals in line with the policies of the professional board. They are able to carry out these tests as long as they have been certified by the Psychometrics Committee of the Professional Board of Psychology. Furthermore, the test user has to comply with whatever limits relevant to that category of user are placed on the test's use. In addition, the tester needs to seek guidance from a psychologist where specific extra input would be advantageous to the testing and understanding of the results. Finally, the tester would need to have received all the training necessary for the carrying out of the tests (Health Professions Act, 1974).

On 2 September 2011, the government notice *Regulations defining the scope of the profession of psychology* (No. R. 704) (<http://www.hpcsa.co.za>), outlined, in a new annexure, additional psychological acts falling into the practice of registered counsellors, psychometrists, clinical psychologists, counselling psychologists, educational psychologists, research psychologists, industrial psychologists, neuro-psychologists and forensic psychology. This allowed for a wider scope of practice for both psychometrists and psychologists. Further additional regulations to govern the control and use of psychological and other similar assessments are expected to be presented at a later date (Health Professions Act, 1974).

(ii) *Test classification*

The American Psychological Association (APA) proposed guidelines for the classification of tests in order to prevent psychological assessment measures from being accessed by individuals not qualified to use or understand them. This classification system was initially followed in South Africa until the Health Professions

Act also gave a mandate to the Psychometrics Committee to provide a psychological assessment classification system. The Psychometrics Committee, therefore, became responsible for classifying the use of all psychometric measures which constitute a psychological act, to regularly revise any measures, and report to the Professional Board in this regard. The classification system identifies whether a predictor is a psychological measure or not, based on whether it will result in a psychological act or not (Barnard, 2010; Foxcroft & Roodt, 2001; Health Professions Act, 1974).

According to the *Policy on the classification of psychometric measuring devices, instruments, methods and techniques* (Form 208) (<http://www.hpcsa.co.za>) of the HPCSA's Professional Board for Psychology, utilising a psychometric measuring tool that assesses a psychological construct and which, in relation to its contents or necessary responses, may cause the test-taker to become anxious or embarrassed, is constituted as being a psychological act. Form 208 also describes two classification categories for psychological tests. The first category refers to psychological tests that can be used to varying extents by psychometrists but need to be under the control of psychologists with regard to selecting, administering, scoring, interpreting and reporting on a test. The second category consists of those tests that can be used by other health professionals.

The *List of tests classified as being psychological tests* (Form 207) (<http://www.hpcsa.co.za>) states that when a test has been classified by the Psychometrics Committee as a psychological test, conditions under which a test can be carried out are somewhat relaxed and therefore the assessment will be more readily available. Furthermore, the test user still has the responsibility to ensure that the test is valid for the purpose for which it is being utilised, that suitable norms are referred to, and that, in the case of an international test, the necessary studies are done to determine if the test is biased with regard to culture and that extra care is taken when interpreting the results.

The *Policy on the classification of psychometric measuring devices, instruments, methods and techniques* (Form 208) (<http://www.hpcsa.co.za>) also reminds the researcher/tester of the history of segregation in South Africa's not so distant past,

which caused psychological testing to be culturally unfair, stereotypical, insensitive and inappropriate. Because of this, there are very few tests that are standardised for all South Africans, and which have taken cultural and other diversity into account. Form 208 further emphasises that development and adaptations of psychological testing must be culturally unbiased and fair. It is also essential for test users to be very familiar with the policies of the tests being used. Moreover, all individuals and their personal rights must be treated with respect, informed consent regarding professional procedures is non-negotiable, no test taker must be discriminated against as the tests must be normed for all types of people, there must be no conflict of interest and, when a psychometrist performs a test, a psychologist should be consulted when the need arises.

*(iii) Registration of test professionals*

The Board makes a clear distinction between psychologists and psychometrists, as well as the requirements for registering in one of these categories. Once registered and, with the necessary training, industrial psychologists are permitted to use all levels of tests and may control all assessment work done by psychometrists. In *Training and examination guidelines for psychometrists* (Form 94) (<http://www.hpcsa.co.za>), the assessments that psychometrists are not permitted to use are outlined along with required core competencies and steps to follow to be registered.

The Employment Equity Act (No 55 of 1998) and Labour Relations Act (No 66 of 1995) were developed to deal with labour matters, both emphasising fair labour practices and the development of equity in the workplace (Mauer, 2000). Relevant legal implications arising from these acts are discussed below.

*2.4.2.2 Labour Relations Act (66 of 1995)*

The Labour Relations Act No 66 of 1995 provides a legal framework for fair labour practices, safeguarding employees from being exploited by employers. It also points out the rights of applicants and employees to be treated fairly (Mauer, 2000). As

such, it describes unfair labour practices, which could be either an act or an omission, and which could include unfair discrimination against an employee or the unfair conduct of an employer regarding certain HR issues (Labour Relations Act, 1995). To treat an employee unfairly could therefore be a legal issue.

#### 2.4.2.3 *Employment Equity Act (55 of 1998)*

The Employment Equity Act No 55 of 1998 requires that tests need to be shown to be fair, unbiased, valid and reliable. This introduces major implications for South African test practitioners if tests have not been investigated for bias and have not been cross-culturally validated (Foxcroft & Roodt, 2001; Mauer, 2000).

The Act aims to achieve equity in the workplace by encouraging equal opportunity and fair treatment of employees by eliminating unfair discrimination, as well as applying measures regarding affirmative action, as per the employment equity plan, in order to ensure equitable representation in all areas of the labour force (Employment Equity Act, 1998).

Unfair discrimination is described by the Act as discriminating or being discriminated against, in any way, direct or indirect, regarding employment policy or practice on grounds such as race, gender, pregnancy, marital status, ethnic/social origin, colour, sexual orientation, age, disability, religion, culture, language, HIV status or political opinion. Because of this, practitioners could be summoned to prove that particular tests are not discriminatory to specific groups of people. The Act also mentions that it is not unfair discrimination to differentiate, reject or take preference for an individual based on the inherent requirement of a role, or to take affirmative action measures aligned with the Act's purpose (Employment Equity Act, 1998). Avoiding unfair discrimination is possible in psychological assessment by using valid, reliable and unbiased selection instruments (Theron, 2007).

The American Psychological Association (APA) code of professional ethics entitled the *Standards for educational and psychological testing* was developed in order to prevent psychological assessments from being misused. The code of ethics and

other issues pertaining to ethics and ethical conduct will be discussed in the next section (Muchinsky et al., 2005).

### **2.4.3 Ethical principles and conduct**

Codes of conduct for assessment are usually designed based on ethical principles that protect the needs of those being assessed (Barnard, 2010). While the APA code of professional ethics provides universal guidelines for the whole psychology profession, the Society for Industrial Psychology in South Africa developed guidelines for the validation and use of personnel selection procedures for the workplace. These relate to the use of results for decision-making. As a result of the legislation in South Africa, the regulations for psychological testing are very strict in comparison to many other countries, with test users requiring specific qualifications and training on specific tools (Muchinsky et al., 2005). The Professional Board for Psychology's *Rules of conduct pertaining specifically to psychology* (<http://www.hpcs.co.za>) outline the rules of conduct including those pertaining to assessment activities. This section of the *Rules of conduct* includes a discussion on assessment within a psychological context, the appropriate use of assessment methods, informed consent in assessments, test development, cultural diversity, communication of results, information for professional uses, interpreting and explaining assessment results, test scoring and interpretation services, release of test data, obsolete and outdated test results, and maintaining test security.

Some pertinent ethical principles involve rights to privacy, confidentiality and informed consent. Regarding privacy, tests should not reveal more information than necessary to make an informed decision. Therefore, tests need to be as specific as possible. Gaining irrelevant information that has no relationship to performance can be viewed as an invasion of privacy. In order to protect confidentiality, those taking the test should be informed about the purpose of the test, how the results will be utilised and who will have access to these. Informed written consent should be a prerequisite for the test-taker to allow specific individuals or departments to get access to the results. Unless written permission is given, all results should be strictly confidential (Barnard, 2010; Muchinsky et al., 2005).

Christensen (2001) goes into more detail regarding the principles of ethics in the South African context as presented by the Psychological Society of South Africa (PsySSA). He points out the professional and scientific responsibility that assessment professionals need to acknowledge and, based on this, act in the individual's best interests and take personal responsibility for their actions. This is complemented by having the competence to carry out tests, which includes keeping up to date, planning and being responsible. Moral and legal standards need to be upheld. The assessor should prioritise the welfare of the individual and groups involved in the assessment. Should the assessment professionals wish to make public statements about their services, they would need to comply with all policy guidelines. In the event of conflict regarding ethics, private information must remain private at all times and conflicts should be sorted out if possible, in the correct moral and legal ways. Christenson (2001) also lists confidentiality, necessary informed consent and that research has to be to the benefit of all involved, as essential.

#### **2.4.4 Fairness and bias**

Even though assessment takes place between the assessment practitioner and the individual being assessed, there are many stakeholders in this process. The assessment practitioner holds a large amount of power because of the knowledge that has been gained from conducting the assessment. Registered professionals, as a result of their registration, have the ability to remove discrimination, reduce inequalities and address the other issues in the South African context by carrying out meaningful, fair and ethical assessment practices, and by not abusing the power that they have in this regard (Pretorius, 2012).

While the responsibility of the assessment practitioner in ensuring fair and ethical assessment practices is of utmost importance, the responsibility of organisations in this regard is increasingly being focused on. This responsibility includes: (i) the employment of competent and adequately-trained assessment practitioners, who are supervised and mentored where necessary; (ii) using valid assessment measures for proper purposes; (iii) using assessment results non-discriminately; (iv) putting all necessary support structures in place for assessment practitioners to build a

research database that will be able to examine the fairness and efficacy of the assessment measures being used; and, most important for this study, (v) that the organisation has an assessment policy, made up of fair and ethical practices, firmly in place. In addition, the organisation should monitor how the assessment policy is being used and carried out, making changes to it where necessary (Foxcroft & Roodt, 2001).

Bias can be described as the presence of factors that create trouble in psychological assessment in cross-cultural contexts and exists where a psychological item or tool does not measure the same construct across different cultural groups (Jensen, 1980; Poortinga, 1989; Van de Vijver, 2002). As the issue of bias cannot be removed until it is properly acknowledged, it is clear that much more research is necessary regarding fairness and bias in South Africa (Huysamen, 2002). Bias in psychological testing has two components, namely, test bias and test fairness (Bergh & Theron, 2004).

Test bias exists when there is an indication that a test has a different validity for different groups of the population and occurs when different groups understand different items of the test differently or certain groups find the test easier than other groups (Bergh & Theron, 2004; Taylor, 1994). Three types of bias occur, namely, construct bias, method bias and item bias (Van de Vijver, 2002; Van de Vijver & Poortinga, 1997). Construct bias occurs when a construct measured across different cultures is not identical. Method bias involves the presence of nuisance variables as a result of factors related to method, which impacts on test fairness (Bergh & Theron, 2004). Lastly, item bias occurs where individuals from different cultures have the same level of a trait but are likely to endorse the trait differently.

There may be test bias between genders or language groups. With regard to language, conducting a test in English could significantly disadvantage many South Africans whose English comprehension and expression are poor compared to communication in their first language. A responsibility exists for test users to ensure that the same construct is measured across different cultural groups (Foxcroft, 2004). In line with the Employment Equity Act, the prohibition of biased assessment measures contributes to combating unfair discrimination. Research on bias is

common and necessary to ensure that the law is complied with (Abrahams & Mauer, 1999; Schaap, 2003; Schaap & Basson, 2003; Theron 2007).

Test unfairness occurs when policies and strategies are unfair, test administration is biased or decisions are made using test scores that are unfairly prejudiced, whether test results turn out to be biased or not (Bedell et al., 1999). This very often leads to discrimination.

An understanding needs to be achieved regarding fair and unfair discrimination (Theron, 2007). The Employment Equity Act regulates the treatment of employees in a fair and equitable manner, which eliminates unfair discrimination (Employment Equity Act, 1998). The purpose of assessments is to differentiate between candidates in order to ensure that the most appropriate individual is identified (Theron, 2007). This can be achieved by understanding the criteria for selection of the job through conducting a thorough job analysis, as well as creating a model of fairness with regard to selection.

While it is advisable to avoid biased assessment measures, it should be understood that bias will never be eliminated completely (Theron, 2007). Even so, steps need to be taken to minimise it as much as possible (Bedell et al., 1999; Bergh & Theron, 2004; Huysamen, 2002; Muchinsky et al., 2005). Therefore, it is necessary to utilise whatever methods are necessary in order to produce valid, reliable and non-biased test results that are non-discriminatory and fair to all South Africans (Foxcroft & Roodt, 2001).

#### **2.4.5 A deeper look at culture-fair assessment in the South African context**

Research on cross-cultural assessment is flourishing, with a constant flow of further findings being reported (Sackett & Lievens, 2008; Van de Vijver, 2002). Much focus regarding psychological assessment is placed on the psychometric areas of reliability, validity, test standardisation and bias. While culture is such an important dictating factor of the knowledge, skills and abilities that one currently has, it does not

necessarily display one's potential (Ferguson, 1956; Geisinger & Carlson, 1998). Furthermore, a misconception with regard to psychological assessments is that the higher the score, the better one's psychological adjustment or the better one's understanding or ability. While this may be true, problems may arise when the assumptions at the core of the test have a cultural bias (Friedman & Schustack, 2003). Therefore, culture plays a vital role in people's current abilities, as it influences the way that we learn, think and behave. Thus, it is clearly an essential component of the environment.

As mentioned above, many personnel policies and legislation have been introduced in the South African context to correct past injustices. In order for these procedures to work effectively for the good of all, both representativeness and merit need to be considered very carefully (Mauer, 2000). However, when a specific selection strategy gives members of a certain group a lower probability of being selected for a role than that of another group, adverse impact occurs, where evidence would be provided by the selection ratio of one group being less than 80% of the ratio of the group with the highest selection ratio (Collins & Morris, 2008; Muchinsky et al., 2005; Theron, 2009).

Test developers and users need to acknowledge that different groups of individuals are at different stages of acculturation as well as the fact that test performance is affected by socio-economic conditions and prior education of individuals. In South Africa, levels of education vary greatly. This gives those who have had a good education a much greater advantage than those who have had a poorer education. As an example, two individuals who have both been at school for the same amount of time may still have very different levels of formal education. In the past, measures reflected the culture of those who designed them and the target audience that they were meant for and those who do not share in this culture could be at a disadvantage (Magwaza, 1995; Owen & Taljaard, 1996).

Nell (1997) explained that a battery of tests should be developed for those with a lower level of education, and it should be explained in many languages. Taylor (1994) explains that many tests available in South Africa measure crystallised abilities, which are influenced strongly by culture and one's schooling, where specific

cultural groups have had a better standard of schooling and development opportunities. As such, a need has arisen in South Africa to not just assess skill, but rather potential for development, despite the skill gaps that may exist due to previous disadvantages (Taylor, 1994).

Language is generally accepted to be the most important moderator of performance in psychological testing and is closely tied to culture (Foxcroft, 2004; Nell, 1997; Van de Vijver & Rothmann 2004). In South Africa, very limited numbers of tests have been developed that can be utilised across cultural and language groups (Foxcroft, 2004). While test adaptation is an option, where the original meaning is also preserved, it is more applicable to a specific context where the same language will be used. Translating is a possible solution, but only in some cases, as there are certain languages that do not possess certain concepts that are central to what is being studied (Berry, Poortinga, Segal & Dasen, 2002). Another issue, particularly in South Africa, is the fact that many South Africans speak in a combination of languages or 'township patois' (Foxcroft & Roodt, 2001).

Another issue relates to assessments developed internationally and then introduced into South Africa without South African norms. In some examples of tests, care has been taken to limit cultural content within tests, which helps to ensure that the test can be relevant to one and all. Because responses will differ amongst cultures, tests and their scores cannot be considered to have the same meaning for people of different cultures and living in different countries – the problem of variations in acculturation. If culturally relevant measures which have norms applicable to all cultures being tested are not put in place, we cannot expect to have testing that is fair (Berry et al., 2002; Van de Vijver & Rothmann, 2004).

Regarding the testing process, the testers need to make sure that the specific test is relevant to the group being tested. When it comes to the comparability of test scores across groups in a multi-cultural society such as ours, it is essential that all criteria are clearly laid out. By testers becoming more culture sensitive and through giving clearer feedback, many current misunderstandings could be resolved (Bedell et al., 1999).

Owen (1991) pointed out that large differences were picked up in mean score performances between different cultural groups taking the Junior Aptitude Tests (JAT), where language was seen to cause a bias (Owen, 1991). The GSAT test was developed with similar structures for different population groups, but when interpreting, past educational opportunities need to be kept in mind. The SSAIS-R, while being a good measure, demands a certain proficiency in either English or Afrikaans languages, which obviously varies amongst race groups and is affected by the large number of official languages and different dialects of each language (Foxcroft & Roodt, 2001). It has been pointed out that a popular test such as the Sixteen Personality Factor Questionnaire (16PF) may also not be clear to non-English speakers, and therefore certain cultures need to be evaluated differently when interpreting (Bedell et al., 1999). With regard to the Thematic Apperception Test (TAT), the stimuli and symbols used are quite westernised and may take on diverse meanings which means that responses could be hard to interpret. There is a dire need to develop a thematic apperception test that is appropriate for the richly culturally-diverse South African context (Foxcroft & Roodt, 2001).

Adverse impact can be reduced by increasing the predictive validity of psychological assessment instruments, allowing access to job opportunities for different groups in the market (Sackett & Ellingson, 1997; Theron, 2007). However, predictors should not be blamed for adverse impact.

Dynamic assessment is an umbrella term for the method of fairer multi-cultural cognitive assessment, measuring a candidate's underlying potential in an engaged, flexible manner, where an individual is given guidance and feedback on their cognitive skills (De Beer, 2006; Elliott, 2003; Grigorenko & Sternberg, 1998; Murphy & Maree, 2009). A test is administered, which is followed by a learning opportunity, after which another test is administered, with learning potential being the construct being measured (De Beer, 2006). Dynamic assessment can be described as a possible relevant alternative to "regular" psychometric assessments in the South African context due to many individuals being disadvantaged as a result of their past. As such, it could potentially limit adverse impact (Murphy & Maree, 2009).

Extensive research has taken place in South Africa around dynamic assessment (De Beer, 2006; Lopes, Roodt & Mauer, 2001; Taylor, 1994; Van Eeden, De Beer & Coetzee, 2001). In the South African context, dynamic assessment measures are used to address the reality of unfair discrimination. The APIL, TRAM-1 and TRAM-2 assess learning potential with material that is novel to each candidate, irrelevant of culture (Bedell et al., 1999). Research has shown success for learning potential measures at predicting training and other criteria (Taylor, 1994). Similarly, the LPCAT has been shown to be a culture-fair screening assessment measure, measuring learning potential without relying on language or academic level. Rather, it utilises non-verbal ability to assess reasoning ability (De Beer, 2006).

