A Psychometric Investigation of the South African Career Interest Inventory-isiXhosa Version among Secondary School Learners using the Rasch Measurement Model

by

Nicola Jansen van Vuuren

Thesis presented in fulfilment of the requirements for the degree of Master of Arts (Psychology) in the Faculty of Arts and Social Sciences at Stellenbosch University

Supervisors:

Dr Stephan Rabie

Prof Anthony V. Naidoo

April 2022

Stellenbosch University https://scholar.sun.ac.za

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

April 2022

Copyright © 2022 Stellenbosch University
All rights reserved

Stellenbosch University https://scholar.sun.ac.za

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following individuals for making this research study possible:

- My supervisors, Dr Stephan Rabie and Professor Anthony Naidoo for their guidance, wisdom, patience, and continuous encouragement through times that the completion of the research seemed out of reach.
- Dr Enoch Teye-Kwadjo for imparting his Rasch model expertise.
- Dr Xander van Lill for his advice on the statistical analysis.
- My friends and family for their support and upliftment.

Stellenbosch University https://scholar.sun.ac.za

SUMMARY

The assessment of interests is an indispensable component of career counselling. John Holland's theory of vocational personalities provides an extensively used theoretical framework in career interest assessment. However, South African research has indicated a lack of validity of imported interest inventories based on Holland's model in the South African context. These findings provided the catalyst for the development of the South African Career Interest Inventory (SACII) – an interest inventory underpinned by Holland's circular order model that takes the unique South African context into consideration. The SACII was developed in English and Afrikaans, even though the majority of the South African population does not speak either English or Afrikaans as a home language. Due to this limitation, and the importance of language as a moderator of assessment performance, Rabie and Naidoo (2019) constructed the SACII-isiXhosa (SACII-X) - the first career interest inventory in an indigenous South African language. Promising initial results for the reliability and validity of the SACII-X have been found among a South African sample. Nevertheless, the SACII-X is in need of further study and refinement. Since the SACII-X has previously only been subjected to Classical Test Theory methodology, Item Response Theory methodology makes a valuable contribution to the study of the SACII-X. The present study aimed to investigate the psychometric properties, and the item functioning of the SACII-X, using the Rasch measurement model. To achieve these aims, secondary data, collected by Rabie (2017) from 266 isiXhosa home language-speaking secondary school learners during the development of the SACII-X, were analysed. The data demonstrated good fit to the Rasch measurement model, thus providing support for the reliability and validity of the SACII-X in the South African context. The analysis of item functioning further highlighted a small

Stellenbosch University https://scholar.sun.ac.za

proportion of poor functioning items across the six SACII-X scales. Recommendations regarding the adaptation or removal of the identified items are made.

Stellenbosch University https://scholar.sun.ac.za

OPSOMMING

Die assessering van belangstellings is ‘n belangrike komponent van beroepsvoorligting. John Holland se teorie van beroepspersoonlikhede is ‘n populêre teoretiese raamwerk in beroepsbelangstelling assessering. Suid-Afrikaanse navorsing het egter ‘n gebrek aan geldigheid van Holland se model aangedui in die Suid-Afrikaanse konteks. Hierdie bevindings was die katalis vir die ontwikkeling van die Suid-Afrikaanse Loopbaanbelangstellingsinventaris (SACII) – ‘n beroepsbelangstellings-toets gebaseer op Holland se sirkelvormige model. Die SACII was slegs ontwikkel in Engels en Afrikaans, alhoewel die meerderheid van die Suid-Afrikaanse populasie nie Engels of Afrikaans as huistaal praat nie. As gevolg van hierdie tekortkoming, en die belangrike rol wat taal in psigometriese assessering speel, het Rabie en Naidoo (2019) die SACII-isiXhosa ontwikkel – die eerste beroepsbelangstellings-toets in ‘n inheemse Suid-Afrikaanse taal. Die SACII-X het belowende resultate in terme van die betroubaarheid en geldigheid van die toets in ‘n Suid-Afrikaanse steekproef gelewer. Nietemin is dit nodig om die SACII-X verder te bestudeer en te verfyn. Aangesien die SACII-X voorheen slegs deur Tradisionele Toets Teorie metodes ondersoek is, sal Item Respons Teorie metodologie, ‘n waardevolle bydrae lewer tot die bestudering van die SACII-X. Die doelstellings van hierdie studie was om die psigometriese eienskappe, en die item-funksionering van die SACII-X te ondersoek, deur gebruik te maak van die Rasch metings-model. Ten einde hierdie doelstellings te bereik, was sekondêre data wat ingesamel is deur Rabie (2017) in ‘n steekproef van 266 isiXhosa huistaal-sprekende sekondêre skool leerlinge, tydens die ontwikkeling van die SACII-X, ondersoek. Die data het goeie ooreenstemming met die vereistes van die Rasch metings-model gewys, en daardeur die betroubaarheid en geldigheid van die SACII-X in die Suid-Afrikaanse konteks, ondersteun. Die analise van item funksionering het verder ‘n klein proporsie items

Stellenbosch University https://scholar.sun.ac.za

aangedui wat swak funksioneer oor die ses skale van die SACII-X. Aanbevelings met betrekking tot die aanpassing of verwydering van die items word gemaak.

Stellenbosch University https://scholar.sun.ac.za

TABLE OF CONTENTS

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

LIST OF TABLES

Stellenbosch University https://scholar.sun.ac.za

LIST OF FIGURES

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

LIST OF APPENDICES

Stellenbosch University https://scholar.sun.ac.za

LIST OF ABBREVIATIONS

15FQ+	15 Factor Questionnaire Plus
16PF	16 Personality Factor Questionnaire
APA	American Psychological Association
ANC	African National Congress
CTT	Classical Test Theory
d.f.	Degrees of freedom
DIF	Differential Item Functioning
EEA	Employment Equity Act
ETS	Educational Testing Service
HPCSA	Health Professions Council of South Africa
ICC	Item Characteristic Curve
IRT	Item Response Theory
ITC	International Test Commission
MnSq	Mean-square
PCA	Principal Component Analysis
P-E	Person-Environment
PsySSA	Psychological Society of South Africa
REC	Research Ethics Committee
S.E.	Standard Error
SACII-E	South African Career Interest Inventory-English version
SACII-SR	South African Career Interest Inventory-Short Research version
SACII-X	South African Career Interest Inventory-isiXhosa version

Stellenbosch University https://scholar.sun.ac.za

SAVII	South African Vocational Interest Inventory
S.D.	Standard deviation
SDS	Self-Directed Search
SII	Strong Interest Inventory
SIOPSA	Society for Industrial and Organisational Psychology of South Africa
t	t-test
Zstd	Z-standardised

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 1

INTRODUCTION

1.1 Introduction

Psychometric assessment, including the assessment of interests, remains one of the most important components of career counselling (Harrington & Long, 2013). Some counsellors use assessments to ultimately enhance clients' self-awareness by exploring clients' interests, aptitudes, and values as they pertain to vocational environments (De Bruin & De Bruin, 2018). The importance of paying attention to interests in career counselling is underpinned by the notion that a high degree of fit between an individual's work-related interests and vocational environment will contribute to life satisfaction and career productivity (De Bruin & De Bruin, 2018; Gregory, 2007; Wiernik, 2016). This degree of interest-environment fit can only be accurately determined if the assessment measures used are reliable, valid, and unbiased[1] towards the test-taker – thus adhering to fair testing practices (Psychological Society of South Africa, 2007).

John L. Holland's (1985, 1997) theory of vocational personalities and work environments has been widely used in vocational interest assessment (Nauta, 2010). Holland's (1997) theory categorises individuals and work environments into six types, namely Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (collectively referred to as RIASEC). The arrangement of the six types in a hexagonal form, or circular order, displays the theoretical relationships between them (Holland, 1997). In South Africa, the Self-Directed Search (SDS; Gevers et al., 1997) and the South African Vocational Interest Inventory (SAVII; Du Toit et al., 1993) are popular interest inventories based on Holland's (1985, 1997) theory. However, the item content of these measures is outdated. As interests may be

[1] In this instance, bias refers to test bias, that is, the differential validity of test scores for groups (Frencher, 2014).

Stellenbosch University https://scholar.sun.ac.za

responsive to changing economic, cultural, and temporal contexts, the outdated item content may not accurately reflect current adolescent career interests (Wiernik, 2016). Furthermore, the reliable and valid use of interest inventories based on the theory have been brought into question in recent years (Du Toit & De Bruin, 2002; Foxcroft et al., 2004; Mintram et al., 2019; Morgan, 2014; Rabie & Naidoo, 2019; Spokane et al., 2017). In the South African context, a study by Du Toit and De Bruin (2002) failed to provide support for the validity of the underlying theoretical structure among seTswana and isiXhosa home language-speaking participants, when using the SDS (Du Toit & De Bruin, 2002). To address this limitation, Morgan et al. (2015) developed the South African Career Interest Inventory (SACII), taking into account the unique South African context (Rabie & Naidoo, 2019). However, even though the majority of South Africans do not speak either English or Afrikaans as a home language (Statistics South Africa, 2018), the SACII was not available in a language other than English and Afrikaans. As research has shown that language is a significant moderator of assessment performance (Foxcroft, 2004; Van Eeden & Mantsha, 2007), and that more reliable and accurate results on interest inventories based on Holland's circular model are obtained when test-takers are assessed in their home language (Einarsdóttir et al., 2002; Rabie & Naidoo, 2019; Šverko & Babarović , 2006; Zhang et al., 2013), it is imperative to provide test-takers with the opportunity to be assessed in either their home language, or their language of choice. Rabie and Naidoo (2019) subsequently adapted and translated the SACII into isiXhosa, thereby developing the SACII-X, the first interest inventory in an indigenous African language. Both the SACII-E[2] and the SACII-X have been shown to be reliable and valid among diverse samples of South African test-takers (Morgan et al., 2015; Rabie & Naidoo, 2019). However, the reliability and validity of the SACII-X have only been

[2] SACII-E will be used to refer to the English version of the SACII.

Stellenbosch University https://scholar.sun.ac.za

investigated using Classical Test Theory (CTT) approaches (Rabie, 2017). Due to the numerous limitations of using CTT in the investigation of the psychometric properties of tests (see Section 4.7.1 for further elaboration on these limitations; Boone & Noltemeyer, 2017), it is important to also use modern test theory approaches, such as Rasch analysis, to investigate the reliability, validity, and item functioning of the SACII-X, before fully incorporating this test into the career counselling field.

Informed by this context, the present study aimed to utilise secondary data to investigate the psychometric properties of the SACII-X, with a focus on the item functioning of the test, to identify potential improvements to the inventory, using the Rasch measurement model, in a sample of isiXhosa home language-speaking secondary school learners.

1.2 Rationale

The indiscriminate use of career interest inventories based on Holland's model in the diverse South African context is contentious, as studies have produced mixed results regarding the construct validity of the model among the South African population. This led to the recent development of the SACII-E (Morgan, 2014) and SACII-X (Rabie & Naidoo, 2019). These interest inventories are underpinned by Holland's model and initial studies have shown promising results for the reliability and validity of these measures among diverse South African and African samples (Mintram et al., 2019; Morgan et al., 2015a, 2015b; Morgan & De Bruin, 2018, 2019; Rabie & Naidoo, 2019). As the SACII was initially only available in English and Afrikaans, the SACII-X provides isiXhosa home language-speaking individuals with the opportunity to be assessed in the language of their choice. Assessing individuals in their home language, or language of choice, is crucial in order to obtain reliable and valid assessment results, as the language of instruction of a psychometric assessment is a significant moderator of test performance (Foxcroft, 2004; Van Eeden & Mantsha, 2007).

Stellenbosch University https://scholar.sun.ac.za

In light of this, the aim of the present study is to investigate the psychometric properties of the SACII-X, and identify potential improvements to the SACII-X. As the SACII-X has previously only been subjected to CTT methodology (Rabie, 2017), and due to the limitations of CTT as a method of investigation, the Rasch measurement model, an Item Response Theory (IRT) methodology was used for the analysis. The Rasch model addresses some of the criticisms raised against CTT methodology (Boone & Rogan, 2005). The model provides the same information as yielded by CTT, but also provides novel, useful information (Boone & Rogan, 2005; Retief et al., 2013). The data acquired in this study provide additional empirical support for the reliability and validity of the SACII-X in the South African context.

1.3 Aims and Hypotheses

The overarching aim of this study was to investigate the psychometric properties of the SACII-X, identify potential improvements to the SACII-X, and investigate differential item functioning across gender groups, among a sample of isiXhosa home language-speaking secondary school learners using the Rasch measurement model. To achieve this, the present study employed quantitative methodology to meet the following objectives:

- To determine the reliability and validity of the SACII-X items through the analysis of the adherence of the data to the Rasch model assumptions and the fit of the data to the Rasch model.
- To investigate the item functioning of the SACII-X using Rasch analysis to identify problematic items that could be adapted or removed.
- To investigate differential item functioning in the SACII-X, across gender groups.

The present study had three hypotheses:

H1: The Rasch measurement model will substantiate the SACII-X to be reliable and valid.

Stellenbosch University https://scholar.sun.ac.za

H2: The Rasch measurement model will identify improvements to items across all six scales of the SACII-X.

H3: Differential item functioning across gender groups will be indicated in some of the SACII-X items.

The rationale for these hypotheses is based on previous research findings on the SACII-E and SACII-X. The SACII-E and SACII-X have previously been shown to be a reliable and valid measure of vocational interests among diverse samples (Mintram et al., 2019; Morgan et al., 2015a, 2015b; Morgan & De Bruin, 2018, 2019; Rabie & Naidoo, 2019), providing the rationale for the first hypothesis. The rationale for the second and third hypothesis is based on research findings in Rabie (2017) that have identified poor functioning items across all six scales, and potentially gender biased items in the SACII-X.

1.4 Demographic and Socio-Economic Context of Kayamandi

The secondary data utilised in the present study were collected in Kayamandi (meaning *nice-* or *sweet home* in isiXhosa; Toms, 2015), a peri-urban community situated on the outskirts of Stellenbosch in the Western Cape province of South Africa (Dube-Addae, 2019). Kayamandi is an impoverished community with an unemployment rate of 22.3% (Rock, 2011; Stellenbosch Municipality LED, 2008). In addition, the community has poor quality of education, with only 28.6% of the population aged 20 and older having completed Grade 12 (Statistics South Africa, 2011b). Other socio-economic problems include limited formal housing (Rock, 2011); overcrowding (Toms, 2015); high levels of crime; substance- and alcohol abuse; and health concerns such as HIV/AIDS, tuberculosis, and teenage pregnancies (Albien, 2013; Toms, 2015).

The 2011 Census reported the population of Kayamandi as 24,645 (Statistics South Africa, 2011). IsiXhosa is the most prevalent home language spoken in Kayamandi (84.9%; Rock, 2011; Statistics South Africa, 2011a; Toms, 2015), and 94.6% of the population

Stellenbosch University https://scholar.sun.ac.za

identify as Black African (Statistics South Africa, 2011a). Kayamandi further has a business centre, police station, a health clinic, multiple churches, a library, a playground for children, and some sport facilities (Albien, 2013; Toms, 2015). The community has three primary schools, and two secondary schools, namely Kayamandi Secondary School and Makupula Secondary School (Toms, 2015). More information on these two schools is provided in Chapter 4.

1.5 Definition of Key Terms

1.5.1 Vocational Interests

In this thesis, the terms *vocation* and *career* are used interchangeably. A vocation can be defined as “the work in which a person is regularly employed; a career” (Allen, 2000a, p. 1578). A career can be defined as “the course of somebody’s working life or a specified part of it” (Allen, 2000b, p. 208). A vocation can also be viewed as a career that a person is especially suited to or qualified for (Allen, 2000a).

Interests can be conceptualised as “stable dispositional attributes” (Low & Rounds, 2007, p. 2). They indicate an individual’s liking for some activities and objects in comparison to others (Aiken & Groth-Marnat, 2006). Thus, when the definitions for vocation and interests are combined, vocational interests specifically refer to an individual’s preference for activities or objects that are common in certain vocational environments (Schermer, 2016).

From a social learning perspective, interests are argued to develop as a result of the reinforcement an individual receives for taking part in some activities in comparison to others, and also through modelling and imitating important others in the individual’s life (Aiken & Groth-Marnat, 2006). From a biological perspective, interests can be viewed as resulting from a person’s genetic make-up, with research suggesting that almost half of the variance in career interests, results from genetic variance (Aiken & Groth-Marnat, 2006; Low & Rounds, 2007). However, gene-environment interaction also plays a role in the development of interests

Stellenbosch University https://scholar.sun.ac.za

(Aiken & Groth-Marnat, 2006; Low & Rounds, 2007), as a person's genotype may influence the environment the person is exposed to (Dick, 2011). Furthermore, the influence of genetics on interest development can also be viewed as an indirect one, as genetics determine one's abilities, personality, and intelligence, and individuals have a tendency to be interested in activities that they are good at, or that they believe they are able to do (Aiken & Groth-Marnat, 2006; Low & Rounds, 2007). Furthermore, contextual factors such as socio-economic status, social class, gender, race or ethnicity, and sexual orientation influence vocational interest development and expression (Fouad, 2007; Low & Rounds, 2007). In addition, vocational interests may change as the nature of work and careers change over time (Wiernik, 2016). Acknowledging the factors that influence vocational interest development is crucial in a study population of adolescents embarking on the career exploration phase of their schooling, while attempting to incorporate their self-concept into their vocational identity (Fouad, 2007).

Holland (1997) proposed that interests are an expression of personality. This conceptualisation of interests as personality traits has been supported by research showing that RIASEC interests show certain qualities of personality traits. For example, the RIASEC types show correlation with the Big Five personality traits (Barrick et al., 2003; Larson et al., 2002; Nauta, 2010; Tsotetsi, 2020); demonstrate stability over time (Nauta, 2012; Stoll et al., 2020); and show self-other agreement correlations, suggesting that RIASEC interests are visible to others (Nauta, 2012). As it has been shown that adolescents' RIASEC interests fit Holland's circular order structure by early high school, and remain relatively stable over the high school years, it can be assumed that vocational interests can be accurately measured in the study population to aid them in making career decisions (Low & Rounds, 2007; Tracey et al., 2005).

Stellenbosch University https://scholar.sun.ac.za

1.5.2 The Distinction between Career Counselling, Psychological Assessment, and Psychological Testing

Career counselling is a process whereby a career counsellor facilitates career exploration and enhances clients' awareness and understanding of their career-related characteristics to assist in making career choices and resolving career-related problems (De Bruin & De Bruin, 2017, 2018; Stead & Subich, 2017). Career counselling utilises psychological assessment and standardised testing to achieve the above-mentioned aims (De Bruin & De Bruin, 2017).

Even though the terms *assessment* and *testing* are often used interchangeably, these words are not equivalent. Psychological assessment is a process whereby a large amount of information about an individual is gathered through various activities such as interviews, exploring an individual's history, and psychological tests (Foxcroft & Roodt, 2013). The assessment results can be used to draw conclusions about an individual and predict the individual's behaviour (Foxcroft & Roodt, 2013; Gregory, 2007). Psychological testing is a particular aspect of the assessment process (Foxcroft & Roodt, 2013). It involves the administration and scoring of standardised tests (also used interchangeably with measures, inventories, scales, or instruments) to collect data about a specific aspect of a person's life (Foxcroft & Roodt, 2013).

In the context of career counselling, testing involves the selection and administration of tests that measure aspects such as aptitude, personality, or interests. Assessment refers to the collection and integration of information relevant to careers from various sources to aid the individual to make meaning of this information and ultimately make a career decision (De Bruin & De Bruin, 2017). Interest inventories, such as those used in the present study, are thus psychological tests measuring individuals' preferences for activities and objects in vocational environments (Gregory, 2007).

Stellenbosch University https://scholar.sun.ac.za

1.5.3 Stable Behaviour

Stable can be defined as something that is "not subject to change or variation; unvarying" (Allen, 2000c, p. 1363). In the context of behaviour, it refers to behaviour that remains relatively constant and is unlikely to change and vary. According to Holland (1985, 1997), individuals whose personality type is congruent with the vocational environment they find themselves in, will display stable behaviour. This means that a person is unlikely to change their personality traits, values, and belief systems, and/or change aspects of the vocational environment, in order to increase congruence between the person and the environment (Holland, 1997). Moreover, the individual will show stability in career choice and employment, and is consequently less likely to change career paths (Holland, 1997).

