

**The construct equivalence of the Customer Contact  
Styles Questionnaire (CCSQ7.2) for proctored and  
unproctored administration groups**

by

Aleksandra Hyra

Submitted in partial fulfilment of the  
requirements for the degree

Magister Commercii (Human Resources Management)

in the

Faculty of Economic and Management Sciences

at the

University of Pretoria

## CONCEPT DECLARATION

I, Aleksandra Maria Hyra, do hereby declare that **“The construct equivalence of the Customer Contact Styles Questionnaire (CCSQ7.2) for proctored and unproctored administration groups”** is my own work. All the resources used in this study are cited and referenced by means of a comprehensive referencing system.

I declare that the content of this thesis has never been used before for any qualification at any tertiary institution.

---

Aleksandra Hyra

---

Date

## **ABSTRACT**

The advent of the Internet in personnel selection has brought with it many opportunities, challenges and concerns. The opportunities afforded to the employer include, among others, increased security, a reduction in the resources needed, reduced geographical constraints, reduced time to hire, immediate scoring and a lowering of costs. Owing to these benefits, the use of the Internet for psychological testing will continue to become more prevalent in personnel selection over the next few years.

The use and acceptance of the Internet has encouraged organisations and assessment practitioners to make use of Internet-delivered testing as the preferred mode of administration, especially in view of the benefits that it offers to the employer, as well as to the candidate. Internet-delivered testing brings with it a few concerns and many organisations are still making use of the paper-and-pencil version of an instrument even though it is available in an online version. This is because of concerns over measurement equivalence, as the paper-and-pencil version is seen as a "safer" and more accepted mode of administration.

There are also many test publishers that would like to take advantage of the practicability of Internet-delivered testing by converting their current selection procedures to an Internet format. This conversion cannot simply be assumed and needs to be examined and documented for each instrument.

The primary purpose of this study is to determine the construct equivalence of the Customer Contact Styles Questionnaire (CCSQ7.2) when it is administered via paper-and-pencil in the presence of a proctor and when it is administered online in the absence of a proctor. The aim is to determine whether the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric property.

The results of the study revealed that the mode in which the CCSQ7.2 is administered does not compromise scale reliabilities and that the relationship between scales of the CCSQ7.2 are not affected by modes of administration. It can thus be said with confidence that the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric property.

## **ACKNOWLEDGEMENTS**

I would like to express thanks to the following individuals and organisations for their contributions, patience and endless support towards the completion of this study:

- Chantal Olckers, for her commitment, patience and support;
- Tina Joubert for all her assistance throughout the duration of this study;
- SHL South Africa, who provided me with the opportunity of carrying out this project;
- My family, who supported and motivated me from start to finish;
- Nicolas Furtak, for his encouragement and continuous support.

## **TABLE OF CONTENTS**

<b>CONCEPT DECLARATION</b>	<b>i</b>
<b>ABSTRACT</b>	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b>	<b>iii</b>
<b>TABLE OF CONTENTS</b>	<b>iv</b>
<b>LIST OF TABLES</b>	<b>vi</b>
<b>LIST OF FIGURES</b>	<b>vi</b>
<b>CHAPTER 1</b>	<b>1</b>
<b>Introduction and overview</b>	<b>1</b>
1.1 Background and motivation	1
1.2 Problem statement	2
1.3 Purpose of the study	3
1.4 Significance of the study	4
1.5 Research question and hypotheses	5
1.5.1 Research questions	5
1.5.2 Hypotheses	5
1.6 Chapter overview	7
1.7 Conclusion	8
<b>CHAPTER 2</b>	<b>9</b>
<b>Review of psychological assessment and Internet-delivered testing</b>	<b>9</b>
2.1 Introduction	9
2.2 Psychological assessment	9
2.3 Personality questionnaires and their role in occupational assessment	12
2.4 Previous research conducted on the CCSQ7.2	15
2.5 Internet-delivered testing	22
2.5.1 Measurement equivalence	25
2.5.2 Advantages of Internet-delivered testing	28
2.5.3 Concerns and challenges of Internet-delivered-testing	32
2.5.4 Best practice guidelines	35
2.6 The Internet and the future	38
2.6.1 Internet interviews	38
2.6.2 Reference checks	38
2.6.3 Assessment and development centres	39
2.7 Conclusion	39
<b>CHAPTER 3</b>	<b>40</b>
<b>Research methodology</b>	<b>40</b>
3.1 Research approach	40
3.2 Research design	42
3.3 Sample	42
3.3.1 Description of the sample	43
3.3.2 Ethical issues surrounding the study	46
3.4 Data-collection procedure	47
3.4.1 Type of data collected	48
3.5 Measurement instrument	49
3.5.1 The Customer Contact Styles Questionnaire (CCSQ)	50
3.5.2 Development of the Customer Contact Styles Questionnaire (CCSQ)	52
3.5.3 Fairness of the instrument	54
3.5.4 Reliability and validity of the CCSQ7.2	55
3.5.5 Validity of the CCSQ7.2	58
3.5.6 Norms of the instrument	59
3.5.7 The use of ipsative measurement	59
3.5.8 Justification for the use of ipsative measurement	60

3.6 Data analysis	64
3.6.1 Comparison of effect sizes	65
3.6.2 Test reliability (Alpha coefficient)	65
3.6.3 Equivalence of covariance structures using SEM with EQS	66
3.7 Conclusion	67
<b>CHAPTER 4</b>	<b>68</b>
<b>Data analysis and results</b>	<b>68</b>
4.1 Descriptive statistics	68
4.2 Comparison of effect sizes	70
4.3 Test reliability (Alpha coefficient)	71
4.4 Equivalence of covariance structures using SEM with EQS	73
<b>CHAPTER 5</b>	<b>74</b>
<b>Discussion, limitations, recommendations and conclusion</b>	<b>74</b>
5.1 Discussion	74
5.2 Limitations and recommendations	75
5.3 Conclusion	77
<b>REFERENCES</b>	<b>78</b>

## LIST OF TABLES

CHAPTER 1	
Table 1-1 Chapter overview	7
CHAPTER 2	
Table 2-1 The uses of psychological testing	11
Table 2-2 The Big Five factors of personality	13
Table 2-3 Benefits of personality assessment	14
Table 2-4 Analysis of the CCSQ7.2 by gender	20
Table 2-5 Analysis of the CCSQ7.2 by age	21
Table 2-6 Internet users in the world by geographic region	24
Table 2-7 Improving candidate reactions to Internet-based testing	30
CHAPTER 3	
Table 3-1 Descriptive statistics for sample group 1 (N=358)	44
Table 3-2 Descriptive statistics for sample group 2 (N=239)	45
Table 3-3 Standardised test administration	47
Table 3-4 Customer Contact Styles Questionnaire model and its relationship to the SPQ and CSQ	53
Table 3-5 Reliability of the CCSQ7.2 (N=617)	56
Table 3-6 Reliability study of the CCSQ7.2 between a South African (N=737) and a British (N=617) sample	57
Table 3-7 Correlations between the two versions of the CCSQ (N=261)	62
Table 3-8 Effect sizes	65
CHAPTER 4	
Table 4-1 Mean, Standard deviation, Minimum, Maximum, Skewness and Kurtosis for the CCSQ7.2 (N=358)	68
Table 4-2 Mean, Standard deviation, Minimum, Maximum, Skewness and Kurtosis for the CCSQ7.2 (N=239)	69
Table 4-3 Mean, Standard deviation and Effect size for both sample groups of the CCSQ7.2	70
Table 4-4 Alpha coefficients and SEM scores for both sample groups of the CCSQ7.2	72
Table 4-5 Differences in reliability scores for both sample groups	73

## LIST OF FIGURES

CHAPTER 2	
Figure 2-1 The evolution of assessment	31
Figure 2-2 Honesty contract	34
Figure 2-3 Better practice steps	36
CHAPTER 3	
Figure 3-1 A conceptual model of quantitative research	41
Figure 3-2 The OPQ model of personality	50
Figure 3-3 Forced choice format	60
Figure 3-4 Percent raw score point change in score on scale	64

# **CHAPTER 1**

## **Introduction and overview**

The purpose of the opening chapter will be to provide the reader with a comprehensive understanding of the objective of this study. Information will be given on the problem statement, the background and the purpose of the study, as well as its significance. The content of each chapter of this study will be discussed in detail and an outline of each given.

### **1.1 Background and motivation**

Bartram (2001:4) states that "the infrastructure is now being built to support a radical change in the way testing is done." The availability and acceptance of computers have increased dramatically and more and more organisations are gravitating towards Internet administration as a means of enhancing personnel practices, decreasing costs, increasing speed and efficiency, diminishing the amount of missing data, and allowing candidates to be assessed in different locations (Bartram, 2001; Ployhart, Weekley, Holtz & Kemp, 2003). For these reasons, the use of the Internet for psychological testing will continue to become more prevalent in personnel selection over the next few years.

Ployhart et al (2003:736) state that "when an organization or test developer contemplates implementing a Web-based version of an existing test, it is important that the psychometric equivalence of the test be documented across the two test formats." For the purposes of this study, the researcher will be testing the psychometric equivalence of the Customer Contact Styles Questionnaire (CCSQ7.2) when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting. It is important to note that a managed mode was used for the sample group that completed the CCSQ7.2 under supervised assessment conditions and a controlled mode was used for the sample group that completed the online version of the CCSQ7.2 (Bartram, 2001).

There are currently few studies that have examined the equivalence of Internet-administered and traditional tests (Potosky & Bobko, 2004). This study will aim to add to the literature on existing equivalence studies, thereby providing individuals and organisations with more evidence that yet another instrument, specifically the CCSQ7.2, can indeed be used across both administration modes without loss of psychometric property.



## **1.2 Problem statement**

Owing to the fact that many individuals have accepted the Internet as an accessible medium with which to interact, it has become an important part of the everyday lives of many individuals. Also, the percentage of people with high-speed Internet access continues to grow, making it more feasible to administer increasingly sophisticated, high fidelity assessments over the Internet (Beaty & Shepherd in Tippins, Beaty, Drasgow, Gibson, Pearlman, Segall, & Shepherd, 2006). It is for this particular reason that the Internet rather than the traditional paper-and-pencil mode of administration will become the medium of choice for testing over the next decade (Bartram, 2001).

The Internet offers many possibilities for improving the practice of personnel selection, including increasing efficiency and lowering costs, reducing the amount of missing data, allowing applicants to be tested from diverse locations and at different times and providing immediate scoring (Ployhart et al, 2003). Despite these possibilities there are still many individuals and organisations who express their concern over the issue of Internet-delivered testing. Many of the challenges for and concerns of these individuals and organisations which are related to Internet-delivered testing include, among others, performance, security, privacy, ethical, legal and fairness issues (Bartram, 2001; Ployhart et al, 2003). An additional concern over, as well as an advantage of, the issue of Internet-delivered testing is the abandonment of the traditional requirement for local supervision (Bartram & Brown, 2004).

There are currently many test publishers that would like to take advantage of the benefits of Internet-delivered testing by converting their current selection procedures to an Internet format. Labouvie, Riordan and Vandenberg (in King & Miles, 1995:644) state that "for an instrument to be considered stable across administration modes it must first have a constant conceptual domain, that is, the construct being measured must be equivalent." It is therefore imperative that any selection procedure that is being converted to an Internet format should be tested for equivalence.

The CCSQ7.2 is available in different administration modes and, in order to demonstrate the constant conceptual domain of this questionnaire, the main purpose of this research study is to establish the construct equivalence of the ipsative version of the CCSQ7.2 when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting.

### **1.3 Purpose of the study**

The main purpose of this research study is to establish the construct equivalence of the CCSQ7.2 when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting.

The researcher will endeavour to provide organisations and individuals with ample evidence that the CCSQ7.2 is indeed stable across both administration modes and that it can be used safely across both modes throughout the entire employee lifecycle.

This research study was conducted by adopting the scientific knowledge approach. Welman and Kruger (2003:3) indicate that "scientific knowledge is not something that is merely accepted on the authority of society or other sources."

There are three core features of scientific knowledge:

1. Systematic observation

According to Welman and Kruger (2003:5), "the scientific approach requires an investigation to be planned in which we use the results of two groups that have actually applied these methods strictly, as the final arbiter". In this instance, the investigation that was conducted consisted of administering the CCSQ7.2 to two different sample groups in two different administration modes.

2. Control

Welman and Kruger (2003:5) state that "scientific knowledge needs to be obtained in a controlled manner and alternative explanations for the obtained results should be eliminated systematically". In this specific research study the CCSQ7.2 was administered to the first sample group in a managed and proctored setting and the second sample group completed the online version of the CCSQ7.2. Control for the online sample group was exercised by providing each candidate in the sample group with a unique username and password.

3. Replication

Welman and Kruger (2003:6) declare that "the manner in which scientific knowledge is obtained needs to be replicable. Comparable results should be obtained by other

researchers, independent of the original/s, involving other research participants and other circumstances, yet still compatible with the same theory". Other researchers should be able to replicate this specific research study by making use of different sample groups under the same circumstances, that is, the CCSQ7.2 can be administered to different or the same sample group(s) in a proctored as well as an unproctored administration setting.

The primary objective for this research study involves the following:

*Determining the construct equivalence of the Customer Contact Styles Questionnaire (CCSQ7.2) when it is administered in two different administration modes, in a proctored and unproctored setting*

#### **1.4 Significance of the study**

By proving that CCSQ7.2 can be administered in two different administration modes, in a proctored as well as an unproctored setting, one can justify the constant conceptual domain of this specific questionnaire thereby encouraging organisations and individuals to make use of the Internet-delivered platform for this and other questionnaires that are available in an Internet version.

Bartram (in Evers, Anderson & Voskuil, 2005:414) states that "despite the rapid growth of the use of the Internet for testing, especially in the recruitment and selection field, the volume of tests delivered in this way is still far less than that delivered through paper-and-pencil". As a result of this, many providers are finding it difficult to substantiate the costs of developing new forms of testing (Bartram in Evers et al, 2005:414). The results of this study will also aim to change the perceptions and concerns that individuals have about the challenges of Internet-delivered testing and thereby encourage the increased use of this particular mode of administration. The increased use of the Internet for testing will also enable providers to justify the costs of developing new forms of testing leading to progress in the field of occupational assessments. The results of this study will also provide evidence to demonstrate that a lack of supervision does not impact on the equivalence of this questionnaire, consequently encouraging the use of unproctored Internet testing in employment settings.

Foxcroft and Roodt (2005:21) state that, in terms of psychological testing, new measures and test development technologies continue to be developed each year and yet there are still individuals who continue to attack and criticise this discipline despite its advantages

to the employer and the employee. The results of this study will aim to provide organisations and individuals with confirmation that there is constant progress in the field of psychological testing despite these criticisms and that there will continue to be progress, and consequently, new measures will open the way for further progress in the field of psychological testing. As a result, many individuals will come to recognise the value of these new measures, as the Internet will continue to become more prevalent in personnel selection over the next few years.

This study will enable individuals and organisations to recognise that the CCSQ7.2, yet another psychometric instrument, is safe to use in an Internet as well as paper-and-pencil version.

## **1.5 Research question and hypotheses**

The need to conduct this study and to answer the research question below arose as a result of the perceptions and concerns that individuals have about Internet-delivered testing.

### **1.5.1 Research questions**

The reason for conducting this research study is to answer the following research question: *Can the Internet version of the Customer Contact Styles Questionnaire (CCSQ7.2) be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting?*

### **1.5.2 Hypotheses**

The main reason for conducting this research study is to determine the construct equivalence of the CCSQ7.2 when it is administered in two different modes. The following hypothesis was explored:

H1: The Internet version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting.

The null hypothesis of this study can then be stated as:

Ho: The Internet version of the CCSQ7.2 cannot be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting.

#### **1.5.2.1 Operational definitions of the variables within this study**

A brief definition of the operational variables that will be used in this specific research study will be given below. This will ensure that the reader has a comprehensive understanding of the variables in this study.

- Unproctored testing

Tippins et al (2006:193) define unproctored testing as "a testing event that is not monitored by a human test administrator and, as a result, a human proctor does not verify candidate identification nor is candidate behaviour during the testing event observed by a human."

- Proctored testing

Tippins et al (2006:193) define proctored testing as:

A testing event that is monitored by a test administrator who typically verifies the identity of the candidate taking the test and observes the candidate while he/she takes the test to ensure that no outside help is given, no discussion occurs among candidates, no items are copied for later sharing, and so forth.

In a proctored testing environment, test conditions are as near as possible or exactly the same every time an assessment is conducted.

- Psychological assessment

Foxcroft and Roodt (2005:4) define psychological assessment as "a process-orientated activity aimed at gathering a wide range of information by using assessment measures and information from many other sources". One of the key elements of the broader evaluative process of psychological assessment is concerned with the affective or non-

intellectual aspects of behaviour and this key element is referred to as psychological testing (Foxcroft & Roodt, 2005:4).

## 1.6 Chapter overview

Table 1-1 provides a detailed description of the outline and content of the chapters of this research study.

**Table 1-1** Chapter overview

Chapter	Chapter title	Chapter content
1	Introduction and overview	Introduction to the objective, purpose, background and significance of the study.
2	Review of psychological assessment and Internet-delivered testing	Review of the existing literature regarding psychological assessment.  Review of existing literature regarding measurement equivalence.  Previous studies conducted on the CCSQ7.2.  Review of existing literature regarding Internet-delivered testing.
3	Research methodology	The research design and methodology, the sample and sampling methods will be discussed. Details of the measurement instrument, data collection procedure as well as the techniques used to analyse the data. The ethical considerations of the study will be mentioned.
4	Research results and data analysis	Results obtained from the study will be presented and conclusions will be made based on the findings.
5	Discussion and conclusion	Summary of the findings as well as conclusions of the study will be discussed.  Limitations of the study and recommendations for future research will be made.

The following chapter will investigate the current literature that exists on the research topic. The research methodology that was used in this study will then be discussed in detail and the chapter subsequent to that will focus on the results that were obtained from this research study. The findings and conclusions that were gained from the results

will be presented and, lastly, limitations and recommendations for future research will be made.

## **1.7 Conclusion**

The aim of this research study is to establish the construct equivalence of the ipsative version of the CCSQ7.2 when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting. The results of this research study will be able to provide many organisations with adequate evidence that the CCSQ7.2 is indeed stable across administration modes and that it can safely be used across both modes in the selection and/or development process. All the methodology that was adopted by the researcher was used with the purpose of answering the above-stated research question and complete details of the methodology and the findings of the study will follow.