Claassen (1995, in Foxcroft & Roodt, 2001) suggests that individuals representing all cultures should be integrally involved in the developing of valid tests. Researchers have also pointed out that there is still a lot of work to be done regarding finding ways to delineate and compile norm groups (Van de Vijver, 2004; Van de Vijver & Rothmann, 2004) and that more cross-cultural studies must be done (Bartram, 2004). A balance also needs to be established where popular and useful international assessment measures can be adapted and normed for the South African context as well as developing new, local measures which relate to our unique cultural context (Van de Vijver & Rothmann, 2004).

In conclusion, it is important to move past the bias and unfairness of the past and to ensure that fair, ethical, reliable, valid measurement tools, free from adverse impact and unfair discrimination, are utilised for selection and other human resources decisions. In order to ensure that a focus is placed on assessments being fair and related to the requirements of the job for which the candidate is being assessed, competency based assessment can be used as a suitable approach for developing tests or for compiling assessment batteries (Erasmus & Arumugan, 1997; McLagan, 1997; Potgieter & Van der Merwe, 2002). Competency based assessment will be discussed in the following section.

## **2.5. COMPETENCY BASED ASSESSMENT**

In our constantly changing world we live in, an organisation has become defined by how competent it and its employees are. Boyatzis (1982, p. 21) defines a job competency as “an underlying characteristic of a person, which results in an effective and/or superior performance in a job.” He goes on to say that “it may be a motive, trait, skill, aspect of one’s self-image or social role, or body of knowledge which he or she uses.” Competencies have three features in common, namely, that they relate to a specific work or role; they are able or have the potential to create high performance; and they contain a number of individual features (Zhao & Du, 2011).

Assessment in organisations is usually centered on competency models and competency based assessment (Bartram, 2004). Individual knowledge, skills, abilities and other characteristics - or combinations of some or all of these - make up the competencies, used to match an individual with a job (Campion et al., 2011; Heinsman, de Hoogh, Koopman & Van Muijen, 2007; Mirabile, 1997). Furthermore, a set or collection of competencies required to perform a specific job effectively, is usually referred to as a competency model. This can be used for many reasons in organisations including hiring new employees, training employees, evaluating performance of employees, promoting employees, developing careers and compensating employees (Bartram, 2004; Campion et al., 2011; Lawler, 1994; Zhao & Du, 2011). Competency modelling aims to develop models which lay the foundation for integrated human resource solutions across the organisation (Bartram, 2004). The Great Eight structure defines eight broad competency factors that make up an organisational performance model, namely (i) Leading and Deciding, (ii) Supporting and Cooperating, (iii) Interacting and Presenting, (iv) Analyzing and Interpreting, (v) Creating and Conceptualizing, (vi) Organizing and Executing, (vii) Adapting and Coping, and (viii) Enterprising and Performing. The Great Eight competency structure or model has been replicated a number of times, including in the analysis of the competencies relating to the Occupational Personality Questionnaire (OPQ32) (Bartram, 2004; 2005).

As indicated previously, information derived from the job analysis is integrated into the competency profile. This reflects all desired performance requirements for the position (Cheetham & Chivers, 1997; Potgieter & Van der Merwe, 2002). Through competency based assessments, which include combinations of different assessment methods, people are tested as to whether they have what is required for a job. The assessment is in line with South African labour legislation and has the ability to cut across cultural boundaries (Bartram, 2004; Potgieter & Van der Merwe, 2002; Zhao & Du, 2011). Competency based assessments have high predictive validity and high face validity because they have a direct effect on many essential parts of jobs (Potgieter & Van der Merwe, 2002). In South Africa especially, these assessments may also help in alleviating discrimination against candidates who do not have a high level of education, but have the necessary skills (Grobbelaar, Roodt & Venter, 2004; Potgieter & Van der Merwe, 2002; Ulrich, 1998).

A competency based assessment process can be carried out in order to identify the most appropriate and/or optimal candidate for a specific position by establishing whether an individual has the ability to perform what is required by a job. It is an integrated process and not just a single test (Hager & Gonczi, 1994; Potgieter & Van der Merwe, 2002). Competency based assessment does not focus on personality but rather on behaviour. This allows for much higher levels of fairness. More than that, competency based assessments meet all requirements of labour legislation in South Africa due to the fact that test takers are evaluated according to the critical job-related competencies required.

In summary, competency based assessments can be good, accurate predictors of how the individual will perform in his/her job and should also be able to offer substantial information on the individual, leading to fairer decisions in the workplace (Heinsman et al., 2007). From this, we can establish that competency based assessments are really important for organisations in today's world and should also be able to point out where training and upgrading is needed in order for the organisations to become world class, competitive, productive and built for excellence.

## **2.6 ESSENTIAL PSYCHOMETRIC PROPERTIES OF PSYCHOLOGICAL ASSESSMENT TOOLS**

Selection errors can have major cost implications for organisations, have an impact on employees and even have far reaching effects on a country's economy (Barnard & Schaap, 2005). Furthermore, the Employment Equity Act highlights the terms reliability and validity, thereby regulating the need to ensure that these essential psychometric properties of psychological assessment tools are prioritised and met (Lopes et al., 2001). Core to IO psychological assessment practice, these psychometric properties of reliability and validity will be discussed. Before that, an initial discussion will be presented on norms and standardisation, moderator variables and some other basic statistical concepts, which can ultimately impact on reliability and validity of assessment measures (Foxcroft & Roodt, 2001).

### **2.6.1 Norms and standardisation**

Norming is described as a type of benchmarking, where the performance or raw scores of a candidate are compared to a related group of individuals (the norm group). This process requires a careful consideration of a number of moderating factors (Paterson & Uys, 2005).

Different types of test norms exist, namely: (i) developmental scales (where certain human traits get stronger with age and experience); (ii) percentile scores (the percentage of the individuals in the sample who scored below the given raw score); (iii) standard scores which include z scores (indicating the individuals deviation from the mean), (iv) linearly transformed standard scores (which can get rid of some of the z score's disadvantage and (v) normalised standard scores (standard scores adjusted to fit normal distribution) (McIntyre & Miller, 2007).

### **2.6.2 Moderator variables**

A moderator variable is a variable that may affect the relationship between two other variables. As such, predictors and criteria for some candidates who have a particular

score on the moderator variable could display different relationships than the predictors and criteria of other candidates, who have an alternative score on the moderator variable. Examples of moderator variables could include gender, age, length of service, and nationality (Spector, 2012).

### **2.6.3 Basic statistical concepts relevant to validation studies**

In quantitative research, statistical analysis is generally used in order to be able to draw conclusions based on a large amount of data. Descriptive statistics, which will be indicated in a table relating to all predictors and criterions, are utilised to summarise data in order to be able to describe it in a meaningful fashion. Measures of central tendency look at the centre of a group of scores, the mean being the most commonly used measure. Other measures of central tendency are the mode and the median. The measure of variability shows how intensely the scores are distributed around the mean. Also important is the correlation coefficient, which is used when determining how one variable relates to another. This statistic can be used to create a regression equation indicating how great the correlation is and, therefore, how accurate the prediction will be (Coetzee, 2010).

By plotting data graphically we are able to work out whether a positive or negative relationship exists between variables. Also apparent is whether the attribute is normally distributed or whether it is somehow skewed, which can explain how easy (negatively skewed) or hard (positively skewed) the test is (Coetzee, 2010).

### **2.6.4 Reliability**

The reliability of a test refers to the consistency with which it measures what it is supposed to be measuring, in that it produces similar results at different specified points in time if the predictive construct has not been altered and similar conditions exist (Barnard, 2010). Barrick et al. (2011) explain that reliability deals with errors of measurement, where a perfectly reliable test would be free of any errors. This scenario is unlikely as selection measures are prone to some kind of error due to the candidate, assessor or the situation. Generally, the larger the amount of

measurement error, the lower the reliability of the test. Therefore, the concept of reliability is best seen as a continuum ranging from minimal consistency of measurement to almost perfect repeatability of outcomes. The majority of psychological tests fall somewhere in this range (Gregory, 2007). Five key types of reliability are identified:

#### *2.6.4.1 Test-retest reliability*

Assessment of the test-retest reliability of a selection measure is done by carrying out the assessment to the same group twice and measuring the reliability after each test application, correlating the two sets of scores in order to obtain the reliability coefficient (Barrick et al., 2011). Barnard (2010) terms the reliability coefficient the “coefficient of stability”, as it offers an indication of the test’s stability over time. The greatest danger of this method is that candidates may score slightly higher on the retest because of practice effects – being able to solve problems more efficiently and effectively the second time around (McIntire & Miller, 2007).

#### *2.6.4.2 Alternate-form reliability*

In this method, which is also known as equivalent-form reliability, two identical forms of the same test are carried out with the same group at different times. This could control the effects of memory on test-retest reliability. The two assessments are required to be truly equivalent, the construction of which is costly and time consuming and, therefore, not suggested (Barnard, 2010; Barrick et al., 2011; Foxcroft & Roodt, 2001).

#### *2.6.4.3 Internal consistency reliability*

Internal consistency reliability, which is also known as split-half reliability, provides an indication of the degree of homogeneity of the items within a specific assessment. The test scores from that assessment are split into two halves and then examined, after which the correlation between the two scores is calculated (Barnard, 2010; Foxcroft & Roodt, 2001). The major challenge here is dividing the group into two

equivalent halves (Gregory, 2007). Coefficient alpha reliability is also a measure of internal reliability. This method is based on consistency of responses to every item in the assessment, where every item is correlated with all other items, resulting in an inter-item correlations matrix (Barnard, 2010; Foxcroft & Roodt, 2001).

#### *2.6.4.4 Inter-rater reliability*

Inter-rater (inter-scorer) reliability can be determined by having two assessment practitioners score all the candidate's assessment responses, where the correlation between the sets of scores is known as the inter-rater reliability coefficient (Foxcroft & Roodt, 2001). This method is usually used when it is difficult to control the standardised testing procedures and conditions (Barnard, 2010). However, it should supplement other reliability methods and not just simply replace them (Gregory, 2007).

#### *2.6.4.5 Factors affecting the reliability of an assessment*

Due to reliability coefficients being estimates, certain factors may have an impact on the size of the coefficient, constituting a measurement error. Sources of measurement error or moderators (as discussed in the previous section) may include individual differences among respondents, the method of estimating reliability, a sample which is not representative of the population, questions which are too difficult, issues regarding administration and scoring of an assessment, restriction of range, or whether speed or power tests are used, amongst other factors (Barnard, 2010; Barrick et al, 2011; Gregory, 2007).

While reliability is important, it may constrain validity, which, in essence, defines the meanings of the test scores. If a test is unreliable, it is not possible for it to be valid. On the other hand, reliability is crucial but is not an adequate precursor to validity (Gregory, 2007). In the next section, the essential psychometric property of validity will be discussed.

## **2.7 VALIDITY**

According to Schmidt (2006), validity is recognised as the key issue in psychological assessment and in IO psychological assessment. Validity is concerned with what is being measured by the test and how well it is being measured (Barnard, 2010; Foxcroft & Roodt, 2001). Putting in time and effort to ensure that assessment instruments are valid in turn ensures that legal requirements are adhered to (Employment Equity Act, 1998) and that the most productive job applicants are likely to be selected into the organisation (Lopes et al., 2001; Pelsier, Bergh & Visser, 2005; Schmidt & Hunter, 1998).

The most popular approach to measuring the relationship between test scores and criterion measures is to calculate the correlation between the criterion and the test, which is known as the validity coefficient (Gregory, 2007). The higher the value of the validity coefficient, the higher the accuracy of the test's prediction of the criterion. Perfect validity does not exist and correlations rarely exceed .80 (Gregory, 2007). This is due to the fact that there are other variables, apart from predictors, that could have an effect on the test results. These may include environmental conditions, the assessment candidate's circumstances or characteristics of the individual assessing. When the predictor is affected in such a way, it is known as criterion contamination. While this is usually inevitable, it is important to try to minimise this as much as possible (Barnard, 2010).

In this section, the different types of validity will be discussed, followed by an exploration of the debates around conceptualising validity. After this, the requirements to determine the validity of a test will be looked at.

### **2.7.1 Types of validity**

Three methods to determine the validity of a test exist, namely, content validity, construct validity and criterion-related validity.

### *2.7.1.1 Content validity*

Content validity is a non-statistical type of validity. It involves examining whether the content of the assessment covers a representative of the sample of the psychological aspect that the assessment was created to sample. This is a useful determinant of validity when much is known regarding the variable to be measured and, in contrast, it becomes more difficult to determine when not much is known about the construct (Barnard, 2010; Foxcroft & Roodt, 2001; Gregory, 2007).

Face validity, which is not an actual psychometric term, is related to content validity but is based on people's impression of how appropriate items tend to be in relation to the psychological constructs being measured. As it is not based on a subject matter expert's opinion, it should not be used in isolation, but rather in addition to content validity. However, it is crucial for face validity to exist in order to avoid doubt or dissatisfaction from those using the assessment measure (Barnard, 2010; Gregory, 2007)

### *2.7.1.2 Construct validity*

Construct validity involves amassing evidence that an assessment measure is based on comprehensive psychological theory as it tests hypotheses about the relationships between assessment measures and the constructs they aim to measure (Barrick et al, 2011; McIntire & Miller, 2007). Some examples of constructs that can be researched using this type of validity test include verbal ability, numerical ability, eye-hand coordination and anxiety tests, amongst others (Foxcroft & Roodt, 2001).

Construct validity is determined by convergent and discriminant validity. Convergent validity exists when other variables, with which an assessment measure should correlate, show a high correlation. Where the correlation is minimal with other variables, from which it theoretically should differ, discriminant validity exists (Foxcroft & Roodt, 2001; McIntire & Miller, 2007).

### 2.7.1.3 *Criterion-related validity*

A criterion is defined as any outcome measure against which an assessment will be validated (Gregory, 2007). Criterion-related validity is a quantitative measure that sets out to determine if an assessment really predicts what it claims to measure (Foxcroft & Roodt, 2001; McIntire & Miller, 2007). Barnard (2010) explains that, in order to do this, a correlation coefficient is calculated between the predictor (the assessment measure) and the criterion (behaviour being predicted). Two kinds of criterion-related validity are described, namely concurrent validity and predictive validity (Gregory, 2007).

#### (i) *Concurrent validity*

In concurrent validity, the test scores are obtained around the same time as (or simultaneously with) the criterion measures, indicating the extent of accuracy with which the test scores estimate a candidate's present position or behaviour regarding the criterion in question (Foxcroft & Roodt, 2001; Gregory, 2007).

#### (ii) *Predictive validity*

Predictive validity, on the other hand, involves collecting data on the predictor first and then, at a later date, collecting data on the criterion (Lopes et al., 2001). This type of validity, therefore, is used to predict future performance-related behaviour and is very useful for selection and other employment assessments, where decisions are likely to be made regarding candidates in terms of appointments (Barnard, 2010; Gregory, 2007). Predictive validity underpins the notion that psychological assessment can be used to make decisions (Foxcroft & Roodt, 2001).

## **2.7.2 The evolving application of validity in IO Psychology practice**

IO Psychology places a large emphasis on criterion-related validity as a central element to predicting future performance in employee selection (Schmidt, 2006). Generally, this type of validity would be carried out by analysing the role, identifying

predictors and criteria, assessing a number of individuals for the same role, correlating the assessment scores with criterion data and, if significant correlations occur, appointing the candidates with the most effective scores (Cascio, 1995). Schmidt (2006) explains a number of considerations that challenge the conceptualisation of validity in terms of IO psychological assessment. These include the inherently normative nature of psychological assessment, conceptual challenges to the exclusive emphasis on criterion-related procedures to demonstrate validity, challenges to the assumption of context-independence of predictor-criterion relationships, and the limited magnitude of predictive meta-analytic concepts.

Framing validity in terms of only technical terms appears to lead to gaps in how it is implemented in practice. On the other hand, framing validity as an action concept allows the effectiveness of assessments to be evaluated in terms of not just how they are used, but how they are used in the organisational context (Anastasi & Urbina, 1997; Guion, 1998; Muchinsky, 2004; Schmidt, 2006). Schmidt (2006) goes on to say that this provides a scientifically accountable foundation for practice, where a wider range of skills is provided by professionals, whereby they are able to act more in correspondence with the values of their profession, and thus validity is highlighted in the workplace. Furthermore, by combining theory and practice into what they do, IO psychologists' actions appear to be consumed by validity.

Another challenge of developing valid selection measurements is to ensure that these measures also add value, do not discriminate unfairly and minimise adverse impact (Theron, 2009). Adverse impact can never be eradicated, but can be reduced by increasing the predictive validity of the selection measures. Moreover, looking for alternative assessment measures to address adverse impact would not be of benefit to South Africa. Rather, adverse impact itself needs to be focused on by developing competency potential and creating opportunities for candidates to be successful in jobs, otherwise adverse impact will remain an issue going forward (Theron, 2007, 2009). Therefore, a proactive approach needs to be taken to address obstacles to valid selection.

In multi-cultural assessment, validity can be enhanced by a number of methods including documenting in the manual how a test has been made suitable for multi-cultural use; exporting international tests to other countries and studying whether bias exists; developing culture-specific norms for members of minority groups; developing new instruments that are cross-culturally fair; and studying the factors that threaten the validity of cross-cultural assessments, bearing in mind that improving their quality is the priority (Van de Vijver & Rothmann, 2004).

### **2.7.3 Validation of a psychological assessment battery**

Using validated employee selection measures is essential for organisational effectiveness and can ultimately lead to greater levels of individual, team and company performance. In addition, they are essential for making legally justifiable selection decisions. While validation work may be challenging, the use of validated procedures is essential for the selection of a productive workforce. The importance, therefore, of using up-to-date methods to validate these tools cannot be downplayed. Moreover, it is vital that the decisions made using the results of these procedures are legally defensible (Van Iddekinge & Ployhart, 2008). Validation studies, therefore, explore the relationships between a selection process and the after-effects of that process.