1.6 Overview of Chapters

The details of this study are presented in the format of six chapters.

Chapter 1

Chapter 1 contextualises the present study. The practice of interest assessment as a component of career counselling research and practice, is introduced. This is followed by a brief introduction of Holland's theory and its influence on interest assessment in the South African context. The issues of the validity of Holland's theory in the South African context, and language as a moderator of assessment performance, which provided the motivation for the development of the SACII-E and SACII-X, is highlighted. The rationale for the study, the study's aims, and hypotheses are provided. Lastly, the context of Kayamandi is described and key terms relevant to the study are defined.

Chapter 2

Chapter 2 discusses the theoretical viewpoints that underlie the SACII-E and SACII-X. A brief overview of the Person-Environment Fit approach to career counselling is firstly provided, followed by a detailed discussion of Holland's theory of vocational personality

Stellenbosch University https://scholar.sun.ac.za

types and work environments. The basic postulates and theoretical assumptions of Holland's theory are lastly discussed.

Chapter 3

In Chapter 3, literature relevant to the research focus is reviewed. Firstly, a discussion of the diversity of South Africa is provided, with a specific focus on language diversity. This is followed by a discussion of English as the language of instruction in South Africa. Issues relating to home-language use, compared to the use of the language of instruction for psychological assessment, are discussed. Hereafter, an overview of fair and ethical assessment practices is provided. This includes a history of psychological testing in South Africa, and language and culture as moderators of assessment performance. Subsequently, a review of the literature relating to the cross-cultural applicability and validity of Holland's theory is provided. The chapter ends with a discussion of research conducted on the SACII-E and SACII-X in South Africa.

Chapter 4

Chapter 4 outlines the methodology employed in this study. It provides information on the quantitative research design and secondary data analysis research method. The demographic characteristics of the sample are provided, followed by a description of the SACII-X and biographical questionnaire, used as data collection measures in the original study. Hereafter, the research procedures employed in the present study is provided. A detailed discussion of the Rasch measurement model as data analysis method is provided. Lastly, ethical considerations in the present study are outlined.

Chapter 5

Chapter 5 presents the study results. The chapter provides information on the overall fit of the data to the Rasch model, and the response category functioning of the SACII-X. Results for the dimensionality of each SACII-X scale is presented, as well as the local independence of the SACII-X items. This is followed by a presentation of item fit and person

Stellenbosch University https://scholar.sun.ac.za

fit, and the person-item targeting of the SACII-X scales. Hereafter, results from differential item functioning analysis are presented. Lastly, the effect of item removal on the fit of the data to the Rasch model is presented.

Chapter 6

Chapter 6, the final chapter of this thesis, provides a discussion and interpretation of the research results. A discussion of the results for each objective is provided. Hereafter, implications of the findings for theory, research, and practice are provided. The chapter concludes with limitations of the present study, and recommendations for practice, and future research.

1.7 Chapter Summary

Interest assessment is an important component of career guidance and counselling. The SACII-E is an interest inventory underpinned by Holland's theory that was recently developed to address the lack of validity of Holland's model in the South African context. However, as language is a significant moderator of test performance, it is crucial to allow test-takers to be assessed in their home language or language of choice. Subsequently, the SACII-X was developed to provide test-takers with this opportunity. This study aimed to determine the psychometric properties and item functioning of the SACII-X in a sample of home-language isiXhosa-speaking adolescents, to inform improvements to the SACII-X. The rationale of the study was presented, followed by the aims and hypotheses, and a definition of key terms relevant to the study. In the ensuing chapter, the theoretical origins of the SACII are discussed.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 2

THEORETICAL ORIGINS OF THE SACII-E AND SACII-X

2.1 Introduction

Career theories provide vocational psychology practitioners and researchers with a framework within which to understand, explain, and predict vocational choice and behaviour (Watson & Stead, 2017). Holland's (1985, 1997) theory of vocational personalities and work environments is a prominent career theory in the vocational psychology field (Spokane et al., 2017). Both the English and isiXhosa versions of the South African Career Interest Inventory (SACII) are informed by Holland's (1985, 1997) theory of vocational personalities and work environments. In this chapter, an outline of Holland's theory is presented. A brief discussion of the Person-Environment (P-E) fit approach to vocational counselling is firstly presented, followed by an overview of Holland's theory, including its foundational principles, and primary and secondary assumptions.

2.2 Person-Environment Fit in Vocational Psychology

Numerous career theories form part of the P-E fit approach in vocational psychology, such as Holland's theory of vocational personalities and work environments (1985, 1997), the Theory of Work Adjustment (Dawis, 2005; Dawis & Lofquist, 1984), Schneider's (1987) Attraction-Selection-Attrition Framework, and the Model of Person-Organisation Fit (Chatman, 1989; 1991). Of these, Holland's theory can be considered as the most well-known (Glosenberg et al., 2019; Spokane et al., 2017). This P-E fit approach to career counselling examines the degree to which an individual's values, personality characteristics, and aspirations match their vocational environment (Glosenberg et al., 2019). The P-E fit approach is underpinned by the assumption that the extent of correspondence between the personality and work environment will determine work-related outcomes (Glosenberg et al., 2019). The P-E fit approach developed out of the trait-and-factor approach to career

Stellenbosch University https://scholar.sun.ac.za

counselling (Su et al., 2015). Whereas the trait-and-factor approach assumes that an individual's traits and vocational environments are relatively stable over time, the P-E fit approach assumes that there is a reciprocal interaction between an individual's traits and vocational environments, which may lead to change over time (Watson & Stead, 2017). Some view the P-E fit approach's conceptualisation of career choice as a developmental process, and its dynamic and reciprocal matching of individuals' traits and career environments as an improvement upon the more static and time-specific matching of the trait-and-factor approach (Su et al., 2015). Within the P-E fit approach, people are therefore viewed as having the agency to both choose their environments, and change environments to increase fit (Su et al., 2015).

Three main assumptions underlie P-E fit theories (Rounds & Tracey, 1990). Firstly, individuals will look for and create occupational environments in which they can express their unique personality traits. Secondly, the degree to which a person fits their vocational environment, has consequences. A person who fits their environment well will experience positive work-related outcomes, such as vocational satisfaction and success, increased productivity, lower levels of stress, and behavioural stability, with the opposite being true for a person who displays poor fit to the vocational environment (Rounds & Tracey, 1990). Lastly, this model maintains that the relationship between the person and the environment is a dynamic and reciprocal process, where both the person and the environment exert an influence on each other with individuals being able to actively shape their vocational environments and environments also shaping individuals (Rounds & Tracey, 1990).

Holland's theory of vocational personalities and work environments originated out of the P-E fit approach (Su et al., 2015). His theory conceptualises and measures P-E fit in terms of congruence, indicating the degree of fit between a person's personality code and work

Stellenbosch University https://scholar.sun.ac.za

environment code (Su et al., 2015). A discussion of Holland's theory is subsequently presented.

2.3 Holland's Theory of Vocational Personalities and Work Environments

John Holland's (1985, 1997) theory of vocational personalities and work environments is viewed as one of the most significant contributors to the field of vocational psychology (Nauta, 2010). Interest inventories informed by his theory have been at the forefront of vocational interest assessment (Nauta, 2010). The popularity of his theory results from its simplicity and user-friendliness, its organising power, the numerous psychometric instruments originating from the theory, and its empirical testability (Gottfredson & Johnstun, 2009; Nauta, 2010). The completion of interest inventories based on Holland's (1997) theory provides the test-taker with a three-letter code known as a Holland code, categorising the test-taker according to the three vocational interest domains they show the closest resemblance to with respect to how high they score on the six interest domains (Sheldon et al., 2020). This Holland code can subsequently be used to find careers that match the individual's interest code, based on the categorisation of careers according to Holland's vocational environment types (Sheldon et al., 2020). A basic tenet of Holland's theory is that a good match between the vocational personality and vocational environment will result in positive work-related outcomes.

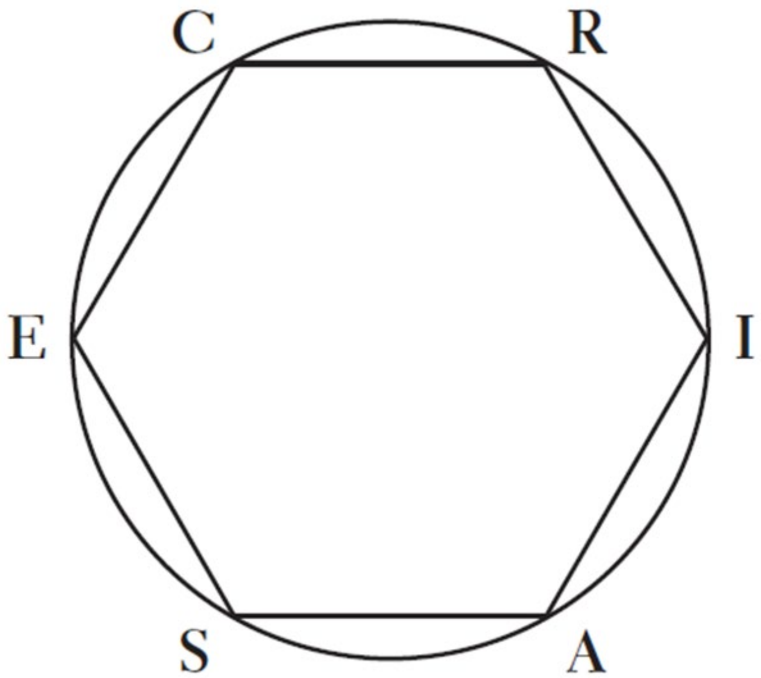
Holland's (1997) theory categorises individuals' vocational personalities, and vocational work environments, into six types, namely: Realistic (R); Investigative (I); Artistic (A); Social (S); Enterprising (E); and Conventional (C), commonly abbreviated as the RIASEC model. Holland proposed that the relationship between the six vocational personality, and vocational environment, types can be represented in a hexagonal form (see Figure 2.1; Darcy & Tracey, 2007; Holland, 1997; Nauta, 2010). Furthermore, one of the main ideas on which Holland's theory rests is the concept of congruence. That is, a high

Stellenbosch University https://scholar.sun.ac.za

degree of similarity between an individual's vocational interests and vocational environment will lead to greater job satisfaction, stability, and success. Holland's theory rests on six foundational principles, and nine theoretical assumptions – four primary assumptions, and five secondary assumptions. In the ensuing section, these principles and assumptions are discussed.

Figure 2.1.

Holland's Model Represented in a Hexagonal and Circular Order Form



Note. From "The Structural Validity of Holland's R-I-A-S-E-C Model of Vocational Personality Types for Young Black South African Men and Women", by R. du Toit and G. P. de Bruin, 2002, *Journal of Career Assessment, 10*(1), p. 62-77. Copyright 2002 Sage Publications.

2.3.1 Foundational Principles of Holland's theory

Holland's first foundational principle states that a person's choice of career is an expression of their personality (Holland, 1997). Secondly, Holland postulated that interest inventories can be considered as personality inventories (Holland, 1997). That is, by

Stellenbosch University https://scholar.sun.ac.za

measuring a person's career interests, one can get a sense of a person's personality, which can be used to suggest suitable career paths. Support for the conceptualisation of RIASEC types as personality traits has been shown by correlating the RIASEC types with the Big Five personality traits (Barrick et al., 2003; Larson et al., 2002; Nauta, 2010; Tsotetsi, 2020). Further support for Holland's postulate is demonstrated by the RIASEC interest profiles showing stability over time (Low et al., 2005; Nauta, 2010), and that vocational interests, as measured by Holland's RIASEC types, and personality traits show similar self-other agreement correlations (Nauta, 2012). The third principle posits that stereotypes relating to certain vocations have social and psychological meanings (Holland, 1997). Fourthly, individuals who work in similar careers, and thus similar vocational environments, will tend to have similar personal development histories and personalities (Holland, 1997). Holland's (1997) fifth principle states that, since people in the same career group have similar personalities, their reactions to problems and situations will be similar. These individuals will further create interpersonal environments that correspond to their similar personalities. Lastly, Holland posits that congruence between an individual's personality and vocational environment, will lead to career satisfaction, and success (Holland, 1997). A more detailed discussion of the congruence hypothesis is presented later in this chapter.

2.3.2 Primary Assumptions of Holland's theory

Holland (1997) proposed four primary theoretical assumptions, which are subsequently discussed.

2.3.2.1 Vocational Personality Typology. Holland's first primary assumption states that individuals can be categorised as resembling six vocational personality types, namely Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. Holland proposed that each personality type has unique preferences, abilities, self-evaluations, and belief

Stellenbosch University https://scholar.sun.ac.za

systems (Gottfredson & Johnstun, 2009; Nauta, 2010). Below follows a brief description of each personality type.

The Realistic Personality Type. Individuals who score high on the Realistic personality type prefer activities that involve working with animals, objects, or tools (Glosenberg et al., 2019; Holland, 1997). These individuals will therefore be attracted to Realistic vocational environments that provide them with the opportunity to exercise their technical, practical, or mechanical abilities (Glosenberg et al., 2019; Holland, 1997).

The Investigative Personality Type. Persons who score highly on the Investigative personality type prefer scientific activities that involve the investigation of cultural-, biological-, or physical phenomena to acquire knowledge and skills (Glosenberg et al., 2019; Holland, 1997). These individuals will be attracted to careers that involve mathematics and science. The investigative person can be described as precise, logical, analytical, intellectual, and curious (Chen & Simpson, 2015).

The Artistic Personality Type. This personality type enjoys creative, ambiguous, self-expressive, and unsystematic activities (Glosenberg et al., 2019; Holland, 1997). Individuals who score highly on this personality type enjoy occupations that involve creativity, music, the arts, and literature (Holland, 1997). The artistic individual is likely to be emotional, introspective, nonconforming, expressive, and intuitive (Chen & Simpson, 2015; Holland, 1997).

The Social Personality Type. Individuals with a strong Social personality type prefer activities that involve caring, training, informing, counselling, and supporting other persons (Holland, 1997). These individuals enjoy vocational environments in which there are regular social and interpersonal interactions, and show an aversion to Realistic careers (Holland, 1997). They may be described as friendly, warm, patient, and empathetic (Chen & Simpson, 2015).

Stellenbosch University https://scholar.sun.ac.za

The Enterprising Personality Type. Persons with an Enterprising personality type use their extroverted, self-confident, ambitious, and energetic personalities, together with their good interpersonal- and persuasive skills, to achieve their economic and organisational goals (Chen & Simpson, 2015; Holland, 1997). These individuals may thrive in leadership positions and in a business environment (Holland, 1997).

The Conventional Personality Type. Lastly, the Conventional personality type enjoys activities that are systematic, structured, and precise, and shows a dislike towards ambiguous and unstructured activities. These individuals are thus drawn to a Conventional vocational environment that involves accounting-, clerical-, business system-, and computing activities – thus systematically manipulating data. Persons who score high on this personality type value rules, precision, structure, and order (Gottfredson & Johnstun, 2009).

2.3.2.2 Vocational Environment Typology. Vocational environments can also be categorised according to their resemblance to the six RIASEC types. These environments include different vocations, university degrees, and also other individuals (Gottfredson & Johnstun, 2009). A brief discussion of each vocational environment type is subsequently given.

The Realistic Vocational Environment. The Realistic vocational environment provides individuals with opportunities to exercise their practical-, technical-, and mechanical abilities. These environments, for example, provide individuals with opportunities to work outside, with manual- and mechanical objects, or with animals. Careers in the Realistic environment include farming, animal training, and engineering (O*Net, 2020).

The Investigative Vocational Environment. The Investigative vocational environment provides persons with the opportunity to investigate cultural-, biological-, or physical phenomena. The environment will enable individuals to gather knowledge, and employ and

Stellenbosch University https://scholar.sun.ac.za

develop their scientific and analytical abilities (O*Net, 2020). Laboratory technologists and scientists can be found in the Investigative environment (O*Net, 2020).

The Artistic Vocational Environment. Individuals in the Artistic vocational environment are able to engage in activities that are aesthetic, liberating, and unsystematic. They will further be able to express their innovative and creative competencies. Examples of careers in this work environment include acting, journalism, and interior designing (O*Net, 2020).

The Social Vocational Environment. Individuals working in the Social environment will be provided with the opportunity to help, support, teach, cure, and counsel others (Holland, 1997). They will be able to express their empathetic and caring personalities. Careers such as nursing, teaching, and psychology form part of the Social environment (O*Net, 2020).

The Enterprising Vocational Environment. The Enterprising vocational environment will afford persons the opportunity to exercise their energetic, optimistic, resourceful, and manipulative personality traits to achieve economic or organisational goals (Holland, 1997). Real estate agents, retail sales workers, and chief executive officers form part of the Enterprising career environment (O*Net, 2020).

The Conventional Vocational Environment. The Conventional environment will provide the opportunity to systematically manipulate data (Holland, 1997). Individuals are able to express their systematic, conforming, and methodical personality traits. Careers such as office clerks, quality controllers, librarians, and accountants form part of the Conventional vocational environment (O*Net, 2020).

2.3.2.3 Person-Environment Interaction. The third assumption states that people will look for environments that correspond with their personality and values, consequently enabling them to perform their skills and capabilities and to express their unique belief system

Stellenbosch University https://scholar.sun.ac.za

(Holland, 1997). Holland (1985) proposed that the vocational environment will reward individuals for expressing values and behaviours that mirror the environment and will subsequently reinforce these values and behaviours. For example, an individual who scores highly on the Investigative personality type, will seek out an Investigative vocational environment as this career environment will reward the individual for expressing their value system, personality traits, and investigative competencies, and will consequently reinforce the individual's characteristics.

2.3.2.4 Person-Environment Interaction Determines Behaviour. Holland's final primary assumption posits that the interaction between the personality and environment determines behaviour. For example, if there is congruence between the individual's personality and vocational environment, the individual will exhibit stable behaviour (Holland, 1997). In contrast, if the individual's personality and vocational environment are incongruent, the individual may change their behaviour to fit into the environment, change the environment, or leave the environment (Holland, 1997).

2.3.3 Secondary Assumptions of Holland's Theory

Holland (1985, 1997) further proposed five secondary theoretical assumptions, namely (a) calculus; (b) consistency; (c) differentiation; (d) congruence; and (e) identity. He proposed that these assumptions influence career decision making and work-related outcomes.

2.3.3.1 Calculus. Holland proposed that vocational personality types and work environments can be represented spatially in a hexagonal structure adhering to the calculus assumption (Darcy & Tracey, 2007; Nauta, 2010). This assumption states that "the distances among the types or environments are inversely proportional to the theoretical relationships between them" (Holland, 1997, p. 5). This means that adjacent types on the hexagon are theoretically more related to each other than alternate types, which are more related than opposite types (Morgan et al., 2015). For instance, in practice, this would mean that the

Stellenbosch University https://scholar.sun.ac.za

Conventional personality type is more similar to the Enterprising and Realistic personality types, than the Artistic personality type, which is situated on the opposite side of the hexagon. The same reasoning can be applied to Holland's vocational environment types.

Research has generally shown support for Holland's calculus assumption (Darcy & Tracey, 2007; Nauta, 2010). However, less support has been shown for the equal distances constraint of the sides of the hexagon (Nauta, 2010). Hence, some scholars now refer to Holland's (1997) hexagonal model as a circumplex model (Darcy & Tracey, 2007; Morgan et al., 2015; Nauta, 2010), or a circular ordering model (Morgan et al., 2015). A circumplex model (see Figure 2.1) means that the vocational environment and personality types can be arranged in a circular form, but the distances (and thus the theoretical relationships) between the types do not have to be equal (Morgan & De Bruin, 2018).