## **CHAPTER 2**

### **Review of psychological assessment and Internet-delivered testing**

#### **2.1 Introduction**

This chapter will give the reader an overview of the existing literature on this topic. Firstly, the concept of psychological assessment will be discussed in order to provide the reader with an understanding of which discipline the Customer Contact Styles Questionnaire (CCSQ7.2) fits into. As the instrument that is being used in this particular study is a personality questionnaire, the advent of the personality questionnaire and its role in occupational assessments will also be discussed. Thereafter, the reliability and validity, as well as the developmental studies that have been conducted on the CCSQ7.2, will be discussed in order to provide the reader with scientific evidence that this particular instrument can be used with confidence throughout the entire employee life cycle. As the CCSQ7.2 is being administered in two different modes, it is imperative for the reader to become familiarised with both modes of administration. Most assessment instruments are administered via a paper-and-pencil mode and this mode of administration has become widely accepted. The advent of Internet-delivered testing has produced many opportunities, challenges and concerns. It is for this particular reason that the Internet as an administration mode, as well as Internet-delivered testing and the challenges, concerns and best practice issues that come along with it, will be discussed in detail.

#### **2.2 Psychological assessment**

In order to provide the reader with a better understanding of which discipline the CCSQ7.2 belongs to, the concept of psychological assessment will be discussed in detail. Psychological testing, which is a key element of psychological assessment, will also be explored.

Foxcroft and Roodt (2005:4) define psychological assessment as "a process-orientated activity aimed at gathering a wide-array of information by using assessment measures (tests) and information from many other sources like interviews and other collateral sources". This information is then evaluated and integrated in order to assist individuals and/or organisations in making their final decision.



Foxcroft and Roodt (2005:4) and Anastasi (1988:17) also mention that one of the key elements of the broader evaluative process of psychological assessment is concerned with the affective or non-intellectual aspects of behaviour and this key element is referred to as psychological testing. Tests designed for this purpose are commonly referred to as personality tests, which measure characteristics such as emotional states, interpersonal relations, motivation, interests and attitudes (Anastasi, 1988:17; Foxcroft & Roodt, 2005:4).

In order to gain a greater understanding of the purpose of psychological tests, Tuckman (in Owen & Taljaard, 1996:12) identifies the general value of psychological tests, and these values can be summed up as follows:

- Objectivity is added to our observations.
- Behaviours are elicited under controlled conditions.
- The behaviour people are capable of may be sampled.
- Progress can be measured based on set standards.
- Insight is gained into aspects that are not visibly observable.
- Characteristics of behaviour may be traced.
- Future behaviour may be predicted.
- Information that will assist decision making may be provided.

Psychological tests provide individuals and organisations with a range of benefits and consequently it is essential to understand what a psychological test is. Murphy and Davidshofer (1991:3) declare that a psychological test is a measurement instrument that has three characteristics:

- A psychological test is a standardised measure of a sample of behaviour.
- The sample of behaviour is obtained under standardised conditions.
- Quantitative information from the sample of behaviour is obtained by means of consistent procedures.

Aiken (1994:12) mentions that psychological testing and assessment are used for many purposes, including evaluating behaviour and other personal qualities that will enable individuals and organisations to make appropriate decisions, judgements and predictions. In particular, psychological tests in conjunction with other sources of information assist individuals and organisations in making selection, development, career guidance and promotional decisions. Table 2-1 gives the different fields where psychological tests are used and their uses within those fields.

**Table 2-1** The uses of psychological testing

<b>Field</b>	<b>Use</b>
Educational testing	Psychological tests are used to aid educational institutions in making admission and advancement decisions.
Personnel testing	Psychological tests are used to aid employers in making selection, development and promotional decisions.
Clinical testing	Psychological tests provide clinical psychologists with data regarding the intellectual performance of individuals.

Adapted from Murphy and Davidshofer (1991:10-11)

According to Murphy and Davidshofer (1991:6), psychological tests fall into three categories:

1. The subject performs a specific task.
2. The subject's behaviour is observed in a particular situation.
3. Self report measure.

The instrument used in this particular study, the CCSQ7.2, falls into the third category. Murphy and Davidshofer (1991:6) describe the psychological tests in this category as "a self-report measure in which the subject describes his/her feelings, attitudes, beliefs, values, opinions or physical or mental state".

The above information has provided the reader with more insight into the field of psychological assessment and specifically psychological testing. At present, the field of psychological testing is in a revolutionary phase. Many employers, globally as well as locally, are making use of personality questionnaires, one of them being the CCSQ7.2, as well as other assessment instruments, to assist them in making occupational selection decisions. As the instrument that is used in this particular study is a personality questionnaire, the subsequent section will focus on the advent of the personality questionnaire and its role in occupational assessments.

## **2.3 Personality questionnaires and their role in occupational assessment**

In any work environment personnel decisions regarding hiring, development, placement, training, performance and promotion need to be made on a daily basis (Murphy & Davidshofer, 1991:331). It is therefore imperative that employers and assessment practitioners make accurate workforce decisions. The decisions they make will have a direct impact on productivity and therefore it is in their own interests for them to have the appropriate information that will help them to predict future behaviour and job performance in order to enable them to make sound workforce decisions.

Murphy and Davidshofer (1991:333) state that psychological tests are able to provide employers and assessment practitioners with information about candidate's skills, abilities, experience and characteristics that will enable them to predict future behaviour and job performance. There are currently many psychological tests that are used in personnel testing to assist employers in obtaining the specific data that they require. Some of these tests may be more appropriate than others, but for the purposes of this study we will be focusing on personality tests and questionnaires as the tool that is used to assist employers and assessment practitioners in making workforce decisions.

Owen and Taljaard (1996:351) state that "personality tests measure personality traits so accurately that specific and reasonably reliable conclusions can be drawn with regard to human functioning". It is for this particular reason that, for the past 15 years, personality measures have been the main focus in personnel selection owing to their ability to predict employee performance thereby aiding employers in making sound workforce decisions (Converse, Oswald, Imus, Hedricks, Roy & Butera, 2008).

The notion that personality questionnaires are able to predict job success and employee performance has not always been supported. Ghiselli, Guion and Gottier, Locke and Hulin, Reilly and Chao, Schmitt, Gooding, Neo and Kirsch (in Barrick & Mount, 1991) have conducted several studies and have thus concluded that the validity of personality as a predictor of job performance is quite low. Barrick and Mount (1991) state that this could be due to the fact that when these studies were conducted a taxonomy for classifying personality traits had not yet existed and it was not possible to determine whether a relationship does in fact exist between particular constructs and performance criterion.

In recent times, this view has changed and many researchers and psychologists have come to the agreement that several personality measures can now be described within

the Big Five factors of personality thereby facilitating meta-analytic cumulation (Barrick & Mount, 1991; Ones & Viswesvaran, 1996). This view has also supported the fact that earlier studies, which concluded that the validity of personality as a predictor of job performance is quite low, are seen as premature owing to the existing validity evidence on this particular topic (Ones & Viswesvaran, 1996). In addition to this, Ones and Viswesvaran (1998:246) state that "the increasing credibility of personality measurement in Industrial/Organisational (I/O) Psychology is the result of large-scale studies and meta-analyses reporting substantial validities for theoretically relevant personality constructs for various criteria". Table 2-2 provides the reader with a description of each of the Big Five factors of personality.

**Table 2-2** The Big Five factors of personality

<b>Dimension</b>	<b>Traits associated with dimension</b>
Extraversion (or Surgency)	Being sociable, gregarious, assertive, talkative and active.
Emotional Stability (or Neuroticism)	Being anxious, depressed, angry, embarrassed, emotional, worried and insecure.
Agreeableness (or Likeability)	Being courteous, flexible, trusting, good-natured, cooperative, forgiving, soft-hearted and tolerant.
Conscientiousness (or Conscience)	Being careful, thorough, responsible, organised, hardworking, achievement-orientated and persevering.
Openness to Experience	Being imaginative, cultured, curious, original, broad-minded, intelligent and artistically sensitive.

Adapted from Barrick and Mount (1996)

The notion that personality questionnaires are able to predict job success and employee performance has been accepted as a result of a landmark paper of Barrick and Mount (1991), who investigated the relation of the Big Five personality dimensions to three job performance criteria for five occupational groups. These job criteria consisted of job proficiency, training proficiency and personnel data and the five occupational groups consisted of professionals, police, managers, sales staff and skilled-semi skilled groups (Barrick & Mount, 1996). The findings of this study revealed that conscientiousness showed reliable relationships with all the job performance criteria for all five occupational

groups (Barrick & Mount, 1996). For the other personality dimensions, estimated true score correlations varied by occupational group and criterion type, and it can thus be concluded that the results of the study reveal the benefits of using the Big Five factor model to accumulate and communicate empirical findings (Barrick & Mount, 1996).

The meta-analysis conducted by Barrick and Mount (1991) opened many doors for future studies and, consequently, Hough et al, Ones et al, Tett, Jackson, Rothstein (in Ones & Viswesvaran, 1996:610) have stated that “meta-analysis in the personality domain has demonstrated that a construct-oriented approach uncovers meaningful relations between personality variables and job performance”.

In conclusion, Ones and Viswesvaran (1996:610) state that “human resource practitioners and researchers are now far more optimistic about the potential of personality variables in personnel selection.” Many organisations assess personality before making any workforce decisions, as personality assessments offer organisations and assessment practitioners many benefits. These benefits are depicted in Table 2-3.

**Table 2-3** Benefits of personality assessment

Provide additional information to support other selection techniques	Personality assessments are able to measure areas that cannot be measured by other selection techniques and they are able to determine how a candidate is likely to perform in a working environment. In addition to this, they provide employers with additional information, which can be used during an interview.
Provide valuable information	Personality assessments can provide organisations with information on the strengths and developmental areas of new recruits.
Cost and time saving	Substantial cost and time savings are made as unsuitable candidates are identified early on in the process.

Adapted from SHL Global Website (accessed on 2008/08/19)

## **2.4 Previous research conducted on the CCSQ7.2**

As SHL is the sole distributor of the CCSQ7.2, it has conducted numerous validity, reliability and developmental studies, internationally as well as nationally, in order to ensure that the CCSQ7.2 can be used throughout the employee life cycle. It is imperative that these studies are carried out and documented, as South African legislation, and more specifically, the Employment Equity Act, 1998 states that:

Psychometric testing and other similar assessments of an employee are prohibited unless the test or assessment being used

- has been scientifically proven to be valid and reliable
- can be applied fairly to all employees and
- is not biased against any employee or group.

In addition to this, Baron and Miles (2002) state that "it is essential to ensure that the tools used to make decisions are fair both because they have a huge impact on a candidate's life and because employers are wary of legal challenges". It is for this particular reason that the following section will focus on the validity, reliability and development studies that have been conducted on the CCSQ7.2.

SHL has conducted a number of studies to establish the concurrent validity of the CCSQ7.2. These studies will now be discussed in detail.

The first concurrent validity study that was carried out involved the testing and collection of job performance data on incumbent retail sales staff. The organisation, Electro-Energy, that was involved in this particular study aimed to identify the characteristics of their top performers to aid them with future recruitment. The sample group consisted of 78 retail sales staff where the majority, 61%, of the sample consisted of females. The mean age of the group was 35 years and the majority of the sample group, 91%, described themselves as white Europeans. The criteria that were used involved assessing job performance on 16 dimensions using the CCSQ7.2. Job incumbents were also given a numerical value, which represented their sale performance from June to November. The findings of the study indicated that the following competencies namely, results orientated, structured, innovative, competitive and flexible, were the aspects of personality that correlated across a number of competency areas as well as a hard measure of sales performance. The findings of the study also showed that sales

performance correlated with the results-driven competency. This meant that the HR team should recruit candidates with the above-mentioned competencies and in this way they are likely to recruit candidates with the right attributes that would enable them to attain good sales results (SHL, 1997).

In the financial industry, SHL conducted another concurrent validity study, which involved RTC Financial Planning, one of the UK's leading home insurance companies. Owing to its expanding workforce and the high turnover that came with it, the HR department aimed to review its recruitment and development process. The HR department identified the following competencies which were essential for the success of a financial planner: convincing, communicating orally, fact finding, organisation, reliability, customer focus, resilience and, lastly, results driven. The sample group consisted of 196 sales staff where the majority of the sample, 86%, was male. The mean age of the group was 38 (SD=11.4) years and the majority of the sample group, 86%, described themselves as white Europeans. The criteria that were used involved assessing job performance on 16 dimensions using the CCSQ7.2. Job incumbents were also given a numerical value, which represented their sales performance. A multiple regression analysis was also used to predict sales performance against targets from the CCSQ7.2. The findings of the study revealed that the following competencies, conscientious, self-control, detail conscious and analytical, were crucial for the success of the financial planner's role. In addition to this the organisation identified a "danger zone profile" which would assist them in the interviewing process. The results from the multiple regression analysis revealed that those job incumbents with the identified pattern of style scores achieved almost double the proportion of their targets, that is, those job incumbents that achieved a high score on the CCSQ7.2 achieved an average of 107% of their sales target (SHL, 1997).

In order to ensure the reliability of the CCSQ7.2, SHL has carried out many reliability studies within the South African environment. Each study has a particular study number and most of these studies have been conducted in South African financial institutions. The following reliability studies will be discussed:

- Study number R011 (February 2000) SA Financial Institution
- Study number R013 (February 2000) SA Financial Institution
- Study number R014 (February 2000) SA Financial Institution
- Study number R015 (April 2000) SA Financial Institution
- Study number R016 (April 2000)
- Study number R025 (November 2000) SA Financial Institution

- Study number R027 (November 2000) SA Financial Institution
- Study number R035 (June 2001) SA Financial Institution

With regards to study number R011, the sample group consisted of 371 employees of which 58% were female and 42% were male. The mean age of the group was 32.03 (SD=8.86) years. The ethnic distribution of the sample consisted of the majority, 41.7%, being coloureds. The majority of the sample group, 60.4%, had Grade 12 as their highest qualification. The findings of the study indicated that the mean scores followed the same pattern as those of international studies and the mean scores differed with less than one standard deviation from one another indicating that these scores compare well with the findings of other international studies. The alpha coefficient scores ranged from 0.68 to 0.88 and it can thus be said that the CCSQ7.2 can be used with confidence among the South African population (SHL, 2002).

The next reliability study, study number, R013, consisted of 641 applicants. This study yielded similar results to those of study number, R011, where the mean scores differed with less than one standard deviation from one another indicating that these scores compare well with the findings of other international studies. The alpha coefficients ranged from 0.75 to 0.91 indicating that the coefficients obtained for this study compare well with coefficients that were obtained by international research. It can thus be concluded that the CCSQ7.2 can be used with confidence among the South African population (SHL, 2002).

Study number R014 involved 1305 employees of which 76.90% were female and 23.1% were male. These employees were also assessed as part of a restructuring process to determine their person-job match. The mean age of the group was 32.76 (SD=8.98) years. The ethnic distribution of the sample consisted of the majority, 40.5%, being whites. The majority of the sample group, 61.5%, had Grade 12 as their highest qualification. The alpha coefficients ranged from 0.76 to 0.90 indicating that the CCSQ7.2 can be used with confidence among the South African population (SHL, 2002).

The next reliability study, study number R015, consisted of 1079 employees. As in study number R014, these employees were also assessed as part of a restructuring process to determine their person-job match. The sample group consisted of 76.90% females and 23.1% males. The mean age of the group was 32.76 (SD=8.98) years. The ethnic distribution of the sample consisted of 49.12% whites and 50.88% blacks. The majority of the sample group, 61.5%, had a Grade 12 as their highest qualification. The findings of this study reveal that the overall mean scores for the two ethnic groups follow the



same pattern and they compare well, with the mean scores differing with less than one standard deviation from one another. The alpha coefficients ranged from 0.72 to 0.91 indicating that the CCSQ7.2 can be used with confidence among the different South African ethnic groups (SHL, 2002).

Study number R016 consisted of 2397 employees as well as applicants. The findings that were reported for this particular study revealed similar results as those of the reliability studies that were mentioned above. The alpha coefficients ranged from 0.74 to 0.90 and the mean scores obtained in this study compare well with those that were obtained by international research. As a result, it can be once again said with confidence that the CCSQ7.2 can be used among the South African population (SHL, 2002).

With regards to study number R025; the sample group consisted of 737 employees. As in study numbers R014 and R015, these employees were also assessed as part of a restructuring process to determine their person-job match. There were also employees in the sample group who applied for particular positions within the company. The sample group consisted of 79.35% females and 20.65% males. The mean age of the group was 32.56 (SD=8.35) years. The ethnic distribution of the sample consisted of 45.79% whites and 54.21% blacks. The majority of the sample group, 61.5%, had Grade 12 as their highest qualification. As in previous reliability studies conducted on the CCSQ7.2, this particular study reported similar results stating that there were no differences exceeding more than one standard deviation between any of the groups and these results are similar to the findings of other studies that were conducted on personality differences between groups (SHL, 2002).

The second last reliability study that was carried out by SHL, study number R027, involved 737 employees and, as in previous studies, these employees were also assessed as part of a restructuring process to determine their person-job match. There were also employees in the sample group who applied for particular positions within the company. The sample group consisted of 79.35% females and 20.65% males. The mean age of the group was 32.56 (SD=8.35) years. The ethnic distribution of the sample consisted of 45.79% white as the majority of the sample group. This particular study reported results showing that there were no differences exceeding more than one standard deviation between any of the groups and these results were similar to the findings of other studies that were conducted on personality differences between groups (SHL, 2002).

The last reliability study that was carried out by SHL, study number R035, involved 2405 employees and, as in study numbers R014, R015, R025 and R027, these employees were

assessed as part of a restructuring process to determine their person-job match. The sample group consisted of 79.88% females and 20.12% males. The ethnic distribution of the sample showed the majority, 49.02%, as being white. The majority of the sample group, 64.12%, had Grade 12 as their highest qualification. The findings of the study revealed that the overall mean scores followed the same pattern as the mean scores obtained in international studies (SHL, 2002). The alpha coefficients ranged from 0.76 to 0.90 (SHL, 2002).

There are also various development studies that have been carried out by SHL on the CCSQ7.2. These studies endeavour to ensure the suitability of the CCSQ7.2 in the South African environment. These development studies will now be discussed in detail.

One particular development study, study number D005, was carried out on the CCSQ7.2. This study was carried out with the intention of determining the mean differences between gender groups on the CCSQ7.2. The sample group consisted of 1310 employees in a South African financial institution where 77.02% of the sample consisted of females and 22.98% of males. The mean age of the group was 32.77 (SD=8.99) years. These participants were assessed as part of a restructuring process to determine their person-job match. The ethnic distribution of the candidates consisted of the majority, 40.69%, being white. Most participants in the sample group had Grade 12 as their highest qualification. The findings of the study indicated that there were no practically significant differences between the two groups. This is similar to findings that were reported on other international studies. These studies indicated that there are small differences between groups on personality scales.