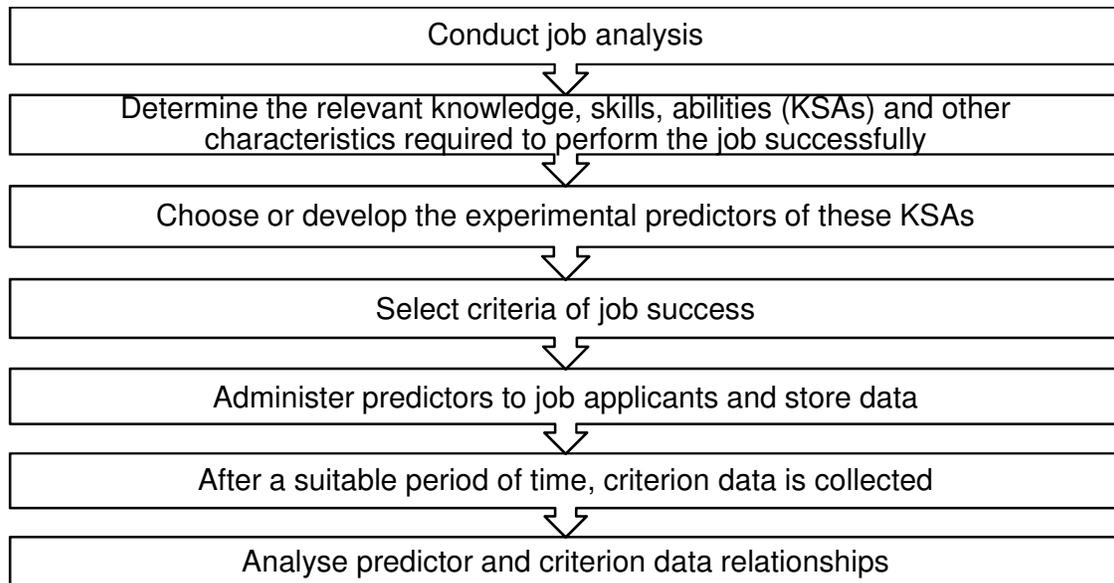
#### *2.7.3.1 Requirements for a criterion-related validity study*

In order to conduct a validation study, certain minimum requirements must be met, as a poor study is worse than having no study at all. A poor validation study could lead to the rejection of selection measures which could have aided in selecting quality candidates, or the choice to use selection measures which do not actually predict future performance of employees (Barrick et al., 2011). According to Barrick et al. (2011), four requirements need to be in place before a validation study can be attempted, namely: (i) the position should be stable and not undergoing change; (ii) a criterion which is relevant and contamination-free should exist or be possible to develop; (iii) the validation study should be based on a sample that represents the people and jobs regarding which the results will be generalised; and (iv) there should

be a large and representative enough sample from which to collect predictor and criterion data.

### 2.7.3.2 Major steps in a predictive validation study

A predictive validation study allows the researcher to collect the assessment results at one time and then collect the criterion data at a later stage (Gregory, 2007). The correlation of the assessment scores and the on-the-job performance can only take place once the other steps of the validation process are followed (Barrick et al., 2011; Spector, 2012;), as summarised in Figure 2 below:



**FIGURE 2 Summary of the major steps undertaken in conducting a predictive validation study** (Barrick et al., 2011)

It may take some time to collect all the data, depending on how many candidates are assessed per month. However, predictive validation is appropriate for employee selection as it studies how candidates will be able to perform once in the job role (Barrick et al., 2011).

## **2.8 CONCLUSION**

In this chapter, a review of relevant literature was discussed, focusing on selection, the difference between psychological assessment and psychological testing, as well as competency based assessment. Psychological assessment in the South African multi-cultural environment was discussed. This was followed by a look at the psychological criteria for assessment measures. Finally, the major steps of a predictive validation study were outlined along with a discussion on selecting predictors for the study. In Chapter 3, a research article which explores the empirical study – methodology, results and a discussion – will be presented, following a short introduction to the literature which has been presented in Chapter 2.

## CHAPTER 3

### RESEARCH ARTICLE

#### ABSTRACT

**Orientation:** Validation of psychological tests is an imperative to ensure efficient, fair and legally defensible selection of human talent.

**Research purpose:** To determine the predictive validity of a psychological assessment battery consisting of ability tests, a personality questionnaire and a competency based interview applied in the selection of customer service agents (CSAs) in a South African commercial airline company.

**Motivation for the study:** Validation studies have taken place in the customer service environment locally and internationally as well as in the airline industry. In South Africa, however, no validation studies regarding selection of CSAs in the airline industry have been conducted to date. Research about the effectiveness of a relevant psychological assessment battery will provide value for managers involved in CSA selection.

**Research design, approach and method:** Validation of the assessment battery was conducted by means of a non-experimental, correlational study. Through non-probability, purposeful sampling, 1223 individuals who had applied for the CSA position were included in the study. All predictor and criterion data was available for 192 appointed CSAs. Criterion data measuring job performance consisted of training course scores, annual performance appraisal ratings and criterion questionnaire ratings.

**Main findings:** The results showed statistically significant relationships between ability test scores, the essential and highly important Customer Contact Competencies (CCC) as identified by the job analysis, the Person Job Match (PJM) score and competency scores of the behavioural interview, and job performance.

**Practical implications:** Human resource professionals and managers in customer service environments need to recognise whether assessment measures predict future job performance of CSAs.

**Contribution/value-add:** The findings contribute to the body of research focussing on the predictive validity of different types of psychological tests in the work place and thus add to existing literature on psychological assessment used in selection. Pragmatically, the study contributes to the efficient selection of CSAs in the airline industry.

### **3.1 INTRODUCTION**

#### **3.1.1 Background to the study**

In the current competitive economic climate, much focus is being placed on the assessment tools utilised in the employee selection process, the aim being to appoint the best possible employees, who will perform optimally and, thereby, contribute to the success of the company. As a result, choosing assessment tools which measure what they are designed to measure and which are fair, unbiased and valid, is essential to finding and selecting the best candidates.

The study contributes theoretically to the bulk of knowledge on selection and the importance of validating selection assessment tools being used in the South African multi-cultural context, which is regulated by legislation and guided by ethical principles and frameworks.

According to Machado and Diggines (2012), organisations have a much better chance of holding on to their clients if they provide them with superior service, which will hopefully, in turn, positively affect their business. Customer service jobs depend on the fulfilment of customers' needs and, subsequently, customers feeling pleased with the service that they receive (Connellan & Zemke, 1993; La Grange & Roodt, 2001). As a result, organisations which provide good customer service are able to gain competitive advantage. Even more important than the service given, is the perception of the service by the customer. This should encourage organisations and their employees to view service from the point of view of the customer (Brink & Berndt, 2008).

In order to ensure customer satisfaction, focus should be placed on customer service employees' job performance and the ways in which to measure potential for future job performance of candidates applying for customer service jobs (Cronbach & Gleser, 1965; McLagan, 1997). This will ensure that the best possible employees are appointed. Customer service agents (CSAs) are appointed in airline companies to provide high standard customer service to promote and support customer loyalty to the company brands. The job performance of staff can be optimised if the qualities of people entering the organisation can be measured (Theron, 2007). The most suitable candidates for selection should be identified only after adequate data is collected from all candidates in the selection process. This requires adequate planning (Barrick, Feild & Gatewood, 2011).

Within a South African commercial airline company, potential CSAs partake in a thorough selection process, the design of which is based on job analysis, so that the required quality standards are upheld. The selection process begins with screening of curriculum vitae (CVs) for gross negative disqualifiers. Thereafter, candidates complete ability tests (verbal, numerical and checking), a personality questionnaire and a competency based interview, after which CCC scores (derived from the ability scores as well as scores from the OPQ32) are calculated, along with a Person-Job Match (PJM) score. Following this, a final decision is made regarding whether the individual is selected.

In the airline industry, high fuel prices and other factors make it difficult for companies to compete for market share, while, at the same time, having to save costs and increase profit. The need, therefore, to perform at the highest level and to motivate staff to give of their best is paramount, not only to their success, but to their survival. One airline company requested a research study be done on the validity of the selection battery they use to select CSAs in order to ascertain whether the selection measures are, in fact, contributing to the prediction of future job performance. The company had not previously conducted a validation study, even though performance data was available.

Before the empirical study took place, a review of the research literature was carried out. The following section includes an exploration of the pertinent topics of selection and psychological testing, fair psychological testing in the multi-cultural South African context, and validity of psychological testing in selection.

### **3.1.2 Trends from the research literature**

#### *3.1.2.1 Selection and psychological testing*

Selection is defined as a process during which information about a potential employee is collected and evaluated in order to decide on the employability of the candidate with regard to a particular job (Barrick et al., 2011). The future interests of the organisation and the individual are central to the selection process. At the same time, all legal and ethical aspects, as well as the reliability, validity, interpretability and practicality of all selection methods must be taken into account (Barrick et al., 2011; Noe, Hollenbeck, Gerhart & Wright, 2003).

According to Novit (1979), the selection process is initiated with the defining of organisational goals. This is followed by job design and, thereafter, explanation of the criteria for job success. The individual traits, skills and qualities required are then defined, after which selection assessments are chosen and applied. During the selection process, enough data regarding future job performance are collected in order to select the most ideal applicants (Barrick et al., 2011).

Thorough job analysis constitutes the foundation of an effective selection process, focusing on all factors relevant to the job (Barrick et al., 2011; Grobler, Warnich, Carrell, Elbert & Hatfield, 2002; Kriek, 2000; Mirabile, 1997; Potgieter & Van der Merwe, 2002). According to Cascio (1991), job analysis results in a job description and job specification, ultimately identifying the relevant dimensions of successful job performance as well as the knowledge, skills and abilities required for the job (Potgieter & Van der Merwe, 2002; Voskuil & Evers, 2008). Every possible action should be taken to ensure that these documents are accurate as they provide a starting point for competency frameworks and a competency profile, which describe

behaviours which should be observed in the role (Grobler et al., 2002; Noe, 2005; Voskuijl & Evers, 2008).

Following this, selection instruments can be decided upon in order to measure the identified constructs and ensure reliable decisions are made. These may include application forms, assessment centre exercises, psychological assessment and interviews (Barrick et al., 2011; Bartram, 2004; Coetzee & Roythorne-Jacobs, 2010; Muller & Schepers, 2003; Schmidt & Hunter, 1998). Hoffman and McPhail (1998) explain that multiple sources of data are reflective of a good selection process and psychological assessment is only one aspect of selection assessment measures. The successive hurdle technique, where only candidates who meet the minimum requirements of a stage of the selection process move on to the next stage, is used by many organisations (Van der Merwe, 2002).

Psychological assessment is distinct from, but incorporates psychological testing. Psychological assessment is a standardised procedure for sampling behaviour, wherein a variety of data is collected through a number of measures, including psychological tests (Bartram, 2004; Branine, 2008; Foxcroft & Roodt, 2001; Gregory, 2007; Shum, O’Gorman & Myors, 2006). As a component of the psychological assessment process, psychological tests used in the past in South Africa were heavily criticised for being biased and unfair (Foxcroft, 1997; Foxcroft & Roodt, 2001). The past decade has however seen a resurgence in the use of psychological tests based on research evidence confirming their validity and objectivity (Branine, 2008; Parkinson, 2009; Paterson & Uys, 2005; Van der Merwe 2002). The value of psychological tests lies in their scientific foundation as most tests possess norms which enable comparison of an individual’s results with a related group of individuals. The tests are also standardised to measure a specific construct (Gregory, 2007; Patterson & Uys, 2005; Spector, 2012). Psychological testing is now therefore generally accepted as increasing the efficiency and value of the selection process, especially when used in combination with other assessment measures (Foxcroft, 1997; Owen & Taljaard, 1996; Parkinson, 2009; Van der Merwe 2002).

Choosing the most appropriate psychological tests for a selection battery is based on job-relevant competencies as well as the psychometric soundness of the test which can confirm it as being unbiased, fair and valid (Potgieter & Van der Merwe, 2002).

### 3.1.2.2 *Fair psychological testing in the multi-cultural South African context*

Early psychological testing in South Africa was administered with little consideration for the impact of culture on non-white candidates. Acknowledging the fact that cultural differences may affect test results and cause systematic error has inspired psychometric research focusing on the issues of bias and fairness in test development and administration. As such, Van de Vijver & Rothmann (2004, p. 6) define multi-cultural assessment as “a new branch of the tree of psychological assessment.” Facilitating an equal opportunity for all remains a main objective of psychological testing in the South African context even today (England, 1991; Theron, 2007).

Fairness alludes to a requirement to responsibly utilise fair and relevant assessment methods and practices, taking into consideration the reason for gathering the data and, most importantly, the needs of every individual involved (Bedell, Van Eeden & Van Staden, 1999; Potgieter & Van der Merwe, 2002; Skarlicki, 2003; Van de Vijver & Rothmann, 2004). While it is essential for organisations to identify valuable candidates, South African legislation demands the fair use of psychometric testing, which by nature is discriminatory, in selection and other human resources functions (Nzama, De Beer & Visser, 2008).

The Health Professions Act, 56 of 1974, discusses professional registration in terms of who can do what, scope of practice and test classification. The use of any psychological assessment which could lead to a candidate becoming anxious or embarrassed, is referred to as a psychological act. These psychological tests are required, by the Act, to be classified. The Act therefore plays its part in ensuring that psychological testing is culturally fair, sensitive and appropriate by outlining these conditions (Health Professions Act, 1974).

The Employment Equity Act, 55 of 1998, regulates the treatment of employees in a fair and equitable manner in order to eliminate unfair discrimination. South African test practitioners could be summoned to provide evidence that particular tests are not discriminatory. The Act outlines that fair discrimination takes place when affirmative action measures are taken or when a particular person is pointed out, favoured or excluded on requirements of a job, consistent with the purpose of the Act (Employment Equity Act, 1998).

Further to the legal requirements, privacy, informed consent and confidentiality are important issues regarding the fairness of psychological assessment, and are governed by ethics (Barnard, 2010; Christenson, 2001; Muchinsky, Kriek & Schreuder, 2005). Assessment professionals are required to act in the best interests of the individual at all times and, as such, the rules of conduct relating to assessment activities in South Africa are outlined in the Professional Board for Psychology's *Rules of conduct pertaining specifically to psychology* (<http://www.hpcsa.co.za>). While this document guides the practitioner on how to avoid abusing power, via unfair and unethical assessment practices, organisations are also responsible for fair and ethical assessment practices through the introduction and use of an assessment policy (Foxcroft & Roodt, 2001).

Apart from the strides made in ensuring fairness in terms of the legal, professional and ethical context in South Africa, fairness is also emphasised by focusing on the psychometric soundness of tests, through the elimination of bias as well as proving reliability and validity.

The presence of factors not measuring the same construct in different cultural groups constitutes bias, which creates trouble in psychological assessment (Jensen, 1980; Poortinga, 1989; Van de Vijver, 2002). Much research has been done relating to issues when culture is a factor in a test (Bedell et al., 1999; Owen, 1991). Huysamen (2002) argued that much more research is necessary regarding bias in South Africa. Consequently, due to the relevance of cross cultural assessment, there are ongoing research studies to understand bias and to take steps to reduce it as much as

possible (Abrahams & Mauer, 1999; Schaap, 2003; Schaap & Basson, 2003; Sackett & Lievens, 2008; Theron, 2007; Van de Vijver, 2002).

The issue of adverse impact, when the methodology for selecting employees gives certain individuals from a certain group much less chance of being selected, is particularly relevant in multi-cultural assessment (Collins & Morris, 2008; Theron 2009). In South Africa, certain groups have received, and currently receive, lower quality education, giving an advantage to those who have received better education. Research studies concerning gender and language specifically, have shown these factors to be common areas for cross-cultural bias (Foxcroft, 2004; Nell, 1997; Theron; 2007; Van de Vijver & Rothmann, 2004). Certain concepts having no meaning in certain languages in South Africa, combinations of languages being spoken and international assessments being introduced without relevant norms can pose further problems for cross-cultural fairness, which could increase bias (Berry, Poortinga, Segal & Dasen, 2002).

In order for tests not to prejudice any cultures, the need to use assessments to deal with these varying education and language levels, and focus on job potential is evident (Magwaza, 1995; Nell, 1997; Sackett & Ellingson 1997; Taylor, 1994). Dynamic assessment, widely researched internationally, measures underlying potential in a flexible fashion (Elliott, 2003; Grigorenko & Sternberg, 1998). In the past two decades, dynamic assessment has also been at the centre of wide research in South Africa (De Beer, 2006; Lopes, Roodt & Mauer, 2001; Taylor, 1994; Van Eeden, De Beer & Coetzee, 2001). Taylor (1994) found that when presenting original ideas, unaffected by culture, to each candidate, evidence was displayed for predicting training and other kinds of success. De Beer (2006) established that non-verbal ability could be used to measure learning potential when a test is administered, followed by learning, after which another test is administered. By measuring learning potential, adverse impact is decreased (Murphy & Maree, 2009).

Competency based assessment has also been able to cut across cultural boundaries (Bartram, 2004; Potgieter & Van der Merwe, 2002; Zhao & Du, 2011) and is consistent with South African labour legislation and ethical practices (Potgieter & Van

der Merwe, 2002). Knowledge, skills, abilities and other characteristics, or combinations of all of these have shown to work together to comprise competencies, which can be used to match an individual to a job (Campion et al., 2011; Heinsman, de Hoogh, Koopman & Van Muijen, 2007; Mirabile, 1997).

Reliability of a test refers to how consistently it measures what it is meant to measure under similar conditions (Barnard, 2010). Internal-consistency reliability and inter-rater reliability are both frequently used in IO psychology (Barnard, 2010; Barrick et al., 2011; Foxcroft & Roodt, 2001; Muchinsky et al., 2005). Reliability may constrain validity, which, in turn, defines the meanings of the test scores. When a test is unreliable, it is also not valid. While reliability is not necessarily a precursor to validity, it is very important to the test being sound (Gregory, 2007).

According to the *List of tests classified as being psychological tests* (Form 207) (<http://www.hpcsa.co.za>), test users still have the responsibility to ensure the validity of psychological tests being used even if they have been classified. Validity relates to what is being measured by the test and how accurately it is being measured (Barnard, 2010; Foxcroft & Roodt, 2001). While three types of validity exist, namely content, construct and criterion-related validity, a large emphasis is placed on criterion-related validity as being central to predictive future job performance. By framing validity as an action concept, theory and practice should be combined into what IO psychologists do. This allows for assessments to be effectively examined in the organisational context and for validity to, in essence, consume IO psychologists' work (Muchinsky, 2004; Schmidt, 2006).

It is not simple to operationalise fairness when utilising psychological assessment in the South African multi-cultural context. However, everything possible needs to be done to ensure that the psychological assessment tools which are used for selection are ethical, fair, reliable and valid. This can be done by ensuring the procedural fairness and psychometric soundness of the psychological tests. Validity is seen as the most important issue in psychological assessment and research has shown that ensuring that assessment measures are valid can lead to the most productive

individuals being selected by the organisation (Lopes et al., 2001; Pelsler, Bergh & Visser, 2005; Schmidt, 2006; Schmidt & Hunter, 1998).

### *3.1.2.3 Validity of psychological testing in selection*

In an effort to ascertain that psychological tests are, in fact, valid for selection purposes, validation studies should be carried out. Even though adverse impact can never be eliminated, it can be reduced by increasing the predictive validity of selection measures (Theron, 2009). Predictive validity, which involves the initial collection of predictor data and, at a later stage, the collection of criterion data, can be utilised on selection assessments to determine their accuracy in predicting future performance (Gregory, 2007).

Many studies have been carried out in the customer services environment, internationally and locally, which have shown that scores from personality assessments, competency based interviews and ability tests are valid predictors of future performance (Bartram, 2004; La Grange & Roodt, 2001; Nicholls, Viviers & Visser, 2009; Nzama et al., 2008; Swanepoel, 1998).