2.3.3.2 Consistency. Consistency refers to the relatedness of the personality and environment types (Holland, 1985, 1997). It measures how similar the personality types within a person's RIASEC profile are (Hirschi & Läge, 2007). Considering Holland's arrangement of the vocational personality and environment types in a hexagonal form, types that are in closer proximity to each other, such as adjacent types, will be more similar, and thus more consistent with each other, compared to non-adjacent or opposite types (Nauta, 2010; Nauta & Khan, 2007). Holland (1997) contends that individuals with more consistent personality profiles will have clearer vocational identities, find it easier to make vocational decisions, show stability in career behaviour, and show career satisfaction (Nauta, 2010; Nauta & Khan, 2007). However, research on the consistency hypothesis has produced equivocal findings (Nauta, 2010; Nauta & Khan, 2007; Tracey et al., 2014). As a result, Holland placed less emphasis on this construct in the 1985 and 1997 versions of his theory (Nauta, 2010). Instead, more emphasis is placed on the differentiation hypothesis in the interpretation of assessment scores. The differentiation hypothesis is subsequently discussed.

Stellenbosch University https://scholar.sun.ac.za

2.3.3.3 Differentiation. Differentiation indicates how different or dissimilar the six scores of an individual's RIASEC interest profile are (Arnold, 2004). A well differentiated interest profile is defined by one dominant interest type, and a large difference in scores between the individual's strongest and weakest interest profile (Nauta & Khan, 2007). An individual who resembles multiple interest types relatively equally will therefore have an undifferentiated profile (Holland, 1997; Nauta & Khan, 2007). Holland posits that individuals with a high degree of RIASEC differentiation will find it easier to make vocational choices and will make more predictable career decisions than those with undifferentiated profiles (Gottfredson & Johnstun, 2009; Nauta, 2010). The same hypothesis can be applied to vocational environments, with Holland hypothesising that consistent and differentiated environments will produce more satisfied and productive employees (Nauta, 2010). Research has, however, failed to produce conclusive evidence in support of this hypothesis (Nauta, 2010; Nauta & Khan, 2007).

2.3.3.4 Congruence. Holland's congruence hypothesis is one of the most widely researched of the secondary assumptions (Hirschi & Läge, 2007; Nauta, 2010). Holland's congruence assumption stipulates that a good degree of fit between an individual's vocational personality and work environment will lead to beneficial outcomes, such as vocational satisfaction, success, and stability (Holland, 1997). For example, an individual who scores highly on the Investigative personality type will be best suited to an Investigative vocational environment type. In contrast, a person who scores highly on the Realistic personality type, will theoretically be highly incongruent in a Social vocational environment, as these types lie on opposite sides of the hexagon. Individuals will be attracted to, and satisfied by, the vocational environment that corresponds to their personality as the environment will provide the individual with opportunities to showcase their competencies and skills, and express their value-system (Gottfredson & Johnstun, 2009; Holland, 1997). The environment will further

Stellenbosch University https://scholar.sun.ac.za

reward and reinforce behaviour that corresponds to the behaviour that is valued in the environment. According to Holland (1997), congruency will result in behavioural stability, and people in a congruent environment will be less likely to leave that environment than people in incongruent environments (Furnham, 2001). Individuals finding themselves in an incongruent environment may leave the environment, try to change the incongruent environment, or change their behaviour to be more congruent with the environment (Holland, 1997). The congruence hypothesis has been extensively researched, and results generally support the hypothesis (Nauta, 2010). However, the size of the relationship between congruence and work-related outcomes tends to be small (Nauta, 2010).

2.3.3.5 Identity. Holland (1997) defined vocational identity as a “picture of one’s goals, interests, and talents” (Holland, 1997, p. 5). He contends that one’s vocational identity influences one’s choice of career path. Individuals with a clearly defined vocational identity may find it easier to choose an occupation and will likely choose a career environment that is congruent with their personality (Holland, 1997). Similarly, a vocational environment with a crystallised identity has a clear idea of its aims, tasks, rewards, and expectations for its workers (Gottfredson & Johnstun, 2009; Nauta, 2010). Studies on the relationship between vocational identity and outcomes such as career satisfaction, well-being, career decision-making self-efficacy, and career choice readiness, have produced mixed findings (Nauta, 2010).

2.4 Chapter Summary

In this chapter, the theory underlying the SACII-E and SACII-X was discussed. Holland’s theory originated out of the P-E Fit approach to career assessment and counselling. This approach assumes that the degree of correspondence between an individual’s work related interests and career environment, will inform career-related outcome. Holland’s theory categorises individuals’ career interests and career environments into six types, namely

Stellenbosch University https://scholar.sun.ac.za

Realistic, Investigative, Artistic, Social, Enterprising and Conventional, and assumes good fit between a person's career interests and career environment will result in career success, satisfaction, and stability. The theoretical assumptions underpinning Holland's theory has attracted much research interest, resulting in the refinement of the theory and its application in practice. The ensuing chapter includes a detailed discussion of the cross-cultural validity of Holland's theory and psychometric instruments informed by the theory.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 3

LITERATURE REVIEW

3.1 Introduction

This chapter contextualises the development of the SACII-X in South Africa, by highlighting some of the considerations that provided the impetus for the development of the instrument. Firstly, the diversity of South Africa, and English as the language of instruction, and its implications for psychological assessment is briefly discussed. Hereafter, the historical context of the practice of psychological assessment in South Africa is provided. This is followed by a discussion of the guidelines and legislation that informs fair and ethical psychological assessment practices in South Africa at present. The influence of language and culture on psychological assessment performance is highlighted through the presentation of relevant research. This is followed by a discussion of the research on the cross-cultural applicability of Holland's model and its validity in South Africa. The chapter ends of with a presentation of the available research on the SACII-E and SACII-X.

3.2 Diversity of South Africa

The diverse South African context calls for particular consideration when developing and using psychological tests (Van Eeden & Mantsha, 2007). Factors in South Africa that may moderate psychological assessment performance include culture, socio-economic status, quality of education, test-wiseness, urbanisation, home language, and English proficiency (Byrne et al., 2009; Foxcroft & Grieve, 2018; Joseph & Van Lill, 2008; Meiring et al., 2006; Swanepoel & Kruger, 2011).

This thesis focuses specifically on the issue of the influence of language in psychological testing. The unique South African multilingual society is distinguished by 11 official languages (Statistics South Africa, 2018). The 2018 General Household Survey (GHS) reported isiZulu as the most prevalent home language, spoken by approximately a quarter of

Stellenbosch University https://scholar.sun.ac.za

the South African population (25.3%; Statistics South Africa, 2018). This is followed by isiXhosa, spoken as home language by 14.8% of the population. English, according to the 2018 GHS, is only spoken as home language by 8.1% of the population, making it the sixth most prevalent home language in South Africa. It is, however, the second most prevalent language spoken outside of the household, with 16.6% of the population speaking English outside of their homes. However, according to the 2018 GHS, among the Black population group, only 1.6% speak English as home language, and only 8.6% speak English outside of the household. In comparison, 18.2% of Black South Africans are isiXhosa home language-speaking. Furthermore, isiXhosa is the second most prevalent home language in the Western Cape (24.7%; Statistics South Africa, 2011a). Despite these numbers, nearly all psychological instruments in South Africa use English as the language of administration (Health Professions Council of South Africa, 2017). This issue is further elaborated on in the subsequent section.

3.3 English as the Language of Instruction in South Africa

In South Africa, psychological tests are almost exclusively administered in English (Health Professions Council of South Africa [HPCSA], 2017; McDonald & Van Eeden, 2014), with a few measures also available in Afrikaans (HPCSA, 2017). Some regard the language of instruction of a psychological assessment instrument as having a particularly salient influence on assessment performance (Foxcroft & Aston, 2006; McDonald & Van Eeden, 2014; Nell, 2000). Consequently, the majority of the population whose home language is neither English nor Afrikaans is potentially being disadvantaged by the lack of assessment instruments available in the other nine official languages and excluded from accessing appropriate career counselling (Rabie & Naidoo, 2019).

However, in South Africa, the lingua franca in business, industry, and most educational institutions is primarily English (De Kock et al., 2013; Desai, 2016; Foxcroft, 2011; Posel & Casale, 2011; Rabie & Naidoo, 2019). As a result, numerous South Africans

Stellenbosch University https://scholar.sun.ac.za

are not educated in their home language, but in English (Desai, 2016). For instance, it is common for African home language-speaking learners to be educated in English from Grade 4 onwards (Desai, 2016). Some thus assume that it is fair to assess all test-takers in English, as the inability to understand, and converse fluently, in English, would deem the test-taker unsuitable for admission into higher education institutions or the workplace context (Foxcroft & Aston, 2006; Foxcroft et al., 2004). One can even argue that test-takers will be disadvantaged if assessed in their home language if there is a disjuncture between their home language and language of education, as learners would have acquired knowledge of constructs in English (Foxcroft & Grieve, 2018; McDonald & Van Eeden, 2014).

In contrast to this, many psychologists in practice and academia see this method of assessing test-takers in a language other than their home language as a cause for concern (Foxcroft et al., 2004; Grobler & De Beer, 2015; Van de Vijver & Rothmann, 2004). It has been indicated that non-English home language-speaking learners experience problems with English as the medium of instruction in schools (Desai, 2016), and many are not adequately proficient in English when completing school (Watson, Davies, & Foxcroft, 2006, as cited in Rabie & Naidoo, 2019). The notion by some is thus that every test-taker should be granted the opportunity to be assessed in their home language to ensure the reliability and validity of assessment results, and to adhere to fair testing practices (Albien & Naidoo, 2016; Foxcroft, 2011; Foxcroft et al., 2004). In South Africa, this implies that psychological measures should be translated into more official languages to enable all test-takers to be assessed in their home language, in turn increasing the reliability and validity of assessment outcomes (Albien & Naidoo, 2016; Foxcroft et al., 2004).

Stellenbosch University https://scholar.sun.ac.za

3.4 Ethical Testing and Test Fairness

3.4.1 History of Psychological Testing in South Africa

Psychological assessment has a contentious history in the South African context and is still a controversial area of psychology (Laher & Cockroft, 2014). A complete understanding of the development and use of psychological assessment instruments in South Africa cannot be obtained without acknowledging the complex context in which the practice developed (Foxcroft et al., 2018). Particularly, apartheid-informed practices and laws had a substantial and formative influence on the development and use of psychological instruments in South Africa (Foxcroft et al., 2018; Joseph & Van Lill, 2008).

During apartheid, the Population Registration Act of 1950 categorised the South African population based on race, namely White, Coloured (people of mixed descent), Indian, and Bantu (Black African)[3] (Finchilescu & Tredoux, 2010). These categories were used to inform practices that excluded and oppressed non-White South Africans (Finchilescu & Tredoux, 2010; Laher & Cockcroft, 2013). However, even prior to the official legalisation of apartheid, the early practice of psychological assessment in South Africa was tainted by the tendency to administer inappropriate and biased tests to groups for which it was not standardised, and consequently reaching inaccurate conclusions about group characteristics (Foxcroft et al., 2018; Laher & Cockcroft, 2014). Assessment measures were often adapted versions of imported instruments from Western contexts such as the United States of America and Europe (De Kock & Foxcroft, 2018), or indigenous measures developed for the White population (Foxcroft et al., 2018; Joseph & Van Lill, 2008). Moreover, psychological instruments were often used as a tool to validate apartheid laws by using the results obtained

---

[3] The researcher acknowledges that the use of the apartheid-informed racial terms *White*, *Coloured*, *Indian*, and *Black*, is sensitive and contentious in South Africa. In using these terms, the researcher does not intend to perpetuate their use or underpinning ideology, but uses it as a way to encapsulate the historical and current experiences, and cultural- and language contexts of the South African population.

Stellenbosch University https://scholar.sun.ac.za

from inappropriate and biased tests to indicate racial group differences (Foxcroft et al., 2018; Joseph & Van Lill, 2008; Laher & Cockcroft, 2014).

After World War II and with the official legalisation of apartheid in 1948, psychological assessment and instrument development were afforded increased focus (Foxcroft et al., 2018). This increase in the use and development of measures can largely be ascribed to the demand of the rapidly developing mining and manufacturing sectors for Black workers frequently lacking formal education (Foxcroft et al., 2018; Joseph & Van Lill, 2008). Test use and development were thus driven by a political agenda (Foxcroft et al., 2018).

During the 1960s and up until the early 1980s, the number of emic psychological tests developed in South Africa increased (Laher & Cockcroft, 2014). This can partially be ascribed to international sanctions that resulted in many imported psychological assessments not being available in South Africa (Laher & Cockcroft, 2014). In addition, during the 1970s and 1980s, the impact of culture on psychological assessment was starting to gain more attention (Bedell et al., 1999). Issues of fairness, discrimination, and bias in the field of psychological testing, thus gained more recognition (Meiring et al., 2005).

However, during apartheid, White people enjoyed job reservation and did not compete for the same jobs as other racial groups. Thus different psychological measures were developed for the different racial groups as categorised by apartheid legislation (Foxcroft et al., 2018; Laher & Cockcroft, 2014). Black people were limited to lower quality formal education compared to White people, which inter alia meant that the different population groups were restricted to different jobs (Naidoo et al., 2017). In addition, few measures were developed for the Black, Coloured, and Indian groups compared to the White group (Foxcroft et al., 2018). As a result, inappropriate measures were administered to groups without the necessary standardisation (Laher & Cockcroft, 2014). A prominent example of the inappropriate use of measures relates to intellectual assessment under the Apartheid regime

Stellenbosch University https://scholar.sun.ac.za

(Laher & Cockcroft, 2014). The inaccurate conclusions drawn from the inappropriate use of intelligence measures for groups for which it was not standardised, were used as validation for the lower quality of education afforded to Black people (Laher & Cockcroft, 2014).

In terms of career guidance and counselling specifically, a pervasive lack of appropriate career guidance and counselling existed for the majority of the South African population (Naidoo et al., 2017). During apartheid, schools were separated according to racial groups, and the approach to career guidance and counselling followed distinct routes of development in these separate schools (Naidoo et al., 2017). In schools for Black learners, career guidance was a low priority practice and often consisted of only group psychological assessment, driven by South Africa's labour force needs (Naidoo et al., 2017). Guidance and counselling practices for Black high school students were thus used as a tool to serve the needs of the apartheid government. This is in stark contrast to the practices and objectives of career guidance for White high school students (Naidoo et al., 2017). From 1981 onwards, guidance started to receive more attention and was implemented in schools. Nonetheless, career guidance in non-White schools was still not prioritised and adequately implemented (Naidoo et al., 2017) and this status quo still persists (Rabie et al., 2021).

The practice of using separate measures for the different racial groups was placed under the spotlight in the 1980s and early 1990s as apartheid laws became less strict, and as the different racial groups began to compete for the same jobs (Foxcroft et al., 2018). Research on, and the development of, psychological assessment measures increased markedly (Laher & Cockcroft, 2014), due to the dire need for suitable and unbiased tests for the multiracial and multilingual population (Joseph & Van Lill, 2008). However, as appropriate psychological instruments were not always available, imported instruments, or instruments developed specifically for White South Africans, were sometimes still used for other racial groups (Foxcroft et al., 2018).

Stellenbosch University https://scholar.sun.ac.za

The misuse of psychological tests for political purposes and as justification for apartheid (Laher & Cockcroft, 2014), resulted in the development of a negative perception of psychological assessment and its value (Foxcroft et al., 2018). After apartheid came to an end, the new ruling African National Congress (ANC) government emphasised the redress of past inequalities, including a focus on psychological test use (Foxcroft et al., 2018). Due to the discriminatory history of psychological testing, trade unions were strongly against testing (Nzimande, 1995, as cited in Laher & Cockcroft, 2014) and group testing in many schools was prohibited, while earlier versions of the Employment Equity Act (EEA) prohibited psychological assessment as a part of the selection process for employment (Foxcroft et al., 2018; Laher & Cockcroft, 2014). Following advocacy by psychologists, the Psychological Society of South Africa (PsySSA), and the HPCSA's Professional Board for Psychology, the EEA was revised allowing the use of psychological tests under certain prerequisites (Laher & Cockcroft, 2014). At present, many researchers and practitioners in South Africa are mindful to consider factors that may contribute to psychological test bias, to contribute to the redress of inequality and discrimination in the past (Laher & Cockcroft, 2014). As a result, psychological assessment is gradually gaining a more positive view among the South African population (Laher & Cockcroft, 2014).

3.4.2 Ethical- and Fair Testing Practices

It is clear from the above discussion that psychological assessment has been a contentious topic in South Africa. This history has partially informed legislation and guidelines for the ethical and fair use of psychological assessment instruments in South Africa. To redress past discriminatory practices in the field of psychological assessment, the EEA (no. 55 of 1998; Republic of South Africa, 1998), the main legislation regulating occupational assessment in South Africa (Society for Industrial and Organisational Psychology of South Africa, 2006), includes a section relating to this. The EEA strictly

Stellenbosch University https://scholar.sun.ac.za

expects that all psychological instruments are proven to be reliable and valid, is not biased against individuals or groups, and requires proof that it can be fairly used with an individual or group. In addition to the EEA, the Health Professions Act (No. 56 of 1974) also informs ethical testing practices in South Africa.

These two Acts, together with professional standards and guidelines set by professional bodies such as the American Psychological Association (APA; APA, 2017), the PsySSA (PsySSA, 2007), the Society for Industrial and Organisational Psychology of South Africa (SIOPSA; 2005, 2006), the HPCSA (HPCSA, 2006), and the International Test Commission (ITC; International Test Commission, 2001, 2005), inform ethical testing practices in South Africa at present.

All of the above-mentioned professional bodies expect practitioners to use only psychological assessments for which adequate technical information such as reliability and validity information is available, to ensure that measures used are fair and bias-free, and to be cognisant of any factors that may bias an assessment or moderate performance, such as language. Language considerations to ensure ethical and fair assessment is a golden thread throughout the different professional guidelines. The influence of language on psychological assessment is subsequently discussed.

3.4.3 Language as a Moderator of Assessment Performance

As discussed previously, language is viewed as a crucial moderator of assessment performance (Nell, 2000), which includes personality assessment (McDonald & Van Eeden, 2014; Van Eeden & Mantsha, 2007). Language can be a source of bias as the inability to sufficiently understand, or the incorrect interpretation, of words and concepts in a psychological test may result in differential item functioning [4] (DIF) (Foxcroft & Grieve,

---

[4] Differential item functioning is present when individuals with the same ability, but from different groups, do not have the same probability of getting the item right (De Kock & Foxcroft, 2018).

Stellenbosch University https://scholar.sun.ac.za

2018). Assessing individuals in a language other than their home language can therefore be problematic, as the DIF negatively influences the reliability and validity of the results obtained from psychological tests, and ultimately the equivalence of scores (Abrahams & Mauer, 1999a; Foxcroft et al., 2004). Crucially, if individuals are assessed in a language in which they are not adequately proficient, it can become difficult to determine whether assessment results are a true reflection of an individual's standing on a trait, or whether language difficulties moderated the results (Foxcroft et al., 2004). This could lead to discrimination against test-takers based on language proficiency and not on the construct being tested (Foxcroft et al., 2004).

It is therefore critical to be cognisant of language influences when assessing the suitability of a psychological test, especially in a multilingual context such as South Africa (Foxcroft & Aston, 2006). South African research has indicated that language proficiency moderates the results obtained from psychological tests such as cognitive and intelligence measures (Foxcroft & Aston, 2006; Grobler & De Beer, 2015; Koch, 2007; Schoeman et al., 2008). Notwithstanding, few studies have focused on the influence of language proficiency on the results obtained from personality- and interest assessment specifically, even though these tests are popular in South Africa (Van de Vijver & Rothmann, 2004; Visser & Viviers, 2010). Among the studies that did focus on the influence of language in personality and interest assessment, some found satisfactory equivalence across language groups (Coetzer & Rothmann, 2007; Visser & Viviers, 2010), while several studies report language bias (Abrahams & Mauer, 1999b; Grobler & De Beer, 2015; Meiring et al., 2006).