During the development of the CCSQ7.2, certain information was collected on the participants who were completing the questionnaire in order for the data to be analysed by gender as well as age (SHL, 1997). Table 2-4 depicts the analysis that was conducted on the CCSQ7.2 by gender.

**Table 2-4** Analysis of the CCSQ7.2 by gender

Scale	Male (n=353)				Female (n=195)			
	Mean	sd	Alpha	SEm	Mean	sd	Alpha	SEm
Persuasive	35.6	6.4	0.79	2.9	33.5	6.4	0.78	3.0
Self-control	42.4	9.2	0.88	3.2	44.0	8.1	0.86	3.1
Empathic	46.0	7.2	0.83	2.9	49.4	6.6	0.85	2.6
Modest	37.8	9.0	0.86	3.3	37.3	9.0	0.87	3.2
Participative	44.2	9.7	0.89	3.2	46.5	10.9	0.92	3.1
Sociable	38.7	7.6	0.80	3.4	41.2	7.1	0.74	3.6
Analytical	39.7	6.7	0.79	3.1	40.0	7.0	0.81	3.1
Innovative	36.7	9.7	0.91	2.9	38.9	9.0	0.89	3.0
Flexible	35.2	5.5	0.79	2.5	36.6	5.4	0.80	2.4
Structured	37.5	7.7	0.85	3.0	39.2	7.6	0.85	3.0
Details Conscious	33.2	7.2	0.84	2.9	35.9	7.0	0.84	2.8
Conscientious	36.1	6.7	0.86	2.5	39.4	5.5	0.85	2.1
Resilience	37.5	8.6	0.82	3.7	36.2	7.9	0.80	3.6
Competitive	30.0	8.2	0.83	3.4	26.4	7.3	0.78	3.4
Results orientated	35.4	6.7	0.81	2.9	38.1	5.4	0.73	2.8
Energetic	33.5	7.1	0.87	2.6	33.4	6.5	0.86	2.4
Consistency	59.5	5.6	0.58	3.6	60.4	4.4	0.35	3.5

Adapted from CCSQ Manual & User's Guide, 1997

The above results reveal that the reliabilities for the scales of the two groups are very similar and consequently it can be concluded that the instrument is performing in the same manner for both gender groups (SHL, 1997).

The next analysis was carried out by age. Table 2-5 depicts the analysis that was conducted on the CCSQ7.2 by age.

**Table 2-5** Analysis of the CCSQ7.2 by age

<b>N=531</b>			
<b>Scale</b>	<b>Mean</b>	<b>sd</b>	<b>Correlation with Age</b>
Persuasive	35.2	6.5	0.13**
Self-control	42.3	8.7	-0.3
Empathic	46.8	7.2	-0.00
Modest	37.6	9.3	0.07
Participative	45.0	10.1	-.017**
Sociable	39.7	7.4	-0.11*
Analytical	39.8	6.8	0.10*
Innovative	39.4	9.5	0.04
Flexible	35.5	5.7	-.017*
Structured	37.5	8.0	0.03
Details Conscious	33.7	7.3	-0.01
Conscientious	36.7	6.0	-0.01
Resilience	37.0	8.4	-0.02
Competitive	29.1	8.0	0.05
Results orientated	36.2	6.2	-0.12**
Energetic	33.5	7.1	-0.09
Consistency	59.6	5.1	-0.04
*Correlation is significant at the 1% level (two tailed)			
** Correlation is significant at the 5% level (two tailed)			

Adapted from CCSQ Manual &amp; User's Guide, 1997

The ages in the sample group ranged from 18 to 64 with a median value of 31, a mean age of 33 and a standard deviation of 10 (SHL, 1997). The results reveal that the correlations between scale scores and age are small and do not reach significance (SHL,

1997). It should also be noted that when this analysis is interpreted the fact that the results are cross-sectional should be taken into account (SHL, 1997).

In 1997, Cartwright and Tidswell also reviewed the CCSQ7.2 independently. These reviewers evaluated the validity and reliability of the CCSQ7.2. Their findings revealed that the CCSQ7.2 has a high level of face validity as it is based on a job analysis of the roles for which the instrument is anticipated to be used and the content validity of the instrument seems to be good. These reviewers also cited the fact that no predictive validity studies have been carried out on the CCSQ7.2. Owing to the lack of predictive validity studies, it would be beneficial for the test publisher to explore this area further in order to ensure this instrument's predictive validity. The reviewers also indicated that, in terms of reliability, there is good evidence of internal consistency within the CCSQ7.2. There is a lack of evidence of the instrument's stability over time and such evidence would be very valuable. The reviewers concluded that the CCSQ7.2 seems to have reliability that is equivalent to the reliability of many other personality measures.

The above section has provided the reader with sufficient evidence of the reliability and validity of the CCSQ7.2 and thus it can be said with confidence that the CCSQ7.2 can be used throughout the entire employee life cycle. In this particular study the CCSQ7.2 is administered in two modes of administration and the next section will focus on the Internet as an administration mode.

## **2.5 Internet-delivered testing**

As this study involves determining the construct equivalence of the CCSQ7.2 when it is administered in two different administration modes, it is essential that the reader gains a greater understanding of both modes of administration. The paper-and-pencil mode of administration has been acknowledged and accepted and consequently the next section will focus on the Internet as an administration mode.

There are still many organisations and assessment practitioners that are administering assessment instruments in a paper-and-pencil version even though the same instrument is available online. This could be because of the belief that the instrument will measure a different construct when it is administered in a different mode and it is for this particular reason that the concept of measurement equivalence, and the importance of establishing and documenting it, will be discussed in detail.

The advent of the Internet and its role in psychometric testing has brought with it many opportunities, challenges and concerns. Bartram (2001:32) claims that "there will be a revolution in how occupational testing is carried out and consequently the Internet rather than the traditional paper-and-pencil mode of administration will become the medium of choice for testing over the next decade". Salgado and Moscoso (2003) declare that the next ten years will be seen as the decade of the Internet in personnel selection. Internet-delivered testing, as well as the opportunities, concerns and best practice issues that accompany it, will now be discussed in more detail.

In order for the reader to become familiarised with the concept of Internet testing, the term needs to be clearly defined. Harris (in Bartram & Hambleton, 2006:116) defines Internet-based testing as "a procedure that uses the Internet or Intranet for administering a test". Internet testing can be proctored or unproctored, but for the purposes of this study, the researcher will be making reference to unproctored Internet testing.

Buchanan and Smith (1999) state that there are two main ways in which Internet-delivered tests differ from their paper-and-pencil counterparts:

1. The format of presentation
2. The nature of the participants and the conditions under which the test is expected to be taken

During a paper-and-pencil assessment session candidates are provided with a standard test booklet and answer sheet. The format of the online version of the same questionnaire will differ as candidates access the assessment via a computer screen. In addition to this, candidates will not undertake the assessment under the same conditions as they would have if they had been completing the assessment in a paper-and-pencil format. Candidates could be sitting in various geographical locations and they could be completing the test at the location of their choice, which could be at the office or even in the comfort of their own home.

Bartram (2001) is of the opinion that, at present, we are witnessing the widespread use of the Internet as a generic communication medium and there are currently more and more individuals who are gaining access to the World Wide Web. This had led to a general increased acceptance of computers, as they are becoming a part of many people's everyday lives (Bartram, 2001). Table 2-6 depicts the most recent statistics of Internet users in seven geographic regions.

**Table 2-6** Internet users in the world by geographic region

<b>Geographic region</b>	<b>Users (in millions)</b>
Asia	578.5
Europe	384.6
North America	248.2
Latin America/Caribbean	139.0
Africa	51.1
Middle East	41.9
Oceania/Australia	20.2

Adapted from Internet World Stats (accessed 19/08/2008)

Kriek (2005) reports on a study that was conducted by SHL in 2005 where a global survey was sent out to 42 country managers of the SHL group to determine the acceptance of Internet testing in their specific country. There was a 45.6% response rate and 19 completed questionnaires were received. The results reveal that Internet testing is well accepted in the following countries:

- Australia
- Hong Kong
- Ireland
- Netherlands
- New Zealand
- Portugal
- Sweden
- Switzerland
- USA

Owing to the convenience and practicability of Internet testing, many individuals and organisations have turned to integrating Internet-delivered testing into their recruitment and selection process. Consequently, many test publishers have taken or are considering taking advantage of the benefits of Internet-delivered testing by converting their current selection procedures to an Internet format. A test publisher cannot simply convert an assessment instrument that is available in a paper-and-pencil mode to an Internet version and thereby assume that both versions are equivalent. Buchanan and Smith (in Salgado & Moscoso, 2003) state that for equivalence to be confirmed each questionnaire should be examined individually. The next section will focus on the concept of measurement equivalence.

### 2.5.1 Measurement equivalence

Bartram (in Evers et al, 2005:402) states that “in relation to selection, the Internet has been used more for personality testing than for ability testing”. Despite the extensive use of online personality questionnaires, there is currently very little literature documenting the equivalence of the online version of these questionnaires and their paper-and-pencil counterparts (Bartram in Evers et al, 2005:402). This view is supported by Potosky and Bobko (2004), who state that there are few studies that have examined the equivalence of Internet-administered tests and their traditional counterparts.

When a researcher and/or test publisher aims to convert a paper-and-pencil instrument into an Internet format, the equivalence of the different versions of the same instrument needs to be examined and documented (Cronbach, 1990; King & Miles, 1995; Buchanan & Smith, 1999; Ployhart et al, 2003).

The International Test Commission (2005) has drafted specific equivalence guidelines that test developers need to adhere to when an Internet test has been developed from a paper-and-pencil version. Clear evidence of the equivalence between the two versions needs to be provided specifically to show that the two versions:

- have comparable reliabilities
- correlate with each other at the expected level from the reliability estimates
- correlate comparably with other tests and external criteria
- produce comparable means and standard deviations or have been appropriately calibrated to render comparable scores.

There are certain criteria that are used in order to establish measurement equivalence. Van de Vijver and Poortinga (in Brown, 2007) state that there are three types of equivalence that can be used, specifically:

1. *Structural/functional equivalence*: The same psychological constructs across groups are obtained demonstrating that patterns of correlations between variables are the same across groups.
2. *Measurement unit equivalence*: The same measurement unit is obtained and individual differences can be compared for both sample groups.
3. *Scalar/full score equivalence*: The same measurement unit and the same origin is obtained thereby allowing scores to be compared across groups.



For the purposes of this study, structural/functional equivalence will be used to determine whether the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties.

There are several equivalence studies that have been conducted by Bartram and Brown (2004), Holtzhausen (2005), Mylonas and Carstairs (2003) and Salgado and Moscoso (2003) on the OPQ32i, OPQ32n, MQ and IP/5F respectively, which reveal that all these instruments can be administered in two different modes without loss of psychometric property. These studies will now be discussed in more detail.

A study conducted by Bartram and Brown (2004) on the ipsative version of the Occupational Personality Questionnaire (OPQ32i) revealed that the instrument is not affected by changes in administration mode. The analysis involved comparing unsupervised web-based administration samples with paper-and-pencil supervised samples for three pairs of matching samples. An overall average alpha score of 0.78 was obtained for the web-based data and an overall average alpha score of 0.80 was obtained for the paper-and-pencil sample group. The mean Standard Error of Measurement obtained for the paper-and-pencil sample group is 2.12 and the mean Standard Error of Measurement obtained for the online sample is 2.03. Effect sizes for the three pairs of matching samples were also computed. The weighted averages for the three pairs of sample groups were no larger than  $d=0.27$ . No differences in scale means, reliabilities or scale intercorrelations were found. A comparison of covariance structures was carried out using Structural Equation Modelling with AMOS and the model that was tested obtained a CFI of 0.949, 0.951, 0.934 and the RMSEA was 0.031, 0.030 and 0.034. These results are indicative of an exceptionally good fit indicating that the instrument is not affected by changes in administration mode.

A similar study conducted by Holtzhausen (2005) on the normative version of the Occupational Personality Questionnaire (OPQ32n) revealed similar results. Alpha coefficients for the supervised paper-and-pencil sample group ranged between 0.69 and 0.88 and Alpha coefficients for the controlled web-based sample group ranged between 0.71 and 0.91. The average Alpha coefficient of 0.80 is the same for both sample groups. The mean Standard Error of Measurement obtained for the paper-and-pencil sample group is 2.05 and the mean Standard Error of Measurement obtained for the online sample is 1.99. Effect sizes for both groups were also computed. The effect sizes ( $d$ ) range from 0.01 to 0.34 indicating that there are only small differences in the scores of the two samples. A comparison of covariance structures was carried out using Structural Equation Modelling with EQS and the model that was tested obtained a CFI of 0.961 and

the RMSEA was 0.026. These results are indicative of an exceptionally good fit indicating that the instrument is not affected by changes in administration mode.

Mylonas and Carstairs (2003) also conducted an equivalence study on the Motivation Questionnaire (MQ). The Expert computerised version of the MQ was administered to the first sample group in a proctored environment and the second sample group completed the Internet version of the MQ in an unproctored environment. The data was analysed using a series of two-way ANOVAs with repeated measures on one factor. Effect sizes for both groups were also computed. The results revealed that a small effect size was obtained for 14 of the 18 scales with the largest effect size being 0.327. A correlation analysis was also carried out and an average correlation of 0.79 was obtained. The Alpha coefficient was also computed and a mean coefficient of 0.75 was obtained. The analysis of the data supports the equivalence of the two administration modes.

Salgado and Moscoso (2003) conducted an equivalence study on a Big Five personality questionnaire (IP/5F). The instrument was administered to the same sample group in a paper-and-pencil and Internet-based version. The results of the study show that both the mean and standard deviation for both sample groups on all five personality dimensions are similar, although a slightly larger standard deviation was obtained for the Internet-based version. Alpha coefficients ranging from 0.71 to 0.89 were obtained for both versions on all personality dimensions. Coefficients of equivalence and stability (CES) obtained for the Big Five personality dimensions were 0.90 for Emotional Stability, 0.91 for Extroversion, 0.86 for Openness, 0.81 for Agreeableness and 0.85 for Conscientiousness therefore supporting the equivalence of the two versions of this questionnaire. The results of an exploratory factor analysis that was carried out revealed that the factor structure is identical for both versions thereby concluding that the magnitude of factor loadings is also very similar. Five types of evidence, namely, descriptive statistics, reliability correlations, factor analysis and congruence were used in this study and indicate that both versions can be used without loss of psychometric property.

There are currently many individuals and organisations that are making use of the Internet as a mode of administration, thereby benefiting from the advantages that it offers. Although Internet-delivered testing offers individuals and organisations many advantages, it does bring along with it some concerns and challenges that still remain unresolved. The next section will focus on the advantages, concerns and challenges of Internet-delivered testing.

### **2.5.2 Advantages of Internet-delivered testing**

Internet-delivered testing brings with it many advantages and benefits, which are not present in the traditional paper-and-pencil mode of administration. These advantages will now be discussed in detail.

- Security

Bartram (in Bartram & Hambleton, 2006:26) points out that the main aspect of the Internet is that the application software as well as all the assessment data resides on a server and not in the public domain which also makes it easier to manage. In addition to that, the intellectual property that is associated with a test in terms of scoring, norms and, amongst others, report generation remains under the strict control of the distributor (Bartram in Bartram & Hambleton, 2006:26). As the distributor has strict control of all the data that is available on the server, this enables updates to be made on a regular basis, which everyone has access to and control of product usage can be monitored and exercised more closely (Bartram in Bartram & Hambleton, 2006:27).

- Resource and geographical constraints

Tippins et al (2006) indicates that an Internet-delivered test can be administered anywhere where the candidate has access to an Internet connection. The candidate has the option to undertake the test at the convenience of his/her own home, at the office or even in a library. Candidates also have the luxury of undertaking the test in their own time. This takes considerable pressure off the employer, as computers do not need to be made available and venues do not need to be booked. Pressure is also taken off the candidate, as he/she does not have to travel to a venue that has been stipulated by the employer. In addition to this, candidates can be assessed in various locations, which enables employers to assess candidates in various countries and even on different continents. Employers make substantial savings in the event where a candidate needs to be assessed in a different country, as the organisation does not have to pay for travel costs.

- Practice tests

Bartram (in Bartram & Hambleton, 2006:19) states that candidates undertaking an online assessment, or any assessment for that matter, should be provided with the opportunity to practise and prepare. In the past candidates were not afforded this

opportunity and only recently have test distributors provided candidates with practice tests and/or practice leaflets which provide candidates with examples of test items (Bartram in Bartram & Hambleton, 2006:19). The Internet makes it possible for test publishers to provide candidates with the opportunity to prepare themselves accordingly by making these practice tests public domain and thereby enabling candidates to access these tests anytime and anywhere. The SHL Direct ([www.shldirect.com](http://www.shldirect.com)) site provides candidates with sample items for verbal, numeric and personality tests and candidates can complete timed practice tests and they are also provided with immediate feedback (Bartram in Bartram & Hambleton, 2006:19).

- The candidate experience

Although Internet-delivered testing is mainly perceived as providing assessment practitioners and organisations with many advantages, it is also important to take note of the advantages and perceptions that candidates experience with online assessments. The main concern of using the Internet as an administration medium is that candidates who do not make use of a computer on a regular basis may experience anxiety with regard to the computer as well as the assessment instrument.

A study conducted by SHL in 2005 shows that people prefer the convenience and comfort afforded by online testing (SHL, 2006). This study surveyed 1300 participants, ranging from those having little or no formal education to those with postgraduate qualification, with ages ranging from 18 through to 70. The results show that 63% of the participants preferred to complete their assessments online (SHL, 2006).

Anderson (in Salgado & Moscoso, 2003) also concludes that positive applicant reactions have been reported universally for Internet-based testing batteries. In a study conducted by Dowdeswell (2006), a survey questionnaire was administered to applicants after the completion of an online assessment to determine their perceptions of the online assessment. The findings of the study reveal that the majority of the respondents were in favour of the online assessment (Dowdeswell, 2006).

Mead (in Evers et al, 2005:405) reported that users experience a high degree of satisfaction with the online administration of the 16PF and this is due to the advantage of remote administration as well as speedy feedback. According to a study reported by Reynolds et al (in Evers et al, 2005:406), applicants experience more positive attitudes towards Internet testing than traditional testing. Bartram and Brown (in Evers et al, 2005:406) also reported on a study where 54 applicants completed feedback forms after

they had completed the OPQ32i online and the evaluations indicate that most applicants would prefer to complete such inventories online despite the technical difficulties that could arise with the Internet.