#### *(i) Cognitive ability assessment*

Ability is a person's capacity to learn or carry out a particular task (La Grange & Roodt, 2001; Spector, 2012). Cognitive ability is a good predictor of work and training performance (Anderson, Lievens, Van Dam & Ryan, 2004; Barnard & Schaap, 2005; Bartram, 2005; Hunter & Hunter, 1984; Salgado & Anderson, 2002) and would, therefore, be relevant to tasks in the customer service industry. Ability tests used in selection include verbal ability tests, numerical ability tests, clerical ability tests and deductive reasoning tests (Bartram 2005; Taylor, 1994). Much research has been based on developing unbiased cognitive tests that can be used by people from different cultures (De Beer, 2006; Ployhart, 2006; Taylor, 1994). Cognitive ability tests are thus fair and possess some predictive validity, yet they are not necessarily sufficient on their own to predict future performance (Bartram, 2004; Murphy, Cronin & Tam, 2003). Combining cognitive ability tests with other assessments such as

personality assessments seems to predict performance more effectively (Oultz, 2002).

*(ii) Personality assessment*

A personality trait is defined by Spector (2012) as the tendency to act in a specific way, across various situations. Research supports the use of personality in predicting performance (Bartram, 2004; 2005; Borman, Penner, Allen & Motowildo, 2001; Goldberg, 1993; La Grange & Roodt, 2001; Tett & Christiansen, 2007). The five-factor model, on which many personality assessments used for selection are based, displays significant relationships with performance (Barrick and Mount, 1991; Hurtz & Donovan, 2002; Ones, Dilchert, Viswesvaran & Judge, 2007; Rothmann & Coetzer, 2003). While some researchers feel that validity could be threatened by reactions of different cultures to personality items (Anderson & Witvliet, 2008; Van de Vijver & Rothmann, 2004), Bartram (2004) argues that personality assessment could limit adverse impact. Barnard (2010) argues that since personality is rather stable over time and since individuals tend to be distinctive in their behaviour, personality assessment should add value in predicting performance. However, according to Huang and Ryan (2011), personality at work can be affected by certain situational influences and using personality assessments in combination with other measures may add more predictive value to assessment results.

*(iii) Competency based interviews*

As a result of an individual's personality constructs not necessarily reflecting in their observable behaviour, competency based assessment focuses on assessing competencies required for the job, with competency based interviews being the most common type (Anderson & Witvliet, 2008; Barnard, 2010; Branine, 2008; Hunter & Schmidt, 1998). As interviews become more structured, the more predictive they become of training success and job performance (Chung-Yan, Hausdorf & Cronshaw, 2005; Nzama et al., 2008; Ployhart, 2006; Salgado & Moscoso, 2002; Schmidt & Rader, 1999; Schmidt & Zimmerman, 2004). An interview allows the interviewer to examine social behaviour and communication (Branine, 2008). There

could, however, be dangers of performance anxiety and subjectivity, the latter of which can be limited by means of a panel interview (McCarthy & Goffin, 2004). Once again, it is not advisable to use just the interview as a predictor but to combine it with other measures.

Researchers have shown confidence in the validity of personnel selection methods which is particularly significant (Robertson & Smith, 2001). In terms of validation studies, studies by Damitz, Manzey, Kleinmann and Severin (2003) and Murdy, Sells, Gavub and Toole (1973) are just two examples of such which have taken place in the airline industry. In South Africa, there have been research studies on the predictive validity of assessment batteries used for pilot selection (see Flotman, 2002; Mnguni, 2011).

Validation studies explore the relationships between the selection process and the after-effects of that process (Van Iddekinge & Ployhart, 2008). Conducting predictive validation studies for selection purposes - and specifically for cognitive, personality and competency based tests - can be instrumental in minimising adverse impact and unfair discrimination. Validation studies are therefore of utmost importance in ensuring that the most suitable candidate can be selected for a job role, in a fair and unbiased manner.

### **3.1.3 Research objectives**

The purpose of this study was to explore the effectiveness of an organisation's selection process for CSAs by determining the relationship between the scores of assessments used in the selection process – namely the Verbal Interpretation Test (VCC1), the Numerical Reasoning Test (NP6.1), the Basic Checking Test (CP7.1), CCCs, PJM and competency based interview – with job performance, and to ascertain whether the use of these tools in the selection process helped in predicting future job performance.

More specifically, the study aimed to achieve the following specific research objectives:

- Explore the empirical relationship between ability testing, personality assessment and the competency based interview respectively with job performance.
- Determine whether the scores of the VCC1, NP6.1, CP7.1C, CCCs, PJM and competency based interview predict job performance of CSAs in the South African airline industry.
- Put forward recommendations to the airline company regarding future selection decisions.

### **3.1.4 Potential value add of the study**

This study is of benefit to Human Resource Practitioners as well as Industrial and Organisational (IO) Psychologists. It should also add value to professionals and managers in fast-paced customer service organisations, giving them a better understanding of the selection tools that are most effective at predicting future job performance in this industry. In order to ensure best practice, validation studies should be carried out consistently across all assessment practices. Therefore, validation studies are always necessary in the IO Psychology field.

The following section will explore the research design as well as the research approach and method. The results will then be presented, followed by a discussion on the findings. The article will conclude with a brief discussion of the study's main conclusions, limitations and recommendations for future research.

## **3.2 RESEARCH DESIGN**

### **3.2.1 Research approach**

This was a quantitative study, where scores were analysed statistically by means of a correlational approach, following a cross-sectional survey design. The scores were derived from psychological assessments and a competency based interview, as well as performance data.

### **3.2.2 Research method**

#### *3.2.2.1 Research participants*

A new psychological assessment battery - including ability tests, a personality questionnaire and a competency based interview - was introduced into a South African commercial airline company in June 2008. In order to maximise the sample sizes for each correlation matrix, the sample used in the study was made up of 1 223 individuals. In the correlational analysis of this research study, the population consisted of all the individuals who were appointed as CSAs in the airline company over a three-year period following the introduction of the new psychological assessment battery in 2008.

While at least some data were collected for all candidates who took part in any single part of the selection process, not all the employees completed all the ability tests or had all the criterion data, and a maximum sample was used for each step. As a result, the sample size differs for each variable used as it decreased after each step in the selection process as additional job requirements were in place. All predictor and criterion data were available for 192 candidates, who were all appointed as CSAs and who had completed their ability tests, personality questionnaire and competency based interview, as well as having performance data (training scores, performance appraisal ratings and criterion questionnaire ratings) available for them. The demographic breakdown of the sample, relating to distribution of the ethnic origin groups and genders, is shown in Table 2 (N=1223).

**TABLE 2. DEMOGRAPHIC BREAKDOWN OF SAMPLE**

<b>Population groups</b>	<b>N</b>	<b>%</b>	<b>Cumulative %</b>
<b>Gender</b>			
Female	881	72.0	72.4
Male	337	27.6	100.0
<b>Total</b>	<b>1223</b>	<b>100.0</b>	
<b>Ethnic Origin</b>			
African	651	53.2	63.5
Coloured	185	15.1	78.6
Indian	158	12.9	91.5
White	104	8.5	100.0
<b>Total</b>	<b>1223</b>	<b>100.0</b>	

### 3.2.2.2 *Measuring instruments*

The independent variables in the study were defined as verbal, numerical and checking ability, personality and competency based interview scores, the measures of which are discussed below. The dependent variable, defined as job performance, was measured by means of training scores and performance data (performance appraisal ratings and criterion questionnaire ratings).

#### *Verbal Interpretation (VCC1)*

The VCC1 measures an individual's ability to comprehend straightforward written information in order to arrive at rational conclusions. This is a task that is relevant to customer service jobs where job incumbents receive product information in a written format, as well as written communication from fellow employees and customers. The test consists of 36 items, presented in a multiple-choice setup and the time limit for the test is 12 minutes (SHL, 2009).

#### *Numerical Reasoning (NP6.1)*

The NP6.1 is a test which measures basic reasoning skills with regards to numbers. Items could involve decimals, graphs and/or fractions, and calculators are not

allowed. There are 30 items in the NP6.1 and the test is 15 minutes long (SHL, 2009).

#### *Basic Checking (CP7.1C)*

The CP7.1C is concerned with measuring both accuracy and speed of checking at a basic level. It is generally used with employees in clerical positions and whose roles have a routine checking component in them. There are 80 questions and candidates have 10 minutes to complete the test (SHL, 2009).

#### *Occupational Personality Questionnaire (OPQ32)*

The OPQ32 is a personality questionnaire for the use in selection and development of people at work. It provides valuable information on 32 dimensions or scales of people's preferred or typical style of behaviour at work. It is particularly appropriate for use with professional and managerial groups, although the content of the OPQ32 model deals with personality characteristics important to a wide variety of roles. The ipsative version of the OPQ32 (OPQ32i) was used in this study. While the OPQ32 was administered, it was not used as a predictor in the study, but was utilised to derive competency scores for CCCs, as well as the PJM score. These scores were then applied as predictor scores. The CCCs which are derived from Saville and Holdsworth Limited's (SHL) Customer Contact Competency Inventory (CCCI) are defined in Table 3 below (SHL, 2000). The rating scale parameters for CCCs are described in Table 4.

**TABLE 3. CUSTOMER CONTACT COMPETENCIES: DEFINITIONS**

<b>COMPETENCY</b>	<b>DEFINITION</b>
<b>People Focus</b>	
Relating to Customers (P1)	Quickly builds rapport and easily establishes relationships with customers. Relates well to different types of customer; listens and gets on with them.
Convincing (P2)	Presents the key points of an argument persuasively. Negotiates and convinces others. Changes people's views and influences their decisions.
Communicating Orally (P3)	Speaks confidently and fluently. Talks at a suitable pace and level. Holds others' attention when speaking.
Communicating in Writing (P4)	Writes fluently, clearly and concisely. Adapts own written communication style to suit others.
Team Working (P5)	Fits in with the team. Develops effective and supportive relationships with colleagues. Is considerate towards them and creates a sense of team spirit.
<b>Information Handling</b>	
Fact Finding (I1)	Knows where to find relevant information. Checks facts and data. Retrieves and absorbs information quickly.
Problem solving (I2)	Identifies potential difficulties and their causes. Generates workable solutions and makes rational judgements.
Business Awareness (I3)	Is aware of competitor activity and market trends. Is profit conscious and appreciates the commercial impact of own work on profits.
Specialist Knowledge (I4)	Has background knowledge and a thorough grasp of products and services. Has expertise on own area.
<b>Dependability</b>	
Quality Orientation (D1)	Provides a quality service. Maintains high professional standards and gets work right first time.
Organisation (D2)	Organises own time effectively and creates own work schedules. Prioritises and prepares in advance. Sets realistic time-scales.
Reliability (D3)	Is reliable; follows directions from supervisors and respects policies and procedures. Shows commitment to the organisation and task completion.
<b>Energy</b>	
Customer Focus (E1)	Puts the customer first and is eager to please them. Works hard to meet customer needs and looks after their interests.
Resilient (E2)	Remains calm and self-controlled under pressure. Reacts well to change and stays positive despite setbacks. Keeps difficulties in perspective.
Results Driven (E3)	Gets results and willingly tackles demanding tasks. Sets and exceeds challenging personal targets.
Using Initiative (E4)	Takes responsibility for own actions and makes decisions without referring to others. Acts on own initiative.

**TABLE 4. CCC SCALE PARAMETERS AND DESCRIPTIONS**

<b>Rating scale</b>	<b>Description</b>
<b>1 Unsatisfactory performance</b>	The employee's performance of the activity is unacceptable, poor and must improve drastically.
<b>2 Below average performance</b>	The employee's performance of the activity is below standard, must still improve and does not always meet expectations.
<b>3 Acceptable performance</b>	The employee's performance of the activity is of acceptable standard and meets expectations.
<b>4 Exceeding expectations performance</b>	The employee's performance of the activity is above standard, of a high standard and exceeds expectations.

*Person Job Match (PJM) score*

If a person has the abilities necessary to perform the inherent requirements of a job, then a good Person Job Fit exists. The PJM score provides an indication of the candidate's 'degree of fit' to a role, with the higher the score the better the person-job fit. The PJM incorporates the ability tests and the OPQ32, with the key behaviours that influence work performance being provided by the CCC scores (SHL, 2013). The overall PJM score is divided into five match bands that are based on a grading principle, and are illustrated in Table 5 (SHL, 2013).

**TABLE 5. PJM MATCH BAND CATEGORIES**

<b>PJM score range</b>	<b>Interpretation</b>
<b>71+ : Extremely strong match</b>	The candidate has a very strong match to the requirements of the job.
<b>&lt;71 : Strong match</b>	The candidate has a strong match to the requirements of the job, but may require development in some areas.
<b>&lt;59 : Moderate match</b>	The candidate has a reasonable match to the requirements of the job, but will require development of competencies in some areas.
<b>&lt;42 : Weak match</b>	The candidate has a poor match to the requirements of the job. S/he may be able to develop the necessary competencies, but considerable effort is likely to be required.
<b>&lt;30 : Extremely weak match</b>	The candidate does not match to the requirements of the job. S/he is unlikely to be able to develop the necessary competencies without considerable time and effort being required.

### *Competency based interview*

Structured competency based interviews were designed for the CSA position to measure the knowledge, skills and experience of the job applicants. All the applicants were asked the same questions. These questions referred to situations that candidates had faced previously and how they handled them. The five competencies measured were Customer Focus, Quality Orientation, Resilience, Impact and Communication. Two questions per competency were asked. The sample sizes for Impact and Communication were too small to include in the study and were therefore left out. An average score for all the competencies measured was also computed and used in the correlations. Following a discussion between raters, candidates were given a score, based on a five-point scale, for each competency assessed. The scale parameters and meanings of the five-point scale are described in Table 6.

**TABLE 6. COMPETENCY BASED INTERVIEW MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS**

<b>Rating scale</b>	<b>Description</b>
<b>1 Unsatisfactory performance</b>	The employee's performance of the activity is unacceptable, poor and must improve drastically.
<b>2 Below average performance</b>	The employee's performance of the activity is below standard, must still improve and does not always meet expectations.
<b>3 Acceptable performance</b>	The employee's performance of the activity is of acceptable standard and meets expectations.
<b>4 Exceeding expectations performance</b>	The employee's performance of the activity is above standard, of a high standard and exceeds expectations.
<b>5 Well above expectations</b>	The employee's performance is exceptional, going well beyond expectations.

### *Performance ratings*

Three sources of performance ratings were utilised in the research study, namely training scores, performance appraisal ratings and criterion questionnaire ratings for the questionnaire which was created for the study.

Up until December 2008, only an Overall Training score (referred to as *Overall Training (until December 2008)*) was given to employees. From January 2009, Self Study Training (referred to as *Training – SS (from January 2009)*) and Passenger Handling Training (referred to as *Training – PH (from January 2009)*) courses were introduced, with scores available for each of them. These, along with Overall Training scores (referred to as *Training Overall (from January 2009)*) were incorporated as performance data. These courses took place after candidates had been selected and had to be passed before the candidate could assume the role of a CSA.

Secondly, existing performance data were made available by the company from performance appraisal ratings. This enables formal appraisal scores to be included in the analysis. The company evaluates the performance of employees twice annually, with an appraisal halfway through the financial year (Half Year) and a full-year appraisal at the conclusion of the financial year (Full Year). Included in the analysis is the most recent score, which was from the Half Year 2011 performance appraisals. All performance objectives are rated on a four-point scale, described in Table 7, after which a final overall average rating is computed.

**TABLE 7. PERFORMANCE APPRAISAL MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS**

<b>Rating scale</b>	<b>Description</b>
<b>1 Unsatisfactory performance</b>	The employee's performance of the activity is unacceptable, poor and must improve drastically.
<b>2 Below average performance</b>	The employee's performance of the activity is below standard, must still improve and does not always meet expectations.
<b>3 Acceptable performance</b>	The employee's performance of the activity is of acceptable standard and meets expectations.
<b>4 Exceeding expectations performance</b>	The employee's performance of the activity is above standard, of a high standard and exceeds expectations.

Finally, SHL designed and developed a criterion questionnaire for the CSA position in the commercial airline industry. The front page of the questionnaire contained a short section for the rater- and the candidate's names, followed by instructions explaining

the purpose of the study and how the supervisor should rate the employee. The criterion questionnaire then posed questions/items on which employees should be rated in four sections, as follows:

- Key Performance Areas (KPA) -10 items;
- Organisational Culture -4 items;
- Time-Keeping and Disciplinary Measures (consisting of Absence Without Leave (AWOL) rating, Sick Leave rating, Time Keeping rating, Disciplinary rating, ICAS referral rating, Serious Offences rating, Problem Employee rating, Open to Feedback rating);
- Key Behaviours (ratings of items loading onto competencies identified as important for a CSA to be successful in the role) -39 items.

Apart from the Time-Keeping and Disciplinary Measures, the other items were based on information contained in the job profile, which was designed using the Work Profiling System (WPS) from SHL. The rating scale parameters for the KPA, Organisational Culture and Key Behaviours items are described in Table 8. The rating scale parameters and coding per item for the items relating to Time-Keeping and Disciplinary Measures are defined in Table 9. Example items for each of the Key Behaviours (competencies) measured in the criterion questionnaire are presented in Table 10.

**TABLE 8. CRITERION QUESTIONNAIRE MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS (KPA, ORGANISATION CULTURE AND KEY BEHAVIOURS ITEMS)**

Rating scale	Description
<b>1 Unsatisfactory performance</b>	The employee's performance of the activity is unacceptable, poor and must improve drastically.
<b>2 Below average performance</b>	The employee's performance of the activity is below standard, must still improve and does not always meet expectations.
<b>3 Acceptable performance</b>	The employee's performance of the activity is of acceptable standard and meets expectations.
<b>4 Exceeding expectations performance</b>	The employee's performance of the activity is above standard, of a high standard and exceeds expectations.