For instance, Abrahams and Mauer (1999b) studied the impact of home language on the results obtained on the 16 Personality Factor Questionnaire (16PF) and reported that test-takers who were not English home language-speaking found it difficult to understand many of the simple English words in the 16PF that would easily be understood by English home

Stellenbosch University https://scholar.sun.ac.za

language-speakers. Similarly, Meiring et al. (2005) found that African language-speaking test-takers struggled to understand some of the English language used in the 15 Factor Questionnaire Plus (15FQ+) resulting in construct bias across different population groups. In response to this finding, Meiring et al. (2006) adapted the 15FQ+ and investigated participants' experience with the language used in the adapted version. They found that test-takers with home languages other than English still found it problematic to understand some of the words in the adapted 15FQ+ (Meiring et al., 2006). Furthermore, Schaap et al. (2003) as well as Berg et al. (2004) partly attributed group differences in the Locus of Control Inventory to the fact that many participants did not complete the measure in their home language. In Grobler and De Beer's (2015) investigation of item bias in the Basic Traits Inventory (BTI) across the 11 official South African languages, they noted that some items in the BTI showed differential interpretation by test-takers from the different language groups. This points to language being a possible moderating factor in this study.

Within this context it is critical to assess test-takers in their home language to ensure fair and ethical assessment (Foxcroft et al., 2004). Nonetheless, in a linguistically diverse context such as South Africa, the decision of what the most suitable language of assessment is, is challenging (Foxcroft et al., 2004). South Africa's 11 official languages make test translation difficult (Van Eeden & Mantsha, 2007). In addition, some languages lack the words for the concepts used in some psychological instruments, or equivalent expressions (Foxcroft & Grieve, 2018).

3.4.3.1 Language and Culture. The potential bias caused by language difficulties is further compounded by the fact that language and culture are intertwined (Foxcroft et al., 2004). Culture has a salient influence on behaviour and the meaning attached to events (Foxcroft & Grieve, 2018), which can influence career decisions (Albien & Naidoo, 2016). The South African population is ethnically and culturally diverse (Statistics South Africa,

Stellenbosch University https://scholar.sun.ac.za

2011a). This cultural diversity complicates the use of imported Western vocational theories in South Africa, as it may not adequately take culture-specific beliefs and values that influence a person's vocational interests and decision-making processes into account (Albien & Naidoo, 2016).

Test developers approach psychological instrument development through a specific cultural lens, and item content is thus a reflection of the culture of the test developers (Foxcroft & Grieve, 2018), irrespective of the measure's language (Van Eeden & Mantsha, 2007). Some contend that culture is an especially important moderating factor in personality assessment (Foxcroft et al., 2004; Wallis & Birt, 2003). Imported measures with item content that is not applicable to the culture in which it is to be used poses a threat to fair assessment (Foxcroft & Grieve, 2018). Accordingly, it is important to ensure that personality- and interest tests based on a Western model of personality, such as Holland's theory, are relevant to other cultures in a multicultural society (Van Eeden & Mantsha, 2007). A simple translation of a test may not be sufficient to neutralise bias. Test adaptation to ensure that test items are applicable and relevant to the target culture is required (De Kock & Foxcroft, 2018). Even though developing culture-free measures is practically impossible, test developers can be cognisant of cultural bias and fairness in assessment measures (Foxcroft & Grieve, 2018). Given this context, the development of the SACII has made a significant contribution to the practice of career counselling in South Africa. Further research on how test-takers respond to the items in a test adapted specifically for their cultural and linguistic group, such as the SACII-X, can make a further valuable contribution to the refinement of the instrument.

3.5 The Cross-Cultural Applicability of Holland's Model

Holland's hexagonal/circular order model has been a research focus since the 1960s in numerous countries (Hedrih et al., 2018; Morgan et al., 2015a; Nagy et al., 2010; Šverko & Pedović, 2018). Tests based on Holland's model have been adapted for various cultural

Stellenbosch University https://scholar.sun.ac.za

contexts and are available in numerous languages such as, English, Icelandic, Croatian, Irish English, Serbian, German, Spanish, Japanese, Chinese, Korean, French, and isiXhosa. Studies on the model predominantly focus on the construct validity of interest inventories based on Holland's model (Hedrih et al., 2018).

3.5.1 Validity of Holland's Model Internationally

The structural validity of Holland's model has been especially well-researched in the USA (Nagy et al., 2010; Vardarlı et al., 2017). Most studies have demonstrated fit of the circular order model among diverse samples (Armstrong et al., 2003; Gupta et al., 2008; Tracey & Robbins, 2005). Kantamneni and Fouad (2011), however, found a lack of fit among samples of African American females and Latino individuals. Wilkins et al. (2013) further demonstrated fit in Jamaica, and Trinidad and Tobago.

Structural validity studies in countries outside of the United States have produced varying results (Hedrih et al., 2018). Two meta-analyses by Tracey and Rounds (1993, 1996) demonstrate the disjuncture in the structural validity of Holland's model between samples from the United States and samples from other countries. Tracey and Rounds (1993) studied 104 correlation matrices published between 1965 and 1989 originating from numerous countries. The structure of vocational personalities showed satisfactory fit to Holland's circular order and circumplex models in these correlation matrices (Tracey & Rounds, 1993). However, significantly better fit to Holland's model was found in correlation matrices from the United States, than in correlation matrices obtained outside of the United States (Tracey & Rounds, 1993). In contrast to the adequate fit found by Tracey and Rounds (1993), Rounds and Tracey (1996) compared international matrices from 18 countries and matrices from ethnic minorities in the United States to 73 United States benchmark matrices. Rounds and Tracey (1996) found a lack of support of Holland's model among the international and ethnic

Stellenbosch University https://scholar.sun.ac.za

minority matrices, with significantly poorer fit found in these matrices compared to the benchmark matrices (Rounds & Tracey, 1996).

In Europe, Holland's model has also been the subject of numerous studies (Tien, 2009). Most of these studies found fit of the data to Holland's model. Support for Holland's model has been shown in Croatia (Šverko & Babarović, 2006; Šverko, 2008), Ireland (Darcy, 2005), Germany (Nagy et al., 2010), and Serbia (Hedrih, 2008). Studies in Iceland demonstrated equivocal findings with Einarsdóttir et al. (2002) indicating support for the model, while Einarsdóttir et al. (2010) found a lack of support. A lack of support for the model was also indicated among a Spanish sample (Elosua, 2007).

In Asia, fit to Holland's model has been indicated in Japan (Long et al., 2006) and Taiwan (Tien, 2009). In China, fit has been indicated by Long et al. (2005) and Zhang et al. (2013), and a lack of fit has been found in studies by Tang (2001), and Leung and Hou (2005). Partial fit has been found in Korea (Tak, 2004), China (Law et al., 2001; Tang, 2009; Wong & Wong, 2006; Yang et al., 2005), and Taiwan (Sung et al., 2016).

Morgan and De Bruin (2018) investigated the structural validity of Holland's circumplex model in Africa, using the Strong Interest Inventory (SII), and obtained mixed results. The inventory was administered to 599 adolescents and young adults above 18 years of age in 28 countries, divided into three regions namely, Southern Africa, Eastern Africa, and Western Africa. Support for the model was indicated in all three regions, but the fit in the Eastern and Western regions was lower than the USA fit indicated by Rounds and Tracey (1996), and the fit in the Eastern African region was less satisfactory. Morgan and De Bruin (2018) theorise that since the SII is an imported measure, the item content of the inventory may not be appropriate to the cultural, and economic and labour context in the different regions, providing a reason for the differential fit found across the regions. Furthermore, English is not the primary language spoken in some of these countries (Morgan & De Bruin,

Stellenbosch University https://scholar.sun.ac.za

2019). Providing support for the suggested reasons for poorer fit, is the fact that Morgan et al. (2015a, 2015b) found better fit to Holland's model in a South African sample, when using an interest inventory developed for the country's cultural- and language context.

Morgan et al. (2015a) and Tracey et al. (2008) suggest further possible reasons for the equivocal findings internationally. Tracey et al. (2008) contend that varying results may be due to the different methods, and methods of analyses used. Morgan et al. (2015a) further highlight that most of the cross-cultural studies of Holland's model have used instruments imported from the United States. Even though these instruments are mostly translated, and in some cases adapted, they are not necessarily appropriate for a country's culture and work environment, as the different cultural and work contexts can influence the meaning attached to activities (Morgan et al., 2015a).

3.5.2 Validity of Holland's Model in South Africa

In South Africa, Du Toit and De Bruin (2002) administered the English version of the Self-Directed Search (SDS) to a large sample of over 1,400 Black South Africans who spoke two indigenous African languages as home language, seTswana and isiXhosa, and found poor fit to Holland's circular order model. Du Toit and De Bruin (2002) ascribe the lack of fit to the administration of the English version of the instrument to samples who are not English home language-speaking, cultural differences in participants' beliefs about vocations, and their beliefs about the role of interests in selecting a career.

In response to the poor structural validity of Holland's model in South Africa specifically, Morgan et al. (2015a) developed an indigenous South African interest inventory based on Holland's RIASEC model – the SACII. The SACII was developed using a constrained emic approach, to ensure applicability of the items in the inventory to the South African context (Morgan et al., 2015a). The development of this inventory has provided momentum for research on the validity of Holland's model in South Africa, with promising

Stellenbosch University https://scholar.sun.ac.za

initial results obtained. Morgan et al. (2015a) examined the structural validity of the SACII-E among multiracial samples of approximately 1,100 adolescents, young adults, and adults, from three higher education institutions, and found support for the circular structure. Nonetheless, it is important to note that their data did indicate that indigenous African home language-speaking participants showed less-promising results than those obtained from English home language-speaking participants (Morgan et al., 2015a).

Rabie and Naidoo (2019) developed the SACII-isiXhosa (SACII-X) to address the limitation of the SACII only being available in English and Afrikaans. Research on the SACII-X has shown promising results in terms of the inventory's reliability and structural validity in the South African context. Rabie and Naidoo (2019) analysed the results obtained from a sample of 266 isiXhosa home language-speaking Grade 9 learners and demonstrated the reliability and validity of the SACII-X in this sample. The promising results found in Morgan et al. (2015a) and Rabie and Naidoo (2019) demonstrate the usefulness of Holland's model as a depiction of vocational interest types in South Africa, if psychometric assessments that take the South African context into account, are used.

3.6 Research on the SACII and SACII-X in South Africa

Additional research to that of Morgan et al. (2015a) and Rabie and Naidoo (2019) on the psychometric properties of the SACII-E and SACII-X, is limited. While three additional studies investigated the psychometric properties of the SACII-E, no other research, other than Rabie and Naidoo (2019), has been conducted on the SACII-X. Moreover, these studies have largely employed CTT methodology in their investigations of the inventory's psychometric properties, and have focused mainly on the inventories' reliability and structural validity.

In Morgan et al.'s (2015a) study, they partially used the Rasch measurement model to reduce the 237-item SACII-E down to 167 items, by identifying items showing poor fit to the data, in a sample of 985 adolescents, young adults, and adults. Hereafter, the reduced version

Stellenbosch University https://scholar.sun.ac.za

of the SACII-E was administered to a sample of 175 adolescents, young adults, and adults, and better fit to the circular order model was obtained with the reduced SACII-E version than the original version (Morgan et al., 2015a).

Morgan and De Bruin (2019) provided support for the structural validity and reliability of the SACII in South Africa. This study indicated satisfactory reliability across the six scales, and fit to a circumplex model among a sample of 1,000 primarily Black university students, using a brief version of the SACII (SACII-SR). However, Morgan and De Bruin (2019) note that these satisfactory findings may be because they obtained data from the SACII database, collected during the initial development of the SACII, and chose items that would increase the probability of producing data that would fit the circumplex model. To address this limitation, Morgan and De Bruin (2019) subsequently investigated the psychometric properties of the SACII-SR in a sample not obtained from the SACII database. They found satisfactory reliability across the six scales, and support for the circumplex model among a sample of 183 Black and White university students, with the majority speaking English as home language (Morgan & De Bruin, 2019).

The development of the SACII-SR partially employed the Rasch (1960) model to identify poor-fitting items in the 169 item-version of the SACII-E. In the Rasch analysis of the 169-item version of the SACII-E, Morgan and De Bruin (2019) found 29 items demonstrating underfit to the model, and 5 items demonstrating overfit to the model. Morgan and De Bruin (2019) classified infit and outfit MnSq values >1.30 as underfitting and values <.70 as overfitting.[5]

Three studies focused specifically on gender equivalence and/or gender-based DIF, with satisfactory gender equivalence demonstrated. Utilising the participant database used in

---

[5] A discussion of infit, outfit, overfit, and underfit is provided in Section 4.7.10.

Stellenbosch University https://scholar.sun.ac.za

the original development of the SACII-E, Morgan et al. (2015b) found that Holland's circumplex model fits equally well for men and women in a sample of 1,000 adolescents (500 men and 500 women), young adults, and working adults. Mintram et al. (2019) investigated the fit and structural equivalence of the model in a sample of 139 men and 268 women made up of mostly Black, university students. They found fit to Holland's model, and also demonstrated the structural similarity between genders in this sample (Mintram et al., 2019). Mintram et al. (2019) did find small mean score differences between genders on the different scales, but these differences mostly mirror those obtained in international research. The study sample is, however, small, and only from two university faculties. Mintram et al. (2019), using a two-way ANOVA analysis, found 18 items on the SACII-E demonstrating DIF across gender groups. The distribution of the items with DIF across the scales were as follows: (R) 5; (I) 3; (A) 2; (S) 2; (E) 4; and (C) 4. Wang et al. (n.d.) also used Item Response Theory methodology to investigate gender-based DIF in the items of the SACII-E. They found that gender-based DIF formed a negligible proportion of the observed gender differences in items of the SACII-E.

The item functioning of the SACII has been investigated by Morgan (2014) and Rabie (2017). Morgan (2014) used both CTT methodology and IRT methodology to investigate the item functioning of the SACII-E. He found three items on the Realistic Scale and two items on the Conventional scale to be problematic in terms of their skewness and kurtosis values. Using the Rasch model, Morgan (2014) found 18 misfitting items on the Realistic scale, eight on the Investigative scale, 10 on the Artistic scale, 10 on the Social scale, 5 on the Enterprising scale, and 11 on the Conventional scale. Rabie (2017), using Classical Test Theory methodology, a Principal Component Analysis, and arc-tangent transformation, found a total of 34 items with poorer functioning compared to the other items on the SACII-X. The distribution of these items across the SACII-X scales were as follows: (R) 1; (I) 2; (A) 8; (S)

Stellenbosch University https://scholar.sun.ac.za

4; (E) 14; and (C) 5. Rabie (2017) also identified poor functioning items using angular location analysis and communality estimates. Rabie (2017) found a total of 31 poor functioning items across the SACII-X scales, distributed as follows: (R) 0; (I) 2; (A) 7; (S) 3; (E) 13; and (C) 6.

In conclusion, preliminary studies have found support for the reliability and validity of the SACII-E and SACII-X among diverse South African samples. However, these South African studies have used mainly CTT analysis methods in their investigations, focusing on the reliability and structural validity of the SACII-E and SACII-X (Morgan & De Bruin, 2019). Few studies included an investigation of item functioning, specifically using IRT methodology. Further investigation of the psychometric properties of the SACII-X using Item Response Theory methodology, instead of Classical Test Theory methodology, and with a specific focus on the item functioning of the SACII-X, makes a valuable contribution to the incorporation of the SACII-X into the career counselling field.

3.7 Chapter Summary

This chapter demonstrates the importance of the current study by grounding the study within the relevant literature. Numerous factors in South Africa can moderate assessment performance, with language being a particularly significant moderator. Despite the majority of the South African population not speaking English or Afrikaans as home language, these form the language of administration of the majority of psychological measures available in South Africa. Given this, in addition to the contentious history of psychological assessment in South Africa, as well as the ethical guidelines informing psychological assessment practices, the need for valid, and culturally and linguistically appropriate vocational interest assessment, can be clearly deduced. The development of the SACII and SACII-X has made great strides towards achieving fair assessment in the South African context. These instruments have shown promising results in terms of their reliability and validity in the South African context.

Stellenbosch University https://scholar.sun.ac.za

In the subsequent chapter, the IRT methodology employed in this study to investigate the psychometric properties and item functioning of the SACII-X, is discussed in detail.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 4

METHODOLOGY

4.1 Introduction

This chapter presents the methodological considerations of the present study. The quantitative research design is discussed, followed by a description of secondary data analysis as research method. Hereafter, a description of the study sample is provided. The development and psychometric properties of the SACII-X, as the measure used in the study, are discussed. Next, details on the research procedure are presented, followed by the data analysis procedures employed in the study. The chapter concludes with a discussion of ethical considerations.

4.2 Aims and Hypotheses

The overarching aim of this study was to investigate the psychometric properties of the SACII-X, identify potential improvements to the SACII-X, and investigate differential item functioning across gender groups, among a sample of isiXhosa home language-speaking secondary school learners. To achieve this, the present study employed quantitative methodology to meet the following objectives:

- To determine the reliability and validity of the SACII-X items through the analysis of the adherence of the data to the Rasch model assumptions and the fit of the data to the Rasch model.
- To investigate the item functioning of the SACII-X using Rasch analysis to identify problematic items that should be adapted or removed.
- To investigate differential item functioning in the SACII-X, across gender groups.

The present study had three hypotheses:

Stellenbosch University https://scholar.sun.ac.za

H1: The Rasch measurement model will substantiate the SACII-X to be reliable and valid.

H2: The Rasch measurement model will identify improvements to items across all six scales of the SACII-X.

H3: Differential item functioning across gender groups will be indicated in some of the SACII-X items.

4.3 Research Design

This study utilised a quantitative research design (Bryman, 2016). Quantitative research is suitable in studies where the testing of hypotheses is prominent (Bless et al., 2013), and involves the collection and analysis of numerical data (Bryman, 2016). Results are objectively examined and interpreted in terms of numbers and statistics, with the aim of generalising from the sample to the larger population (Bless et al., 2013). A quantitative approach was thus applicable in this study, as the researcher wanted to test particular hypotheses and relied on statistical data to achieve the study aims. Further, this study can be categorised as a psychometric study as it aimed to investigate the psychometric properties of a psychological assessment (De Bruin & Buchner, 2010). The research method in this study involved the analysis of secondary data. The researcher utilised cross-sectional data that were collected by Rabie (2017) to investigate different research aims, to achieve the new aims of this study (Bless et al., 2013).

4.4 Study Sample

The study sample comprised of 266 isiXhosa home language-speaking Grade 9 learners from two secondary schools in Kayamandi – an impoverished community in the Cape Winelands district in the Western Cape province of South Africa. Both high schools have approximately 800 learners from Grade 8 to Grade 12 (Albien, 2018). One high school has a science focus offering subjects such as physical sciences and life sciences, while the other

Stellenbosch University https://scholar.sun.ac.za

high school has a business focus, offering subjects such as accounting and economics (Albien, 2018).

This chosen sample is a subsample of 628 Grade 9 learners from five secondary schools in the Cape Winelands district that were used as the normative research sample in the development of the SACII-X (Rabie & Naidoo, 2019). The sample of 266 learners comprised the total number of isiXhosa home language-speaking learners in the normative sample.

Rabie (2017) used a convenience sampling method with universal inclusion to recruit research participants. As convenience sampling is a non-probability sampling method, it cannot be assumed that the sample is representative of the study population. All Grade 9 learners from the two secondary schools were invited to participate in the study (Rabie, 2017). As inclusion criteria, learners had to be in Grade 9 and had to provide both informed consent and assent before completion of the biographical questionnaire and SACII-X.

The data of the whole sample of 266 learners were included in the data analysis of the current study. This sample of 266 learners is composed of 100 males and 166 females, with the majority of participants identifying as Black (n = 250; 94.0%), and 16 participants identifying as Coloured (6.0%). Participants had an age range of 13 to 20 years with a mean age of 15 years ([*SD*] = 1.02). At both secondary schools, the primary language of instruction is isiXhosa, and the majority of learners are isiXhosa home language-speaking (Rabie, 2017).

4.5 Measure

4.5.1 Biographical Questionnaire

A biographical questionnaire forms part of the SACII-E and SACII-X. Questions pertain to the participant's race, gender, parents' occupations, and home circumstances.