Table 2-7 provides assessment practitioners and organisations who are making use of Internet-based testing with some suggestions for improving candidate reactions to Internet-based testing.

**Table 2-7** Improving candidate reactions to Internet-based testing

Clarify the purpose of the testing	Candidates need to be informed of the reasons for taking the test and information should be provided on why their data is being gathered.
Data information	Candidates need to be provided with information on their rights to the data, where it is being stored and the individuals who have access to it.
Practice tests	Candidates should be provided with information that will aid them in preparing for the test.
Contact person	Candidates need to be notified of who they can contact should they encounter any difficulties during the test session.

Adapted from Harris (in Bartram & Hambleton, 2006:129-130)

- Reducing time to hire

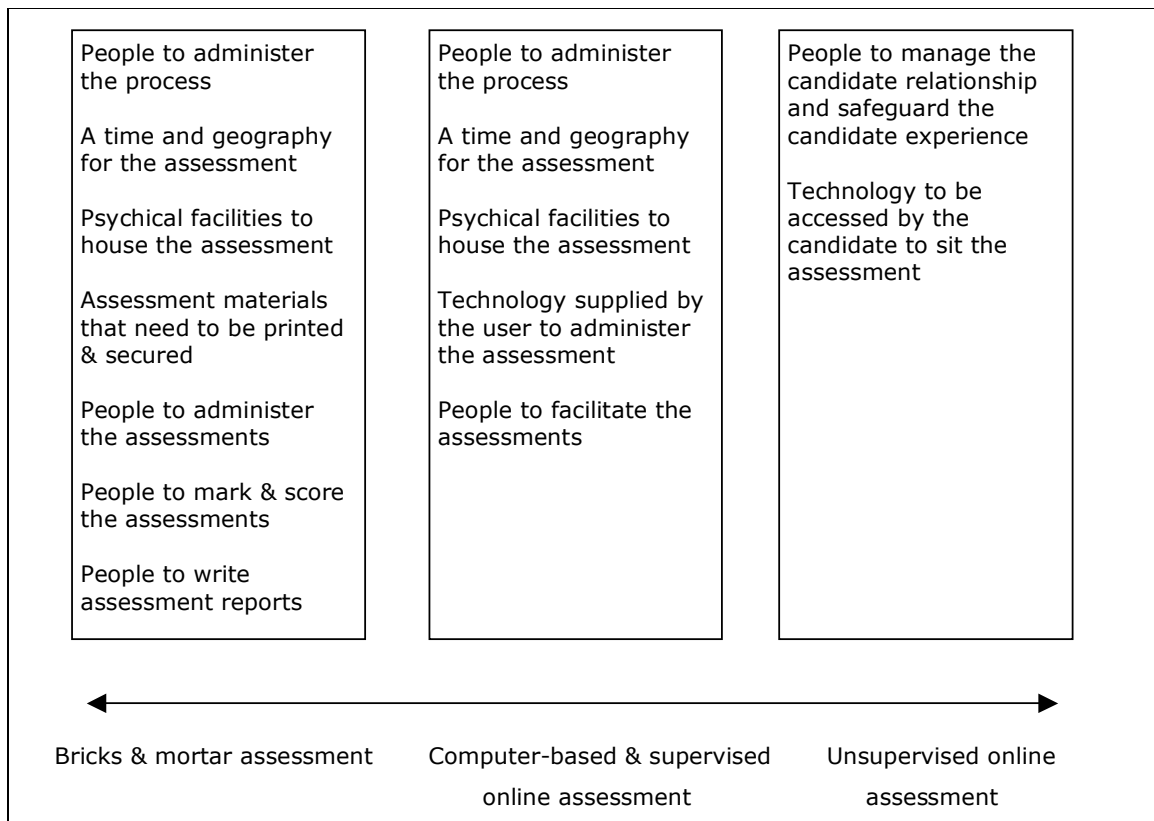
Gibson in (Tippins et al, 2006) and Bartram (in Bartram & Hambleton, 2006:20) state that one of the key features of unproctored Internet testing is enabling employers to pre-screen candidates more effectively, thereby enabling them to reduce their time to hire. Candidates who are dispersed in various geographical locations can be pre-screened before being asked to attend an interview. Employers can make substantial savings during the pre-screening phase of the recruitment process as candidates in various locations who are unsuitable for the specific job do not have to be asked to attend an interview at the employer's expense. In addition to this, the recruitment cycle time is significantly shortened as employers do not need to make logistical arrangements with candidates to attend an assessment session.

- Immediate scoring and lowering of costs

As mentioned previously, one of the main features of the Internet is that assessment data resides on the distributor's server thereby making provision for immediate scoring (Kriek, 2005; Bartram in Bartram & Hambleton, 2006:26). The employer receives the report in a shorter turnaround time thereby taking full advantage of shortening the recruitment cycle time (Gibson in Tippins et al, 2006). Bartram (in Evans, 2002) states that Internet-delivered testing also holds benefits for smaller employers as they are able to pre-screen candidates more effectively by having access to highly sophisticated assessment procedures without having to make enormous investment costs.

Figure 2-1 below clearly illustrates the benefits of unsupervised online assessment. This form of assessment reduces the time, resource and geographical constraints that are present in the traditional paper-and-pencil, as well as computer-based and supervised online assessment, modes of administration.

**Figure 2-1** The evolution of assessment



Adapted from SHL (2006)

### **2.5.3 Concerns and challenges of Internet-delivered-testing**

Although Internet-delivered testing offers many advantages, it also brings along with it a small number of concerns and challenges. These concerns and challenges will now be discussed in greater detail.

- Performance

Bartram (in Bartram & Hambleton, 2006:25) states that "it is in the area of performance that the major limitations of the Internet are to be found". According to Bartram (2001), and Bartram (in Bartram & Hambleton, 2006:25), the Internet, as delivery medium, needs to meet two requirements: firstly, the timing of delivery needs to be controlled and, secondly, it should be robust and not fail halfway through a test. These two requirements are not always met when a candidate is completing an Internet-delivered test. This could be due to a number of reasons. The candidate's server capacity and the bandwidth of the server connection to the Internet may not be big enough (Bartram, 2001; Kriek 2005). In addition to this, the candidate may be accessing the server via a slow connection; a bottleneck could occur or a test session could be terminated owing to a lost Internet connection (Bartram, 2001; Bartram in Bartram & Hambleton 2006:25). These performance issues cannot always be managed and a major drawback occurs in those parts of the world where the Internet is in relatively early stages of development (Bartram, 2001). Bartram (2001) states that reference is being made here to those parts of the world outside North America, Europe and Asia-Pacific, as such countries currently do not have widespread availability of the Internet.

One suggestion for overcoming the above-mentioned problems is to download the crucial material as an applet and thereby reducing the reliance on the Internet (Bartram, 2001). In order to minimise the aforementioned performance issues, the requirements for completing the Internet-delivered test need to be specified by the test distributor. SHL aims to combat these performance issues by sending candidates a document containing technical guidelines and an FAQ (frequently asked questions) document that outlines any technical questions that candidates may possibly have with regards to an online assessment. These documents will provide candidates with support by ensuring that they have the appropriate software installed to enable them to complete the online assessment. Bureau operators are also available during working hours in the event that a candidate needs assistance with any technical difficulties that arise during an online assessment.

- Consistency of appearance

There are currently two Internet browsers that are used by most individuals and organisations when accessing the World Wide Web; they are Internet Explorer and Netscape Navigator. According to Bartram (2001) these browsers do not display information in a similar way, thereby resulting in different candidate experiences of the same test. An additional issue, which the test distributor has no control over, is the screen size and resolution of each candidate's computer (Bartram, 2001). Segall (in Tippins et al, 2006) and Buchanan and Smith (1999) also state that each candidate who will be completing the same test might complete it using a variety of hardware including different monitor sizes and other input devices. Downloading the test within an applet is recommended for resolving the above-mentioned issues (Bartram, 2001).

- Cheating/lack of supervision

The advent of Internet-delivered testing has brought with it an abandonment of the traditional requirement for local supervision thereby increasing the issues of candidate authentication and giving candidates an opportunity to cheat (Bartram & Brown, 2004; Tippins et al, 2006). Gibson (in Tippins et al, 2006:195) claims that "candidate authentication in most common applications of Internet-delivered testing will remain an intractable problem". Beaty and Shepherd (in Tippins et al, 2006) state that when it comes to unproctored Internet testing, cheating will always be a concern as candidates have effectively cheated on non-Internet tests (paper-and-pencil tests) and will do so on Internet tests and it is for this reason that test distributors need to focus on trying to minimise the effects of cheating. One way of minimising the act of cheating is by conducting the assessment via a managed mode and thereby providing test takers with a unique username and password (Bartram, 2001). Another way of minimising cheating is by making use of an honesty contract (SHL, 2006). In addition to this, an organisation can inform the candidate that they will have to go through an authentication procedure and/or that they will have to undertake a proctored verification test (SHL, 2006). Figure 2-2 provides the reader with an example of what an honesty contract should state:



**Figure 2-2** Honesty contract

*The expectation that the candidate will take the assessment honestly and without the help of others*

*The expectation that the candidate will respect the confidentiality of the assessment and not share its contents with others*

*The expectation that the candidate will respect copyright, trademarks and other legal rights over the content of the assessment*

*That the candidate accepts that verification will be undertaken and that the candidate may be required to sit a psychometric verification at some point in the process*

Adapted from SHL (2006)

- Faking on personality measures

Owing to the fact that personality testing has become highly conventional in employment selection settings, the concerns of faking have heightened (Converse et al, 2008). Paulhus (in Converse et al, 2008:156) defines faking as "impression management where individuals consciously misrepresent themselves in order to create a positive image". Dilchert, Ones, Viswesvaran and Deller (2006:211) state that "when assessments are made for the purpose of making decisions that affect the employment of test takers, those assessed have an incentive to produce a score that will increase the likelihood of the desired outcome". In addition to this, Hogan in Dilchert et al (2006) claims that the tendency to mislead is a mechanism that helps individuals gain access to life's resources.

It is also important to note that not all candidates who conduct personality measures via the Internet complete the test with the aim of faking. Research suggests that making use of forced-choice personality items and warning test takers against misrepresenting themselves and informing them that cheating can be detected will reduce the act of cheating (SHL, 2006; Drasgow in Tippins et al, 2006; Converse et al, 2008). Rothstein and Goffin (in Converse et al, 2008) claim that these two methods, using forced choice items and warning test takers, are the most hopeful means of reducing and dealing with faking. In a study conducted by Ones and Viswesvaran (1998) it was found that social desirability does not influence the validity of personality measures.

It is for this reason that test distributors need to ensure that their instruments are cheat resistant. An assessment instrument qualifies as being cheat resistant when it is a self-report questionnaire (SHL, 2006). Self-report measures reduce the risk of cheating as they have no right or wrong answers and scores from questionnaires are combined (SHL, 2006). The CCSQ7.2, which is the instrument that is being used in this particular study, is a self-report measure which also has an accuracy check built into the questionnaire to reduce faking. The administration instructions warn candidates of this particular accuracy check and they are advised to answer the questionnaire as honestly as they can. The CCSQ7.2 will be discussed in more detail in Chapter 3.

It should also be noted that the concept of faking on personality measures is not an issue that is specific to online assessments and that it often occurs during paper-and-pencil administration.

Murphy and Davidshofer (1991:15) claim that, given that psychological tests are being used more often, there is potential for their abuse and for this particular reason professional and legal standards and guidelines have been developed for psychological testing. The next section will focus on best practice guidelines that should be adopted and enforced by organisations and assessment practitioners that are making use of psychological tests and more specifically unproctored Internet-delivered tests.

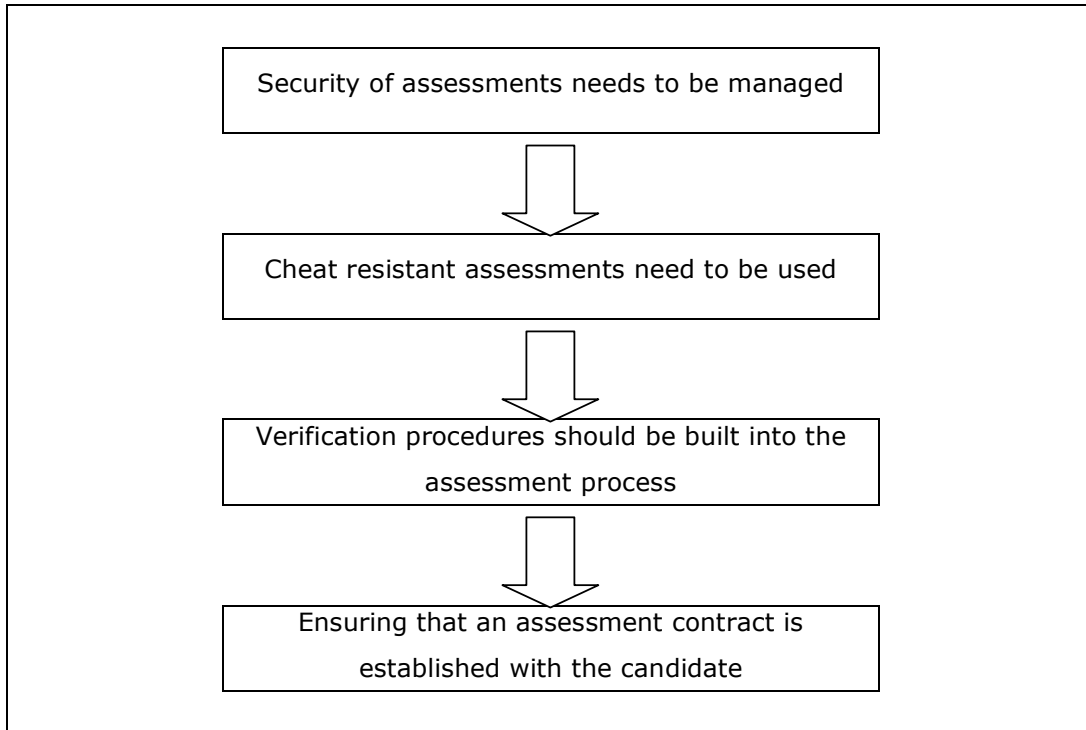
#### **2.5.4 Best practice guidelines**

Better practice can be defined as “a response to the apparent dilemma of addressing legitimate concerns while also meeting the need for unsupervised on-demand assessment without compromising the validity of assessment” (SHL, 2006:5). Key steps need to be followed in order to ensure best practice, and it is these key steps that are defined as better practice (SHL, 2006).

The unsupervised online environment has left many assessment practitioners and organisations feeling cautious and it is for this reason that these best practice guidelines have been designed (SHL, 2006). These guidelines also aid an organisation in avoiding disputes by following a set of justifiable procedures (SHL, 2006). It is important to note that the world of online assessment is in an innovative phase and these guidelines will not remain applicable forever and therefore need to be amended constantly.

SHL (2006) has identified the following four steps that assessment practitioners and organisations making use of occupational assessments can adopt in order to ensure that best practice guidelines are followed. These four steps are depicted in Figure 2-3.

**Figure 2-3** Better practice steps



Adapted from SHL, 2006

#### Step 1: The security of assessments needs to be managed

This entails formulating an assessment policy. This policy needs to include the reasons for the policy as well as the individuals, systems and materials that are going to be affected by the specific policy (SHL, 2006). There are also certain risk factors that need to be covered by the assessment policy, namely, that the assessment practitioner and/or organisation that is making use of assessment instruments is seen as a sound user of those specific products (SHL, 2006). In addition to this, the appropriate technology needs to be in place in order to support the assessment and the parties involved need to ensure that they are conducting the assessments in accordance with the laws that are stipulated for that specific country (SHL, 2006). In South Africa, assessment practitioners and organisations need to adhere to the guidelines that are stipulated by the Health Professions Council (HPCSA). Lastly, the policy needs to stipulate how assessment data is stored and protected, individuals responsible for the policy need to be notified of their

specific responsibilities and the implementation of the policy needs to be communicated to all (SHL, 2006).

#### Step 2: Cheat-resistant assessments need to be used

An assessment instrument is cheat resistant when the following requirements are met: Firstly the questionnaire needs to be a self-report questionnaire and, secondly, it needs to minimise the opportunity to obtain and to learn an answer if the assessment is an ability test by using methodology like item response theory (SHL, 2006). The CCSQ7.2 is an example of a cheat resistant measure as it is a self-report questionnaire, it has an accuracy check built into the questionnaire and it has been developed in an ipsative response style.

#### Step 3: Verification procedures should be built into the assessment process

A verification procedure can be defined as "a check on the consistency of scores and other information that support or question the evidence available to make a decision regarding the candidate" (SHL, 2006:13). Procedures that could be used include consistency checks built into the questionnaire, making use of other assessment instruments that can aid the decision-making process and, lastly, using other information that is related to the assessment to make a final decision (SHL, 2006).

#### Step 4: Ensuring that an assessment contract is established with the candidate

Candidates need to be informed about the objectives of the assessment, why the specific instrument has been chosen and its relevance for the job that the candidate is applying for, the procedure that will be followed once the assessment has been completed and lastly the rules of the assessment should be stipulated (SHL, 2006). Candidates should also be provided with an opportunity to conduct a practice test before undergoing the assessment and candidates need to be given technical guidelines in the event that problems occur during the assessment session (SHL, 2006).

Foxcroft and Roodt (2005:188) give South African assessment practitioners the following guidelines that should be adhered to when making use of the Internet as a mode of administration:

- The practitioner who is going to be making use of the Internet as a mode of administration should be competent in being able to use Internet-delivered tests.

- The potential effectiveness of the test should be established.
- The equivalence of the assessment instrument needs to be established.
- Test-takers should be prepared by offering them a practice test, which will enable them to become familiarised with the assessment instrument.
- Assessment results need to be stored securely.

As previously mentioned, the field of psychological testing is currently in an evolutionary phase and owing to the prevalence of the Internet, new developments in the field are occurring more regularly. The next section will focus on the role of the Internet in future developments.

## **2.6 The Internet and the future**

Owing to the evolving nature of online assessments, new developments are occurring in the field on a regular basis. Bartram (in Bartram & Hambleton, 2006) claims that the following areas will be affected in the future:

### **2.6.1 Internet interviews**

The main benefit of an interview is the interaction that is experienced between the interviewer and the applicant (Bartram in Bartram & Hambleton, 2006:23). The world is currently becoming smaller and smaller and, as a result of the geographical dispersion of applicants and organisations, video-conferencing provides an alternative to a formal interview thereby saving costs and time (Bartram in Bartram & Hambleton, 2006:23). Bartram (in Bartram & Hambleton (2006:23) claims that by the middle of this decade, domestic digital TV will enable high fidelity interviewing with built in cameras being used as video-phones.