**TABLE 9. CRITERION QUESTIONNAIRE MEASUREMENT SCALE PARAMETERS AND DESCRIPTIONS (TIME-KEEPING AND DISCIPLINARY MEASURES ITEMS)**

<b>Description</b>	<b>Question asked to rater</b>	<b>Measurement Scales and Coding</b>
<b>AWOL rating</b>	Has this person been absent without leave?	Never (Coded to 1) – <b>good</b> Occasionally (once) (Coded to 2) – <b>average</b> Frequently (more than once) (Coded to 3) – <b>poor</b>
<b>Sick Leave rating</b>	Has there ever appeared to be an issue with this person with regards to sick leave?	Never (Coded to 1) – <b>good</b> Occasionally (Coded to 2) – <b>average</b> Frequently (Coded to 3) – <b>poor</b>
<b>Time Keeping rating</b>	Does this person have time-keeping issues (e.g. coming late / leaving early)?	Never (Coded to 1) – <b>good</b> Occasionally (1-2 times) (Coded to 2) – <b>average</b> Frequently (more than twice) (Coded to 3) – <b>poor</b>
<b>Disciplinary rating</b>	Has this person had any disciplinary issues in the past six months? <i>(Can include counselling, verbal warning, written warning, final written warning or suspension without pay)</i>	Yes (Coded to 2) No (Coded to 1)
<b>ICAS Referral rating</b>	Has this person been referred to ICAS for any reason?	Yes (Coded to 0) No (Coded to 1)
<b>Serious Offences rating</b>	Are there any serious offences which this person has committed since the beginning of their employment (e.g. fraud, dishonesty, any form of violence, etc.)?	Yes (Coded to 2) No (Coded to 1)
<b>Problem Employee rating</b>	Despite this person's formal disciplinary record, do you believe that he/she is a problem employee?	Yes (Coded to 3) Maybe (Coded to 2) No (Coded to 1)
<b>Open to Feedback rating</b>	Is this person open to feedback?	Yes (Coded to 3) Maybe (Coded to 2) No (Coded to 1)

**TABLE 10. EXAMPLE ITEMS FOR KEY BEHAVIOURS MEASURED IN CRITERION QUESTIONNAIRE**

<b>Key Behaviour</b>	<b>Example Item</b>
<b>Customer Focus</b>	Works hard to meet customer needs Is driven by customer demands
<b>Quality Orientation</b>	Produces high quality work Is highly accurate in their work
<b>Relating to Customers</b>	Listens to customers Easily establishes relationships with customers
<b>Reliability</b>	Follows directions from superiors Respects company policies and procedures
<b>Resilient</b>	Remains positive despite setbacks (e.g. delays, dealing with difficult and angry passengers) Stays calm under pressure

### *3.2.2.3 Research procedure*

A job analysis was carried out prior to assessments for CSAs commencing in June 2008, using the WPS from SHL. Through this process, a competency profile report was generated for the position of CSA, which consisted of the purpose of the job, job objectives, job specific requirements, competency requirements and suggested assessment methods, which included the VCC1 for verbal ability, the NP6.1 for numerical ability, the CP7.1C for checking ability and the OPQ32 to assess personality. Seven Customer Contact Competencies (Customer Focus, Relating to Customers, Resilient, Quality Orientation, Reliability and Using Initiative) were identified as essential for the job and four (Communicating Orally, Team Working, Fact Finding and Specialist Knowledge) were identified as having high importance.

The decision regarding the selection of CSAs involved a hurdle-approach, where only individuals who met the minimum requirements in the ability-testing phase, advanced to the personality assessment and interview phases.

The assessments were administered by the company's psychometrists, intern psychometrists and intern industrial psychologists, according to the best practice

guidelines suggested by SHL (SHL, 2008a). The practitioners are registered with the HPCSA (Form 94, Form 208) (<http://www.hpcsa.co.za>). While there was no specified order of administration, generally the verbal ability test (VCC1) was carried out first, followed by the numerical ability test (NP6.1) and then the checking test (CP7.1C). While the whole sample completed numerical tests, they were not all assessed on the NP6.1, which is the reason why the size of the sample for the NP6.1 is smaller than those of the other ability tests. The candidates were assessed under standard test conditions, generally in one of the airline's test administration or training centres. Candidates were required to sign an informed consent form and were read instructions for each test, which were followed by several practice questions, during which time the candidates could ask any outstanding questions to clarify any uncertainties. The same process and testing conditions were used for every individual or group being assessed.

If the minimum requirements in the ability test phase were met, candidates were invited for a competency based interview. If the candidates successfully met the minimum requirements of the interview, they would be invited to complete the OPQ32 personality questionnaire. As a result of correlation coefficients being dependent on the variability of scores in a sample, correlation coefficients in this study may be lower than if candidates were selected without regard to the test scores. This outcome is known as restriction of range.

CCC scores and a PJM score were then derived from a combination of the ability test data and personality questionnaire data. The PJM score, a single indicator based on the important CCCs identified in the job analysis, demonstrates the individual's potential for success in the role. If the candidate's PJM score indicated at least a moderate match to the requirements of the job, the candidate was appointed as a CSA and was invited to the training programme.

The data for the research were obtained from various sources involved in the selection process. The ability test scores, personality questionnaire scores and PJM scores were gathered from the available assessment reports by the selection consultants who manage the assessment process. The selection interviews were

carried out by a panel made up of airline employees, and included an HR officer and a Customer Service Supervisor/Manager from the airport operations. Following the interviews, the panel would discuss the scores they had initially given for the measured competencies and together decide on final scores, per competency.

In terms of the dependent variables, training course results were obtained from the Commercial Training department. The HR Division and specifically, the Organisational Development (OD) department provided the performance appraisal ratings. The criterion questionnaires were distributed to the Customer Service Supervisors and Managers who had supervised the candidates that had been appointed and, subsequently, employed. Prior to receiving the questionnaires, the supervisors were given instructions, on how to complete the questionnaires and information on why the study was being carried out. The items were rated, after which the questionnaires were returned to the OD department.

#### *3.2.2.4 Statistical analysis*

A descriptive and exploratory design was used. Pearson's Product Moment Correlations were used in the study, where correlation coefficients were calculated in order to provide an analysis of the relationship between the different predictors and the criterion scores (Clark-Carter, 2004). IBM's SPSS Statistics 20 was used to analyse the full data set.

Magnitudes of the effect sizes of the correlations were interpreted by making use of the guidelines recommended by Cohen (1988), where an effect size between .0 and .2 is described as a small correlation, between .2 and .5 is described as a moderate correlation, and over .5 can be described as a large correlation. In Table 11, the descriptive statistics for the study are explained.

**TABLE 11. DESCRIPTIVE STATISTICS FOR PREDICTORS AND CRITERIA**

<b>VARIABLES</b>	<b>N</b>	<b>M</b>	<b>SD</b>	<b>MINIMUM</b>	<b>MAXIMUM</b>
<b>Ability Scores</b>					
VCC1	842	14.03	4.948	1	33
NP6.1	600	9.25	3.303	1	20
CP7.1C	840	59.43	8.381	26	80
<b>CCC Scores</b>					
Relating to Customers (P1)	338	6.90	1.961	1	10
Convincing (P2)	338	6.77	1.938	2	10
Communicating Orally (P3)	338	7.29	1.981	1	10
Communicating in Writing (P4)	338	6.06	1.989	2	10
Team Working (P5)	338	7.36	1.971	2	10
Fact Finding (I1)	338	5.27	1.684	1	10
Problem Solving (I2)	338	5.07	1.966	1	10
Business Awareness (I3)	338	5.49	2.065	1	10
Specialist Knowledge (I4)	338	4.94	1.882	1	10
Quality Orientation (D1)	338	7.03	1.499	3	10
Organisation (D2)	338	6.68	1.811	1	10
Reliability (D3)	338	7.62	1.659	2	10
Customer Focus (E1)	338	7.42	1.967	2	10
Resilient (E2)	338	7.18	2.153	1	10
Results Driven (E3)	338	7.06	2.047	1	10
Using Initiative (E4)	338	5.54	1.677	2	10
<b>PJM Score</b>					
PJM Score	338	54.50	14.413	17	95
<b>Interview Ratings</b>					
Customer Focus	546	3.16	.627	1	5
Quality Orientation	429	3.07	.710	1	5
Resilience	426	3.11	.687	1	5
Interview Average Score	550	3.087	.5909	1.0	5.0
<b>Training Scores</b>					
Overall Training (until December 2008)	133	90.57	7.064	33	99
Training - SS (from January 2009)	217	88.90	12.078	43	100
Training - PH (from January 2009)	196	87.78	6.992	55	98
Training Overall (from January 2009)	192	91.19	4.724	81	99
<b>Performance Appraisal Ratings</b>					
Half Year 2011	189	2.69	.506	2	4

**TABLE 11 CONTINUED. DESCRIPTIVE STATISTICS FOR PREDICTORS AND CRITERIA**

<b>Criterion Questionnaire Ratings</b>					
Key Performance Areas	270	2.87	.503	1	4
Organisational Culture	270	2.82	.581	1	4
AWOL Rating	270	1.33	.596	1	3
Sick Leave Rating	270	1.46	.708	1	3
Time Keeping Rating	270	1.54	.677	1	3
Disciplinary Rating	268	1.49	.501	1	2
ICAS Referral	96	.09	.293	0	1
Serious Offences Rating	268	1.94	.237	1	2
Problem Employee Rating	267	1.36	.765	1	3
Open to Feedback Rating	259	2.63	.758	1	3
Competency_Customer Focus	270	2.92	.517	1	4
Competency_Quality Orientation	270	2.89	.579	1	4
Competency_Relating to Customers	270	2.93	.472	1	4
Competency_Reliability	270	2.88	.588	1	4
Competency_Resilient	270	2.81	.514	1	4

### **3.3 RESULTS**

Correlations provide insight into the relationship between different variables. Based on the objectives of the research and in order to test the research hypotheses, a number of correlations were calculated. Correlations were calculated between predictor variables and criterion measures and these correlations are reported in the following sub-sections. P-values of  $\leq 0.01$  and  $\leq 0.05$  were considered as statistically significant.

#### **3.3.1 Correlations between training scores and predictors**

Correlations between training scores and predictors (VCC1, NP6.1, CP7.1C, CCC scores, the PJM score and interview competency scores) are represented in Table 12. A key can be found for the competency codes in Table 10 (above).

**TABLE 12. CORRELATIONS BETWEEN TRAINING SCORES AND PREDICTORS**

		Abilities						CCCs					
		CP 7.1C	NP 6.1	VCC 1	P1	P2	P3	P4	P5	I1	I2	I3	I4
<b>Overall Training (Until December 2008)</b>	Corr.	<b>.259**</b>	.292	<b>.201*</b>	.089	-.075	-.165	.281	-.061	.322	.307	.056	.299
	N	121	36	121	36	36	36	36	36	36	36	36	36
<b>Training - SS (From January 2009)</b>	Corr.	.089	.057	.101	.086	.015	-.005	-.137	.011	-.040	-.022	-.088	-.019
	N	130	130	132	107	107	107	107	107	107	107	107	107
<b>Training - PH (From January 2009)</b>	Corr.	<b>.198*</b>	.179	<b>.265**</b>	<b>.206*</b>	-.037	-.034	.087	<b>.256*</b>	.047	.074	<b>-.254*</b>	.055
	N	119	119	121	100	100	100	100	100	100	100	100	100
<b>Training Overall (From January 2009)</b>	Corr.	<b>.259**</b>	<b>.210*</b>	<b>.214*</b>	<b>.207*</b>	-.067	-.061	.043	<b>.226*</b>	.039	.097	-.170	.085
	N	116	116	118	98	98	98	98	98	98	98	98	98

**TABLE 12 CONTINUED. CORRELATIONS BETWEEN TRAINING SCORES AND PREDICTORS**

		CCCs (continued)							PJM	Interview			
		D1	D2	D3	E1	E2	E3	E4	Score	E1	D1	E2	Average
<b>Overall Training (Until December 2008)</b>	Corr.	.243	.055	-.041	.166	-.156	-.099	-.129	.222	.117	.132	.016	.158
	N	36	36	36	36	36	36	36	36	88	38	40	88
<b>Training - SS (From January 2009)</b>	Corr.	.024	-.043	.079	-.062	.121	-.017	.116	.053	.070	-.025	.048	.023
	N	107	107	107	107	107	107	107	107	210	210	210	210
<b>Training - PH (From January 2009)</b>	Corr.	<b>.279**</b>	.017	.170	<b>.222*</b>	-.091	.185	.048	<b>.272**</b>	.124	.128	.143	<b>.159*</b>
	N	100	100	100	100	100	100	100	100	189	189	189	189
<b>Training Overall (From January 2009)</b>	Corr.	<b>.199*</b>	.064	.129	.131	-.046	.106	-.013	<b>.226*</b>	.033	.053	.137	.087
	N	98	98	98	98	98	98	98	98	185	185	185	185

\*\* . Correlation is significant at the 0.01 level (2-tailed); \* . Correlation is significant at the 0.05 level (2-tailed).

In terms of correlations between the ability test scores and the training scores, moderate positive correlations were observed between the VCC1 and scores for Overall Training (until December 2008), Passenger Handling Training (from January 2009) and Overall Training (from January 2009). No statistical significance was observed between the NP6.1 and the specific training course scores but a moderate positive correlation was observed with Overall Training (from January 2009). This could be due to the size of the sample. With regard to the CP7.1C, a small positive correlation was observed with Passenger Handling Training (from January 2009), as well as moderate positive correlations with both Overall Training (until December 2008) and Overall Training (from January 2009).

Regarding the relationship between CCC scores and training scores, moderate positive relationships were displayed between Passenger Handling Training (from January 2009) and the CCCs of Relating to Customers, Team Working, Quality Orientation and Customer Focus, as well as a moderate negative relationship with Business Awareness. Significant small to moderate positive correlations were also observed between Overall Training (from January 2009) and the CCCs of Relating to Customers, Team Working and Quality Orientation. A moderate positive correlation was found between the overall PJM Score and both the Passenger Handling Training and Overall Training (from January 2009) scores. However, no statistical significance was found between Overall Training (until December 2008) and Self Study Training (from January 2009) and the CCCs. Although only five of the CCCs and the PJM score obtained small to moderate correlations with the training results, all 5 were identified as extremely important or of high importance to the role by the job analysis.

The competency based interview data found a moderate positive correlation between the average interview score and the Passenger Handling Training (from January 2009) scores (.159\*, N=189). No other statistical significance was found between interview data and training results.

### **3.3.2 Correlations between performance appraisal ratings and predictors**

Correlations between Half Year 2011 performance appraisal ratings and predictors, (VCC1, NP6.1, CP7.1C, CCC scores, the PJM score and interview competency scores) are presented in Table 13.

**TABLE 13. CORRELATIONS BETWEEN PERFORMANCE APPRAISAL RATINGS AND PREDICTORS**

		Abilities			CCCs								
		CP 7.1C	NP 6.1	VCC 1	P1	P2	P3	P4	P5	I1	I2	I3	I4
<b>Half Year 2011 rating</b>	Corr.	<b>.165*</b>	.075	.036	-.030	.061	.051	.207	.072	<b>.262*</b>	.182	-.140	.185
	N	146	97	147	83	83	83	83	83	83	83	83	83

**TABLE 13 CONTINUED. CORRELATIONS BETWEEN PERFORMANCE APPRAISAL RATINGS AND PREDICTORS**

		CCCs				PJM				Interview			
		D1	D2	D3	E1	E2	E3	E4	PJM Score	E1	D1	E2	Average Score
<b>Half Year 2011 rating</b>	Corr.	<b>.351**</b>	.066	.153	<b>.281*</b>	-.147	.137	<b>.250*</b>	<b>.279*</b>	.117	-.010	<b>.308**</b>	<b>.175*</b>
	N	83	83	83	83	83	83	83	83	152	123	123	152

\*\* . Correlation is significant at the 0.01 level (2-tailed); \* . Correlation is significant at the 0.05 level (2-tailed).

Regarding correlations between Half Year 2011 performance appraisal ratings and predictors, the CP7.1C showed a small to moderate correlation with the Half Year 2011 performance appraisal ratings. However, no statistically significant relationships were observed between the performance appraisal ratings with the NP6.1 or VCC1.

Small to moderate positive correlations were observed between the performance appraisal ratings and both the scores for the interview competency of Resilience and the Interview Average Score.

In terms of the relationship between performance appraisal ratings and CCCs, moderate positive correlations were observed with the competencies Fact Finding, Quality Orientation, Customer Focus and Using Initiative, all extremely or highly important competencies for the job. In conclusion, when correlated with the performance appraisal ratings, the overall PJM score also showed a moderate positive correlation.

### **3.3.3 Correlations between criterion questionnaire ratings and predictors**

Correlations between criterion questionnaire ratings and predictors (VCC1, NP6.1, CP7.1C, CCC scores, the PJM score and interview competency scores) are presented in Table 14.

**TABLE 14. CORRELATIONS BETWEEN CRITERION QUESTIONNAIRE RATINGS AND PREDICTORS**

		Abilities				CCCs							
		CP 7.1C	NP 6.1	VCC 1	P1	P2	P3	P4	P5	I1	I2	I3	I4
<b>Key Performance Areas</b>	Corr.	.104	.094	.019	-.004	.039	.027	-.135	.044	-.105	-.121	-.022	-.070
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Organisational Culture</b>	Corr.	-.022	.088	-.050	.029	.053	.036	-.093	.049	-.089	-.105	-.014	-.081
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>AWOL</b>	Corr.	-.106	-.073	-.007	-.102	.047	.097	-.026	-.123	-.065	.011	.104	-.042
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Sick Leave</b>	Corr.	-.048	-.090	-.058	-.111	.124	.145	-.023	<b>-.251</b>	-.008	.047	<b>.225</b>	.021
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Time Keeping</b>	Corr.	-.094	-.044	.038	-.031	.048	.050	-.082	-.090	-.061	.023	.130	.014
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Disciplinary Rating</b>	Corr.	<b>-.149</b>	-.064	.039	-.088	.056	.099	.047	-.152	.057	.067	.141	.061
	N	201	134	203	118	118	118	118	118	118	118	118	118
<b>ICAS Referral</b>	Corr.	-.081	-.111	.013	-.039	-.225	-.090	-.005	-.003	-.128	-.079	-.059	-.150
	N	92	59	93	47	47	47	47	47	47	47	47	47
<b>Serious Offences</b>	Corr.	.112	.033	.060	.008	-.060	-.060	-.054	.090	-.040	-.074	.013	-.051
	N	202	135	204	119	119	119	119	119	119	119	119	119
<b>Problem Employee</b>	Corr.	<b>-.153</b>	.083	.059	-.080	.013	.092	.141	-.075	.054	.087	.036	.099
	N	200	133	202	117	117	117	117	117	117	117	117	117
<b>Open to Feedback</b>	Corr.	.105	-.054	.109	.017	.151	.112	.047	.017	.008	-.007	-.012	-.031
	N	192	128	194	113	113	113	113	113	113	113	113	113
<b>Competency_ Customer Focus</b>	Corr.	.083	.113	.039	.026	.019	.022	-.140	.052	-.126	-.113	-.018	-.109
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Competency_ Quality Orientation</b>	Corr.	.111	.048	.004	-.031	.048	.025	-.098	.013	-.049	-.094	-.014	-.036
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Competency_ Relating to Customers</b>	Corr.	.087	<b>.184</b>	.047	.036	.064	.051	-.138	.065	-.104	-.096	.009	-.075
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Competency_ Reliability</b>	Corr.	.091	.006	-.053	.040	.011	-.007	-.066	.092	-.071	-.089	-.056	-.074
	N	203	136	205	119	119	119	119	119	119	119	119	119
<b>Competency_ Resilient</b>	Corr.	<b>.139</b>	<b>.186</b>	.002	.001	.105	.100	-.159	.015	-.093	-.087	.050	-.049
	N	203	136	205	119	119	119	119	119	119	119	119	119