4.5.2 SACII-X – South African Career Interest Inventory (isiXhosa)

The SACII-X is the first career interest inventory in an indigenous South African language (Rabie & Naidoo, 2019). The test was developed through a translation and

Stellenbosch University https://scholar.sun.ac.za

adaptation of the English version of the SACII developed by Morgan (2014). The first step in the development of the SACII-X encompassed a judgemental committee forward-translation to translate the English version of the SACII to the target language of isiXhosa. The translations were collaboratively done by two translators. Both translators were isiXhosa home language-speaking Master's students, one in Psychology, and one in African languages, fluent in both English and isiXhosa (Rabie, 2017). The collaboration between the two translators from different departments ensured that translations were not only linguistically and grammatically accurate, but that important career content was not lost during the translation process (Rabie, 2017). Consideration was also afforded to the South African isiXhosa-speaking population's diversity in terms of age, geographic location, and educational background (Rabie, 2017).

To overcome the challenge of the non-equivalence of some English and isiXhosa terms, some items could not be directly translated to isiXhosa. These items were presented in the SACII-X as a description of the meaning of the English concepts (Rabie, 2017). Furthermore, some items had to be rephrased and restructured to be grammatically and linguistically correct in isiXhosa (Rabie, 2017).

Upon completion of the forward translation, an independent home-language isiXhosa-speaking Master's African Languages student back-translated the newly translated SACII-X into English. The equivalence of the SACII and SACII-X was determined by comparing the original SACII with the back-translated English version of the SACII-X. The translation process was repeated until the original and back-translated versions showed a high degree of correspondence (Rabie, 2017). As a final quality control procedure, the SACII-X was professionally edited to ensure the quality of the language used, and the accuracy of the translation (Rabie, 2017).

Stellenbosch University https://scholar.sun.ac.za

The SACII-X is composed of six scales corresponding to the six RIASEC personality types, with a total of 142 items (Rabie & Naidoo, 2019). The six scales each contain the following number of items: (R) 23; (I) 21; (A) 26; (S) 22; (E) 25; and (C) 25. Rabie and Naidoo (2019) reported acceptable Cronbach alpha reliability coefficients of >.85 in the sample of 266 isiXhosa home language-speaking secondary school learners. The SACII-X further displayed good construct validity in this sample across its six scales (Rabie & Naidoo, 2019).

4.6 Research procedure

The researcher obtained ethical clearance to conduct the study from Stellenbosch University's Research Ethics Committee (REC; PSY-2021-23176; Appendix A). The original study also obtained ethical clearance from Stellenbosch University's REC (SU-HSD-000640; Appendix B) as well as from the Western Cape Education Department (WCED; Appendix C) to conduct research in schools, and from the school governing bodies and/or principals of the participating schools. Written permission to access the secondary data was obtained from the primary investigator of the original study, Rabie (2017; Appendix D). Hereafter, the de-identified data was transferred from the primary investigator of the original study to the current researcher via encrypted email.

The data were received as an Excel file containing participants' anonymised responses to each SACII-X item, and their gender, age, ethnicity, and home language, as coded data. The primary investigator of the original study provided the researcher with the accompanying labels of the codes. The data file contained no identifying information of participants.
The researcher did a thorough review of the data collection procedures to ensure the data's accuracy. The SACII-X was administered in a single session of 45 minutes in the Grade 9 learners' classrooms (Rabie, 2017). Attention was paid to issues of measurement and experimenter bias, and procedures were implemented to control for these (Rabie, 2017). For

Stellenbosch University https://scholar.sun.ac.za

example, the SACII-X administrators were appropriately trained to ensure that they were familiar with the test instructions, format, and purpose, and that they maintained an impartial and professional attitude during the test administration period (Rabie, 2017). Furthermore, bilingual students who could converse in both English and isiXhosa were recruited as assessment assistants in the classes of isiXhosa-speaking learners. Where this was not possible, teachers assisted in providing test instructions to learners (Rabie, 2017). Rabie (2017) further aimed to optimise assessment results by administering the SACII-X during the second school period for the vast majority of participants, to control for possible learner fatigue or late arrivals at school. It was further ensured that classrooms allowed for suitable testing conditions with a desk for each learner, a moderately quiet environment, adequate ventilation and lighting, and space for the assessment assistants to move between the learners (Rabie, 2017). However, factors such as hunger and fatigue could not be controlled for by the researcher (Rabie, 2017).

4.7 Data Analysis

4.7.1 Rasch Rating Scale Model

The data were statistically analysed to investigate the psychometric properties, and item functioning of the SACII-X. The Rasch Rating Scale measurement model (Andrich, 1978; Rasch, 1960, 1980) was employed to determine the fit of the SACII-X items to the Rasch model to identify items that resulted in erratic response patterns or unexpected responses (De Bruin & De Bruin, 2011). All analyses were conducted in *Winsteps* version 5.1.0.0 (Linacre, 2021a). Validity was investigated through the analysis of the unidimensionality of each scale, local independence, and fit statistics (item and person infit and outfit mean-squares; Aryadoust et al., 2021). Reliability analysis focused on the analysis of reliability- and separation coefficients for both persons and items (Aryadoust et al., 2021).

Stellenbosch University https://scholar.sun.ac.za

The Rasch Rating Scale Model was introduced by Andrich (1978) as an extension of the logistic model proposed for dichotomous items by George Rasch (1960; Aryadoust et al., 2021). The Rating Scale Model is suited to psychometric measures such as the SACII-X with several ordered response categories generating polytomous data (Wright & Masters, 1982). The Rating Scale Model assumes that all the items share the same rating scale, for example, *Strongly Disagree*, *Disagree*, *Neutral*, *Agree*, and *Strongly Agree* (Teye-Kwadjo & De Bruin, 2021).

The Rasch measurement model is a probabilistic model (Aryadoust et al., 2021) that predicts person responses to test items (Rasch, 1960, 1980; Wright & Stone, 1979). In other words, based on a person's location on the latent trait, their probability of choosing a particular response category is predicted (Panayides et al., 2010). In the Rating Scale Model, the probability that person 'n' would endorse response category 'j', of item 'i', P_{nij}, is estimated based on three variables: (1) the ability of the person, B_n; (2) the endorsability/difficulty or location of the item, D_i; and (3) the location of the category threshold, F_j (Boone & Rogan, 2005; Linacre, 2021b; Teye-Kwadjo & De Bruin, 2021). This relationship can be mathematically expressed as (Boone & Rogan, 2005; Teye-Kwadjo & De Bruin, 2021):

$$P_{nij} = \frac{e^{(B_n - (D_i + F_j))}}{1 + e^{(B_n - (D_i + F_j))}}$$

The Rasch measurement model is a useful tool to develop, validate, or refine psychometric assessments (Boone & Rogan, 2005; Kean et al., 2017). Technically, the Rasch measurement model is not subsumed under IRT methodology, even though this is often assumed to be the case (Aryadoust et al., 2021). The Rasch model differs from other IRT models in that it aims to determine the fit of the data to the model, and not to find a best fit model for the data (Aryadoust et al., 2021; Bond & Fox, 2015).

Stellenbosch University https://scholar.sun.ac.za

The Rasch model addresses many of the criticisms of CTT. One of the main criticisms of using CTT methodology for psychometric investigations is that it assumes that rating scale data (such as Likert-scale data) are linear and therefore treats it as interval data, instead of ordinal data (Boone & Noltemeyer, 2017). For example, the Likert scale in the SACII-X contains five response categories named *Strongly Disagree*, *Disagree*, *Somewhat Disagree/Somewhat Agree*, *Agree*, and *Strongly Agree*. CTT models incorrectly assume that the distances between each category is equal, thus treating the data as linear data (Boone, 2016; Boone & Noltemeyer, 2017; Boone & Rogan, 2005). However, raw scores from Likert-scales are non-linear, that is, the distances between consecutive categories do not necessarily represent equal intervals (Boone, 2016; Boone & Noltemeyer, 2017; Boone & Rogan, 2005). The only certainty is that each successive response category indicates a greater degree of preference for the item, but the degree of the change between each category is not equal (Boone, 2016; Boone & Rogan, 2005). The Rasch Rating Scale Model takes into account that the distances between each category are not necessarily equal, correctly treating the data as ordinal data (Boone, 2016). Rasch analysis accomplishes this by transforming raw test scores from rating scales into logarithmic values, creating linear interval scale data (Boone, 2016; Wright & Masters, 1982). In other words, the Rasch measurement model transforms the polytomous, ordinal scale data from the SACII-X into linear, interval scale data (Boone, 2016; Wright & Masters, 1982).

Moreover, Rasch analysis takes into account that items differ in their level of endorsability or *difficulty* (Boone, 2016). Therefore, an answer of *Strongly Disagree* on one item, may not indicate the same level of disagreement as an answer of *Strongly Disagree* on another (Boone, 2016). With Rasch analysis, the test-taker’s raw test scores are expressed on a linear scale that accounts for the variability in item difficulties across the test (Boone, 2016). This, in turn, means that the Rasch model considers how people with different *ability* levels,

Stellenbosch University https://scholar.sun.ac.za

or with different locations on the latent trait, perform on the different test items – this is not accounted for in CTT (Teye-Kwadjo & De Bruin, 2021). Other advantages of the Rasch measurement model include that the response category functioning of a psychometric test can be investigated (Teye-Kwadjo & De Bruin, 2021), and that the model can be applied to samples as small as 30 respondents, and still produce useful person measures (Linacre, 1994).

Lastly, the Rasch model is a model of objective measurement (Program Committee of the Institute for Objective Measurement, 2000; Teye-Kwadjo & De Bruin, 2021). This states that the estimated person measures and reliability and separation indices are independent of the specific items used, and estimated item measures and reliability and separation indices are independent of the person sample responding to the items (Boone & Rogan, 2005; Teye-Kwadjo & De Bruin, 2021). If this principle holds, persons should have invariant measures across different test partitions, and items should have invariant measures across different groups in a sample, within a linear transformation (for example, item measures should remain invariant whether the items are administered to low ability test-takers or high ability test-takers; Wright & Masters, 1982). This principle of objective measurement only holds if certain assumptions are met, namely unidimensionality, local independence, monotonicity or equivalent slopes of the empirical Item Characteristic Curves, and invariant item locations across groups of persons (Kean et al., 2017; Teye-Kwadjo & De Bruin, 2021). IRT models therefore clearly have many advantages compared to CTT models but have more stringent assumptions and criteria that need to be met. The assumptions of unidimensionality, local independence, monotonicity, and invariant person locations are discussed in more detail later. If these assumptions are violated and data do not fit the model, parameter estimation (person- and item measures) is affected and may be inaccurate (Chou & Wang, 2010).

Stellenbosch University https://scholar.sun.ac.za

4.7.2 Item- and Person Measures

The Rasch model calculates item measures and person measures, expressed in logits, using a maximum-likelihood approach (De Bruin & Taylor, 2005). As mentioned previously, the person parameter estimation process and item parameter estimation process are independent of each other (Chou & Wang, 2017). Logit values theoretically range from positive to negative infinity, but logit values frequently range from −3.00 to 3.00 logits in the Rasch output (Boone, 2016; Boone & Noltemeyer, 2017). Item measures reflect the endorsability of items, that is, the relative difficulty of agreeing with the items (Aryadoust et al., 2016; Boone & Noltemeyer, 2017; De Bruin & De Bruin, 2011). Item measures are often referred to as item location parameters, as the item measures reflect the location of the items along the latent trait (De Bruin & De Bruin, 2011). In the Rasch analysis process, the item location mean is set at 0 logits, and items with higher measures are more difficult to endorse than those with lower measures (Boone & Noltemeyer, 2017; De Bruin & De Bruin, 2011).

In the same way, person measures or parameters reflect the ability level, or the relative standing of respondents on the latent trait measured in the scale (Boone & Noltemeyer, 2017). For instance, a person with a higher measure on the Realistic Scale of the SACII-X possesses more of the Realistic personality type than a person with a lower measure on the scale (De Bruin & De Bruin, 2011). An advantage of the Rasch measurement model is that both person and item measures are expressed on the same logit scale, enabling one to easily compare persons and items directly (De Bruin & De Bruin, 2011). From the person and item parameters, one can calculate the expected score of any person on any item (De Bruin & De Bruin, 2011).

4.7.3 Rating Scale Structure

As a starting point to confirm model fit, the Likert-scale response category functioning of the SACII-X was investigated. Investigations of rating scale functioning focused on

Stellenbosch University https://scholar.sun.ac.za

whether the response categories are correctly ordered for each item (Jafari et al., 2012). This was achieved through the investigation of the percentage endorsement of each category, category average person measures, Rasch-Andrich threshold ordering, category fit statistics, and category probability curves.

An important assumption of the Rasch model is that a higher standing on the latent trait should correspond with higher ratings on the rating scales, and vice versa (i.e., monotonicity; Bond & Fox, 2015). Average person measures are thus expected to increase monotonically with each successive category (Bond & Fox, 2015; Boone & Noltemeyer, 2017). It was therefore firstly determined whether average person measures increased with each successive category and its correspondence with the expected average person measures of respondents who endorse the category, determined from respondents' total scores on the scale (Boone & Noltemeyer, 2017).

The Rasch-Andrich threshold refers to the point where the probability of choosing a higher or lower category, increases. Ideally, these thresholds should also increase monotonically with the response categories (Jafari et al., 2012). For instance, the threshold between categories 4 and 5, should be higher than the threshold between categories 3 and 4 (De Bruin et al., 2013). Linacre (2002b) suggests that rating category average measures and Rasch-Andrich thresholds should increase by a minimum of 1.1 logits to conclude that categories function adequately and increase monotonically. Rasch-Andrich threshold disordering occurs if the category never becomes the most probable category to be observed for a portion of the latent trait (Bond & Fox, 2015). This can be seen on Category Probability Curves showing the probability of endorsing a particular response category for every particular person ability (Bond & Fox, 2015). If the rating categories are functioning well, one should observe a peak for each rating category on the Category Probability Curves (Bond & Fox, 2015). This would mean that each respective category is the most probable category to

Stellenbosch University https://scholar.sun.ac.za

be chosen for some portion of the latent trait (Bond & Fox, 2015). Disordered Rasch-Andrich thresholds do not necessarily indicate a problem. If no category disordering is observed (average person measures increase with each successive category), the researcher can pass a judgement on whether the Rasch-Andrich threshold disordering is a threat to the interpretation or reporting of the data (E. Teye-Kwadjo, personal communication, September 11, 2021).

The investigation of category functioning further included determining the percentage endorsement of each response category and the fit statistics (infit & outfit MnSq) of each category. One would want the percentage endorsement of each category to be at least 10%, to conclude that the category is a useful response option (E. Teye-Kwadjo, personal communication, September 11, 2021). Infit and outfit MnSq fit statistics of the categories are expected to be between 0.70 and 1.30 to indicate satisfactory functioning (E. Teye-Kwadjo, personal communication, September 11, 2021).

Based on the Rasch-Andrich threshold disordering and other indicators of category functioning the researcher decided whether to collapse categories or rescore categories. This decision is dependent on the goal of the research. If categories do not function properly, categories can be collapsed and rescored before continuing with the Rasch analyses (Bond & Fox, 2015; De Bruin et al., 2013). Collapsing categories that do not function well, may improve the model fit, and reduce the required effort for the test-taker to complete the questionnaire (Jafari et al., 2021). However, collapsing categories can decrease the precision of measurement (Jafari et al., 2012).

4.7.4 Global Model Fit

Winsteps provides statistics to determine the overall fit of the data to the Rasch model. Satisfactory fit is indicated by mean standardised residual correlations close to 0, standard deviations of the mean averaging near 1, and average infit and outfit mean-square values close to 1 (Linacre, 2021b). *Winsteps* calculates the average infit and outfit mean-square values by

Stellenbosch University https://scholar.sun.ac.za

omitting respondents who obtained perfect scores or scores of 0 on a scale, termed *extreme persons* (Linacre, 2021b).

4.7.5 Reliability and Separation

Reliability is another quality control mechanism to evaluate the functioning of the SACII-X (Aryadoust et al., 2016). The Rating Scale Model provides four indices to demonstrate the reliability of each SACII-X scale, namely the person separation index, person reliability index, item separation index, and item reliability index (Boone & Noltemeyer, 2017). These indices reflect how accurately person and item parameters can be estimated, and how stable person and item parameter estimation is (Boone & Noltemeyer, 2017). In other words, the reliability indices indicate how reproducible the estimated item- and person parameters will be, if the test is administered to another sample from the same population, or if the same sample is tested again (Bond & Fox, 2015).

The person separation index (PSI) indicates how well the items in an instrument are able to separate persons with different ability levels (Aryadoust et al., 2016). It therefore depicts the variability of the person locations on the latent variable (De Bruin et al., 2013). A high PSI provides an indication that test-takers have a wide range of ability (Aryadoust et al., 2016). The item separation index indicates how well a sample of people can separate items with different endorsability levels (Boone & Noltemeyer, 2017). A high item separation indicates that the items have a wide range of endorsability (Aryadoust et al., 2016). The two separation indices range from 0 to infinity, with higher values indicating better separation (Boone & Noltemeyer, 2017). In this thesis, separation values of more than 2 for both persons and items indicate that the SACII-X can distinguish adequately between different person ability levels, and between item endorsability levels, respectively (Aryadoust et al., 2021; Linacre, 2021b).

The person reliability index reflects the accuracy with which a person's standing on the latent trait can be predicted from their test score (Boone & Rogan, 2005). The person

Stellenbosch University https://scholar.sun.ac.za

separation reliability index can be equated to Cronbach's (1951) alpha or KR-20 indices in CTT (Bond & Fox, 2015; Boone & Rogan, 2005). A high person reliability index indicates a high probability that a person who scored high on the measure indeed has a high person location, and that a person who scored low on the measure indeed has a low person measure (Aryadoust et al., 2021).

Item separation reliability indicates the reproducibility of the item parameters in the item hierarchy across other samples (Boone & Rogan, 2005). This index has no equivalent in CTT (Boone & Rogan, 2005). The higher the item reliability value, the more confidence can be placed in the item parameter estimation. Both reliability indices range from 0 to 1 (Boone & Rogan, 2005). In this thesis, item and person reliability indices >.08 are deemed as indicating satisfactory reliability and thus that confidence can be placed in the person and item location predictions (Kean et al., 2017; Linacre, 2021b).

The reliability indices provided through Rasch analysis is an improvement compared to reliability indices such as KR-20 and Cronbach alpha computed in CTT (Boone & Noltemeyer, 2017). A notable advantage of the Rasch model's reliability indices compared to reliability indices in CTT, is that it is sample-free. Hence, it is not dependent on the sample completing the test, as the standard error is not influenced by sample variance (Boone & Rogan, 2005).

4.7.6 Unidimensionality

Since an important requirement of the Rasch Rating Scale Model is that scales are unidimensional, the six SACII-X scales were independently analysed on *Winsteps*. Unidimensionality means that item responses are only influenced by the intended measured latent trait and no other unintended secondary dimensions (Aryadoust et al., 2021; Kean et al., 2017). If this is not the case, Rasch estimated person and item measures may not be sufficiently reliable and accurate (Aryadoust et al., 2021). Hence, it should be shown that the

Stellenbosch University https://scholar.sun.ac.za

scale is unidimensional before relying on the Rasch computed person and item parameters (Aryadoust et al., 2021).

Unidimensionality is investigated through a Principal Component Analysis (PCA) of the standardised Rasch residuals to determine whether there are any substantive or meaningful secondary dimensions in the data (Aryadoust et al., 2021). The PCA identifies whether there are any relationships between items that is not due to the latent variable being measured (Kean et al., 2017). Of note is that the PCA analyses relationships in the Rasch residuals, not in the raw data (Teye-Kwadjo, 2021). The Rasch residuals are the differences between observed responses and those expected by the Rasch model (Aryadoust et al., 2021). Items without a perfect fit to the Rasch model will therefore produce residuals (Kean et al., 2017). For unidimensionality to hold, it is expected that there will be no structure in the residuals (Teye-Kwadjo & De Bruin, 2021). Therefore, after controlling for the latent variable, variance in the data is only related to random noise that is uncorrelated between the items (Panayides et al., 2010). If the PCA indicates structure in the residuals, this points to the possible presence of a secondary dimension in the data (Aryadoust et al., 2021).