### **2.6.2 Reference checks**

A study conducted by ADP in 2001 found that out of 2.6 million applicants in the US, 44% of these applicants lied about their work histories, 41% lied about their educational background and finally 23% falsified their qualifications (SHL, 2006). It is for this reason that it is essential for organisations to carry out reference checks. Bartram (in Bartram & Hambleton, 2006:23) states that formal checking of medical, criminal and credit records could become automated and would then be available to individuals with the necessary access permission on databases via the Internet.

### **2.6.3 Assessment and development centres**

The advent of online assessments has brought with it the advantage of a lack of supervision thereby enabling organisations to assess candidates in various geographical locations. Candidates who are applying for a higher-level position may be required to undertake group or individual assessment exercises (Bartram in Bartram & Hambleton, 2006:23). Bartram (in Bartram & Hambleton, 2006:24) is of the opinion that there are currently web-based in-basket exercises that are available but the potential lies in making these group and individual exercises Internet based thereby removing geographical constraints.

## **2.7 Conclusion**

The primary purpose of this literature study was to provide the reader with insight into the concepts of psychological assessment, the role of personality questionnaires in occupational assessments and, more specifically, to discuss the Internet as an administration mode. Previous validity and reliability studies that have been conducted on the CCSQ7.2 needed to be discussed in order to provide the reader with assurance of this particular instrument's suitability in occupational settings.

The subsequent chapter will focus on the research methodology that was adopted by the researcher. Aspects such as the research approach, research design, data collection method and a description of the sample will be discussed. The method that was used to analyse the data will also be discussed.

## **CHAPTER 3**

### **Research methodology**

The main purpose of this chapter will be to focus on the research design and research methodology that were adopted by the researcher to answer the particular research question. The reader will be provided with more information on the research design, the sample and the sampling method that was used. The measurement instrument, the data collection procedure and the techniques that were used to analyse the data will be discussed in more detail. The techniques used to interpret the data will also be investigated and the ethical considerations that were adhered to by the researcher throughout the duration of the study will also be mentioned.

The research question aims to establish the construct equivalence of the ipsative version of the Customer Contact Styles Questionnaire (CCSQ7.2) when it is administered online, in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting.

All the methodology that was adopted by the researcher was used with the purpose of answering the above-stated research question.

#### **3.1 Research approach**

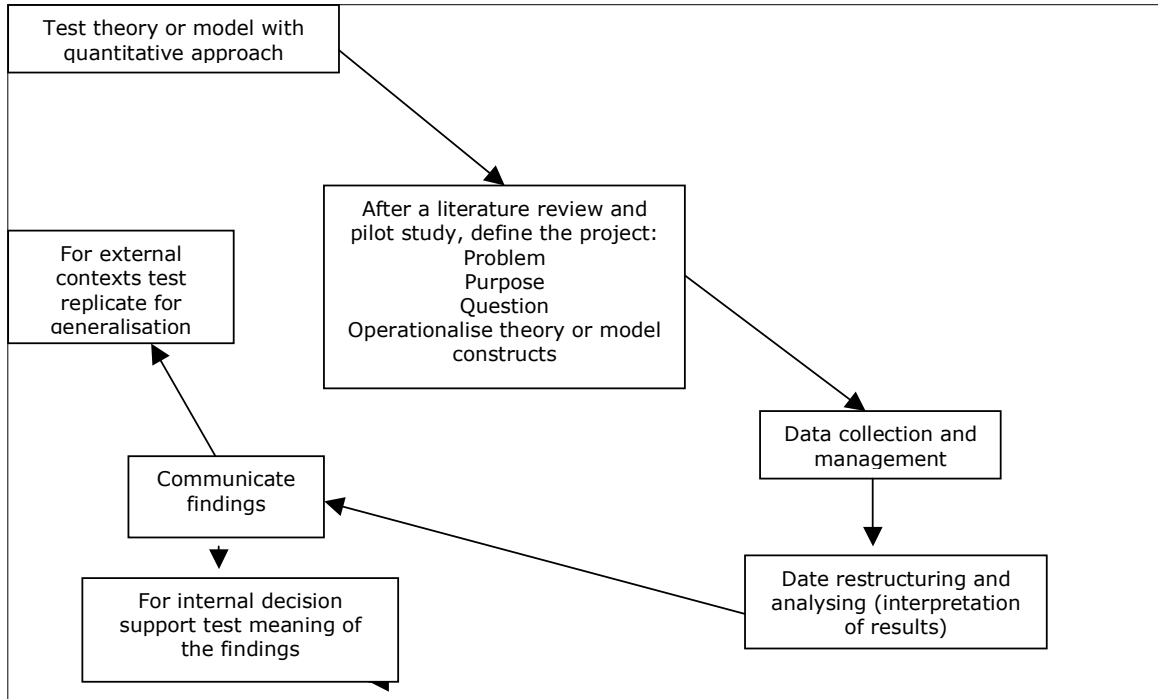
A quantitative research approach will be followed in this particular study. Borland (2001:8) indicates that "the purpose of quantitative research is to describe, predict, and control. Specific variables are isolated through control of the environment to eliminate the effects of confounding variables and testing their relationship to various behaviours".

The researcher adopted a quantitative approach for this particular research study. Before the research question could be answered, the researcher planned the methodology, the data collection instrument, the data collection method and ways in which the data would be analysed before the start of the research process. The researcher also ensured that the data that was going to be used for the research study could be quantified, as statistical analyses of the data were going to be carried out. The above-mentioned steps are characteristic of a research approach that is quantitative in nature.

Characteristics of the quantitative paradigm:

- **Who is studied?** The institutional researcher does not study individual human beings but rather seeks to recognise relationships between variables that explain behaviours that define specific populations of individuals (Borland, 2001). The keys are to define the population of interest and to opt for a sample that accurately represents it (Borland, 2001).
- **What is studied?** The institutional researcher needs to define variables that operationalise the constructs being studied (Borland, 2001). Once these measures have been operationally defined, instruments are created to record the measures on the individuals being studied and methodologies are developed to facilitate the collection of data (Borland, 2001).
- **What is the role of the researcher?** The study is planned by the researcher, including the design of the data collection instruments, data collection and management methodologies, and data analysis, before the start of the research process (Borland, 2001). Prior research plays a significant role in guiding the research design (Borland, 2001).

**Figure 3-1** A conceptual model of quantitative research



Adapted from Borland (2001)



### **3.2 Research design**

The research design that will be used in this research study is a non-experimental research design involving measurements at a single time where no follow-up session is required from participants. Random assignment, treatment or prediction does not take place in non-experimental research and the research will be conducted in a natural setting.

The researcher will be making use of a survey to conduct the research. According to Ruane (2005:87) "a survey is a research instrument that allows the researcher to gather critical information by posing questions." In a survey design the researcher can make use of either an interview or a questionnaire.

Since the researcher will be aiming to investigate the construct equivalence of the CCSQ7.2 when it is administered in two different modes, a non-experimental research design is the most suitable design for this research study, as this study entails assessing two sample groups of candidates by means of the CCSQ7.2 at a single point in time.

### **3.3 Sample**

Owing to the fact that, in this research study, the CCSQ7.2 has to be administered in two different modes, two sample groups were needed. Two different sample groups were used as there was no data available that had originated from the same sample group completing the CCSQ7.2 in both administration modes.

As there are two sample groups, the researcher decided that two sampling strategies would have to be used for this particular study.

The online sample is a convenience sample. Ruane (2005:117) states that "this form of sampling is the oldest sampling strategy and it involves finding convenient or available individuals in an unstructured manner and the sample is selected because they are convenient". This sampling strategy has a few shortcomings: firstly as individuals who are not "conveniently" located have no probability of being selected for such samples and, secondly, the exclusion of all but the most conveniently accessed elements in a population greatly undermines the representativeness of a convenience sample (Ruane, 2005:117). Convenience sampling is the most suitable for this specific research study as a sample of candidates was needed who had completed the online version of the

CCSQ7.2 in an unproctored setting. Once this sample of candidates was found, another sample group of candidates was needed who had completed the paper-and-pencil version of the CCSQ7.2 in a proctored setting. The paper-and-pencil, proctored sample group was drawn from an existing paper-and-pencil database and it is a random sample. Random sampling constitutes drawing a sample of people for a study from a population (Trochim, 2004). A random sample is the most suitable for this specific study as a sample of candidates was needed that would match the convenience sample.

### **3.3.1 Description of the sample**

Two sample groups were dealt with in this particular study. The first sample group completed the CCSQ7.2 for selection purposes and the second sample group completed the CCSQ7.2 for developmental purposes. As mentioned previously, two sample groups were required for this particular study as two sample groups were needed that had completed the CCSQ7.2 on different administration modes. At the time of this study, there was no data available that consisted of the same candidates who had completed the CCSQ7.2 via paper-and-pencil and candidates who had completed the online version.

Although the sample groups do not consist of the same participants, the candidates in both sample groups are comparative and similar. The paper-and-pencil sample group was randomly drawn to reflect the following biographical information of the online sample:

- Age
- Ethnicity
- Gender
- Education

The paper-and-pencil, proctored sample group was drawn from an existing paper-and-pencil database. These candidates were assessed for various positions in different industry sectors. Random samples were drawn from the database to reflect the biographical data of the online sample group. The sample groups were compared for significant differences in order to ensure that biographical information did not act as a moderator variable. For the variables age group, gender, education and ethnicity a chi-square analysis was conducted. No significant differences between the biographical variables of the two sample groups were found.

The biographical information of both sample groups will now be discussed.

The first sample group consisted of 358 candidates in different industry sectors across various job functions who had completed the CCSQ7.2 for selection purposes. The assessment was conducted via paper-and-pencil and in the presence of a proctor under standardised assessment conditions. This sample group was drawn from an existing paper-and-pencil database.

**Table 3-1** Descriptive statistics for sample group 1 (N=358)

<b>Gender</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
Male	204	56.98%	56.98%
Female	154	43.02%	100.00%
<b>Race</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
White	228	63.69%	63.69%
African	33	9.22%	72.91%
Indian	56	15.64%	88.85%
Coloureds	41	11.45%	100.00%
<b>Educational level</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
Degree	5	1.40%	1.40%
Post-matric certificate	8	2.23%	3.63%
Grade 12	345	96.37%	100.00%
<b>Mean age</b>	<b>Min</b>	<b>Max</b>	<b>SD</b>
36.81	20	63	9.97

The gender distribution of the candidates consisted of 204 (56.98%) males and 154 (43.02%) females. Both gender groups are well represented, with more males than females.

The ethnic distribution of the sample group consisted of 228 (63.69%) Whites, 33 (9.22%) Africans, 56 (15.64%) Indians and 41 (11.45%) Coloureds.

With regards to the educational qualifications of the sample group, the majority of the sample group were in possession of a Grade 12 certificate, making up 345 (96.37%) of the sample. The remainder of the sample included eight (2.23%) candidates with a post-matric certificate and five candidates (1.40%) with a degree.

The mean age of the sample group ranged from 20 to 63, with the average age being 36.81 (SD= 9.97) years.

The second sample group consisted of 239 service advisors in the motor industry who had completed the CCCSQ7.2 for developmental purposes. The assessment was conducted online and in the absence of a proctor.

**Table 3-2** Descriptive statistics for sample group 2 (N=239)

<b>Gender</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
Male	149	62.34%	62.34%
Female	90	37.66%	100.00%
<b>Race</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
White	171	71.55%	71.55%
African	13	5.44%	76.99%
Indian	34	14.23%	91.22%
Coloureds	19	7.95%	99.17%
Other	2	0.83%	100.00%
<b>Educational level</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Cumulative %</b>
Degree	3	1.26%	1.26%
Post-matric certificate	6	2.51%	3.77%
Grade 12	230	96.23%	100.00%
<b>Mean age</b>	<b>Min</b>	<b>Max</b>	<b>SD</b>
37.06	20	66	10.43

The gender distribution of the candidates consisted of 149 (62.34%) males and 90 (37.66%) females. Both gender groups are well represented, with more males than females.

The ethnic distribution of the sample group consisted of 171 (71.55%) Whites, 13 (5.44%) Africans, 34 (14.23%) Indians, 19 (7.95%) Coloureds and two (0.83%) candidates from other ethnic groups.

With regards to the educational qualifications of the sample group, the majority of the sample group were in possession of a Grade 12 certificate, making up 230 (96.23%) of the sample. The remainder of the sample included six (2.51%) candidates with a post-matric certificate and three candidates (1.26%) with a degree.

The mean age of the sample group ranged from 20 to 66, with the average age being 37.06 (SD= 10.43) years.

The fact that both sample groups completed the CCSQ7.2 for different purposes (selection and development) poses a limitation on the study. Cohen's d-statistic was used to determine if there is any practical significance between the consistency scores of the two sample groups. Cohen (1988) defines the effect size index, *d*, as "the difference between the means, divided by the pooled standard deviation of both groups." An effect size of 0.20 can be considered small, an effect size of 0.50 can be considered medium and an effect size of 0.08 can be considered large (Cohen, 1988). Here the effect size (using Cohen's *d*) is 0.31. This is indicative of a small effect and thus it can be concluded that there is no practical significance between the consistency scores of the two sample groups. Even though the selection group was more inclined to try and fake whilst completing the questionnaire, the consistency score of the selection group is similar to that of the development group.

### **3.3.2 Ethical issues surrounding the study**

There are several ethical issues that must be taken into account when conducting research that will involve participants who are human beings.

The researcher ensured that the following aspects were taken into consideration throughout the duration of the research study:

- The researcher disclosed the proposed methodology that was used during the research process and the techniques that were used to analyse the data.
- The institution whose data was used will not be mentioned in the research study and no direct references will be made to any individuals who took part in the assessment process. The institution mentioned here provided me with the data for the online sample.
- The confidentiality of respondents was respected at all times. No direct references were made to candidates and no information was used that could breach the privacy of respondents.

- The anonymity of respondents was ensured throughout the duration of the research study. Respondents' responses could not be identified.
- No changes were made to the research data. The researcher ensured that fabrication or falsification of the data did not occur and that the data was used only for the purposes for which it was intended.
- The researcher ensured that any contributions that were made by external participants in the execution of the research study were referenced appropriately.

### 3.4 Data-collection procedure

The first sample group completed the paper-and-pencil version of the CCSQ7.2 under standardised assessment conditions, and in the presence of a trained test administrator. Standardised assessment conditions imply that test conditions are as near as possible or exactly the same every time an assessment is conducted. Table 3-3 provides a description of the various elements that need to be standardised during an occupational assessment session:

Table 3-3 Standardised test administration

Testing conditions	The testing environment should be quiet and free from interruptions.
Timing	Time limits stated in the test instructions must be strictly adhered to.
Scoring and interpretation	These need to be carried out correctly and accurately by trained users.
Instructions	Instructions should be read to candidates exactly as they are presented on the administration card.

Adapted from SHL Occupational Testing Course Notes (2004)

The second sample group completed the online version of the CCSQ7.2. A test administrator was not present and candidates completed the questionnaire in their own time and at a location where they were granted Internet access. All candidates were issued with their own personal logins and passwords prior to the time. These usernames and passwords were sent to the candidates via e-mail and they were also given instructions on what they were required to do in order to complete the questionnaire. These instructions included details of programs and/or software that needed to be installed by the candidates in order to ensure that the questionnaire could be completed.

The online version of the CCSQ7.2 enables candidates to continue the questionnaire at a later period or date if disruptions should arise during the session. Online support is also available to candidates should any technical difficulties arise in the duration of the session.

Bartram (2001) indicates that there are four modes of test administration that can be used for occupational assessments:

1. *Open mode*: These are conditions where the test taker cannot be identified and there is no human supervision.
2. *Controlled mode*: This mode is similar to the open mode in that no human supervision of the test session is assumed. The test is only made available to known test takers and for the Internet this is controlled through the requirement for the test taker to be provided with a unique logon username and password.
3. *Supervised mode*: For this mode a level of human supervision is assumed whereby the identity of the test taker can be authenticated and test-taking conditions validated. For Internet testing, this mode is achieved by requiring the test administrator to log-in the candidate and to confirm that the testing was completed correctly at the end of the session.
4. *Managed mode*: This is a mode where a high level of human supervision is assumed and there is also control over the test-taking environment.

According to Bartram (2001) a managed mode was used for the first sample group, as these candidates completed the CCSQ7.2 under supervised assessment conditions. A controlled mode was used for the second sample group, with these candidates completing the online version of the CCSQ7.2.

### **3.4.1 Type of data collected**

The following biographical information was collected from the CCSQ7.2 for each candidate:

- Age
- Ethnicity
- Gender
- Education

During a proctored assessment session, candidates are prompted to complete the above-mentioned biographical information on their physical answer sheet. The test administrator is responsible for ensuring that all the required information is filled in correctly by each candidate. When candidates complete the online version of the CCSQ7.2, they are prompted to complete the above-mentioned biographical information on their desktop or laptop screen, as this information needs to be completed before the candidate can proceed with the assessment.

The data that was collected from the CCSQ7.2 comes in the form of raw scores taken from each scale of the CCSQ7.2. The raw scores reflect a candidate's true preference on each scale of the CCSQ7.2. "A raw score represents simple counts of the behaviours sampled by a test or measuring procedure" (Murphy & Davidshofer, 1991:51). Raw scores obtained by test-takers on psychological measures have little or no meaning and in order to make interpretation more meaningful, these raw scores are converted to normal scores by means of statistical transformation (Foxcroft & Roodt, 2005:39).

### **3.5 Measurement instrument**

The CCSQ7.2 is the measurement instrument that was used in this particular research study and was selected for the convenience sample group by one of SHL's registered psychologists. The specific instrument was chosen on the basis of the outcome of a job profiling session that was conducted by one of SHL's registered psychologists. A job profiling session entails determining the most important competencies of a job with a subject matter expert; this can be a line manager and/or job incumbent who is familiar with the job. These competencies refer to any knowledge, skills and level of qualification that are necessary to perform well at the specific job. In accordance with guidelines that are stipulated by the Health Professions Council of South Africa (HPCSA), the instrument was chosen by a registered psychologist.

According to the rules and regulations of the Health Professions Council of South Africa (HPCSA), registered psychologists need to control the use of all tests, questionnaires and instruments that measure psychological constructs.