**TABLE 14 CONTINUED. CORRELATIONS BETWEEN CRITERION QUESTIONNAIRE RATINGS AND PREDICTORS**

		CCCs (continued)				PJM				Interview			Average Score
		D1	D2	D3	E1	E2	E3	E4	PJM Score	E1	D1	E2	
Key Performance Areas	Corr.	<b>.234*</b>	-.005	<b>.212*</b>	<b>.202*</b>	.013	.080	.109	.162	.055	.143	.053	.125
	N	119	119	119	119	119	119	119	119	224	185	186	225
Organisational Culture	Corr.	.080	-.029	.107	.084	-.061	.026	.058	.041	.083	.110	.017	.123
	N	119	119	119	119	119	119	119	119	224	185	186	225
AWOL	Corr.	-.175	.083	-.057	<b>-.227*</b>	.150	.030	.080	-.106	-.028	<b>-.159*</b>	-.007	-.101
	N	119	119	119	119	119	119	119	119	224	185	186	225
Sick Leave	Corr.	-.124	.107	.016	<b>-.237*</b>	<b>.183*</b>	.105	-.020	-.042	-.120	<b>-.157*</b>	-.041	-.137
	N	119	119	119	119	119	119	119	119	224	185	186	225
Time Keeping	Corr.	<b>-.213*</b>	.010	-.112	<b>-.246*</b>	.165	.030	-.046	-.112	.012	-.078	-.007	-.031
	N	119	119	119	119	119	119	119	119	224	185	186	225
Disciplinary Rating	Corr.	-.062	.098	.137	-.010	.092	.148	-.006	.071	-.091	<b>-.229*</b>	-.066	<b>-.142*</b>
	N	118	118	118	118	118	118	118	118	222	183	185	223
ICAS Referral	Corr.	-.027	.106	.008	-.205	.043	.266	.019	-.143	-.020	.061	.005	-.098
	N	47	47	47	47	47	47	47	47	86	67	63	86
Serious Offences	Corr.	.111	-.099	.054	.139	-.073	.134	-.019	.038	.060	.118	.009	.104
	N	119	119	119	119	119	119	119	119	222	183	184	223
Problem Employee	Corr.	-.083	.005	-.074	-.146	.130	-.154	-.101	-.022	.025	-.055	.002	-.016
	N	117	117	117	117	117	117	117	117	221	182	183	222
Open to Feedback	Corr.	.139	.065	.083	<b>.189*</b>	-.040	.139	.173	.163	.039	-.020	-.097	.016
	N	113	113	113	113	113	113	113	113	214	176	178	215
Competency_ Customer Focus	Corr.	<b>.207*</b>	-.075	.139	.140	-.075	.073	.081	.078	.062	.107	.056	.118
	N	119	119	119	119	119	119	119	119	224	185	186	225
Competency_ Quality Orientation	Corr.	<b>.206*</b>	.001	<b>.184*</b>	.119	-.108	.078	.074	.077	.021	.125	.037	.092
	N	119	119	119	119	119	119	119	119	224	185	186	225
Competency_ Relating to Customers	Corr.	<b>.196*</b>	-.049	.100	.114	-.014	.051	.107	.110	.098	.111	.120	<b>.174*</b>
	N	119	119	119	119	119	119	119	119	224	185	186	225
Competency_ Reliability	Corr.	.169	-.066	.129	<b>.233*</b>	-.147	.070	.065	.089	.007	.076	.032	.097
	N	119	119	119	119	119	119	119	119	224	185	186	225
Competency_ Resilient	Corr.	.110	-.087	.070	.004	.029	-.007	.076	.056	<b>.138*</b>	<b>.161*</b>	.116	<b>.197*</b>
	N	119	119	119	119	119	119	119	119	224	185	186	225

\*\* . Correlation is significant at the 0.01 level (2-tailed); \* . Correlation is significant at the 0.05 level (2-tailed).

Small to moderate correlations were shown between the CP7.1C and criterion questionnaire ratings, where the higher the score on the CP7.1C, the less the likelihood of disciplinary issues (disciplinary ratings) and/or the appointing of a 'problem employee'. In addition, small positive correlations were found between the NP6.1 and the criterion questionnaire ratings for the competencies termed Relating to Customers and Resilient, as well as between the CP7.1C and the criterion questionnaire rating for the competency termed Resilient.

In terms of statistically significant relationships observed between CCCs and the ratings by supervisors on the criterion questionnaires, a number of correlations were displayed. Regarding the relationships between ratings on criterion questionnaires relating to Key Performance Areas and the CCCs, small to moderate positive correlations were shown with Quality Orientation, Reliability and Customer Focus.

Ratings were also allocated to occurrences and/or issues regarding employees' record relating to AWOL, Sick Leave and Time Keeping as follows: (i) Good performance in terms of AWOL showed a small to moderate positive correlation with Customer Focus; (ii) Good performance in terms of sick leave track record showed small to moderate positive correlations with Team Working and Customer Focus, and small to moderate negative correlations with Business Awareness and Resilient; (iii) Good performance in terms of time keeping track records showed small to moderate positive correlations with Quality Orientation and Customer Focus. The ratings which indicate good, average and poor performance for the criterion questionnaire items are clarified in Table 8, above.

Using the criterion questionnaires, candidates were rated by supervisors on the key behaviours related to the competencies Customer Focus, Quality Orientation, Resilient, Relating to Customers and Reliability. The ratings for these competencies were correlated with scores for the CCCs measured in the selection process. In terms of the relationship between the ratings on Customer Focus (criterion questionnaire) and the CCCs, a small to moderate positive correlation was observed between Customer Focus and Quality Orientation (CCCs). Small to moderate positive correlations were displayed between supervisor ratings for items related to Quality Orientation with the CCCs of Quality Orientation and Reliability. The last two

behavioural competencies rated by supervisors were those of Relating to Customers and Reliability. The supervisor ratings for the competency Relating to Customers showed a small to moderate positive correlation with Quality Orientation (CCC). Finally, supervisor ratings for items measuring Reliability showed a small to moderate positive correlation with the CCC Customer Focus. Despite correlations existing between CCCs and criterion questionnaire ratings, there did not appear to be any significant statistical relationships between these supervisor ratings and the overall PJM score.

Furthermore, in terms of scores for the interview, the Interview Average Score and Quality Orientation competency score showed small to moderate negative correlations with the supervisor ratings of incidents around AWOL, sick leave and disciplinary ratings. Finally, in terms of the relationship between interview data and supervisor ratings of behaviour, small to moderate positive correlations were present with the competencies Customer Focus and Quality Orientation, as well as with the Interview Average Score.

Now that the results of the study have been presented, a discussion on the findings follows in the next section of this article.

## **3.4 DISCUSSION**

### **3.4.1 Findings**

The aim of this study was to determine the predictive validity of the assessment battery being utilised in the selection process of CSAs in a South African commercial airline company. The sample of 192 candidates with scores for all predictor and criterion variables was an adequate sample size to ensure that statistically significant results were established (Guadagnoli & Velicer, 1988).

Regarding the predictive validity of the ability tests, the VCC1 showed statistically significant relationships with training scores; the NP6.1 with training scores and criterion questionnaire ratings (competency ratings); and the CP7.1C with training scores, performance appraisal ratings and criterion questionnaire ratings

(disciplinary, problem employee and competency ratings). The positive, statistically significant correlations found between candidates' ability scores and training results indicate that the candidates who performed better in the ability tests were inclined to perform better in training. Furthermore, individuals who did well in the checking ability test seemed to obtain higher performance appraisal ratings. Restriction of range due to the hurdle process followed during the selection process influence the results of the analysis, and the effect size of the correlation coefficients between the ability tests and the criterion data, although statistically significant, are small.

A statistically significant relationship was present between the participants' most recent performance appraisal scores and their overall PJM scores, as well as for the essential or highly important CCC competencies Fact Finding, Quality Orientation, Customer Focus and Using Initiative. This indicates that the participants identified by the PJM score as having a stronger fit to the requirements of the CSA position did tend to obtain higher performance appraisal ratings.

Furthermore, a number of CCCs displayed statistically significant relationships with training and criterion questionnaire ratings (KPAs, AWOL, sick leave, time keeping, open to feedback and competency ratings). An example of this includes candidates who displayed higher potential for Customer Focus also seemed to have fewer time-keeping, sick leave or AWOL issues and received higher performance ratings from their supervisors in terms of the key performance indicators. The overall PJM score was observed as having a significant positive relationship with training scores.

In conclusion, the essential and highly important competencies identified by the WPS during the job analysis and the overall PJM score contribute to identifying candidates who are more likely to be better performers in the role. The correlations obtained in this study are lower than expected. Schmidt and Hunter (1998) did a meta-analysis where they investigated 85 years of research involving the relationship between selection procedures and training performance. They found that general ability tests have a predictive validity ( $r$ ) of 0.51 and personality tests, represented by Conscientiousness, a predictive validity ( $r$ ) of 0.31. Although certain of the correlation coefficients approach 0.30 (e.g. the overall PJM and overall training performance ( $r=0.226$ ) and Quality Orientation ( $r=0.351$ ) and Customer Focus ( $r=0.281$ ), the

correlations obtained in this study are lower than that found by Schmidt and Hunter (1998). Bartram (2005) also did a meta-analysis using 29 validation studies involving the relationship between selection instruments and job performance and, using the Universal Competency Framework (UCF), determined that ability and personality data together obtain operational validities ranging from 0.20 to 0.44 for the competencies. One possible explanation for the correlation coefficients observed in this study being lower than expected is likely range restriction, which is discussed further in the next section.

### **3.4.2 Limitations**

While every effort was made to make the broadest range of data for job performance available, the sample was skewed in that it consisted only of candidates who had been successful through every stage of the selection process and did not represent individuals that were not selected by the company through the selection process. Therefore, restriction of range is likely to have occurred and correlation results should be interpreted with care. In addition, while instructions were clearly explained to supervisors, only one supervisor rated the items for each candidate in the criterion questionnaire. As such, there was the possibility that potential sources of error including the halo effect, central tendency and level of strictness impacted on the performance ratings.

### **3.4.3 Recommendations**

Examining the extent to which each predictor contributes towards predicting future job performance was not a focus of the research study. Determining this would add further value. The current study focused on one entry-level position in the airline company, the CSA. It would be advisable to conduct similar validation studies on the selection batteries for Cabin Attendants and Call Centre Agents. Finally, while the study yielded positive results, it is important to reassess whether, over time, the assessment measures that are being utilised for the selection of CSAs are actually successful in predicting future job performance. As such, it is suggested to repeat this study at a future date when more data are available. In order for this to be more efficient, it is recommended that all data be captured on a central database in future,

rather than hard copies of information having to be captured before the data can be analysed.

#### **3.4.4 Conclusion**

Despite the limitations and recommendations outlined above, the results of the research study indicate that the selection battery currently being used is a fairly valid predictor of future job performance by CSAs in the airline industry in South Africa and that the organisation can continue to use it for selection of CSAs in the future. Even though the number and effect size of the correlations were limited, the significant correlations obtained are between the criteria and the competencies identified by the WPS as essential or highly important, thereby providing confirmation of the requirements of the role as identified through the job analysis. The correlations obtained between the various selection instruments and the selected work performance criteria are consistent with previous findings in the literature review. As such, the research provides more evidence that personality testing, ability scores (verbal, numerical and checking) and competency based interviews can be used to predict future job performance. Even though small to moderate correlations were reported in most cases, they did represent statistically significant results.

While this research study has added value for the specific airline company for which it was carried out, the results can be of benefit to Human Resource Practitioners and IO Psychologists, as well as professionals and managers in fast-paced customer service organisations, in that it provides an understanding of selection tools which are effective at predicting future job performance in this industry.

### 3.5 REFERENCES

- Abrahams, F. & Mauer, R. (1999). The comparability of the constructs of the 16PF in the South African context. *SA Journal of Industrial Psychology, 25*, 53-59.
- Anderson, N., Lievens, F., Van Dam, K. & Ryan, A.M. (2004). Future perspectives on employee selection: Key directions for future research and practice. *Applied Psychology: An International Review, 53*(4), 487-501.
- Anderson, N. & Witvliet, C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal and Singapore. *International Journal of Selection and Assessment, 16*(1), 1-13.
- Barnard, A. (2010). Psychological Assessment: Predictors of human behaviour. In M. Coetzee & D. Schreuder (Eds.). *Personnel Psychology: An Applied Perspective* (134-169). Cape Town, South Africa: Oxford.
- Barnard, M. & Schaap, P. (2005). The concurrent validity of the technical test battery as an indicator of work performance in a telecommunications company. *SA Journal of Industrial Psychology, 31*(2), 71-77.
- Barrick, M., Feild, H.S. & Gatewood, R.D. (2011). *Selection in Human Resource Management*. (7<sup>th</sup> ed.). Canada: South-Western.
- Barrick, M.R. & Mount, M.K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1-26.
- Bartram, D. (2004). Assessment in organisations. *Applied Psychology: An International Review, 53*(2), 237-259.
- Bartram, D (2005). The Great Eight Competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90*(6), 1185-1203.

- Bedell, B., Van Eeden, R. & Van Staden, F. (1999). Culture as moderator variable in the psychological test performance: Issues and trends in South Africa. *SA Journal of Industrial Psychology*, 25(3), 1-7.
- Berry, J.W, Poortinga, Y.H., Segal, M.H. & Dasen, P.R. (2002). *Cross-Cultural Psychology: Research and Applications* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Borman, W.C., Penner, L.A., Allen, T.D., & Motowildo, S.J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9(1-2), 52-69.
- Branine, M. (2008). Graduate recruitment and selection in the UK: A study of the recent changes in methods and expectations. *Career Development International*, 13(6), 497-513.
- Brink, A. & Berndt, A. (2008). *Relationship Marketing and Customer Relationship Management*. Cape Town: Juta.
- Campion, M.A., Fink, A.A., Ruggeberg, B.J., Carr, L., Phillips, G.M. & Odman, R.B. (2011). Doing competencies well: Best practices in competency modelling. *Personnel Psychology*, 64, 225-262.
- Cascio, W.F. (1991). *Applied Psychology in Personnel Management*. London: Prentice-Hall.
- Christenson, L.B. (2001). *Experimental methodology*. Boston, MA: Allyn & Bacon.
- Chung-Yan, G.A., Hausdorf, P.A. & Cronshaw, S.F. (2005). A criterion-related validation study of transit operators. *International Journal of Selection and Assessment*, 13(2), 172-177.
- Clark-Carter, D. (2004). *Quantitative psychological research: A student's handbook*. (2nd ed.). East Sussex, UK: Psychology Press.

- Coetzee, M. & Roythorne-Jacobs, H. (2010). Job analysis and criterion development. In M. Coetzee & D. Schreuder (Eds.), *Personnel Psychology: An Applied Perspective* (88-133). Cape Town, South Africa: Oxford.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Collins, M.W. & Morris, S.B. (2008). Testing for adverse impact when sample size is small. *Journal of Applied Psychology*, 93(2), 463-471.
- Connellan T.K. & Zemke, R. (1993). *Sustaining knock your socks off service*. New York: Amacom.
- Cronbach, L.J. & Gleser, G.C. (1965). *Psychological tests and personnel decisions* (2<sup>nd</sup> ed.). Urbana, IL: University of Illinois Press.
- Damitz, M., Manzey, D., Kleinmann, M. & Severin, K. (2003). Assessment center for pilot selection: Construct and criterion validity and the impact of assessor type. *Applied Psychology: An International Review*, 52(2), 193-212.
- De Beer, M. (2006). Dynamic testing: Practical solutions to some concerns. *SA Journal of Industrial Psychology*, 32(4), 8-14.
- Elliott, J.G. (2003). Dynamic assessment in educational settings: realising potential. *Educational Review*, 55(1), 15-32
- Employment Equity Act, No 55 (1998). Government Gazette, 400 (19370). Cape Town. 19 October 1998.
- England, L.J. (1991). *Personality assessment: A review with special reference to South Africa*. Pretoria: HSRC.

- Flotman, A.P. (2002). *The predictive validity of the selection battery for trainee pilots in the South African Air Force* (Unpublished masters dissertation). University of South Africa, Pretoria, South Africa.
- Foxcroft, C. (1997). Psychological testing in South Africa: Perspectives regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13(3), 229-235.
- Foxcroft, C. (2004). Planning a psychological test in the multicultural South African context. *SA Journal of Industrial Psychology*, 30(4), 8-15.
- Foxcroft, C. & Roodt, R. (2001). *An introduction to psychological assessment in the South African context*. (5<sup>th</sup> ed.). Cape Town: Oxford.
- Goldberg, L.R. (1993). The structure of phenotypic personality traits: Authors' reactions to the six comments. *American Psychologist*, 48, 1303-1304.
- Gregory, R.J. (2007). *Psychological Testing: History, Principles and Applications*. (5<sup>th</sup> ed.). Boston, MA: Allyn and Bacon.
- Grigorenko, E.L. & Sternberg, R.J. (1998). Dynamic testing. *Psychological Bulletin*, 124, 75-111.
- Grobler, P.A., Warnich, S, Carrell, M.R., Elbert, N.F. & Hatfield, R.D. (2002). *Human Resource Management in South Africa* (2<sup>nd</sup> Ed.). London: Thomson.
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-275.
- Health Professions Act, No 56 (1974). Government Gazette, 717 (29079). Cape Town. 9 October 1974.

- Heinsman, H., De Hoogh, A.H.B., Koopman, P.L. & Van Muijen, J.J. (2007). Competencies through the eyes of psychologists: A closer look at assessing competencies. *International Journal of Selection and Assessment*, 15(4), 412-427.
- Hoffman, C.C. & McPhail, S.M. (1998). Exploring options for supporting test use in situations precluding local validation. *Personnel Psychology*, 51(3-4), 987-1003.
- HPCSA (2010a). *Form 94. Training and examination guidelines for psychometrists*. Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].
- HPCSA (2010b). *Form 207. List of tests classified as being psychological tests*. Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].
- HPCSA (2010c). *Form 208. Policy on the classification of psychometric measuring devices, instruments, methods and techniques*. Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].
- Huang, J.L. & Ryan, A.M. (2011). Beyond personality traits: A study of personality states and situational contingencies in customer service jobs. *Personnel Psychology*, 64, 451-488.
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96(1), 72-98.
- Hurtz, G.M. & Donovan, J.J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology*, 85(6), 869-879.
- 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.