All secondary dimensions are not threatening to the measurement per se, and do not necessarily mean that the data is substantively multidimensional (Linacre, 2021b). Secondary dimensions can either be auxiliary dimensions, or nuisance dimensions (Aryadoust et al., 2021). Nuisance dimensions are adverse and irrelevant to the latent trait being measured. This causes variance in the data, threatening the unidimensionality of the scale (Aryadoust et al., 2021). Auxiliary dimensions can be benign and may indicate secondary strands in the data that are still relevant to the latent trait being measured (Aryadoust et al., 2021; Linacre, 2021b). These may just refer to items with different content areas still relevant to the main variable under investigation (Linacre, 2021b; Panayides et al., 2010). Therefore, even if some secondary dimensions are present, if there is one dominant dimension, the test can still be viewed as

Stellenbosch University https://scholar.sun.ac.za

unidimensional enough, and valid for the test purpose (Aryadoust et al., 2021; Linacre, 2009; Panayides et al., 2010). The researcher can decide if the scale is unidimensional enough for the purposes of the instrument and does not purely have to rely on statistical indications (Linacre, 2021b).

Rasch analysis produces values for the variance accounted for by the Rasch measures, named the Rasch dimension, and raw unexplained variance in five *contrasts* (Teye-Kwadjo, 2021). The term *contrast* used in the Rasch output refers to a secondary dimension, or a component of the PCA (Teye-Kwadjo, 2021). In this thesis, the terms *component* and *dimension* are used to refer to the term contrast. The raw unexplained variance in the first component was focused on and compared to the raw explained variance in the measures. In addition, the Eigenvalue of the unexplained variance in the first component was inspected. An Eigenvalue indicates the strength of one item. An Eigenvalue <2.0 in the first component indicates unidimensionality (Teye-Kwadjo, 2021). An Eigenvalue of 2 means that the secondary dimension in the data has the strength of 2 items (Aryadoust et al., 2021). Since one cannot build a scale with less than 2 items, Eigenvalues of <2.0 can be ignored.

Scales on which Eigenvalues of more than 2 were found were further investigated with reference to the percentage variance explained by the Rasch dimension compared to the percentage variance explained by the first component (secondary dimension). Ideally, the variance explained by the Rasch dimension should be substantially larger than the raw, unexplained variance in the first contrast. Further investigation also entailed an examination of the disattenuated correlations. Correlations close to 1 indicate that the dimensions are very similar and probably measuring the same latent trait (Linacre, 2021b). Disattenuated correlations closer to 0 indicate that the items in contrasting clusters might be measuring different things, pointing to the presence of a secondary dimensions. According to Linacre

Stellenbosch University https://scholar.sun.ac.za

(2021b), disattenuated correlations above 0.57 provide an indication that clusters measure the same latent trait.

The researcher also inspected the content of the items contrasting and subjectively decided whether the items measure different things or can be subsumed under one dimension (Linacre, 2021b). Contrasting items were determined from the scree-plot (standardised residual contrast plot) produced by *Winsteps*. The scree-plot shows the loadings of item residuals on contrasting dimensions (Linacre, 2021b). One compares the items at the top of the map with those at the bottom of the map (Linacre, 2021b). The most contrasting items are denoted as *A*, *B*, *C* at the top of the map, contrasting against *a*, *b*, *c* at the bottom of the map. The size of the loadings is not of significance, but vertical differences between contrasts are important. On the plot, one can determine whether a cluster of items, vertically far removed from the other items, is present (Linacre, 2021b).

4.7.7 Local Independence

Closely influencing unidimensionality is the assumption of local independence (Linacre, 2009; Panayides et al., 2010). For items to be locally independent, the response on one item should not influence the response on another item, once the latent trait is accounted for (Wright, 1996). Stated differently, items should only correlate due to the latent trait being measured (Aryadoust et al., 2021; Wright, 1996). If local dependence is present, it means that unexplained variance in the data influencing item responses, correlate with each other (Aryadoust et al., 2021). Items should approximate local independence, for an accurate estimation of Rasch measures (Linacre, 2009; Wright, 1996).

There are two types of local independence, namely trait independence (or unidimensionality) and response independence (Retief et al., 2013). As discussed previously, trait dependence or multidimensionality occurs if test items measure something additional to the latent trait under investigation (Retief et al., 2013). Response dependence occurs when the

Stellenbosch University https://scholar.sun.ac.za

response to one item is dependent on the response to another item (i.e., one item influences the response to another item) after controlling for the latent trait (Linacre, 2009; Teye-Kwadjo, 2021). When response dependence is present, the item may not be contributing unique information to the test that other items do not already measure (Retief et al., 2013).

Both unidimensionality and local independence were investigated through the analysis of the correlations between standardised residuals of the test items (Aryadoust et al., 2021). Standardised residual correlations between item pairs should be close to 0 (Teye-Kwadjo & De Bruin, 2021). Standardised residual correlations of more than .30 between pairs of items are assumed to indicate local dependence. A value of .30 was decided upon to correspond to recommendations by Aryadoust et al. (2021), and other similar studies (Teye-Kwadjo, 2021; Teye-Kwadjo & De Bruin, 2021). The presence of local dependence is a threat to the reliable and accurate prediction of person and item measures and could exaggerate reliability coefficients if a number of highly similar items are included in the test (Aryadoust et al., 2021; Du Plessis & De Bruin, 2015; Edwards et al., 2018).

4.7.8 Wright map

The Wright map (also called a person-item map) is an especially valuable tool for researchers using the Rasch measurement model (Boone & Noltemeyer, 2017). In the Wright map, item difficulties (in this case, degree of endorsability of the items) are calculated and expressed on a logarithmic scale (Boone, 2016). In the same way, test-takers' ability (i.e., person measures) are calculated and expressed on this same logarithmic scale (Boone, 2016). The fact that person measures and item measures are expressed on the same scale, allows the researcher to compare these measures directly and visually see person-item relationships (Boone, 2016; Boone & Noltemeyer, 2017).

On the map, the person and item measure hierarchy can also be seen. Items are placed from most difficult to endorse at the top of the map, to least difficult to endorse at the bottom

Stellenbosch University https://scholar.sun.ac.za

of the map (Boone, 2016). Similarly, persons are placed from highest ability (person measure) at the top to lowest ability at the bottom of the map. A test-taker would have a higher likelihood of answering items falling below their person measure on the map correctly, and a lower likelihood of answering items falling above their person measure correctly (Boone, 2016).

This representation can illuminate the person-item targeting of the instrument. This refers to the spread of the items compared to the sample's standing on the latent trait (Boone, 2016). Ideally, one would want an instrument where items span the whole range of the construct, and where items are well targeted towards all levels of person ability to be able to separate persons with different ability levels well (Boone, 2016). With good person-item targeting, the item pool is suited to the person sample in terms of the endorsability of the items and the person locations of the latent variable (Boone, 2016). For example, overall, items should not be too difficult or too easy to endorse for the sample (Boone, 2016); there should be items wherever persons are located (Kean et al., 2017), and there should be no gaps between the items (Boone, 2016). With good person-item targeting, the precision of measurement is increased (Kean et al., 2017). Good targeting will increase the person separation and person reliability indices (Kean et al., 2017).

On the map, the person- and item means are denoted by an *M*, one standard deviation from the mean is denoted by an *S*, and two standard deviations from the mean is denoted by an *T* (Boone & Rogan, 2005). A comparison of the distance between the person mean and item mean can provide an indication of the person-item targeting (Boone, 2016). If the two means are close together, this suggests good person-item targeting, and that the items are not too easy or too difficult for the respondents (Boone, 2016).

The Wright map also provide an indication of the usefulness or redundancy of the test items (Boone & Noltemeyer, 2017). Ideally, a test developer would want test items to span

Stellenbosch University https://scholar.sun.ac.za

the entire construct (Boone & Noltemeyer, 2017). Items that are at the same level of endorsability, do not necessarily contribute any novel information to the test, and may be redundant (Boone & Noltemeyer, 2017). Through the identification of such items, a psychometric test can be refined (Boone & Noltemeyer, 2017). Nevertheless, one would want a number of items close to the person mean to accurately distinguish between the different ability levels of test-takers (Aryadoust et al., 2011).

4.7.9 Fit

Next, an investigation of the fit of the persons and test items to the Rasch model was conducted (Boone, 2016). Fit statistics are useful to identify persons who respond inconsistently or erratically, or items that behave idiosyncratically (Boone & Rogan, 2005). During Rasch analysis, the person's expected responses and observed responses are compared and person responses that consistently differ from what is predicted by the Rasch model, or items that consistently produce responses that deviate from what the Rasch model predicts, are flagged as misfitting (Boone, 2016). If items fit the Rasch model well, this provides an indication of content and construct validity. Fit analysis in this thesis relied on fit statistics, point-measure correlations, and the inspection of Item Characteristic Curves (ICCs).

Rasch analysis reports two forms of fit statistics, namely mean-squared (MnSq) and z-standardised (Zstd) (Boone & Noltemeyer, 2017). In samples larger than 250 persons, Zstd can become artificially inflated, while MnSq values are sample size-independent and less prone to error in larger samples (Aryadoust et al., 2021; Boone & Noltemeyer, 2017). Infit and outfit MnSq values were therefore relied upon for fit investigations in this thesis.

The MnSq value is the mean of the squared residuals of an item (Bond & Fox, 2015). This value indicates the size of the misfit in the data (Aryadoust et al., 2021; Linacre, 2002a). No data will ever fit the Rasch model perfectly (Wright & Linacre, 1994), so the decision of fit depends on whether your item fits well enough to be useful for measurement (De Bruin &

Stellenbosch University https://scholar.sun.ac.za

Taylor, 2005). Infit (inlier-sensitive, or more accurately, information-weighted fit; Linacre, 2002a) is sensitive to test-takers' unexpected responses to items that are close to their ability level (Boone & Rogan, 2005; Linacre, 2002a), whereas outfit (outlier-sensitive fit; Linacre, 2002a) is sensitive to test-takers' responses to items that lie further away from the person's ability level (Boone & Rogan, 2005; Linacre, 2002a). For example, outfit indicates a person's idiosyncratic responses to an item that is much harder or much easier than their ability level (Boone & Rogan, 2005). Due to this distinction, the infit Mnsq value is often considered more important than the outfit MnSq value (Aryadoust et al., 2011; Boone & Rogan, 2005; De Bruin & Taylor, 2005). That is, the unexpected responses of a person to an item with an endorsability near the person's ability level, are more important than the person's unexpected response to an item far away from their ability level (Aryadoust et al., 2011; Boone & Rogan, 2005; De Bruin & Taylor, 2005).

MnSq fit statistics are positive values ranging from 0 to infinity (Linacre, 2002a). An item that fits the model perfectly will have a MnSq value of 1 (Linacre, 2002a). Values >1 indicate underfit to the model (i.e., unpredictability of item responses), and values <1 indicate overfit to the model (i.e., item responses are too predictable; Linacre, 2002a). Underfit can be due to noise in the data or an item measuring something unexpected (something different to the other items; Bond & Fox, 2015; Linacre, 2002a). For example, a MnSq statistic of 1.30 indicates 30% noise in the data (Linacre, 2002a). Aryadoust et al. (2021) contend that the lower bound of the MnSq value is of less practical significance than the upper bound. Nevertheless, overfit could indicate local dependence or redundant items not contributing new information to the test (Aryadoust et al., 2021; Bond & Fox, 2015; Linacre, 2002a).

As underfit is a greater cause for concern than overfit, a general rule of thumb is to investigate outfit MnSq before infit MnSq (Boone & Noltemeyer, 2017), and high mean square values before low mean-square values (Boone, 2016; Wright & Linacre, 1994). There

Stellenbosch University https://scholar.sun.ac.za

is no universal rule for the acceptable range of fit (Aryadoust et al., 2021), but it is generally accepted that MnSq values should be between 0.5 and 1.5 to be useful for measurement (Linacre, 2002a). Values less than 0.5 or between 1.5 and 2 are not productive for measurement, but not degrading to measurement (Linacre, 2002a). In contrast, values above 2 are degrading for measurement and especially a cause for concern (Linacre, 2002a). In this thesis, infit MnSq and outfit MnSq values of >1.30 and <0.70 indicate misfit, to correspond with recommendations and other similar studies (Aryadoust et al., 2021; Boone, 2016; Du Plessis & De Bruin, 2015; Morgan & De Bruin, 2019; Teye-Kwadjo & De Bruin, 2021).

The fit of the items was first investigated by looking at the point-measure correlations (Kean et al., 2017). This statistic reflects how well the item correlates with the latent trait (Aryadoust et al., 2011). Point-measure correlations should all be positive, but the size of the correlation is not of as much practical significance compared to fit statistics (Kean et al., 2017). Thus, infit and outfit MnSq statistics of each item were inspected and items that misfit, were flagged and noted. In addition to the MnSq fit indices, Item Characteristic Curves were visually inspected to further investigate item functioning (Teye-Kwadjo & De Bruin, 2021). The Item Characteristic Curves indicate the uniformity of the slopes of the empirical and model expected Item Characteristic Curves (Teye-Kwadjo & De Bruin, 2021). Underfit can be seen if empirical slopes are flatter than the model expected and overfit can be seen if the empirical slope is steeper than the model expected (Teye-Kwadjo & De Bruin, 2021).

After the initial investigation of item fit, persons exhibiting inconsistent or erratic responses were identified. Similar to item fit statistics, infit MnSq flags the idiosyncratic response patterns of persons with average person measures, while outfit MnSq flags the idiosyncratic response patterns of persons with extremely low or high person measures (Aryadoust et al., 2021). For example, the outfit MnSq fit statistic will identify a person with a very high person measure unexpectedly answering in a low response category for an item

Stellenbosch University https://scholar.sun.ac.za

that is easy to endorse (Aryadoust et al., 2016). This index can therefore identify careless responding (Linacre, 2021a).

The raw item scores of persons who misfit were visually inspected across the items of the SACII-X to identify response patterns that were of particular concern. Persons who purposefully responded inaccurately or carelessly can be removed from further analysis if it is believed that the inaccurate responding is not due to the items (Boone & Rogan, 2005). After the removal of the most misfitting persons, item fit can be investigated again to determine whether the removal of misfitting persons had any effect on item fit. Hereafter, items that misfit were iteratively removed, starting with the most misfitting item. The impact of item removal was investigated through person and item separation and reliability indices, and item fit statistics (Du Plessis & De Bruin, 2005).

Items that do not fit the Rasch model, may not be useful for measurement (Boone, 2016). Possible reasons for items misfitting include multidimensionality, poor item quality, test-takers randomly answering or guessing, and items that are misunderstood (Boone, 2016; Boone & Noltemeyer, 2017; Kean et al., 2017). This could all lead to, for example, a difficult item unexpectedly being answered correctly by lower ability test-takers, or an easy item unexpectedly being answered incorrectly by high ability test-takers (Boone, 2016). If a person misfits, the person may be different from the target population in some way (Boone & Noltemeyer, 2017).

Fit analysis not only indicates whether test items only measure one trait, but importantly indicates whether person measures can be accurately computed (Boone & Noltemeyer, 2017). Hence, misfitting persons and items should always be investigated both qualitatively and quantitatively for reasons for the misfit, before further action is taken such as removing persons from an analysis or removing items from a test (Retief et al., 2013). The action taken depends

Stellenbosch University https://scholar.sun.ac.za

on the goals of the researcher (Boone & Noltemeyer, 2017). *Winsteps* provides the researcher with ways to analyse the reasons for misfit in detail (Boone & Noltemeyer, 2017).

4.7.10 Differential Item Functioning

DIF analysis focused on the analysis of both uniform DIF and non-uniform DIF. Uniform DIF was inspected through the *Winsteps*-provided DIF Contrast value and the Rasch-Welch probability value. The DIF Contrast value signifies the size of the DIF and is the difference in item endorsability between the male and female group (Linacre, 2021b). The Rasch-Welch probability value indicates the probability that the difference in endorsability occurred by chance (Linacre, 2021b). For the DIF to be considered meaningful, the difference in endorsability between males and females should have a substantive influence on the assessment results, and the DIF should be highly unlikely to have occurred by chance (Linacre, 2021b). In other words, the DIF should have both a substantive size as well as statistical significance. In this thesis, recommendations by the Educational Testing Service (ETS) and Linacre (2021b) for the identification of DIF were followed. According to these recommendations, a Rasch-Welch probability value ≤0.05, in conjunction with a DIF Contrast value ≥0.43 logits indicate slight to moderate DIF, and a Rasch-Welch probability value ≤0.05 in conjunction with a DIF Contrast value ≥0.64 indicate moderate to large DIF (Linacre, 2021a). Non-uniform DIF was investigated through the visual inspection of non-uniform DIF Item Characteristic Curves for males and females (Linacre, 2021b). Item Characteristic Curves with non-equal slopes between males and females show non-uniform DIF.

Although there are more advanced and complex analytical techniques to investigate the presence of DIF, the current study used the above-mentioned statistics and the inspection of Item Characteristic Curves to remain within the scope of the research. Furthermore, DIF analysis is notorious for lacking replicability (Linacre, 2021b). This, in conjunction with the small sample size of this study is a limitation of the DIF analysis, and results should be

Stellenbosch University https://scholar.sun.ac.za

interpreted with caution. The results of the DIF analysis were therefore not considered in item elimination decisions.

4.8 Ethical considerations

Permission was obtained from Stellenbosch University's Research Ethics Committee to conduct the proposed research (PSY-2021-23176; Appendix A). The primary study also obtained permission to collect the data used in this study from the Stellenbosch University Research Ethics Committee (SU-HSD-000640; Appendix B). In addition, the primary study obtained permission from the WCED (Appendix C) to conduct research in the two schools, as well as from the principals and other role players in the two schools (Rabie, 2017). The researcher obtained written permission from Rabie (2017) to access and use the secondary data, to adhere to the protection of the data during the analyses and publication process (Appendix D). The permission further pertained to the protection of participants' privacy, anonymity, and/or the confidentiality of the data.

Rabie (2017) obtained passive informed consent from participants' caregivers, and informed assent from participants before the commencement of data collection. It was made clear that participation was voluntary and that participants were free to withdraw from the research at any point during the process without negative consequences (Rabie, 2017). Participants were further ensured that their test responses will be anonymised, and biographical information will be confidential (Rabie, 2017). The purpose of the research, as well as benefits and risks were discussed with the participants before they provided assent (Rabie, 2017). There were no significant risks associated with participation in the primary study. There were, however, potential benefits as participation may lead to career exploration (McIlveen, 2007). The researcher of the original study ensured that ethical research standards were adhered to, and that the privacy and confidentiality of the test-takers were maintained

Stellenbosch University https://scholar.sun.ac.za

during all stages of the research process (APA, 2017; Rabie, 2017). Data was stored in a locked office and on a password-protected computer.

The study is a minimal risk study, as it only entails the analyses of secondary data and not the collection of primary data. Furthermore, participant responses and biographical information were anonymised. The researcher received the data from Rabie (2017) via encrypted electronic transfer. The data was fully anonymised and contained no names or other identifying information from test-takers. The data file only contained information pertaining to participants' responses to each SACII-X item, and participants' age, sex, ethnicity, and language. The electronic data file was stored on the password-protected personal computer of the current researcher, and only the current researcher had access to the data. An original version of the data file was stored on the researcher's personal Google Drive and remained unedited throughout the research process, while the researcher made a copy of the data file from which the analyses was run. The data file will be kept for five years, whereafter it will be destroyed.

4.9 Chapter Summary

This chapter discussed the methodology used in the present study. The study used a quantitative research design with the analysis of secondary data as research method. In this study, the Rasch Rating Scale model was used to determine the fit of the SACII-X items to the Rasch model, to identify items that behaved unexpectedly. The Rasch measurement model addresses many of the criticisms raised against CTT methodology. The most notable being that CTT methodology erroneously assumes that Likert-scale data are linear, and treats it as interval data, rather than ordinal data. This chapter further presented a discussion of sample characteristics, the development and psychometric properties of the SACII-X, and ethical considerations. The ensuing chapter presents the results of the data analysis.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 5

RESULTS

5.1 Introduction

In this chapter, the results of the study are presented. The chapter begins with a discussion of the overall fit of the data to the Rasch Rating Scale Model. Hereafter, the results of the response category functioning analysis of the SACII-X are provided. This is followed by dimensionality and local independence analysis of each SACII-X scale. Next, separation- and reliability indices for both persons and items are provided. A summary of the item measures of each SACII-X scale is provided thereafter. Subsequently, the item fit of the SACII-X is discussed with reference to fit statistics, point-measure correlations, and Item Characteristic Curves. This is followed by a brief overview of person fit. Hereafter, the person and item targeting of the sample group and item pool is discussed and the Wright map for each scale is presented. This is followed by the DIF results. Lastly, the removal of items, and its impact on the reliability and fit of the items to the Rasch Rating Scale Model, is discussed.