The control of tests relates to

- selecting tests
- administering tests
- scoring tests and selecting norms
- interpreting tests and setting cut-off scores



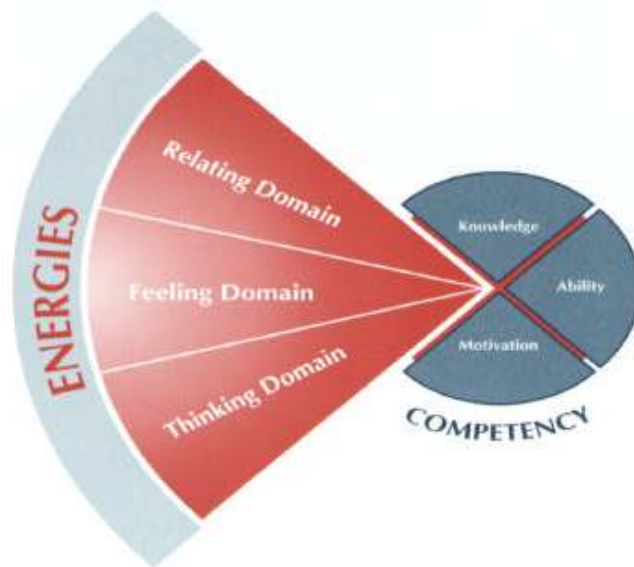
- reporting on results

### 3.5.1 The Customer Contact Styles Questionnaire (CCSQ)

The Customer Contact Styles Questionnaire (CCSQ) is a self-report questionnaire, which provides information about individuals' preferred or typical way of behaving in a work context (SHL, 1997). It can be used for selection, training and development, performance management, counselling, organisational change and redesign as well as placement decisions. Candidates are assessed for their suitability in non-supervisory sales and customer service roles (SHL, 1997).

The CCSQ was originally developed by SHL (Pty) Ltd in the United Kingdom with international collaboration and the model for the CCSQ was based on the OPQ model of personality (SHL, 1997). Figure 3-2 depicts the OPQ model of personality. This model of personality is based on the premise that people differ in behaviour. The situation that individuals find themselves in will have some impact on their behaviour; however general styles of behaviour will tend to show a degree of stability across time. Individuals typically have a preferred style of behaviour which comes naturally to them (SHL, 1997).

**Figure 3-2** The OPQ model of personality



Adapted from CCSQ Manual & User's Guide (1997)

The CCSQ looks at 16 dimensions of personality and is available in two formats. Both formats measure the same set of 16 scales, and personality is broken down into three domains:

<b>Relationships with People</b>	Persuasive Self-control Empathic Modest Participative Sociable
<b>Thinking Style</b>	Analytical Innovative Flexible Structured Detail Conscious Conscientious
<b>Emotions and Control</b>	Resilience Competitive Results orientated Energetic

Adapted from CCSQ Manual & User's Guide (1997)

Personality is concerned with three main areas or domains. Firstly, "Relationships with People" looks at how an individual relates to others. Secondly, "Thinking Style" covers traits such as analytical thinking, innovation, organising, and conscientiousness. Next there are the feelings and energies, the "Emotions" domain (SHL, 1997).

The normative version, CCSQ5.2, has 136 questions and candidates are asked to answer questions using a five-point Likert scale (strongly agree to strongly disagree). The questionnaire is not timed and it usually takes candidates 25 minutes to complete (SHL, 1997).

The ipsative version, CCSQ7.2, has 32 questions. Each question comprises four statements and candidates rate each question on a five-point Likert scale and then

determine which one from the set of four statements is most like them and which is least like them. The questionnaire has no time limit and it usually takes candidates approximately 30 minutes to complete (SHL, 1997).

Both versions of the CCSQ have an accuracy check built into the questionnaire. The normative version (CCSQ5.2) has a "Social Desirability" scale (extent to which individuals have been critical of themselves in their response to questions) and the ipsative version (CCSQ7.2) has a "Consistency" scale (extent to which individuals have answered in a consistent manner across the questionnaire). This scale was built into the questionnaire to reduce faking (SHL, 1997).

### **3.5.2 Development of the Customer Contact Styles Questionnaire (CCSQ)**

The development of the CCSQ dates back to two normative versioned customer-orientated questionnaires. The first is the Sales Personality Questionnaire (SPQ), which was launched in 1988; this questionnaire was designed for sales representatives. The second is the Customer Service Questionnaire (CSQ), which was launched in 1991 and was designed for lower-level jobs with direct customer contact. Feedback given from clients, as well as SHL staff using these two instruments, suggested that the two questionnaires were valid predictors but an ipsative version was considered necessary in order to reduce faking. Clients indicated that a single instrument was considered necessary in order to evaluate staff for their fitness for sales as well as customer service roles (SHL, 1997). The aim of the project was to construct a single model, which is able to integrate both customer service and sales elements (SHL, 1997).

**Table 3-4** Customer Contact Styles Questionnaire model and its relationship to the SPQ and CSQ

	<b>SPQ Scale</b>	<b>CSQ Scale</b>	<b>*CCSQ Scale</b>	
<b>Relating to people</b>	Persuasive	Need to Control	Persuasive	
	Empathy	Emotional	Self-Control	
	Confidence	Understanding of People	Sensitivity	Empathic
			Group Orientation	Modest
			Sociability	Participative
				Sociable
<b>Thinking Style</b>	Creative	Attitude to Change	Analytical	
	Forward Planning	Approach to	Innovative	
	Systematic	Organising	Flexible	
	Conscientious	Mental Awareness	Structured	
		Attitude to Authority	Detail Conscious	
			Conscientious	
<b>Emotions and drives</b>	Resilient/Relaxed	Need for Results	Resilience	
	Observant	Need for Social	Competitive	
	Results Orientated	Approval	Results Orientated	
			Energetic	
		Social Desirability		

\* The final scale names are used here for clarity, although initial working titles were often somewhat different.

Adapted from CCSQ Manual & User's Guide (1997)

A pool of items was created by experienced psychologists within SHL and was then reviewed by other SHL specialists. The items were then narrowed down to the best 18 items per scale. These items were then reviewed by a group of international experts (SHL, 1997).

A normative style was used for the first trial of the questionnaire. Three hundred candidates were used; these candidates came from various organisations as well as various industrial sectors. Three versions of the questionnaire were developed for the final trial. The first version was the CCSQ5.2 (normative), which consisted of 136 items. The second version was the ipsative version, which consisted of 40 sets of items. The third version consisted of a combination of the ipsative and the normative format. The

three versions of the questionnaire were administered to more than 1500 candidates in sales as well as customer service staff in the period of November 1995 and April 1996 (SHL, 1997).

Thirty-two of the 40 quads of the ipsative format of the questionnaire were selected which provided good statistical properties. This became the ipsative version (CCSQ7.2) of the questionnaire (SHL, 1997).

### **3.5.3 Fairness of the instrument**

“Fairness refers to the perceived fairness of the actions taken due to the result of the test scores” (SHL, 2006:9). Research indicates that there are reasonably small differences in the average response patterns of different ethnic groups. Gender differences occur on some scales but these differences reflect known differences in style between men and women. Various studies have been conducted in the South African environment to assess the fairness of the CCSQ. A particular study, study number R015 (2000), consisted of 737 applicants from various industry factors. The outcome revealed that very small-to-small effect sizes were found when a comparison was made between black and white ethnic groups (SHL 2006).

The CCSQ has been developed with a view to reducing the effects of various kinds of unintentional bias, which could misrepresent the meaningfulness of test scores. In particular the questionnaire:

- has comprehensive and carefully phrased instructions
- contains example questions to help candidates become familiar with the test content and to show candidates what is expected of them
- is free of unclear or ambiguous questions
- contains questions that have been designed to have a high degree of content relevance to jobs
- provides a sufficiently reliable measure of a particular job relevant attitude
- has been checked for material which is less familiar for some groups
- has been trialled on groups that include minorities (SHL, 1997)

The design of the instrument cannot ensure fair selection procedures; the administration procedures and the interpretation of results can affect fairness (SHL, 1997). According to the Health Professions Council of South Africa (HPCSA) the CCSQ7.2 has to be administered by a trained test administrator that has undergone the Test Administration

(TA) or the Occupational Testing Course (OT) presented by SHL and results may only be interpreted by trained users.

In addition to this, the HPCSA also states that:

Only persons registered with the Professional Board for Psychology under the auspices of the HPCSA may administer, score, interpret and give feedback on psychological tests.

### **3.5.4 Reliability and validity of the CCSQ7.2**

The Employment Equity Act, 1998 states that:

Psychometric testing and other similar assessments of an employee are prohibited unless the test or assessment being used:

- Has been scientifically proven to be valid and reliable

The reliability and the validity of the Customer Contact Styles Questionnaire (CCSQ7.2) will now be discussed in detail.

#### **3.5.4.1 Reliability of the CCSQ7.2**

Foxcroft and Roodt (2005:28) state that reliability is concerned with the consistency of measurement and can thus be defined as “the consistency with which it measures whatever it measures”.

There are many factors that can affect the reliability of a questionnaire, namely, the content of the measure, the candidate’s state, administration procedures, conditions and scoring (SHL, 1997).

There are three main methods that are used for estimating reliability:

- *Test-retest reliability*: This form of reliability is determined by administering the same questionnaire, on two separate occasions, to the same group of participants and thereafter correlating the results of the first set of scores with the second (Murphy & Davidshofer, 1991:78; SHL 1997).

- *Inter-rater reliability*: This form of reliability is determined by having all the test-takers' test protocols scored by two assessment practitioners (Foxcroft & Roodt, 2005:30).
- *Internal consistency reliability*: This form of reliability is a measure of the accuracy or consistency with which a set of test or questionnaire items measures one particular ability, personality trait or performance dimension (SHL, 1997).

The reliability of a scale of a questionnaire is expressed as a coefficient with a value between 0 and 1. Reliability coefficients of between 0.60 and 0.80 for personality questionnaires are generally considered a minimum for use in selection contexts (SHL, 1997). Table 3-5 depicts a reliability study done on the CCSQ7.2. The reliabilities range from 0.79 to 0.90 with a median of 0.84. The Social Desirability scale has a reliability of 0.66 (SHL, 1997).

**Table 3-5** Reliability of the CCSQ7.2 (N=617)

Scale	Alpha	Mean	sd	SEm raw	SEm Sten Score
<b>CR1 Persuasive</b>	0.79	34.86	6.54	3.0	0.9
<b>CR2 Self Control</b>	0.87	42.86	8.76	3.1	0.7
<b>CR3 Empathetic</b>	0.84	47.07	7.09	2.8	0.8
<b>CR4 Modest</b>	0.87	37.72	9.21	3.3	0.7
<b>CR5 Participative</b>	0.90	45.15	10.09	3.2	0.6
<b>CR6 Sociable</b>	0.79	39.56	7.54	3.5	0.9
<b>CT1 Analytical</b>	0.79	39.88	6.76	3.1	0.9
<b>CT2 Innovative</b>	0.90	39.29	9.49	2.9	0.6
<b>CT3 Flexible</b>	0.80	35.74	5.62	2.5	0.9
<b>CT4 Structured</b>	0.85	37.98	7.79	3.0	0.8
<b>CT5 Detail Conscious</b>	0.84	34.11	7.27	2.9	0.8
<b>CT6 Conscientious</b>	0.86	37.29	6.38	2.4	0.8
<b>CE1 Resilience</b>	0.81	36.90	8.33	3.6	0.9
<b>CE2 Competitive</b>	0.82	28.74	8.00	3.4	0.9
<b>CE3 Results Orientated</b>	0.79	36.42	6.30	2.9	0.9
<b>CE4 Energetic</b>	0.87	33.43	7.01	2.5	0.7
<b>CON Consistency</b>	0.52	59.90	5.2	3.6	1.4

Adapted from CCSQ Manual & User's Guide (1997)

There are ranges of reliability studies that have been conducted in the South African environment which demonstrate that the CCSQ7.2 can be used with confidence among

the South African population. All of these studies have been discussed in Chapter 1. A particular study, study number R011 (2000), reported alpha coefficients ranging between 0.68 and 0.88. Nine of the scale's alpha coefficients exceeded 0.80, five of the scales fell between 0.70 and 0.79 and only one scale fell below 0.70. The sample group consisted of 371 employees in the insurance environment. Another particular study, study number R016 (2000), reported alpha coefficients ranging from 0.74 to 0.90. Coefficients of 0.70 and higher were obtained for all the scales. The sample group consisted of 2397 employees.

Table 3-6 depicts a reliability study done on the CCSQ7.2 involving a South African and a British sample. The South African sample consisted of 737 employees in the financial sector. 79.35% were female and 20.65% were male, with a mean age of 32.56 years. The British sample consisted of 617 telesales and customer service staff in various industry sectors. Forty percent were female and 60% were male, with a mean age of 33 years.

**Table 3-6** Reliability study of the CCSQ7.2 between a South African (N=737) and a British (N=617) sample

CCSQ Scales	South African Sample (N=737)			UK Sample (N=617)		
	Mean	SD	Coefficient Alpha	Mean	SD	Coefficient Alpha
<b>R1 Persuasive</b>	29.37	7.81	0.83	34.86	6.54	0.79
<b>R2 Self Control</b>	46.20	7.79	0.85	42.86	8.76	0.87
<b>R3 Empathetic</b>	51.48	5.77	0.81	47.07	7.09	0.84
<b>R4 Modest</b>	39.25	8.49	0.83	37.72	9.21	0.87
<b>R5 Participative</b>	51.54	9.26	0.90	45.15	10.09	0.90
<b>R6 Sociable</b>	38.65	8.50	0.86	39.56	7.54	0.79
<b>T1 Analytical</b>	40.98	6.10	0.79	39.88	6.76	0.79
<b>T2 Innovative</b>	40.00	9.10	0.88	39.29	9.49	0.90
<b>T3 Flexible</b>	37.54	5.51	0.83	35.74	5.62	0.80
<b>T4 Structured</b>	40.32	6.89	0.84	37.98	7.79	0.85
<b>T5 Detail Conscious</b>	38.64	5.43	0.81	34.11	7.27	0.84
<b>T6 Conscientious</b>	40.16	5.04	0.83	37.29	6.38	0.86
<b>E1 Resilience</b>	36.13	7.41	0.75	36.90	8.33	0.81
<b>E2 Competitive</b>	28.23	8.48	0.84	28.74	8.00	0.82
<b>E3 Results Orientated</b>	37.26	5.76	0.82	36.42	6.30	0.79
<b>E4 Energetic</b>	34.04	6.46	0.87	33.43	7.01	0.87

Adapted from the Customer Contact Styles Questionnaire in South Africa: Psychometric Research Perspectives, SHL (2006)



### **3.5.5 Validity of the CCSQ7.2**

Babbie and Mouton (2006:122) define validity as “the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration”.

There are a number of criteria that can be used to establish validity:

- Face validity

Face validity is concerned with whether a questionnaire appears to have relevance for a particular job for inexperienced users and candidates and it should not be disregarded as it is typically managerial staff that will determine this form of validity (SHL, 1997). Foxcroft and Roodt (2005:33) also mention that face validity is mainly concerned with what an instrument appears to measure. The transparent nature of the items of the CCSQ clearly indicates their work relatedness (SHL, 2006).

- Content validity

Foxcroft and Roodt (2005:33) define content validity as the process of “determining whether the content of the measure covers a representative sample of the behaviour domain/aspect to be measured”. In the development of the CCSQ, SHL paid particular attention to assessing personality characteristics in the world of work only and thus, in order to ensure ample content validity for the CCSQ, it was attempted to include a broad range of item content that represents the domain of work-related behaviour (SHL, 2006).

- Criterion-related validity

Criterion-related validity consists of two main types; concurrent (potential effectiveness of assessment methods is evaluated on current job holders) and predictive (impact of an assessment method is evaluated by following up the job performance of selected individuals some months after initial testing) validity (SHL, 1997).

- Construct validity

Construct validity refers to the extent to which a personality dimension measures some underlying theoretical construct or trait and it is based on the logical relationships among variables (SHL, 1997; Babbie & Mouton, 2006:123). Correlation coefficients with other personality questionnaires or factor analyses are often used to investigate the construct

validity of personality questionnaires and, for this purpose; the CCSQ scores were correlated with the OPQ Concept Model scores (SHL, 2006).

There are also various development studies that are carried out by SHL on an ongoing basis on the CCSQ7.2. These development studies endeavour to ensure the suitability of the CCSQ7.2 in the South African environment. These studies have been discussed in detail in Chapter 1.

### **3.5.6 Norms of the instrument**

The CCSQ7.2 has been translated and is accessible in more than 20 different languages. There are seven norm groups that are available for the CCSQ7.2 ranging from the Finance to Mixed Industry Sectors. Sample sizes range between 255 and 6580. Anastasi (1988:72) indicates that "scores on psychological tests are most commonly interpreted by reference to norms, which represent the test performance of the standardisation sample". The norms are therefore established by determining what the persons in a representative group actually do on the test (Anastasi, 1988:72).

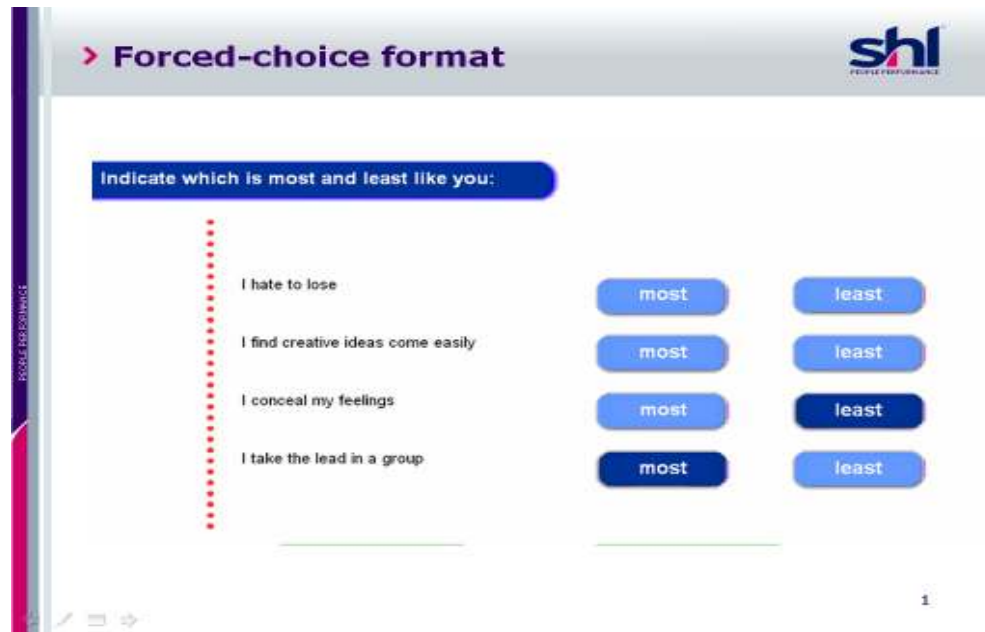
### **3.5.7 The use of ipsative measurement**

A large amount has been written about the difficulties of using ipsative measurement and forced-choice response styles in the assessment of personality using multi-scale questionnaires (Baron, 1996). The use of ipsative data in personality assessments has become a infamous topic with numerous authors arguing that the problems of ipsative data are such that it is impossible to analyse or interpret them using standard procedures owing to their psychometric properties (Baron, 1996). Thus far no agreement has been reached concerning the way in which ipsative test scores should be treated statistically (Martinussen, Richardsen & Varum, 2001). Saville and Willson (1991) contend that ipsative scaling methods are used for two main reasons: for the better control of response sets and to reflect the position that life is about choices.