- Jensen, A.R. (1980). *Bias in mental testing*. London: Methuen.
- Kriek, H. (2000). *Guidelines for best practice in occupational assessment in South Africa*. Pretoria: SHL
- La Grange, L. & Roodt, G. (2001). Personality and cognitive ability as predictors of the job performance of insurance sales people. *SA Journal of Industrial Psychology, 27*(3), 35-43.
- Lopes, A., Roodt, G. & Mauer, R. (2001). The predictive validity of the Apil-B in a financial institution. *SA Journal of Industrial Psychology, 27*(1), 61-69.
- Machado, R. & Diggines, C. (2012). *Customer Service*. Cape Town: Juta.
- Magwaza, S. (1995). Humanising and democratising assessments. *People Dynamics, 13*(5), 12-15.
- McCarthy, J. & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology, 57*(3), 607-637.
- McLagan, P.A. (1997). Competencies: The next generation. *Training and Development, 51*(5), 41-47.
- Mirabile, R.J. (1997). Everything you wanted to know about competency modelling. *Training and Development, 15*(8), 73-77.
- Mnguni, V.Q. (2011). *The predictive validity of a psychological test battery for the selection of cadet pilots in a commercial airline* (Unpublished masters dissertation). University of South Africa, Pretoria, South Africa.
- Muchinsky, P.M. (2004). When the psychometrics of test development meets organizational realities: A conceptual framework for organizational change, examples and recommendations. *Personnel Psychology, 57*(1), 175-209.

- Muchinsky, P.M., Kriek, H.J. & Schreuder, D. (2005). *Personnel Psychology* (3<sup>rd</sup> ed.). Cape Town: Oxford.
- Muller, J. & Schepers, J. (2003). The predictive validity of the selection battery used for junior leader training within the South African National Defence Force. *SA Journal of Industrial Psychology*, 29(3), 87-98.
- Murdy, L.B., Sells, S.B., Gavin, J.F. & Toole, D.L. (1973). Validity of personality and interest inventories for stewardesses. *Personnel Selection*, 26, 273-278.
- Murphy, K.R., Cronin, B.E. & Tam, A.P. (2003). Controversy and consensus regarding the use of cognitive ability testing in organizations. *Journal of Applied Psychology*, 88(4), 660-671.
- Murphy, R. & Maree, D.J.F. (2009). Revisiting core issues in dynamic assessment. *South African Journal of Psychology*, 39(4), 420-431.
- Nell, V. (1997). Science and politics meet at last: The insurance industry and neuropsychological norms. *South African Journal of Psychology*, 27(2), 43-49.
- Nicholls, M., Viviers, A.M. & Visser, D. (2009). Validation of a test battery for the selection of call centre operators in a communications company. *South African Journal of Psychology*, 39(1), 19-31.
- Noe, R.A. (2005). *Employee Training and Development* (3<sup>rd</sup> ed.). New York, NY: McGraw-Hill.
- Noe, R.A., Hollenbeck, J.R., Gerhart, B. & Wright, P.M. (2003). *Human Resource Management* (4<sup>th</sup> ed.). New York, NY: McGraw-Hill.
- Novit, M.S. (1979). *Essentials of personnel management*. Englewood Cliffs, NJ: Prentice Hall.

- Nzama, L., De Beer, M. & Visser, D. (2008). Predicting work performance through selection interview ratings and psychological assessment. *SA Journal of Industrial Psychology*, 34(3), 39-47.
- Ones, D.S., Dilchert, S., Viswesvaran, C. & Judge, T.A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60, 995-1027.
- Owen, K. (1991). Test bias: the validity of the Junior Aptitude Tests (JAT) for various population groups in South Africa regarding the constructs measured. *South African Journal of Psychology*, 21(2), 112-118.
- Owen, K. & Taljaard, J.J. (1996). *Handbook for the use of psychological and scholastic assessments of the HSRC*. Pretoria: K.L. Van Schaik.
- Outtz, J.L. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15(1/2), 161-171.
- Parkinson, M. (1999). *Using business in psychology*. Brookfield, VT: Gower.
- Paterson, H. & Uys, K. (2005). Critical issues in psychological test use in the South African workplace. *SA Journal of Industrial Psychology*, 31(3), 12-22.
- Pelser, M., Bergh, Z.C. & Visser, D. (2005). The concurrent validity of learning potential and psychomotor ability measures for the selection of haul truck operators in an open-pit mine. *SA Journal of Industrial Psychology*, 31(2), 58-70.
- Ployhart, R.E. (2006). Staffing in the 21<sup>st</sup> century: New challenges and strategic opportunities. *Journal of Management*, 32(6), 868-897.
- Poortinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, 24, 737-756.

- Potgieter, T.E. & Van der Merwe, R.P. (2002). Assessment in the workplace: A competency based approach. *SA Journal of Industrial Psychology*, 28(1), 60-66.
- Ree, M.J. & Carretta, T.M. (1996). Central role of g in military pilot selection. *The International Journal of Aviation Psychology*, 6(2), 111-123.
- Robertson, I.T. & Smith, M. (2010). Personnel Selection. *Journal of Occupational and Organizational Psychology*, 74 (4), 441-472.
- Rothmann, S. & Coetzer, E.P. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1), 67-84.
- Sackett, P.R. & Ellingson, J.E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50(3), 707-721.
- Sackett, P.R. & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, 59, 419-450.
- Salgado, J.F. & Anderson, N. (2002). Cognitive and GMA testing in the European community: Issues and evidence. *Human Performance*, 15, 75-96.
- Salgado, J.F. & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11(3), 299-324.
- Schaap, P. (2003). The construct comparability of the PIB/SPEEX stress index for job applicants from diverse cultural groups in South Africa. *South African Journal of Psychology*, 33(2), 95-101.
- Schaap, P. & Basson, J.S. (2003). The construct equivalence of the PIB/SPEEX motivation index for job applicants from diverse cultural backgrounds. *SA Journal of Industrial Psychology*, 29(2), 49-59.

- Schmidt, C. (2006). Validity as an action concept in IO psychology. *SA Journal of Industrial Psychology*, 32(4), 59-67.
- Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.
- Schmidt, F.L. & Rader, M. (1999). Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52(2), 445-464.
- Schmidt, F.L. & Zimmerman, R.D. (2004). A counterintuitive hypothesis about employment interview validity and some supporting evidence. *Journal of Applied Psychology*, 89(3), 553-561.
- SHL (2000). *The Customer Contact Styles Questionnaire manual and user's guide*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2008a). *Guidelines for Best Practice in Occupational Assessment in South Africa*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2009). *Test administration course notes V1.1*. Thames Ditton, UK: SHL Group Ltd.
- SHL (2013). *Person-Job Match (PJM) Research Evidence*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- Shum, D., O'Gorman, J. & Myers, B. (2006). *Psychological testing and assessment*. Melbourne: Oxford University Press.

- Skarlicki, D.P. (2003). HRM special issue: Fairness and human resources management. *Human Resources Review*, 13, 1-5.
- Spector, P.E. (2012). *Industrial and Organisational Psychology: Research and Practice* (6<sup>th</sup> ed.). Singapore: Wiley.
- Swanepoel, S. (1998). The predictive validity of selection criteria for personnel management students at a tertiary institution. *SA Journal of Industrial Psychology*, 24(2), 1-8.
- Taylor, T.R. (1994). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 24(4), 1-16.
- Tett, R.P. & Christiansen, N.D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology*, 60, 967-993.
- Theron, C. (2007). Confessions, scapegoats and flying pigs: Psychometric testing and the law. *SA Journal of Industrial Psychology*, 33(1), 102-117.
- Theron, C. (2009). The diversity-validity dilemma: In search of minimum adverse impact and maximum utility. *SA Journal of Industrial Psychology*, 35(1), 183-195.
- Van de Vijver, F.J.R. (2002). Cross-cultural assessment: Value for money? *Applied Psychology: An International Review*, 51(4), 545-566.
- Van de Vijver, A.J.R. & Rothmann, S. (2004). Assessment in multicultural groups: The South African Case. *SA Journal of Industrial Psychology*, 30(4), 1-7.
- Van der Merwe, R.P. (2002). Psychometric testing and human resource management. *SA Journal of Industrial Psychology*, 28(2), 77-86.

- Van Eeden, R., De Beer, M. & Coetzee, C.H. (2001). Cognitive ability, learning potential and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education* 15(1), 171-179.
- Van Iddekinge, C.H. & Ployhart, R.E. (2008). Developments in the criterion-related validation of selection procedures: A critical review and recommendations for practice. *Personnel Psychology*, 61, 821-925.
- Voskuijl, O.F. & Evers, A. (2008). Job analysis and competency modeling. In S. Cartwright & C.L. Cooper (Eds.), *The Oxford Handbook of Personnel Psychology* (139-162). Oxford, UK: Oxford.
- Zhao, S. & Du, J. (2011). The application of competency based talent assessment systems in China. *Human Systems Management*, 30, 23-37.

## **CHAPTER 4**

### **CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS**

#### **4.1 INTRODUCTION**

In this chapter, conclusions are drawn from the research findings and limitations of the research are highlighted. The chapter is concluded with recommendations for future research purposes.

#### **4.2 CONCLUSIONS**

Chapter 1 outlined the importance of recognising and selecting candidates who are likely to perform on a high level as customer service agents. It was argued that this was a priority for managers as organisations which provide exceptional customer service have a higher possibility of retaining their customers and, as a result, a greater chance of being successful (Machado & Diggines, 2012). Therefore, it is essential to use the most effective selection methods, whilst also meeting legal requirements in terms of psychological assessment practice.

The general aim of the study, as presented in Chapter 1, was to validate the selection battery for CSAs in a South African commercial airline company by establishing whether ability and competency scores derived from the selection tools (VCC1, NP6.1, CP7.1C, OPQ32 and a competency based interview) correlate with job performance. In addition, the specific aims of the study were outlined in terms of the literature review, as well as the empirical study. The literature review focused on gaining an understanding of ability testing, personality assessment and competency based interviews and how these can be operationalised as measurement constructs. It further conceptualised psychological assessment in the South African context and provided clarity on the actual assessment tools utilised in the study.

Chapter 2 presented a literature review of selection from an IO perspective in the South African multi-cultural context. It presented a discussion on psychological assessment and selection, and on the selection process, which begins with the job

analysis, the foundation to competency based assessment. The difference between psychological assessment and psychological testing was discussed, with a particular discussion around psychological constructs, which can be measured through psychological assessment. The South African context was explored, taking into account labour legislation, ethics, fairness and bias.

The literature review further covered competency based assessment, which meets the requirements of South African labour legislation (Employment Equity Act, 1998). It then covered a discussion on essential psychometric principles, with specific emphasis on reliability and validity. The importance of conducting validation studies on psychological assessment tools being used in selection decision making was discussed, with a focus on predictive validity, the measure of validity used in the research study.

Within the research article in Chapter 3, correlations were calculated between the predictors and the criterion data.

Small to moderate statistically significant correlations were found between the ability test scores and the criteria, as follows:

- For the VCC1, statistically significant correlations were observed with training scores (Overall Training – until December 2008; Training – PH – from January 2009; Training Overall – from January 2009).
- For the NP6.1, statistically significant correlations were found with training scores (Training Overall – from January 2009) and the criterion questionnaire ratings for the CCCs of Relating to Customers and Resilient.
- For the CP7.1C, statistically significant correlations were found with training scores (Overall Training – until December 2008; Training – PH – from January 2009; Training Overall – from January 2009), performance appraisal ratings and criterion questionnaire ratings (Disciplinary Rating, Problem Employee Rating, Resilient Competency score).

The reported correlations can therefore be used to deduce that the first hypothesis, “The ability scores of the VCC1, NP6.1 and CP7.1C significantly predicts job performance of a CSA in the airline industry” was accepted.

Small to moderate statistically significant correlations were found between a number of the scores for the essential and highly important CCCs and the performance criteria. These findings are discussed below:

- Small to moderate correlations were found between Relating to Customers, Team Working, Business Awareness, Quality Orientation and Customer Focus, and training scores.
- Small to moderate correlations were found between Fact Finding, Quality Orientation, Customer Focus and Using Initiative, and performance appraisal ratings.
- Small to moderate correlations were found between Team Working, Business Awareness, Quality Orientation, Reliability, Customer Focus and Resilient, and certain ratings on the criterion questionnaire.

As a result of the above correlations, the second research hypothesis, “The competency scores of the CCCs significantly predict job performance of a CSA in the airline industry” was accepted.

In terms of the relationship between the PJM score and the criterion data, small to moderate statistically significant correlations were found with training scores (Training – PH – from January 2009; Training Overall – from January 2009) and performance appraisal ratings. Consequently, the third research hypothesis, “The PJM score significantly predicts job performance of a CSA in the airline industry” was accepted.

Small to moderate statistically significant correlations were displayed between the competency scores of the competency based interview and the criterion data. The main findings of this predictor are outlined as follows:

- A small correlation was found between the Interview Average score and the training score for Training – PH (from January 2009).
- Small to moderate correlations were found between the score for Resilient and the Interview Average Score, and performance appraisal ratings.

- Small to moderate correlations were found between the scores for Customer Focus, Quality Orientation and the Interview Average Score, and certain ratings on the criterion questionnaire.

As a result, the fourth hypothesis for the research study “The competency scores of the competency based interview significantly predict job performance of a CSA in the airline industry” was accepted.

As highlighted in the results section of Chapter 3 and the above research summary, a number of correlations were displayed between the predictors and the criterion data. While the scope of the research study was limited, evaluating the relationships only in terms of the predictors and the criterion data, the research questions have been addressed and the aims of the research were met.

### **4.3 LIMITATIONS**

A number of limitations were present in the research study. These should be noted and caution should be exercised when interpreting the results.

First of all, every effort was made to ensure that the largest amount of job performance data was made available for the research study. A hurdle approach was used in the selection process, whereby performance data were only available for employees who were successful through all stages of the selection process. Consequently, this may have resulted in some restriction of range within the scores. Restriction of range takes place when employees need to “pass” each stage of the selection process before actually being appointed (Shavelson, 1988) and affects the correlation scores by bringing them down somewhat (Ree & Carretta, 1996). This could be a reason why small to moderate correlations, as opposed to large correlations, were observed.

With regards to the collection of criterion data through the criterion questionnaire, it was difficult, at first, to get the buy-in of some customer service supervisors to be a part of the study. Even after the purpose of the study was explained, there might still have been a certain amount of scepticism. In terms of rating the items on the criterion

questionnaires, instructions were explained clearly to the supervisors rating the CSAs. Due to the fact that only one supervisor (as opposed to a multi-rater review) rated each CSA, there was a possibility that potential sources of error occurred (e.g. halo effect, central tendency, level of strictness) when supervisors were assigning ratings to a CSA's performance. In addition, there were certain CSAs who had recently been appointed and, as a result, all the performance data were not available for them because they had not yet had a performance appraisal or there were no supervisors who felt comfortable to rate them. This further limited the size of the sample.

A third limitation of the study was that the correlations with performance appraisal ratings were based on the 2011 Half Year appraisal ratings for candidates. This was as a result of sample sizes for previous performance appraisal ratings not being large enough and the 2011 Full Year ratings not yet being available.

The study provides three effective measures for measuring job performance. Even though limitations to the study were identified, the study provides evidence of relationships between some of the predictors and the criterion data. Recommendations for future research will be presented in the section which follows.

#### **4.4 RECOMMENDATIONS**

The extent to which each predictor contributed to predicting future job performance was not explored. Carrying out a regression analysis in order to determine this would add further value to the understanding of the current selection process for CSAs.

The focus of the research was on one entry-level customer service position in the specific airline company. Similar validation studies could be carried out on the selection process of call centre agents and cabin attendants. This would allow the company to ensure that the psychological assessment batteries for those positions are also valid. In addition, further studies could provide further understanding of the determinants of future job performance in customer service jobs in airline companies and beyond.

While psychological tests appear to be valid predictors of job performance of CSAs in the airline company, they are only one source of information in the selection process, as discussed in Chapter 2. More comprehensive validation studies relating to other aspects of the selection process, such as work experience, reference checking and application forms could add value to improving the entire selection process. In addition, focusing on the effect of moderator variables, such as gender, ethnic orientation, geographical location and tenure, could add further clarity to the results.

In conclusion, while this study provided evidence of predictive validity of the psychological assessment battery, it is suggested that a subsequent study be carried out when more data are available in order to be able to generalise the results more effectively, as well as to ensure that the psychological assessment battery is still relevant for the requirements of the job. It is therefore recommended that the company immediately capture all data electronically in a central location. This would save time in the future in terms of having to capture hard copies of information before analysing the data.

#### **4.5 CONCLUSION**

In this chapter, the main findings of the research were summarised and discussed to determine whether the aims of the study were addressed. Thereafter, the limitations of the study were outlined and recommendations for future research were proposed.

## REFERENCES

- Aamodt, M.G. (2010). *Industrial/Organisational Psychology: An Applied Approach*. (6<sup>th</sup> ed.). Belmont, CA: Wadsworth.
- Abrahams, F. & Mauer, R. (1999). The comparability of the constructs of the 16PF in the South African context. *SA Journal of Industrial Psychology*, 25, 53-59.
- Advisory, Conciliation and Arbitration Service (ACAS) Online. Retrieved 10 March 2009 on World Wide Web. <http://www.acas.org.uk>. [Internet Document].
- Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice-Hall.
- Anderson, N., Lievens, F., Van Dam, K. & Ryan, A.M. (2004). Future perspectives on employee selection: Key directions for future research and practice. *Applied Psychology: An International Review*, 53(4), 487-501.
- Anderson, N. & Witvliet, C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal and Singapore. *International Journal of Selection and Assessment*, 16(1), 1-13.
- Barnard, A. (2010). Psychological Assessment: Predictors of human behaviour. In M. Coetzee & D. Schreuder (Eds.). *Personnel Psychology: An Applied Perspective* (134-169). Cape Town, South Africa: Oxford.
- Barnard, M. & Schaap, P. (2005). The concurrent validity of the technical test battery as an indicator of work performance in a telecommunications company. *SA Journal of Industrial Psychology*, 31(2), 71-77.
- Barrick, M., Feild, H.S. & Gatewood, R.D. (2011). *Selection in Human Resource Management*. (7<sup>th</sup> ed.). Canada: South-Western.

- Barrick, M.R. & Mount, M.K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1-26.
- Bartram, D. (2004). Assessment in organisations. *Applied Psychology: An International Review*, 53(2), 237-259.
- Bartram, D (2005). The Great Eight Competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90(6), 1185-1203.
- Bedell, B., Van Eeden, R. & Van Staden, F. (1999). Culture as moderator variable in the psychological test performance: Issues and trends in South Africa. *SA Journal of Industrial Psychology*, 25(3), 1-7.
- Bergh, Z.C. & Theron, L.T. (2004). *Psychology in the work context*. (2<sup>nd</sup> ed.). Cape Town: Oxford.
- Berry, J.W, Poortinga, Y.H., Segal, M.H. & Dasen, P.R. (2002). *Cross-Cultural Psychology: Research and Applications* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Bogdan, R.C. & Biklen, S.K. (1998). *Qualitative research for education: An introduction to theory and methods*. (3<sup>rd</sup> ed.). Boston: Allyn and Bacon.
- Borman, W.C., Penner, L.A., Allen, T.D., & Motowildo, S.J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9(1-2), 52-69.
- Boyatzis, R. (1982). *The competent manager: A model for effective performance*. New York, NY: John Wiley and Sons.
- Branine, M. (2008). Graduate recruitment and selection in the UK: A study of the recent changes in methods and expectations. *Career Development International*, 13(6), 497-513.