5.2 Global Model Fit

As a point of departure, the overall fit of the data to the Rasch model was determined. All scales demonstrated satisfactory overall fit to the Rasch model with mean standardised residual correlations equal to or close to 0 and infit and outfit MnSq values close to 1. The mean standardised residual correlations of the six scales were as follows: (R) .00; (I) .00; (A) .00; (S) .01; (E) .00; and (C) .00. Table 5.1 depicts the infit and outfit MnSq values, and standard deviations, of persons and items for each SACII-X scale.

Stellenbosch University https://scholar.sun.ac.za

Table 5.1

Average Mean-Square Fit Statistics and Standard Deviations of Persons and Items for the SACII-X Scales

Scale	Person Mean				Item Mean			
	Infit		Outfit		Infit		Outfit	
	MnSq	*SD*	MnSq	*SD*	MnSq	*SD*	MnSq	*SD*
Realistic	1.06	.57	1.04	.57	1.01	.12	1.04	.24
Investigative	1.06	.58	1.05	.58	1.00	.10	1.05	.10
Artistic	1.05	.53	1.04	.51	1.00	.10	1.04	.15
Social	1.05	.52	1.04	.52	1.01	.13	1.04	.21
Enterprising	1.06	.58	1.04	.56	1.01	.16	1.04	.24
Conventional	1.05	.59	1.05	.59	1.00	.11	1.05	.17

Note. MnSq = Mean-square

5.3 Response Category Functioning

Table 5.2 presents the response category functioning across the SACII-X scales. The analysis of response category functioning indicated that all categories were useful across the six SACII-X scales, with the endorsement of each category >10%. Infit and outfit MnSq values for all categories were within the pre-determined range of >.70 and <1.30, and thus satisfactory. Across the six scales, the observed average person measures within each category were close to the model expected values. In addition, across all scales, average person measures increased monotonically with each successive category. Therefore, persons with higher person measures were more likely to endorse higher categories and vice versa. However, a 1.1 logit increase for each successive category, as recommended by Linacre (2002b), was not observed.

Across all six scales, Rasch-Andrich threshold disordering was observed between Categories 3 and 4 (*Somewhat Disagree/Somewhat Agree* and *Agree*), and 4 and 5 (*Agree* and *Strongly Agree*). This disordering was due to respondents' consistent lower endorsement of Category 4 (*Agree*) compared to Category 5 (*Strongly Agree*). This lower endorsement may signify that this category defines a smaller range of the latent trait.

Stellenbosch University https://scholar.sun.ac.za

Table 5.2

Response Category Functioning of the SACII-X Scales

Category Label	% Endorsement	Observed Average Measure	Expected Average Measure	MnSq Infit	MnSq Outfit	Rasch-Andrich Threshold
		Realistic				
1	32	-1.06	-1.05	1.06	1.05	NONE
2	26	-.47	-.48	.88	.82	-.56
3	18	-.15	-.18	.88	.82	.05
4	11	.06	.07	1.05	1.21	.39
5	13	.39	.41	1.03	1.13	.13
		Investigative				
1	25	-.68	-.67	1.05	1.08	NONE
2	24	-.31	-.33	.94	.94	-.53
3	20	-.11	-.09	.96	.94	-.03
4	14	.12	.14	1.00	1.05	.36
5	16	.53	.51	.99	1.08	.20
		Artistic				
1	23	-.50	-.48	1.00	1.02	NONE
2	21	-.19	-.22	.98	.99	-.26
3	20	-.02	-.03	.86	.79	-.06
4	16	.15	.15	.99	1.03	.28
5	20	.37	.37	1.04	1.14	.04
		Social				
1	20	-.58	-.59	1.06	1.10	NONE
2	17	-.17	-.17	.94	.93	-.26
3	19	.07	.08	.86	.86	-.15
4	18	.27	.31	1.03	1.02	.29
5	25	.72	.70	1.01	1.12	.12
		Enterprising				
1	24	-.75	-.75	1.07	1.09	NONE
2	24	-.31	-.31	.91	.89	-.55
3	21	-.05	-.05	.83	.81	-.04
4	14	.20	.18	.91	.91	.42
5	17	.47	.49	1.06	1.19	.17
		Conventional				
1	26	-.80	-.80	1.07	1.06	NONE
2	25	-.40	-.38	.83	.85	-.60
3	21	-.12	-.13	.86	.87	-.05
4	14	.13	.10	.95	.97	.36
5	14	.42	.45	1.06	1.22	.29

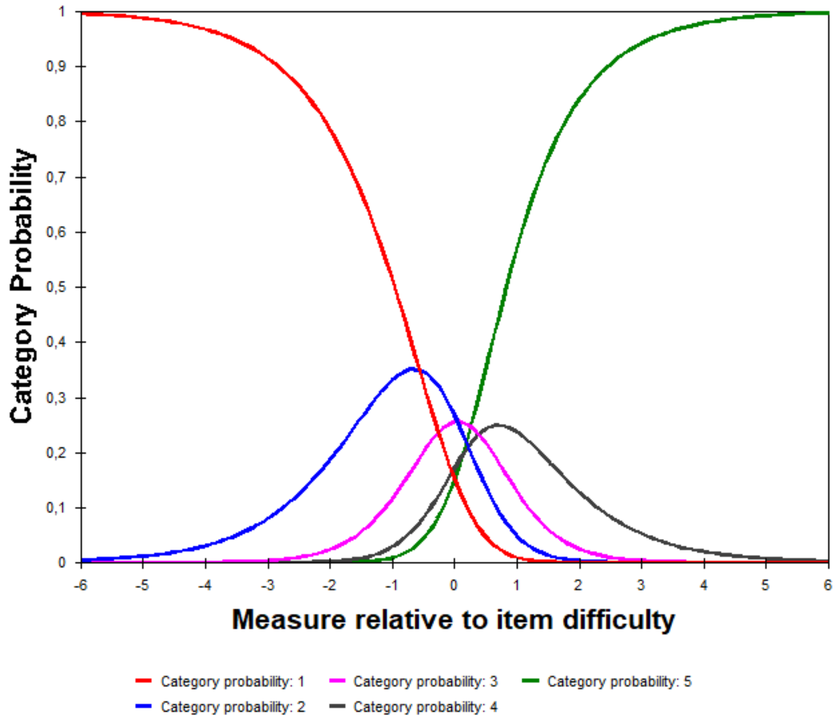
Note. Observed average measure refers to the average person measure of persons who endorsed the category. Expected average measure refers to the Rasch model expected average measure of persons who endorse the category.

Stellenbosch University https://scholar.sun.ac.za

Category Probability Curves also reflect this disordering, as the Rasch-Andrich thresholds between Categories 3 and 4, and 4 and 5, are disordered (see Figure 5.1 for an example). In addition, and as can be seen in the example Category Probability Curve in Figure 5.1, Category 4 (*Agree*) never emerged as a modal category for any of the SACII-X items, while Category 3 (*Somewhat Disagree/Somewhat Agree*) only emerged as a modal category for a very narrow portion of the trait. This Rasch-Andrich threshold disordering was present in every SACII-X item.

Figure 5.1

Example of Rasch-Andrich Threshold Disordering in the SACII-X Items



Note. Graph directly imported from *Winsteps*.

It was decided not to collapse the rating scale for several reasons, namely: (1) No category disordering was observed (i.e., average person measures increased monotonically with categories); (2) the threshold disordering was due to a slightly lower endorsement of Category 4 (*Agree*) compared to Category 5 (*Strongly Agree*); (3) Category 4's (*Agree*) endorsement was still >10%, indicating that this category is still useful as a response option; and (4) infit and outfit MnSq statistics of every category was satisfactory (<1.30; >.70).

Stellenbosch University https://scholar.sun.ac.za

5.4 Dimensionality and Local Independence

5.4.1 Realistic Scale

The PCA of standardised residuals indicated that the raw variance explained by the Rasch measures was 36.4%, which was similar to the 37.6% expected by the Rasch model. This explained variance is approximately six times greater than the unexplained variance in the first component of 6.0%. The unexplained variance in the first component had a strength of 2.20 Eigenvalues. Since this was slightly stronger than 2 Eigenvalues, the contrasting items were further investigated to rule out the presence of a secondary dimension in the data.

Inspection of the standardised residual contrast plot revealed that there were no items clearly clustering together that would create concern regarding an unwanted relationship between the items (Figure 5.2). The standardised residual contrast plot shows the loadings of the residuals of items on contrasting dimensions (Linacre, 2021b). Contrasting items are denoted as *A*, *B*, *C* at the top of the map, contrasting against *a*, *b*, *c* at the bottom of the map. The size of the loadings is not of significance, but vertical differences between contrasts are important. Ideally, one would want to see a random pattern of items with few high loadings (Linacre, 2021b). On the plot, one can determine whether a cluster of items have high loadings and are vertically far removed from the others.

Further inspection of the content areas of the two contrasting clusters, revealed no significant differences between the groups of items. Furthermore, the disattenuated correlations between clusters were not concerning, as the smallest disattenuated correlation was 0.64. In this case, the clusters still correlate sufficiently to assume that the two clusters measure something similar. In the context of this data, the presence of a secondary dimension was ruled out. Further support for unidimensionality was provided by the fact that all items were locally independent with standardised residual correlations <.30 (Table 5.3).

Stellenbosch University https://scholar.sun.ac.za

Figure 5.2

Realistic Scale: Standardised Residual Contrast Plot

```
        -5          -3          -1           1           3           5
        -+-----------+-----------+-----------+-----------+-----------+- COUNT  CLUSTER
     .7 +                               A |                           + 1      1
        |                                 |                           |
     .6 +                                 |                           +
        |                               B |                           | 1      1
  C  .5 +                                 |                           +
  O     |                                 |                           |
  N  .4 +                                 |                           +
  T     |                                 C                           | 1      1
  R  .3 +                                 |D                          + 1      1
  A     |                              E  |                           | 1      2
  S  .2 +                                 |                           +
  T     |                                 G                           |
     .1 +                                IHF                          + 4      2
  1     |                              J  |                           | 1      2
     .0 +---------------------------------K---------------------------+ 1      2
  L     |                                 |                           |
  O -.1 +                                k| L                         + 2      2
  A     |                                 |j                          | 1      3
  D -.2 +                                i| h                         + 2      3
  I     |                                g|                           | 1      3
  N -.3 +                                 fe                          + 2      3
  G     |                                 d                           | 1      3
    -.4 +                                 c                           + 1      3
        |                                 | b                         | 1      3
    -.5 +                                 | a                         + 1      3
        |                                 |                           |
        -+-----------+-----------+-----------+-----------+-----------+-
        -5          -3          -1           1           3           5
                                   ITEM MEASURE
```

Note. Figure directly imported from *Winsteps*.

Stellenbosch University https://scholar.sun.ac.za

Table 5.3

Standardised Residual Correlations between Item Pairs

Realistic			Investigative			Artistic		
Correlation	Item	Item	Correlation	Item	Item	Correlation	Item	Item
.27	RR18	RR7	.24	I6	I8	**.54**	**A1**	**A2**
.25	RR25	RR26	.20	I25	I4	**.37**	**A1**	**A6**
.24	RR19	RR8	.18	I16	I2	**.33**	**A2**	**A6**
.23	RR24	RR28	.17	I25	I9	.24	A13	A28
.22	RR21	RR24	-.23	I13	I23	.23	A15	A18
.20	RR14	RR4	-.23	I4	I6	.23	A6	A7
.19	RR10	RR26	-.22	I25	I6	.22	A8	A9
.18	RR27	RR28	-.22	I16	I8	.21	A23	A27
-.24	RR19	RR4	-.21	I2	I21	-.27	A21	A6
-.24	RR14	RR19	-.20	I23	I8	-.24	A24	A8
-.22	RR18	RR4	-.19	I21	I25	-.24	A1	A26
-.21	RR19	RR27	-.19	I20	I4	-.23	A1	A5
-.20	RR16	RR24	-.18	I10	I11	-.23	A2	A26
-.20	RR14	RR18	-.18	I24	I4	-.23	A23	A24
-.18	RR19	RR28	-.17	I16	I19	-.22	A16	A2
-.18	RR15	RR8	-.17	I2	I6	-.22	A2	A21
-.18	RR19	RR24	-.17	I22	I25	-.22	A2	A28
-.18	RR16	RR25	-.17	I13	I17	-.21	A21	A7
-.18	RR16	RR28	-.17	I16	I21	-.20	A1	A22
-.17	RR11	RR19	-.16	I2	I24	-.20	A18	A6

Social			Enterprising			Conventional		
Correlation	Item	Item	Correlation	Item	Item	Correlation	Item	Item
.22	S11	S24	**.32**	**E11**	**E24**	**.35**	**C27**	**C8**
.22	S24	S27	.23	E10	E9	.24	C17	C19
.21	S23	S7	.20	E5	E8	.23	C19	C2
.19	S20	S6	.20	E27	E28	.20	C7	C8
-.26	S16	S24	.18	E16	E19	.19	C1	C24
-.23	S24	S28	-.22	E27	E5	-.24	C2	C5
-.23	S3	S5	-.22	E14	E28	-.23	C12	C21
-.22	S12	S7	-.21	E10	E24	-.21	C19	C3
-.21	S22	S6	-.21	E24	E9	-.21	C26	C9
-.20	S18	S4	-.21	E19	E3	-.20	C15	C4
-.20	S10	S12	-.21	E12	E3	-.19	C16	C7
-.20	S28	S5	-.20	E1	E12	-.19	C1	C7
-.20	S2	S26	-.20	E19	E22	-.19	C25	C8
-.19	S12	S23	-.20	E15	E24	-.19	C16	C4
-.19	S11	S18	-.19	E18	E7	-.19	C27	C29
-.19	S7	S8	-.19	E17	E23	-.19	C12	C27
-.19	S12	S4	-.19	E11	E18	-.19	C17	C25
-.18	S27	S28	-.19	E22	E8	-.18	C11	C7
-.17	S20	S22	-.19	E12	E18	-.17	C29	C8

Stellenbosch University https://scholar.sun.ac.za

Social			Enterprising			Conventional		
Correlation	Item	Item	Correlation	Item	Item	Correlation	Item	Item
-.17	S11	S7	-.18	E17	E21	-.17	C26	C3

Note. Item pairs with standardised residual correlations >.30 are presented in boldface.

5.4.2 Investigative Scale

Rasch dimensionality analysis of the Investigative scale, indicated that the unexplained variance in the first contrast was close to two Eigenvalues (1.96). Even though the observed and Rasch expected values for the raw variance explained by the Rasch dimension were similar, the raw variance explained by the measures was 33.3%; only 5.37 times larger than the unexplained variance in the first contrast of 6.2%. Contrasting clusters were further investigated to determine whether a possible secondary dimension is present.

Disattenuated correlations were all approximating one with the smallest disattenuated correlation equal to 0.75. This provides support for clusters measuring the same latent trait. On the standardised residual contrast plot, no clear cluster of items at the top of the plot emerged. The contrasting clusters contained items such as items I6 ("Research animals in a laboratory"), I21 ("Study plants in the environment"), I8 ("Research fish in an aquarium"), I24 ("Study wildlife in a game reserve") in one cluster, and items I25 ("Test chemicals in a laboratory"), I2 ("Develop new medicines in a laboratory"), I4 ("Research a cure for a disease"), and I16 ("Study diseases") in its contrasting cluster. These two clusters may potentially measure two different variables, one being interested in physical objects such as plants, fish, wildlife, and animals, and the other referring more to an interest in the chemical and medical sciences. However, it seems like these items can still be subsumed under the overarching latent trait of Investigative personality type. Based on the subjective content analysis of items, and the inspection of the disattenuated correlations and contrast plot, it was decided that the scale is sufficiently unidimensional. All items were locally independent with all residual correlations between items <.30 (Table 5.3).

Stellenbosch University https://scholar.sun.ac.za

5.4.3 Artistic Scale

The PCA of the Artistic Scale showed that the unexplained variance in the first component was equal to 2.92. This means that there is a possible secondary dimension in the data that is approximately three items strong. The raw variance explained by the Rasch measures was equal to 28.6%, whereas the unexplained variance in the first component was 8.0% (i.e., the explained variance is only 3.8 times larger than the unexplained variance). The disattenuated correlations between clusters further supported the presence of a possible secondary dimension with a disattenuated correlation between Clusters 1 and 3 of 0.31. This provides a further indication that the secondary dimension possibly measures something different to the Rasch dimension.

The standardised residual contrast plot visually showed that three items on the scale clearly clustered together (Figure 5.3). These three items, namely, items A1 ("Act in a play"), A2 ("Audition for a play"), and A6 ("Direct a play") contrasted against A21 ("Use art therapy to help people who need counselling"), A8 ("Illustrate stories for children's books"), and A18 ("Teach art to people").

The analysis of local independence between items, further brought the unidimensionality of the scale into question. Standardised residual correlations between items indicated that three pairs of items were locally dependent on each other. Items A1 ("Act in a play") and A2 ("Audition for a play") had a residual correlation of .54; Items A1 and A6 ("Direct a play") had a residual correlation of .37, and Items A2 and A6 had a residual correlation of .33 (Table 5.3). These three items were also the same items contrasting against the other items in the Artistic scale. This further raises suspicion regarding a possible secondary dimension in the data that warrants further investigation.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.3

Artistic Scale: Standardised Residual Contrast Plot

```
       -3            -1            1             3             5             7
       -+-------------+-------------+-------------+-------------+-------------+- COUNT  CLUSTER
        |                     |                                                 |
    .7 +                     AB                                                + 2      1
        |                      C                                                | 1      1
C   .6 +                      |                                                +
O       |                      |                                                |
N   .5 +                     D|                                                + 1      1
T       |                      |                                                |
R   .4 +                      |                                                +
A       |                      E                                                | 1      2
S   .3 +                      |F                                               + 1      2
T       |                      |G                                               | 1      2
    .2 +                      |                                                +
1       |                      |                                                |
    .1 +                      |                                                +
L       |                     H|                                                | 1      2
O   .0 +-------------------I-----------------------------------------------+ 1      2
A       |                     LJK                                               | 3      2
D  -.1 +                   m M                                                 + 2      3
I       |                      |                                                |
N  -.2 +                      lkj                                              + 3      3
G       |                     ghi                                               | 3      3
   -.3 +                     ecfd                                              + 4      3
        |                      b                                                | 1      3
   -.4 +                      |                                                +
        |                      a                                                | 1      3
       -+-------------+-------------+-------------+-------------+-------------+-
       -3            -1            1             3             5             7
                                  ITEM MEASURE
```

Note. Figure directly imported from *Winsteps*.

5.4.4 Social Scale

The PCA indicated that the variance explained by the measures was about six times that of the unexplained variance in the first component (36.2% and 6.0%). The observed value of the variance explained by the measures of 36.2%, was also highly similar to the Rasch expected value of 36.6%. The unexplained variance in the first component was 2.07 Eigenvalues strong. The reason for the high Eigenvalue was further investigated, to rule out the presence of a secondary dimension influencing item responses.

No clear concerning cluster of items could be seen on the standardised residual contrast plot. An inspection of the item content of contrasting clusters on the map also showed no indication of items measuring something different to the latent trait. The decision was

Stellenbosch University https://scholar.sun.ac.za

made that the scale is unidimensional enough, based on the inspection of the above-mentioned items and the fact that the smallest disattenuated correlation was 0.89, indicating that the contrasting clusters measure something similar. Furthermore, no local dependence was present, as all residual correlations between item pairs were below .30 (Table 5.3).

5.4.5 Enterprising Scale

Dimensionality analysis of the Enterprising scale showed that the Rasch dimension explained 33.8% of the variance in the data. This compares well to the Rasch expected value of 34.2%. The first component had an Eigenvalue of 2.28. As this is larger than the preferred cut-off of two Eigenvalues, further investigation was conducted to determine the dimensionality of the scale. The possible secondary dimension contributed 6.0% to the total raw unexplained variance in the data. Therefore, the variance explained by the Rasch dimension was just over five times more than the variance explained by the first component. The disattenuated correlations, however, all approximated one. This indicates that the clusters correlated highly with each other, providing evidence that they measure the same latent trait.