Baron (1996:49) states that "scores on a multi-scale measure are alleged to be ipsative when raw scale scores sum to a constant for any individual"; in other words the sum of the total raw score will be the same for every individual. Ipsative scores can be generated by forced-choice questionnaire formats, where respondents order sets of items loading on different scales and in effect a constant number of points are distributed between the different scales (Baron, 1996).

It is important to note that the term “ipsative” refers to the type of data that is derived from a questionnaire. The format of the questionnaire (ipsative or normative) does not determine the property of the data; the property of the data in actual fact determines the format of the questionnaire. Thus it can be said that a forced-choice format does not necessarily present ipsative data (Brown, 2008). Figure 3-3 provides an illustration of a forced choice format.

**Figure 3-3** Forced choice format



Adapted from Brown (2008)

### 3.5.8 Justification for the use of ipsative measurement

Firstly, it is argued that ipsative scores constitute an ordinal level of measurement and thus do not meet the criteria for standard parametric analysis, which assumes an interval level of measurement (Baron, 1996). The scores represent an intra-individual ordering of the scales of a test (Baron, 1996). Baron (1996) argues that normative scales are not, in fact, true interval measures either, as individual item responses are at the ordinal level and there is no measurement that is capable of determining whether the difference between two adjacent points at the centre of the scale is in any sense equal to the distance between the last two points on the scale. In other words, is the difference between agree and strongly agree the same as the difference between disagree and neither agree nor disagree? Thus the postulation that normative scales are in fact interval measures has been glossed over in the literature and the matter has not been discussed further (Baron, 1996). Bowen, Martin and Hunt (2002) state that historically

psychologists have treated normative measures with great tolerance: normative scales were constantly treated as interval measures, though every psychologist knows that it is more prudent and defensible to treat a normative five-point Likert scale as an ordinal scale.

Secondly, normative measures also have definite weaknesses although they are often used owing to the fact that they are simple to complete and easier to analyse than their ipsative counterparts (Brown, 2008). Saville and Willson (1991) state that normative scales are more inclined to response biases than ipsative measures, such as central tendency responding or failure to use the full range of available response alternatives, social desirability, the halo effect and candidates finding them easier to manipulate. Forced-choice formats are designed to reduce such biases. Converse et al (2006) mention that studies have established that forced-choice measures appear to be more resistant to score inflation than their Likert-scale counterparts. A study conducted by Bowen et al (2002) confirms that ipsative measures are less susceptible to faking though they are not totally free from faking, implying that the possibility of faking cannot be completely eliminated but is reduced.

In forced-choice formats, respondents are forced to rate some items higher than others. They cannot agree with all the items represented and have to choose between two seemingly equally desirable items and their choices cannot be influenced by social desirability (Baron, 1996, Bowen et al, 2002). This supports the view of Saville and Willson (1991), who state that ipsative measures reflect the actual choices that individuals make in everyday life and that individuals have to constantly choose between options. Brown (2008) also argues that ipsative measures are also more cognitively demanding. Bowen et al (2002) state that the use of ipsative measures in selection can be justified by observing that many performance criteria such as voluntary turnovers are fundamentally about making choices between alternatives, so it theoretically makes sense to predict these criteria using measures that require people to make a choice.

The normative and ipsative version of the CCSQ cannot be seen as equivalent forms of the same instrument owing to the differences in item structure, presentation, questionnaire instruction, answering, scoring algorithms and statistical assumptions of the measures (Bowen et al, 2002). The two different versions do, however, measure the same set of traits in different ways. The correlation between the normative and ipsative formats should be regarded as an index of convergent validity between different methods measuring the same construct (Bowen et al, 2002).

Table 3-7 depicts correlations between the CCQ5.2 (normative) and the CCSQ7.2 (ipsative) version. Baron (1996) indicates that when ipsative scores correlate with their normative counterparts then internal consistency reliabilities can be generated for larger sets of scales. The sample consisted of 261 employees in customer service and sales industries from various organisations. The correlations range from 0.49 to 0.81 with a median value of 0.73 (SHL, 1997).

**Table 3-7** Correlations between the two versions of the CCSQ (N=261)

Scale		Correlations between CCSQ5.2 and CCSQ7.2
CR1	Persuasive	0.49
CR2	Self Control	0.64
CR3	Empathetic	0.59
CR4	Modest	0.77
CR5	Participative	0.75
CR6	Sociable	0.75
CT1	Analytical	0.74
CT2	Innovative	0.81
CT3	Flexible	0.64
CT4	Structured	0.72
CT5	Detail Conscious	0.70
CT6	Conscientious	0.71
CE1	Resilience	0.79
CE2	Competitive	0.79
CE3	Results Orientated	0.79
CE4	Energetic	0.70

Adapted from CCSQ Manual & User's Guide (1997)

Johnson, Wood and Blinkhorn (in Saville & Willson, 1991) state five 'uncontroversial' facts regarding ipsative data:

- They cannot be used for comparing individuals on a scale-by-scale basis.
- Correlation between scales cannot be legitimately factor analysed in the usual way.
- Reliabilities of ipsative tests overestimate the actual reliabilities of the scales.
- Validities overestimate their utility.
- Basic statistics such as means, standard deviations and correlation coefficients derived from ipsative test scales are not independent and cannot be interpreted and further utilised in the usual ways.

Saville and Willson (1991) on the other hand conclude and recognise that individuals can be as validly compared on a scale-by-scale basis on ipsative scales as normative, that ipsative data can be factor analysed legitimately and that neither reliabilities nor validities appear to be overestimated. It is unlikely that conclusions with ipsative data, based at least on a relatively large number of scales, are any less valid than those based on the normative (Saville & Willson, 1991).

As the literature points out, both normative and ipsative measures do have their restrictions. The limitations and benefits of normative and ipsative measures continue to be discussed in the literature and both measures can be utilised effectively if these restrictions and benefits are recognised.

Baron (1996) points out that the artificial interdependence of ipsative scores does affect their psychometric properties mainly where an instrument has few scales. A much larger number of scales is needed before factor analytic results resemble those provided by normative data and for the purpose of this study factor structures are not going to be compared (Baron, 1996). Brown (2008), on the other hand, confirms that problems associated with ipsative data are irrelevant when more than 20 scales are used. Figure 3-4 illustrates and supports this point made by Brown (2008) and it can be seen that with large numbers of scales (20 or more); the constraints that scores place on scales are not substantive and have minimal impact in practice.

**Figure 3-4** Percent raw score point change in score on scale



Brown (2008)

### 3.6 Data analysis

EQS 6.1 for Windows (Build 88) was the software package that was used in this study. This statistical program was used to analyse the data and it also enabled the researcher to make definite inferences from the data. The software package was developed by Peter M Bentler and Eric JC Wu and is distributed by Multivariate Software, Inc. It is described on their website ([www.mvsoft.com](http://www.mvsoft.com)) as follows:

*EQS provides researchers and statisticians with a simple method for conducting the full range of structural equations models including multiple regression, multivariate regression, confirmatory factor analysis, structured means analysis, path analysis, and multiple population comparisons.*

The data in this research study was analysed and interpreted using descriptive statistics. Descriptive statistics, according to Ruane (2005:179), are "a set of techniques that organize, summarize and provide a general overview of data". Descriptive statistics

enable the researcher to summarise the information obtained from a sample and to reduce the quantitative data by describing what the data shows, as well as to make comparisons, conclusions and recommendations.

In particular the following descriptive statistics were used to answer the research question:

- Mean
- Minimum and maximum values
- Standard deviation
- Skewness
- Kurtosis

The following three statistical methods were used to examine the relationship between all the scales of the CCSQ7.2 for both sample groups.

### **3.6.1 Comparison of effect sizes**

The two sample groups were compared by using effect sizes. Cohen (1988) defines the effect size index,  $d$ , as “the difference between the means, divided by the pooled standard deviation of both groups”. The standard deviation of either group could be used when the variances of the two groups are homogeneous (Cohen, 1988). Table 3-9 depicts the various effect sizes; an effect size of 0.20 can be considered small, an effect size of 0.50 can be considered medium and an effect size of 0.80 can be considered large (Cohen, 1988).

**Table 3-8** Effect sizes

<b>Small</b>	<b>Medium</b>	<b>Large</b>
0.20	0.50	0.80

Cohen (1988)

### **3.6.2 Test reliability (Alpha coefficient)**

The internal consistency reliability for both sample groups that completed the CCSQ7.2 was calculated using Cronbach’s Alpha. The coefficient alpha method of assessing internal consistency reliability is recognised as being a stringent test of the reliability of a scale (SHL, 2006). Internal consistency reliability is a measure of the accuracy or



consistency with which a set of questionnaire items measures a particular scale of personality (SHL, 2006). The internal consistency reliability of the CCSQ7.2 was calculated, as scales with very high alpha coefficients can be too narrow in their focus and lack bandwidth, which has the effect of reducing the validity for measuring broad personality traits and therefore optimum alpha coefficients for personality questionnaires lie in the range between 0.60 and 0.80 (SHL, 2006).

### **3.6.3 Equivalence of covariance structures using SEM with EQS**

A comparison of covariance structures was carried out using Structural Equation Modeling through EQS. Pui-Wa and Qiong (2007:34) define Structural equation modeling as "a general term that has been used to describe a large number of statistical models used to evaluate the validity of substantive theories with empirical data". Statistically, it represents an extension of general linear modelling (GLM) procedures, such as the ANOVA and multiple regression analysis (Pui-Wa & Qiong, 2007). In addition to this, Pui-Wa and Qiong (2007) mention that one of the primary advantages of SEM is that it can be used to study the relationships among latent constructs that are indicated by multiple measures and it is also applicable to both experimental and non-experimental data, as well as cross-sectional and longitudinal data.

There are a number of statistics that can be used to measure how adequately the hypothesised model describes the sample data (Bartram & Brown, 2004). The Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) were used in this particular study.

The CFI (Bentler, 1990) is an example of an incremental fit index, which provides a measure of complete covariation in the data. This type of index compares the improvement of the fit of the researcher's model over a restricted model, called an independence or null model, which specifies no relationships among variables (Weston & Gore, 2006). A value of  $>0.90$  was originally considered representative of a well-fitting model, but a revised cut-off value close to 0.95 has recently been advised (Hu & Bentler, 1999).

Another fit measure is the root mean square of approximation (RMSEA). This index has recently been recognised as one of the most informative criteria in covariance structure modelling, with values less than 0.05 indicating a good fit (Byrne, 2001).

### **3.7 Conclusion**

The purpose of this chapter was to elaborate on the method of investigation that was adopted for this particular research study. The researcher provided the reader with detailed information on the design of the research, the sampling strategy that was used, the sample, the data-collection procedure that was used and the ways in which the data was analysed. The subsequent chapter will focus on the presentation of the actual results and findings of the research study.

## CHAPTER 4

### Data analysis and results

The main objective of this chapter will be to provide the reader with an indication and a clear understanding of the results that were obtained from this research study. The results will be presented in a tabular format and a detailed explanation will follow.

#### 4.1 Descriptive statistics

Table 4-1 provides the reader with the mean score, the standard deviation, the minimum and maximum scores, the skewness and the kurtosis that was computed for each scale of the CCSQ7.2 for the paper-and-pencil sample group that completed the CCSQ7.2 in a proctored administration setting.

**Table 4-1** Mean, Standard deviation, Minimum, Maximum, Skewness and Kurtosis for the CCSQ7.2 (N=358)

Scale	Mean	Std.Dev.	Min	Max	Skewness	Kurtosis
<b>Consistency</b>	58.58	5.14	46	71	-0.11	-0.40
<b>Resilience</b>	36.53	7.31	15	58	0.09	-0.28
<b>Competitive</b>	28.87	8.14	9	48	-0.02	-0.49
<b>Results orientated</b>	36.70	5.80	17	49	-0.52	0.25
<b>Energetic</b>	34.39	6.53	14	48	-0.39	0.14
<b>Persuasive</b>	32.58	7.76	8	48	-0.21	-0.38
<b>Self-control</b>	45.10	8.02	16	62	-0.61	0.58
<b>Empathetic</b>	49.78	6.68	23	63	-0.65	0.79
<b>Modest</b>	38.66	8.18	15	60	-0.19	-0.00
<b>Participative</b>	49.06	9.71	20	69	-0.52	-0.03
<b>Sociable</b>	39.68	8.79	14	56	-0.50	-0.23
<b>Analytical</b>	41.15	6.13	18	54	-0.42	0.34
<b>Innovative</b>	40.70	9.09	19	62	0.02	-0.49

<b>Flexible</b>	35.91	6.26	15	47	-0.70	0.54
<b>Structured</b>	40.09	6.98	18	55	-0.49	0.20
<b>Detail Conscious</b>	37.94	5.72	15	49	-0.79	1.17
<b>Conscientious</b>	39.20	5.30	24	49	-0.54	0.04

The above results clearly signify that the data obtained for the skewness and kurtosis statistic were smaller than or very close to one for all the scales, thus demonstrating that the data does not deviate to a large extent from the normal distribution.

Table 4-2 provides the reader with the mean score, the standard deviation, the minimum and maximum scores, the skewness and the kurtosis that was computed for each scale of the CCSQ7.2 for the online sample group that completed the CCSQ7.2 in the absence of a proctor.

**Table 4-2** Mean, Standard deviation, Minimum, Maximum, Skewness and Kurtosis for the CCSQ7.2 (N=239)

<b>Scale</b>	<b>Mean</b>	<b>Std.Dev.</b>	<b>Min</b>	<b>Max</b>	<b>Skewness</b>	<b>Kurtosis</b>
<b>Consistency</b>	57.00	4.91	46	68	0.01	-0.60
<b>Resilience</b>	36.79	5.98	22	55	-0.01	-0.38
<b>Competitive</b>	29.23	7.39	11	46	0.15	-0.54
<b>Results orientated</b>	36.95	4.71	24	47	-0.00	-0.57
<b>Energetic</b>	35.64	5.41	19	47	-0.36	0.20
<b>Persuasive</b>	31.20	6.04	13	46	-0.11	-0.34
<b>Self-control</b>	43.79	7.90	19	62	-0.44	0.23
<b>Empathetic</b>	51.63	6.06	32	63	-0.72	0.57
<b>Modest</b>	39.02	7.35	20	58	-0.28	-0.05
<b>Participative</b>	49.53	8.77	21	68	-0.76	0.45
<b>Sociable</b>	39.68	6.94	20	55	-0.24	-0.05
<b>Analytical</b>	40.18	5.22	19	53	-0.15	0.73

<b>Innovative</b>	41.77	6.81	21	57	-0.21	0.01
<b>Flexible</b>	35.57	5.08	18	47	-0.46	0.52
<b>Structured</b>	41.46	5.97	21	55	-0.36	0.14
<b>Detail Conscious</b>	38.93	4.58	24	49	-0.33	0.12
<b>Conscientious</b>	39.69	4.66	27	49	-0.24	-0.55

The above results clearly signify that the data obtained for the skewness and kurtosis statistic were less than one for all the scales, thus demonstrating that the data does not deviate to a large extent from the normal distribution.

#### 4.2 Comparison of effect sizes

Table 4-3 provides the reader with the means and the standard deviations for both sample groups. The effect sizes for both groups were also computed. Effect size (d) is computed by taking the difference between the means of both sample groups and dividing it by the pooled standard deviation of both groups. According to Cohen (1988), an effect size of 0.20 can be considered small, an effect size of 0.50 can be considered medium and an effect size of 0.80 can be considered large. The effect sizes (d) range from 0.00 to 0.31, with four scales having an effect size of above 0.20. These effect sizes are characteristic of a small effect, which clearly indicates that there is no practical statistical significance between the scale scores of both sample groups.

**Table 4-3** Mean, Standard deviation and Effect size for both sample groups of the CCSQ7.2

Scale	Paper and pencil (N=358)		Online (N=239)		Combined group
	Mean	SD	Mean	SD	Effect size (d)
<b>Consistency</b>	58.58	5.14	57.00	4.91	0.31
<b>Resilience</b>	36.53	7.31	36.79	5.98	-0.04
<b>Competitive</b>	28.87	8.14	29.23	7.39	-0.05
<b>Results orientated</b>	36.70	5.80	36.95	4.71	-0.05
<b>Energetic</b>	34.39	6.53	35.64	5.41	-0.20

<b>Persuasive</b>	32.58	7.76	31.20	6.04	0.19
<b>Self-control</b>	45.10	8.02	43.79	7.90	0.16
<b>Empathetic</b>	49.78	6.68	51.63	6.06	-0.29
<b>Modest</b>	38.66	8.18	39.02	7.35	-0.04
<b>Participative</b>	49.06	9.71	49.53	8.77	-0.05
<b>Sociable</b>	39.68	8.79	39.68	6.94	0.00
<b>Analytical</b>	41.15	6.13	40.18	5.22	0.17
<b>Innovative</b>	40.70	9.09	41.77	6.81	-0.13
<b>Flexible</b>	35.91	6.26	35.57	5.08	0.06
<b>Structured</b>	40.09	6.98	41.46	5.97	-0.21
<b>Detail Conscious</b>	37.94	5.72	38.93	4.58	-0.19
<b>Conscientious</b>	39.20	5.30	39.69	4.66	-0.10

#### 4.3 Test reliability (Alpha coefficient)

Table 4-4 depicts the Alpha coefficients and the Standard Error of Measurement (SEm) that were computed for both sample groups. The internal consistency reliability for both sample groups that completed the CCSQ7.2 was calculated using Cronbach's Alpha. Generally, scales with very high alpha coefficients can be too narrow in their focus and lack bandwidth, which has the effect of reducing the validity for measuring broad personality traits and therefore optimum alpha coefficients for personality questionnaires lie in the range between 0.60 and 0.80 (SHL, 2006). Alpha coefficients for the paper-and-pencil sample group ranged from 0.75 to 0.91 with 15 scales above 0.80. Only one scale fell below the 0.80 mark. The mean alpha obtained for the paper-and-pencil sample group was 0.84 and the median 0.85. Alpha coefficients for the online sample group ranged from 0.64 to 0.89 with eight scales above 0.80. The mean alpha obtained for the online sample group was 0.79 and the median 0.81. The mean Standard Error of Measurement obtained for the paper-and-pencil sample group is 2.83 and the mean Standard Error of Measurement obtained for the online sample is 2.78.