- Brink, A. & Berndt, A. (2008). *Relationship Marketing and Customer Relationship Management*. Cape Town: Juta.
- Campion, M.A., Fink, A.A., Ruggeberg, B.J., Carr, L., Phillips, G.M. & Odman, R.B. (2011). Doing competencies well: Best practices in competency modelling. *Personnel Psychology*, 64, 225-262.
- Carretta, T.R. & Ree, M.J. (2000). General and specific cognitive psychomotor abilities in personnel selection: The prediction of training and job performance. *Ability and Personnel Selection*, 8(4), 227-236.
- Cascio, W.F. (1991). *Applied Psychology in Personnel Management*. London: Prentice-Hall.
- Cascio, W.F. (1995). Whither industrial and organizational psychology in a changing world of work? *American Psychologist*, 50(11), 928-939.
- Cascio, W.F. (2001). Knowledge creation for practical solutions appropriate to a changing world of work. *SA Journal of Industrial Psychology*, 27(4), 14-16.
- Cattell, R.B. (1971). *Abilities: their structure, growth and action*. Boston, MA: Houghton Mifflin.
- Cheetham, G. & Chivers, G. (1997). Towards the holistic model of professional competence. *Journal of European Industrial Training*, 20(5), 20-31.
- Christenson, L.B. (2001). *Experimental methodology*. Boston, MA: Allyn & Bacon.
- Chung-Yan, G.A., Hausdorf, P.A. & Cronshaw, S.F. (2005). A criterion-related validation study of transit operators. *International Journal of Selection and Assessment*, 13(2), 172-177.
- Clark-Carter, D. (2004). *Quantitative psychological research: A student's handbook*. (2nd ed.). East Sussex, UK: Psychology Press.

- Coetzee, M. & Roythorne-Jacobs, H. (2010). Job analysis and criterion development. In M. Coetzee & D. Schreuder (Eds.), *Personnel Psychology: An Applied Perspective* (88-133). Cape Town, South Africa: Oxford.
- Coetzee, S. (2010). Psychological assessment: Predictors of human behaviour. In M. Coetzee & D. Schreuder (Eds.). *Personnel Psychology: An Applied Perspective* (134-169). Cape Town, South Africa: Oxford.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Collins, M.W. & Morris, S.B. (2008). Testing for adverse impact when sample size is small. *Journal of Applied Psychology*, 93(2), 463-471.
- Connellan T.K. & Zemke, R. (1993). *Sustaining knock your socks off service*. New York: Amacom.
- Cronbach, L.J. & Gleser, G.C. (1965). *Psychological tests and personnel decisions* (2<sup>nd</sup> ed.). Urbana, IL: University of Illinois Press.
- Damitz, M., Manzey, D., Kleinmann, M. & Severin, K. (2003). Assessment center for pilot selection: Construct and criterion validity and the impact of assessor type. *Applied Psychology: An International Review*, 52(2), 193-212.
- Dash, N.K. (1993). Research paradigms in education: Towards a resolution. *Journal of Indian Education*, 19(2), 1-6.
- De Beer, M. (2006). Dynamic testing: Practical solutions to some concerns. *SA Journal of Industrial Psychology*, 32(4), 8-14.
- Elliott, J.G. (2003). Dynamic assessment in educational settings: realising potential. *Educational Review*, 55(1), 15-32

- Employment Equity Act, No 55 (1998). Government Gazette, 400 (19370). Cape Town. 19 October 1998.
- Erasmus, P.F. & Arumugan, S. (1997). ABLE: The key to adequate job performance. *People Dynamics*, 15(5), 36-37.
- Ferguson, G. (1956). On transfer and the abilities of management. *Canadian Journal of Psychology*, 10(1), 121-131.
- Flotman, A.P. (2002). *The predictive validity of the selection battery for trainee pilots in the South African Air Force* (Unpublished masters dissertation). University of South Africa, Pretoria, South Africa.
- Foxcroft, C. (1997). Psychological testing in South Africa: Perspectives regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13(3), 229-235.
- Foxcroft, C. (2004). Planning a psychological test in the multicultural South African context. *SA Journal of Industrial Psychology*, 30(4), 8-15.
- Foxcroft, C. & Roodt, R. (2001). *An introduction to psychological assessment in the South African context*. (5<sup>th</sup> ed.). Cape Town: Oxford.
- Friedman, H.S. & Schustack, M.W. (2003). *Personality: Classic Theories and Modern Research*. (2<sup>nd</sup> ed.). Boston, MA: Allyn and Bacon.
- Geisinger, K.F. & Carlson, J.F. (1998). *Test interpretation and diversity*. Washington, DC: American Psychological Association.
- Goldberg, L.R. (1993). The structure of phenotypic personality traits: Authors' reactions to the six comments. *American Psychologist*, 48, 1303-1304.
- Gregory, R.J. (2007). *Psychological Testing: History, Principles and Applications*. (5<sup>th</sup> ed.). Boston, MA: Allyn and Bacon.

- Grigorenko, E.L. & Sternberg, R.J. (1998). Dynamic testing. *Psychological Bulletin*, 124, 75-111.
- Grobbelaar, R.L., Roodt, G. & Venter, A. (2004). Critical competencies to promote a customer core capability in a human resources call centre. *SA Journal of Human Resources Management*, 2(3), 16-24.
- Grobler, P.A., Warnich, S, Carrell, M.R., Elbert, N.F. & Hatfield, R.D. (2002). *Human Resource Management in South Africa* (2<sup>nd</sup> Ed.). London: Thomson.
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-275.
- Guion, R.M. (1998). Some virtues of dissatisfaction in the science and practice of personnel selection. *Human Resource Management Review*, 8(4), 351-365.
- Hager, P. & Gonczi, A. (1994). General issues about assessment of competence. *Assessment and Evaluation in Higher Education*, 19(1), 3-16.
- Health Professions Act, No 56 (1974). Government Gazette, 717 (29079). Cape Town. 9 October 1974.
- Heinsman, H., De Hoogh, A.H.B., Koopman, P.L. & Van Muijen, J.J. (2007). Competencies through the eyes of psychologists: A closer look at assessing competencies. *International Journal of Selection and Assessment*, 15(4), 412-427.
- Hoffman, C.C. & McPhail, S.M. (1998). Exploring options for supporting test use in situations precluding local validation. *Personnel Psychology*, 51(3-4), 987-1003.
- HPCSA (2010a). *Form 94. Training and examination guidelines for psychometrists*. Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].

- HPCSA (2010b). *Form 207. List of tests classified as being psychological tests.* Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].
- HPCSA (2010c). *Form 208. Policy on the classification of psychometric measuring devices, instruments, methods and techniques.* Retrieved 4 March 2013 on World Wide Web. <http://www.hpcsa.co.za>. [Internet document].
- Huang, J.L. & Ryan, A.M. (2011). Beyond personality traits: A study of personality states and situational contingencies in customer service jobs. *Personnel Psychology, 64*, 451-488.
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*(1), 72-98.
- Hurtz, G.M. & Donovan, J.J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*(6), 869-879.
- 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 (1999). *International guidelines for test-use* (Version 2000). Prepared by D. Bartram. Available at [http://cwis.kub.nl/~fsw\\_1/itc/](http://cwis.kub.nl/~fsw_1/itc/).
- Janz, T. (1982). Initial comparisons of patterned behaviour description interviews versus unstructured interviews. *Journal of Applied Psychology, 67*(5), 577-580.
- Jensen, A.R. (1980). *Bias in mental testing*. London: Methuen.
- Kriek, H. (2000). *Guidelines for best practice in occupational assessment in South Africa*. Pretoria: SHL

- La Grange, L. & Roodt, G. (2001). Personality and cognitive ability as predictors of the job performance of insurance sales people. *SA Journal of Industrial Psychology, 27*(3), 35-43.
- Labour Relations Act, No 66 (1995). Government Gazette, No.1877. Cape Town. 13 December 1995.
- Lawler, E.E. (1994). From job-based to competency based organizations. *Journal of Organizational Behavior, 15*, 3-15.
- Lopes, A., Roodt, G. & Mauer, R. (2001). The predictive validity of the Apil-B in a financial institution. *SA Journal of Industrial Psychology, 27*(1), 61-69.
- Machado, R. & Diggins, C. (2012). *Customer Service*. Cape Town: Juta.
- Magwaza, S. (1995). Humanising and democratising assessments. *People Dynamics, 13*(5), 12-15.
- Mauer, K.F. (2000). *Psychological test use in South Africa*. Pretoria: University of South Africa.
- McCarthy, J. & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology, 57*(3), 607-637.
- McIntire, S.A. & Miller, L.A. (2007). *Foundations of psychological testing: A practical approach* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- McLagan, P.A. (1997). Competencies: The next generation. *Training and Development, 51*(5), 41-47.
- Mirabile, R.J. (1997). Everything you wanted to know about competency modelling. *Training and Development, 15*(8), 73-77.

- Mnguni, V.Q. (2011). *The predictive validity of a psychological test battery for the selection of cadet pilots in a commercial airline* (Unpublished masters dissertation). University of South Africa, Pretoria, South Africa.
- Mount, M.K. and Barrick, M.R. (1998). Five reasons why the big five article has been frequently cited. *Personnel Psychology*, 51(4), 849-857.
- Mount, M.K., Witt, L.A. & Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and Five Factor personality constructs. *Personnel Psychology*, 53(1-2), 299-323.
- Mouton, J. & Marais, H.C. (1992). *Basic concepts in the methodology of social sciences*. Pretoria: HSRC.
- Muchinsky, P.M. (2004). When the psychometrics of test development meets organizational realities: A conceptual framework for organizational change, examples and recommendations. *Personnel Psychology*, 57(1), 175-209.
- Muchinsky, P.M., Kriek, H.J. & Schreuder, D. (2005). *Personnel Psychology* (3<sup>rd</sup> ed.). Cape Town: Oxford.
- Muller, J. & Schepers, J. (2003). The predictive validity of the selection battery used for junior leader training within the South African National Defence Force. *SA Journal of Industrial Psychology*, 29(3), 87-98.
- Murdy, L.B., Sells, S.B., Gavin, J.F. & Toole, D.L. (1973). Validity of personality and interest inventories for stewardesses. *Personnel Selection*, 26, 273-278.
- Murphy, K.R., Cronin, B.E. & Tam, A.P. (2003). Controversy and consensus regarding the use of cognitive ability testing in organizations. *Journal of Applied Psychology*, 88(4), 660-671.
- Murphy, R. & Maree, D.J.F. (2009). Revisiting core issues in dynamic assessment. *South African Journal of Psychology*, 39(4), 420-431.

- Nell, V. (1997). Science and politics meet at last: The insurance industry and neuropsychological norms. *South African Journal of Psychology*, 27(2), 43-49.
- Nicholls, M., Viviers, A.M. & Visser, D. (2009). Validation of a test battery for the selection of call centre operators in a communications company. *South African Journal of Psychology*, 39(1), 19-31.
- Noe, R.A. (2005). *Employee Training and Development* (3<sup>rd</sup> ed.). New York, NY: McGraw-Hill.
- Noe, R.A., Hollenbeck, J.R., Gerhart, B. & Wright, P.M. (2003). *Human Resource Management* (4<sup>th</sup> ed.). New York, NY: McGraw-Hill.
- Novit, M.S. (1979). *Essentials of personnel management*. Englewood Cliffs, NJ: Prentice Hall.
- Nzama, L., De Beer, M. & Visser, D. (2008). Predicting work performance through selection interview ratings and psychological assessment. *SA Journal of Industrial Psychology*, 34(3), 39-47.
- O'Leary, Z. (2004). *The essential guide to doing research*. London: Sage.
- Ones, D.S., Dilchert, S., Viswesvaran, C. & Judge, T.A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60, 995-1027.
- Owen, K. (1991). Test bias: the validity of the Junior Aptitude Tests (JAT) for various population groups in South Africa regarding the constructs measured. *South African Journal of Psychology*, 21(2), 112-118.
- Owen, K. & Taljaard, J.J. (1996). *Handbook for the use of psychological and scholastic assessments of the HSRC*. Pretoria: K.L. Van Schaik.

- Oultz, J.L. (2002). The role of cognitive ability tests in employment selection. *Human Performance, 15*(1/2), 161-171.
- Parkinson, M. (1999). *Using business in psychology*. Brookfield, VT: Gower.
- Paterson, H. & Uys, K. (2005). Critical issues in psychological test use in the South African workplace. *SA Journal of Industrial Psychology, 31*(3), 12-22.
- Pelser, M., Bergh, Z.C. & Visser, D. (2005). The concurrent validity of learning potential and psychomotor ability measures for the selection of haul truck operators in an open-pit mine. *SA Journal of Industrial Psychology, 31*(2), 58-70.
- Ployhart, R.E. (2006). Staffing in the 21<sup>st</sup> century: New challenges and strategic opportunities. *Journal of Management, 32*(6), 868-897.
- Poortinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology, 24*, 737-756.
- Potgieter, T.E. & Van der Merwe, R.P. (2002). Assessment in the workplace: A competency based approach. *SA Journal of Industrial Psychology, 28*(1), 60-66.
- Pretorius, G. (2012). Reflections on the scope of practice in the South African profession of psychology: a moral plea for relevance and a future vision. *South African Journal of Psychology, 42*(4), 509-521.
- Ree, M.J. & Carretta, T.M. (1996). Central role of g in military pilot selection. *The International Journal of Aviation Psychology, 6*(2), 111-123.
- Robertson, I.T. & Smith, M. (2010). Personnel Selection. *Journal of Occupational and Organizational Psychology, 74* (4), 441-472.

- Rothmann, S. & Coetzer, E.P. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1), 67-84.
- Sackett, P.R. & Ellingson, J.E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50(3), 707-721.
- 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.
- Sackett, P.R. & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, 59, 419-450.
- Salgado, J.F. & Anderson, N. (2002). Cognitive and GMA testing in the European community: Issues and evidence. *Human Performance*, 15, 75-96.
- Salgado, J.F. & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11(3), 299-324.
- Schaap, P. (2003). The construct comparability of the PIB/SPEEX stress index for job applicants from diverse cultural groups in South Africa. *South African Journal of Psychology*, 33(2), 95-101.
- Schaap, P. & Basson, J.S. (2003). The construct equivalence of the PIB/SPEEX motivation index for job applicants from diverse cultural backgrounds. *SA Journal of Industrial Psychology*, 29(2), 49-59.
- Schmidt, C. (2006). Validity as an action concept in IO psychology. *SA Journal of Industrial Psychology*, 32(4), 59-67.

- Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.
- Schmidt, F.L. & Rader, M. (1999). Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52(2), 445-464.
- Schmidt, F.L. & Zimmerman, R.D. (2004). A counterintuitive hypothesis about employment interview validity and some supporting evidence. *Journal of Applied Psychology*, 89(3), 553-561.
- Seijts, G.H. & Kyei-Poku, I. (2009). The role of situational interviews in fostering positive reactions to selection decisions. *Applied Psychology: An International Review*, 59(3), 431-453.
- Shavelson, R.J. (1988). *Statistical reasoning for the behavioural sciences* (2<sup>nd</sup> ed.). Boston: Allyn & Bacon.
- SHL (1992). *Personnel Test Battery manual & user's guide*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2000). *The Customer Contact Styles Questionnaire manual and user's guide*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2004). *Validity Study. Study No. V032. Personnel Test Battery*. Retrieved 4 March 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2006). *OPQ32 Technical Manual*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2008a). *Guidelines for Best Practice in Occupational Assessment in South Africa*. Pretoria: SHL.

- SHL (2008b). Reliability study. Study number R061. *Customer Contact Aptitude Series*. Retrieved 4 March 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2008c). Reliability study. Study number R066. *Personnel Test Battery*. Retrieved 4 March 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- SHL (2009). *Test administration course notes V1.1*. Thames Ditton, UK: SHL Group Ltd.
- SHL (2013). *Person-Job Match (PJM) Research Evidence*. Retrieved 27 May 2013 on World Wide Web. <http://www.shl.com>. [Internet document].
- Shum, D., O’Gorman, J. & Myers, B. (2006). *Psychological testing and assessment*. Melbourne: Oxford University Press.
- Society of Industrial and Organisational Psychology of South Africa (2005). *Guidelines for the Validation and Use of Assessment Procedures for the Workplace*. Pretoria: SIOPSA
- Skarlicki, D.P. (2003). HRM special issue: Fairness and human resources management. *Human Resources Review*, 13, 1-5.
- Spector, P.E. (2012). *Industrial and Organisational Psychology: Research and Practice* (6<sup>th</sup> ed.). Singapore: Wiley.
- Struwig, F.W. & Stead, G.B. (2001). *Planning, designing and reporting research*. Cape Town: Pearson.
- Swanepoel, S. (1998). The predictive validity of selection criteria for personnel management students at a tertiary institution. *SA Journal of Industrial Psychology*, 24(2), 1-8.

- Swinscow, T.D.V. & Campbell, M.J. (1999). *Statistics at Square One* (9<sup>th</sup> ed.). London: BMJ.
- Taylor, T.R. (1994). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 24(4), 1-16.
- Tett, R.P. & Christiansen, N.D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology*, 60, 967-993.
- Theron, C. (2007). Confessions, scapegoats and flying pigs: Psychometric testing and the law. *SA Journal of Industrial Psychology*, 33(1), 102-117.
- Theron, C. (2009). The diversity-validity dilemma: In search of minimum adverse impact and maximum utility. *SA Journal of Industrial Psychology*, 35(1), 183-195.
- Ulrich, D. (1998). *Delivering results: A new mandate for human resource professionals*. Boston, MA: Harvard Business School Publishing.
- Van de Vijver, F.J.R. (2002). Cross-cultural assessment: Value for money? *Applied Psychology: An International Review*, 51(4), 545-566.
- Van de Vijver, F.J.R. & Poortinga, Y.H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13, 29-37.
- Van de Vijver, A.J.R. & Rothmann, S. (2004). Assessment in multicultural groups: The South African Case. *SA Journal of Industrial Psychology*, 30(4), 1-7.
- Van der Merwe, R.P. (2002). Psychometric testing and human resource management. *SA Journal of Industrial Psychology*, 28(2), 77-86.

- Van Eeden, R., De Beer, M. & Coetzee, C.H. (2001). Cognitive ability, learning potential and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education* 15(1), 171-179.
- Van Iddekinge, C.H. & Ployhart, R.E. (2008). Developments in the criterion-related validation of selection procedures: A critical review and recommendations for practice. *Personnel Psychology*, 61, 821-925.
- Voskuijl, O.F. & Evers, A. (2008). Job analysis and competency modeling. In S. Cartwright & C.L. Cooper (Eds.), *The Oxford Handbook of Personnel Psychology* (139-162). Oxford, UK: Oxford.
- Zhao, S. & Du, J. (2011). The application of competency based talent assessment systems in China. *Human Systems Management*, 30, 23-37.