**Table 4-4** Alpha coefficients and SEM scores for both sample groups of the CCSQ7.2

Scale	Paper and pencil (N=358)		Online (N=239)	
	Alpha	SEm	Alpha	SEm
Resilience	0.75	3.66	0.64	3.59
Competitive	0.85	3.15	0.82	3.13
Results orientated	0.81	2.53	0.74	2.40
Energetic	0.87	2.36	0.82	2.30
Persuasive	0.85	3.01	0.75	3.02
Self-control	0.85	3.11	0.83	3.26
Empathetic	0.85	2.59	0.83	2.50
Modest	0.81	3.56	0.76	3.60
Participative	0.91	2.91	0.89	2.91
Sociable	0.88	3.04	0.82	2.95
Analytical	0.81	2.67	0.76	2.56
Innovative	0.89	3.01	0.82	2.89
Flexible	0.86	2.34	0.77	2.44
Structured	0.85	2.70	0.83	2.46
Detail Conscious	0.82	2.42	0.75	2.29
Conscientious	0.83	2.18	0.79	2.14
Mean	0.84	2.83	0.79	2.78
Median	0.85	2.81	0.81	2.72

Table 4-4 depicts the reliabilities for the paper-and-pencil and the online sample group. These reliabilities for the different scales are very similar for both sample groups; this can be deduced by looking at the mean and median alpha coefficients of the scales for both sample groups. These coefficients are similar for the paper-and-pencil and the online sample group and therefore it can be said with confidence that the mode in which the CCSQ7.2 is administered in does not compromise scale reliabilities.

It should be noted that the reliabilities obtained from the online sample are lower than the reliabilities obtained from the paper-and-pencil sample group. Although the reliabilities differ somewhat, it can be deduced that the scale with the smallest reliability score for the online sample group is equal to the smallest reliability scale for the paper-and-pencil sample group and the same applies to the scale with the highest reliability score for both sample groups. An understandable example of this is depicted in Table 4-5 where the scale Resilience has a low reliability score in both sample groups and the Participative scale has the highest reliability score in both sample groups.

**Table 4-5** Differences in reliability scores for both sample groups

Scale	Online reliability	Scale	Paper and pencil reliability
Resilience	0.64	Resilience	0.75
Participative	0.91	Participative	0.89

#### 4.4 Equivalence of covariance structures using SEM with EQS

A comparison of covariance structures was carried out using Structural Equation Modelling through EQS.

There are a number of statistics that can be used to measure how adequately the hypothesised model describes the sample data (Bartram & Brown, 2004). In this study the CFI (Bentler, 1990) was used; it is an example of an incremental fit index, which provides a measure of complete covariation in the data. A value of >0.90 was originally considered representative of a well-fitting model, but a revised cut-off value close to 0.95 has recently been advised (Hu & Bentler, 1999). Another fit measure that was used in this study is the root mean square of approximation (RMSEA). This index has recently been recognised as one of the most informative criteria in covariance structure modelling with values less than 0.05 indicating a good fit (Byrne, 2001).

The model that was tested obtained a CFI of 0.986 and the RMSEA was 0.029. These results are indicative of an exceptionally good fit and consequently it can be deduced that this model fits equally well for both samples and the relationship between scales of the CCSQ7.2 is not affected by modes of administration.



## **CHAPTER 5**

### **Discussion, limitations, recommendations and conclusion**

#### **5.1 Discussion**

The primary objective of this research study was to determine the construct equivalence of the Customer Contact Styles Questionnaire (CCSQ7.2) when it is administered via paper-and-pencil in the presence of a proctor and when it is administered online in the absence of a proctor. The aim was to establish whether the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric property when it is administered in two different administration modes.

The results of the study revealed that the mode in which the Customer Contact Styles Questionnaire (CCSQ7.2) is administered does not compromise scale reliabilities and that the relationship between scales of the CCSQ7.2 is not affected by modes of administration. It can thus be said with confidence that the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric property and that there is no distortion to the instrument itself.

In this particular study, three statistical methods were used to examine the relationship between all the scales of the CCSQ7.2 for the two sample groups used. The first sample group completed the paper-and-pencil version of the CCSQ7.2 in a proctored environment and the second sample group completed the online version of the CCSQ7.2 in the absence of a proctor. Firstly, a comparison of effect sizes was carried out for both sample groups and the results reveal effect sizes ( $d$ ) that ranged from 0.00 to 0.31, indicating that there are only small differences in the scores of the two samples. The effect sizes that were computed in this particular study are comparable to the effect sizes that were found in the studies that were conducted by Bartram and Brown (2004); Holtzhausen (2005) and Mylonas and Carstairs (2003).

The internal consistency reliability for both sample groups that completed the CCSQ7.2 was calculated using Cronbach's Alpha. The results of this study reveal Alpha coefficients for the supervised paper-and-pencil sample group ranging between 0.75 and 0.91 and Alpha coefficients for the controlled web-based sample group ranging between 0.64 and 0.89. The mean alpha obtained for the paper-and-pencil sample group was 0.84 and the mean alpha obtained for the online sample group was 0.79. The mean Standard Error of Measurement obtained for the paper-and-pencil sample group was 2.83 and the mean Standard Error of Measurement obtained for the online sample was 2.78. The alpha

coefficients, mean alpha coefficients and mean Standard Error of Measurement for both sample groups were similar to the reliabilities that were found in studies conducted by Bartram and Brown (2004); Holtzhausen (2005); Mylonas and Carstairs (2003) and Salgado and Moscoso (2003).

A comparison of covariance structures was carried out using Structural Equation Modelling with EQS and the model that was tested obtained a CFI of 0.986 and an RMSEA of 0.029. These results are indicative of an exceptionally good fit indicating that the instrument is not affected by changes in administration mode. Studies that were conducted by Bartram and Brown (2004) and Holtzhausen (2005) reported results that are similar to the results obtained in this study.

In conclusion, the null hypothesis can therefore be rejected and consequently it can be deduced that the Internet version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting. In addition to this, the results of this study reveal that neither the mode of administration nor the presence or absence of a proctor had an effect on the psychometric property of this instrument.

The results of this study also add to the existing literature on existing equivalence studies, thereby providing individuals and organisations with more evidence that yet another instrument, specifically the CCSQ7.2, can indeed be used across both administration modes without loss of psychometric property.

In summary, it is clear that the CCSQ7.2 can be used with confidence via paper-and-pencil in a proctored setting and online in an unproctored setting throughout the entire employee life cycle. Organisations and assessment practitioners can use either versions of this instrument when making workforce decisions. The results of this study also provide evidence demonstrating that a lack of supervision does not impact on the equivalence of this questionnaire, consequently encouraging the use of the Internet as an administration mode as well as the use of unproctored Internet testing in employment settings.

## **5.2 Limitations and recommendations**

After this research study had been conducted the researcher identified a number of limitations.

Firstly the CCSQ7.2 was administered to two different sample groups. Two different sample groups were used as there was no data available that consisted of the same sample group completing the CCSQ7.2 in both administration modes. Although two sample groups were used, the researcher did ensure that the samples were equalised on the basis of biographical information and no significant differences between the biographical variables of the two sample groups were found. It is therefore suggested that future research on measurement equivalence could be conducted on a sample group of the same participants who will complete a specific measurement instrument in different administration modes.

Secondly, the two sample groups completed the CCSQ7.2 for different purposes. The first sample group that completed the CCSQ7.2 via paper-and-pencil in the presence of a proctor completed the questionnaire for selection purposes. The second sample group that completed the CCSQ7.2 online in the absence of a proctor completed the questionnaire for developmental purposes. This places a limitation on the study as participants who are completing a measurement instrument for selection purposes are more inclined to try and fake as the assessment that they are undertaking will most likely affect their chances of employment and they will try to produce a score that will increase the likelihood of the desired outcome (Dilchert et al, 2006). Although it was found that there is no practical significance between the consistency scores of the two sample groups, it is still recommended that future research on measurement equivalence should be conducted on participants who are completing a specific questionnaire for the same purpose.

In order to establish measurement equivalence, there are three types of equivalence method that can be used. Van de Vijver and Poortinga (in Brown, 2007) identify them as, structural/functional equivalence, measurement unit equivalence and scalar/full score equivalence. For the purposes of this study only the first type of equivalence, structural/functional equivalence, was used to determine whether the online version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric property. For the purposes of future research, all three types of equivalence method can be used when determining the measurement equivalence of a specific measurement instrument.

### **5.3 Conclusion**

The main aim of the research study was to determine the construct equivalence of the CCSQ7.2 when it is administered in two different administration modes.

The results that were obtained in this study confirm that the Internet version of the CCSQ7.2 can be considered equivalent to its paper-and-pencil counterpart without loss of psychometric properties when it is administered online in an unproctored setting and when it is administered via paper-and-pencil in a proctored setting thus rejecting the null hypothesis. The results of this study also confirm that the online version of this specific instrument can be used with confidence when making workforce decisions.

The data obtained in this study provides evidence of the construct equivalence of the CCSQ7.2 and the results of this study compare well with results that were obtained in similar studies indicating that the relationship between scales of the CCSQ7.2 are not affected by modes of administration.

## REFERENCES

- Aiken, L.R. 1994. *Psychological testing and assessment*. Massachusetts: Allyn and Bacon.
- Anastasi, A. 1988. *Psychological testing* (6<sup>th</sup> ed). New York: Macmillan.
- Babbie, E. & Mouton, J. 2006. *The practice of social research*. Cape Town: Oxford University Press Southern Africa.
- Baron, H. 1996. Strengths and limitations of ipsative measurement. *Journal of Occupational and Organizational Psychology*, 69, 49-56.
- Baron, H. & Miles, A. 2002. *Personality questionnaires: Ethnic trends and selection*. Paper presented at the BPS Occupational Psychology Conference, 2002.
- Barrick, M.R. & Mount, M.K. 1991. The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-25.
- Bartram, D. 2001. *The impact of the Internet on testing for recruitment, selection and development*. Presented at the 4<sup>th</sup> Australian I/O Psychology Conference, 21-24 June.
- Bartram, D. 2001. *Testing through the Internet: Mapping the issues for managing the future*. A practitioner workshop presented at the Occupational Psychology Conference 4-5 January 2001.
- Bartram, D. & Brown A. 2004. Online testing: Mode of administration and the stability of OPQ32i scores. *International Journal of Selection and Assessment*, 12(3), 278-284.
- Bartram, D. & Hambleton, R.K. 2006. *Computer-based testing and the Internet*. Chichester: John Wiley.
- Bentler, P.M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246.
- Borland, K.W. 2001. Qualitative and quantitative research: A complementary balance. *New Directions for Institutional Research*, 2000(112), 5-13.

- Bowen, C., Martin B.A. & Hunt, S.T. 2002. A comparison of ipsative and normative approaches for ability to control faking in personality questionnaires. *International Journal of Organizational Analysis*, 10(3), 240-260.
- Brown, A. 2007. *Quantitative procedures for establishing measurement invariance/equivalence*. Presented at the DOP Annual conference, Bristol, 12 January 2007.
- Brown, A. & Dowdeswell, K. 2008. *Normative-ipsative debate: Old arguments and new answers*. Pre-conference workshop, SIOPSA conference Pretoria, 21 May 2008.
- Buchanan, T. & Smith, J.L. 1999. Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology*, 90(1), 125-145.
- Byrne, B.M. 2001. *Structural equation modeling with AMOS: Basic concepts, applications and programming*. New Jersey: Lawrence Erlbaum.
- Cartwright, S. & Tidswell, K. 1997. *Customer Contact Styles Questionnaire (CCSQ)*. United Kingdom: Saville & Holdsworth.
- Cohen, L.J. 1988. *Statistical power analysis for the behavioral sciences*. New York: Lawrence Earlbaum.
- Converse, P.D., Oswald, F.L., Imus, A., Hedricks, C., Roy, R. & Butera, H. 2008. Comparing personality test formats and warnings: Effects on criterion-related validity and test-taker reactions. *International Journal of Selection and Assessment*, 16(2), 155-169.
- Cronbach, L.J. 1990. *Essentials of psychological testing* (5<sup>th</sup> ed). New York: Harper Collins.
- Dilchert, S., Ones, D.S., Viswesvaran, C. & Deller, J. 2006. Response distortion in personality measurement: Born to deceive, yet capable of providing valid self-assessments? *Psychology Science*, 48(3), 209-225.
- Dowdeswell, K.E. 2006. *Applicant perceptions of the fairness of online assessment in the selection process*. Masters Dissertation, University of Pretoria.

- Evans, T. 2002. *Equivalence between Internet-based and paper-and-pencil cognitive ability tests and factors associated with the test mode effect*. Masters Dissertation, Birkbeck College.
- Evers, A., Anderson, N. & Voskuijl, O. 2005. *The Blackwell handbook of personnel selection*. Malden: Blackwell.
- Foxcroft, C. & Roodt, G. 2005. *Psychological assessment* (2nd ed). Cape-Town: Oxford University Press South Africa.
- Holtzhausen, G. 2004. *Mode of administration and the stability of the OPQ32n: comparing Internet (unsupervised) and paper-and-pencil (supervised) administration*. Masters Dissertation, University of Pretoria.
- Hu, L. & Bentler, P.M. 1999. Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6 (1), 1-55 [Online] Available from: <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=3384345&site=ehost-live&scope=site> [Accessed: 2008-02-28].
- International Test Commission. 2005. *International guidelines on computer-based and Internet delivered testing*. Available from: [www.intestorg.com](http://www.intestorg.com)
- Internet World Stats. 2008. *Internet users in the world by geographic regions*. [Online] Available from: [www.internetworldstats.com/images/world2008users.png](http://www.internetworldstats.com/images/world2008users.png) [Accessed: 2008-09-04].
- King, W.C. & Miles, E.W. 1995. A quasi-experimental assessment of computerizing noncognitive paper-and-pencil measurements: A test of measurement equivalence. *Journal of Applied Psychology*, 80(6), 643-651.
- Kriek, H.J. 2005. *International trends and Internet-based assessments*. Paper presented at the 8<sup>th</sup> Annual Industrial Psychology Conference, Pretoria, Gauteng, SA.
- Martinussen, M., Richardsen, A.M. & Varum, H.W. 2001. Validation of an ipsative personality measure (DISCUS). *Scandinavian Journal of Psychology*, 42, 411-416.

Multivariate Software. 2005. EQS 6.1 for Windows (Build 88). Encino, USA: Multivariate Software, Inc. [Online] Available from: [www.mvsoft.com](http://www.mvsoft.com) [Accessed: 2008-05-05].

Murphy, K.R. & Davidshofer, C.O. 1991. *Psychological testing: Principles & applications*. New Jersey: Prentice-Hall.

Mylonas, G. & Carstairs, J. 2003. *Comparison of a computer motivation questionnaire under supervised and unsupervised conditions*. MSc Project, Macquarie University.

Ones, D.S. & Viswesvaran, C. 1996. Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, 17, 609-626.

Ones, D.D. & Viswesvaran, C. 1998. The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, (2/3), 245-269.

Owen, K. & Taljaard J.J. 1996. *Handbook for the use of psychological and scholastic tests of the HSRC*. Pretoria: Production Printers.

Ployhart, R.E., Weekley, J.A., Holtz, B.C. & Kemp, C. 2003. Web-based and Paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, 56, 733-752.

Potosky, D. & Bobko, P. 2004. Selection testing via the Internet: Practical considerations and exploratory empirical findings. *Personnel Psychology*, 57(4), 1003-1035.

Professional Board of Psychology. 2006. *Policy on the classification of psychometric measuring devices, instruments, methods and techniques*. Available from: [www.hpcs.org](http://www.hpcs.org) Published under the auspices of the Health Professions Council of South Africa.

Pui-Wa, L. & Qiong, W. 2007. Introduction to Structural Equation Modeling: Issues and practical considerations. *Educational Measurement: Issues and Practice*, 26(3), 33-43 [Online] Available from: <http://0-www.blackwell-synergy.com.innopac.up.ac.za/action/showFullText?submitFullText=Full+Text+HTML&doi=10.1111%2Fj.1745-3992.2007.00099.x> [Accessed: 2008-02-28].



Ruane, J.M. 2005. *Essentials of research methods: A guide to social science research*. Malden: Blackwell.

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

Salgado, F.S. & Moscoso, S. 2003. Internet-based personality testing: Equivalence of measures and assesses' perceptions and reactions. *International Journal of Selection and Assessment*, 11(2/3), 194-205.

Saville & Holdsworth. 1997. *CCSQ manual and user's guide*. Thames Ditton, Surrey: Saville & Holdsworth.

Saville, P. & Willson, E. 1991. The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, 64, 219-238.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R011. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R013. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R014. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R015. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R016. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R025. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R027. United Kingdom: SHL Group Plc.

SHL. 2000. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number D005. United Kingdom: SHL Group Plc.

SHL. 2001. Customer Contact Styles Questionnaire (CCSQ7.2): Study Number R035. United Kingdom: SHL Group Plc.

SHL. 2004. *SHL Training Academy: Occupational testing course notes*. United Kingdom: Saville & Holdsworth.

SHL South Africa. 2006. *The Customer Contact Styles Questionnaire in South Africa: Psychometric research perspectives*. Pretoria: SHL South Africa.

SHL. 2006. *Better practice for unsupervised online assessment*. SHL White paper 2006. United Kingdom: SHL Group plc.

SHL Group. 2008. *Key benefits of objective personality assessment*. [Online] Available from: [http://www.shl.com/SHL/en-int/Products/Access\\_Personality/personality-assessment](http://www.shl.com/SHL/en-int/Products/Access_Personality/personality-assessment) [Accessed: 2008-08-19].

Tippins, N.T., Beaty, J., Drasgow, F., Gibson, W.M., Pearlman, K., Segall, D.O. & Shepherd, W. 2006. Unproctored Internet testing in employment settings. *Personnel Psychology*, 59(1), 189-225.

Trochim, W.M. 2004. *The research methods knowledge base* (2<sup>nd</sup> ed). [Online] Available from: <http://trochim.human.cornell.edu/kb/index.htm> [Accessed: 2008-20-05].

Welman, J.C. & Kruger, S.J. 2003. *Research methodology for the business and administrative sciences* (2<sup>nd</sup> ed). Cape-Town: Oxford University Press South Africa.

Weston, R. & Gore, P.A. 2006. A brief guide to Structural Equation Modeling. *The Counseling Psychologist*, 34, 719.