Contrasting clusters of items on the standardised residual plot contain items E18 (“Network with clients”), E8 (“Increase efficiency of a business”), E1 (“Advise people in a community on financial matters”), E5 (“Give presentations on how to make a business better”) in the one cluster, with items E11 (“Manage a business”), E12 (“Manage a hotel”), E17 (“Negotiate salaries on behalf of people”), and E14 (“Mentor staff at a company”). These two clusters do not seem clearly different, and all these items can still be subsumed under the latent trait of Enterprising. Based on the above-mentioned considerations, the Enterprising scale was determined to be unidimensional enough for its purpose. Only 1 item pair showed a standardised residual correlation of more than .32 indicating local dependence of these items (Table 5.3). These were items E11 (“Manage a business”) and E24 (“Start a business”).

Stellenbosch University https://scholar.sun.ac.za

5.4.6 Conventional Scale

The raw variance explained by the Rasch measures was 33.3%; similar to the model expected value of 33.4%. The raw unexplained variance in the first contrast was 5.2%, with a strength of 1.95 Eigenvalues. To rule out the presence of a dimension other than the Rasch dimension influencing the data, further output of the dimensionality analysis was inspected. Disattenuated correlations all approximated one, with the smallest disattenuated correlation 0.82. Contrasting items were items C2 (“Archive documents for a company”), C19 (“Monitor income and expenditure records of a company”), and C26 (“Prepare contracts for a company”) in the first cluster, with items C3 (“Calculate financial premiums for clients”), C4 (“Calculate wages and salaries of employees”), and C9 (“Control the petty cash in an office”). These items do not seem meaningfully different from each other. Based on the available information, this scale is unidimensional enough for the purpose of the SACII-X. One item pair showed a standardised residual correlation >.30 (0.35), namely items C27 (“Prepare financial records for a company”) and C8 (“Control accounting records”) (Table 5.3).

5.5 Separation and Reliability

Separation and reliability indices for all scales were satisfactory. Table 5.4 demonstrates the separation and reliability indices for each scale. Person separation indices ranged from 2.40 on the Investigative scale to 2.63 on the Social scale. The items of the SACII-X were thus able to separate persons into more than two different ability levels for each scale. Item separation indices ranged from 2.38 to 3.78, with the lowest index observed on the Artistic Scale and the highest index observed on the Conventional Scale. The person sample was therefore able to separate the items of the SACII-X into approximately two to three endorsability levels. Person reliability indices were similar across the scales, with person reliability of .87 across all scales, except the Investigative scale with a person reliability of .85. Item reliability indices ranged from .85 on the Artistic Scale to .93 on the Social,

Stellenbosch University https://scholar.sun.ac.za

Enterprising, and Conventional scales. The high person and item reliability indices indicate the high accuracy with which the SACII-X can predict person ability and item endorsability, respectively.

Table 5.4

Person- and Item Separation and Reliability Indices for each SACII-X Scale

Scale	Person Separation	Person Reliability	Item Separation	Item Reliability
Realistic	2.53	.87	3.32	.92
Investigative	2.40	.85	3.70	.90
Artistic	2.57	.87	2.38	.85
Social	2.63	.87	3.65	.93
Enterprising	2.62	.87	3.74	.93
Conventional	2.54	.87	3.78	.93

5.6 Item Measures

Appendix E provides a table with item measures, standard errors, point-measure correlations, and expected point-measure correlations for all items of the SACII-X. The point-measure correlations of all the SACII-X items were positive, indicating that each item relates well to the latent trait being measured. Item measures on each scale ranged from: (R) −.46 to .30; (I) −.38 to .33; (A) −.25 to .26; (S) −.37 to .34; (E) −.40 to .53; and (C) −.40 to .37.

5.7 Item Fit

Overall, the SACII-X items showed good fit to the Rasch model. Only 12 items of the SACII-X showed misfit to the Rasch model (breakdown per scale: Realistic=2; Investigative=1; Artistic=1; Social=3; Enterprising=3; Conventional=2). All the misfitting items displayed underfit to the model. Table 5.5 shows the infit and outfit MnSq values of all the SACII-X items per scale, with infit and outfit MnSq values that are above the cut-off of 1.30 presented in bold.

Stellenbosch University https://scholar.sun.ac.za

Table 5.5

Infit and Outfit MnSq Fit Indices for Each SACII-X Item

Realistic			Investigative			Artistic		
Item	Infit	Outfit	Item	Infit	Outfit	Item	Infit	Outfit
RR02	.98	.92	I01	.94	.99	A01	1.12	1.17
RR04	.79	.80	I02	1.07	1.07	A02	1.01	1.14
RR07	.93	.95	I04	1.12	1.19	A04	1.11	1.12
RR08	1.01	1.06	I05	.82	.84	A05	1.05	1.08
RR09	1.03	.98	I06	1.02	1.03	A06	1.08	1.10
RR1	.99	.99	I08	1.05	1.03	A07	1.19	**1.47**
RR10	1.01	.98	I09	.95	1.03	A08	1.13	1.20
RR11	.95	.87	I10	1.08	1.10	A09	1.02	1.06
RR13	1.17	1.13	I11	.88	1.00	A10	1.06	1.17
RR14	.86	.92	I12	.99	1.12	A11	1.11	1.18
RR15	1.12	1.21	I13	.91	.90	A13	1.07	1.13
RR16	1.14	1.19	I16	1.07	1.12	A15	.82	.85
RR17	1.09	1.27	I17	1.17	1.12	A16	.82	.82
RR18	1.03	1.02	I19	.82	.81	A18	.90	.92
RR19	**1.31**	**1.85**	I20	.97	.95	A19	.96	.97
RR20	1.12	**1.49**	I21	1.05	1.04	A20	1.09	1.11
RR21	1.07	.95	I22	.96	1.27	A21	1.05	1.03
RR23	.82	.88	I23	.97	.93	A22	.92	.87
RR24	1.13	.99	I24	.96	.94	A23	1.02	1.16
RR25	.91	.84	I25	.96	1.05	A24	1.00	1.03
RR26	1.01	.91	I26	1.20	**1.59**	A25	.94	.91
RR27	.92	.86				A26	.85	.87
RR28	.87	.77				A27	.93	.90
						A28	.99	.97
						A29	1.02	1.04
						A30	.82	.79

Social			Enterprising			Conventional		
Item	Infit	Outfit	Item	Infit	Outfit	Item	Infit	Outfit
S02	.90	.89	E01	.86	.84	C01	.90	.87
S03	.94	.92	E02	1.02	.97	C03	.87	.93
S04	.83	.78	E03	.92	.87	C04	1.10	1.08
S05	.97	.97	E05	.87	.81	C05	.97	.91
S06	.89	.86	E07	1.03	1.02	C07	1.13	1.20
S07	1.02	.97	E08	.85	.82	C08	1.01	1.00
S08	1.01	1.02	E09	1.03	1.02	C09	1.09	1.03
S10	1.02	.99	E10	.82	.77	C11	.79	.84
S11	1.17	1.19	E11	1.11	1.17	C12	1.10	1.28
S12	1.26	**1.35**	E12	1.36	1.49	C14	1.16	**1.51**
S15	.88	.85	E13	.88	.82	C15	**1.31**[a]	1.26
S16	.99	1.02	E14	1.24	1.27	C16	.92	1.08

Stellenbosch University https://scholar.sun.ac.za

S17	1.00	1.16	E15	.94	.90	C17	.91	.86
S18	1.19	1.21	E16	1.02	1.17	C18	1.08	1.28
S20	.82	.75	E17	**1.36**	**1.61**	C19	.91	.90
S22	1.24	**1.56**	E18	.79	.76	C02	.86	.84
S23	.79	.85	E19	1.05	1.02	C20	1.06	1.03
S24	1.00	1.01	E20	.98	.98	C21	.96	1.21
S25	1.17	1.24	E21	.80	.81	C22	.98	1.02
S26	1.11	**1.40**	E22	1.09	1.25	C24	.93	.87
S27	.86	.87	E23	.91	.87	C25	1.08	1.22
S28	1.06	1.13	E24	1.12	1.12	C26	1.07	1.14
			E25	.97	1.24	C27	.90	.85
			E27	1.25	**1.56**	C28	1.00	.94
			E28	.86	.85	C29	.93	.98

Note. Items with Infit- and/or Outfit mean-square values >1.30 are in boldface.

[a]The misfit in item C15 was ignored.

5.7.1 Realistic Scale

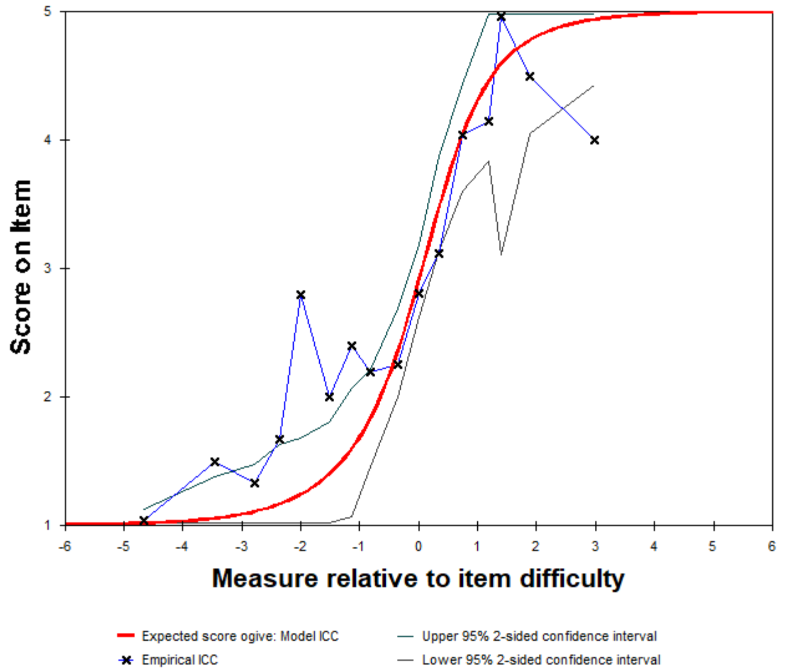
On the Realistic scale, only two items displayed misfit with MnSq values above 1.30, namely item RR19 (“Operate heavy machinery”; outfit MnSq=1.85; infit MnSq=1.31) and RR20 (“Plan the layout of cables”; outfit MnSq=1.49). This was corroborated by Item Characteristic Curves where one could visually see that the majority of the items fit the model well while items RR19 (Figure 5.4) and RR20 (Figure 5.5) attracted unexpected responses, especially for low ability and high ability persons.

Item RR19 was one of the easiest to endorse items with an item measure of −0.26 (Appendix E). The misfit was due to the unexpected responses of 15 persons to the item. Item RR20 was the seventh most difficult to endorse item with an item measure of 0.08 (Appendix H). Six persons provided unexpected responses to this item.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.4

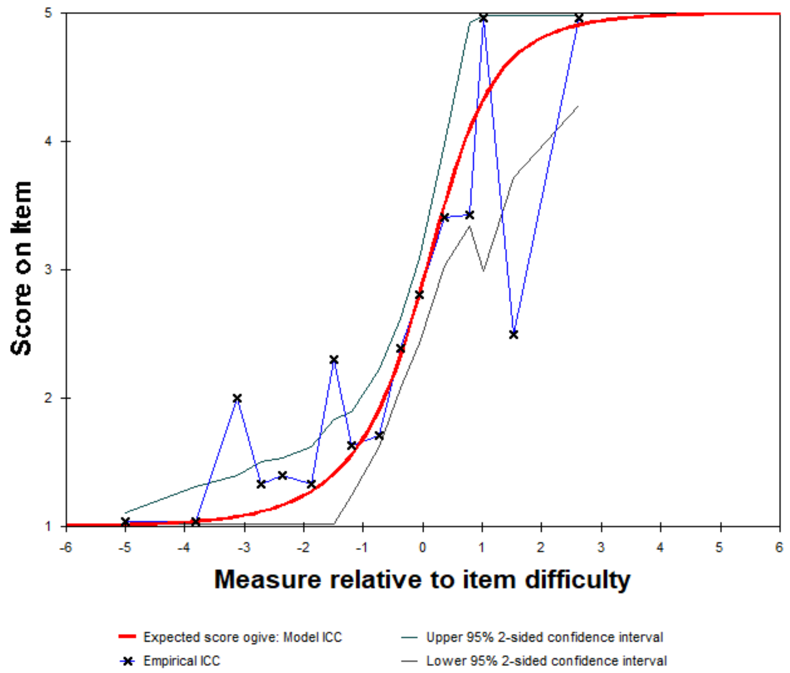
RR19 Expected and Empirical ICC



Note. Graph directly imported from *Winsteps*.

Figure 5.5

RR20 Expected and Empirical ICC



Note. Graph directly imported from *Winsteps*.

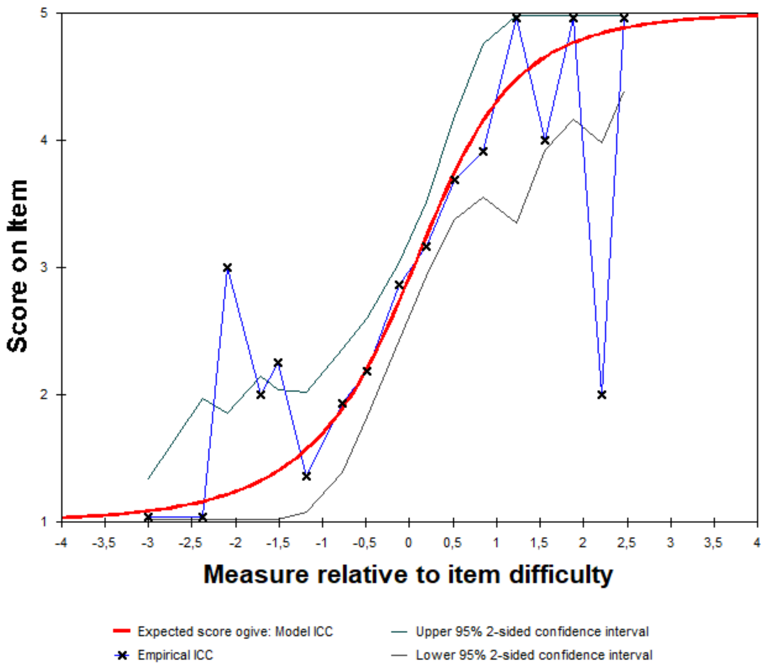
Stellenbosch University https://scholar.sun.ac.za

5.7.2 Investigative Scale

Only one item showed misfit, namely item I26 ("Use new ways of removing diseases from crops"; outfit MnSq=1.59). Item I26 was the fifth easiest to endorse item on this scale with an item measure of −0.14 (Appendix E). Six people responded unexpectedly to this item. The Item Characteristic Curve of I26 showed that I26 generated unexpected responses especially for persons with low- and high person measures (Figure 5.6).

Figure 5.6

I26 Expected and Empirical ICCs



Note. Graph directly imported from *Winsteps*.

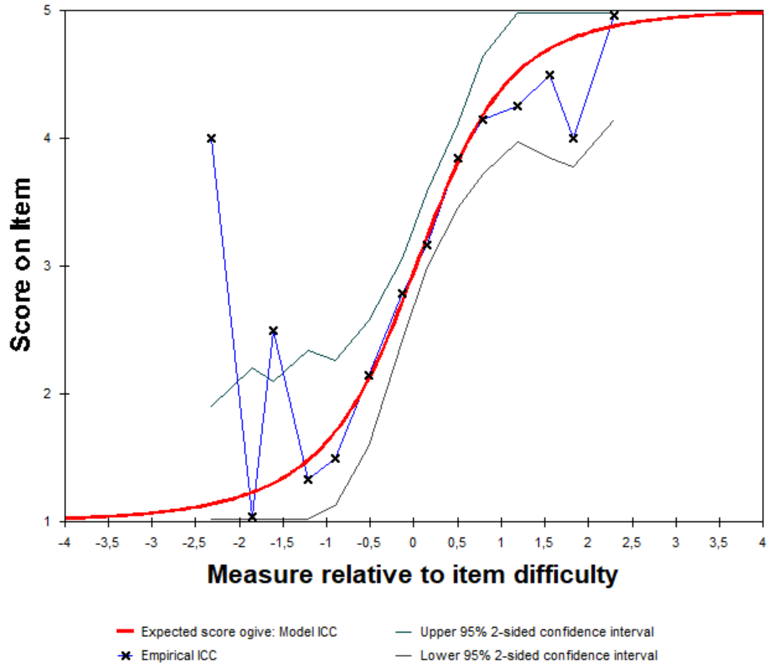
5.7.3 Artistic Scale

Item A7 ("Go to theatre productions") was the only item showing misfit on the Artistic Scale (outfit MnSq=1.47). Item A7 was the third easiest to endorse item on the Artistic Scale with an item measure of −0.18 (Appendix E). The misfit in item A7 was caused by the unexpected responses of nine persons to the item. The Item Characteristic Curve of A7 showed that the item functioned especially poorly for persons with low person measures (Figure 5.7).

Stellenbosch University https://scholar.sun.ac.za

Figure 5.7

A7 Expected and Empirical ICC



Note. Graph directly imported from *Winsteps*.

5.7.4 Social Scale

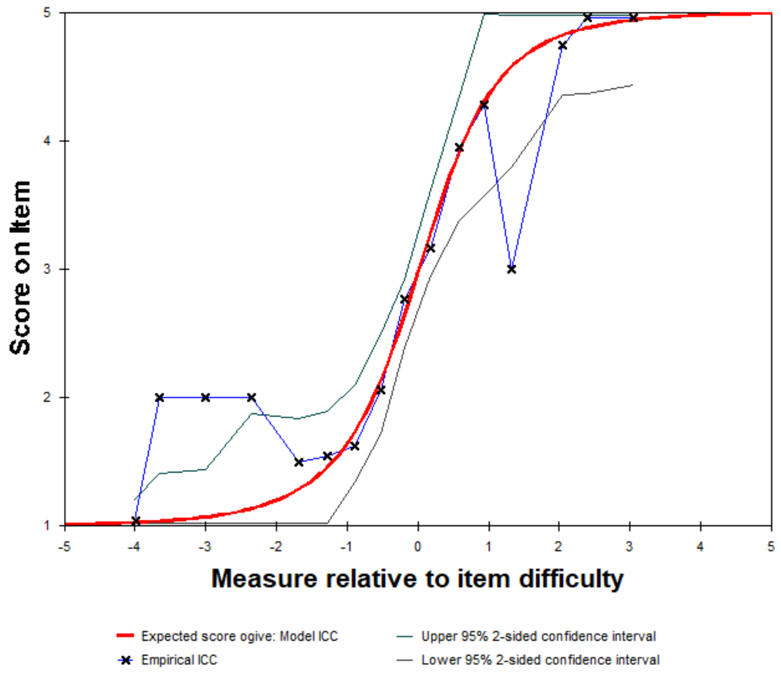
The Social scale included three misfitting items: Item S22 ("Teach at a school"; outfit MnSq=1.56), S26 ("Teach people how to do public speaking"; outfit MnSq=1.40) and S12 ("Model clothes"; outfit MnSq=1.35). Item S22 and S26 was the third hardest to endorse items, both with an item measure of 0.25 (Appendix E). S12 was the seventh hardest to endorse item with an item measure of 0.12 (Appendix E). The misfit in item S22 was due to the unexpected responses of 10 test-takers to the item. The misfit in item S26 was due to the unexpected responses of nine persons, while 12 persons responded unexpectedly to item S12.

The Item Characteristic Curve of item S22 showed that the item functions poorly for low performers and high performers (Figure 5.8). The Item Characteristic Curve of item S26 shows that the item showed slight misfit for low performers and a relative misfit for high performers, as well as a slight misfit for moderate performers (Figure 5.9). The misfit of item S12 is not severe. The item shows slight misfit for moderate performers (Figure 5.10).

Stellenbosch University https://scholar.sun.ac.za

Figure 5.8

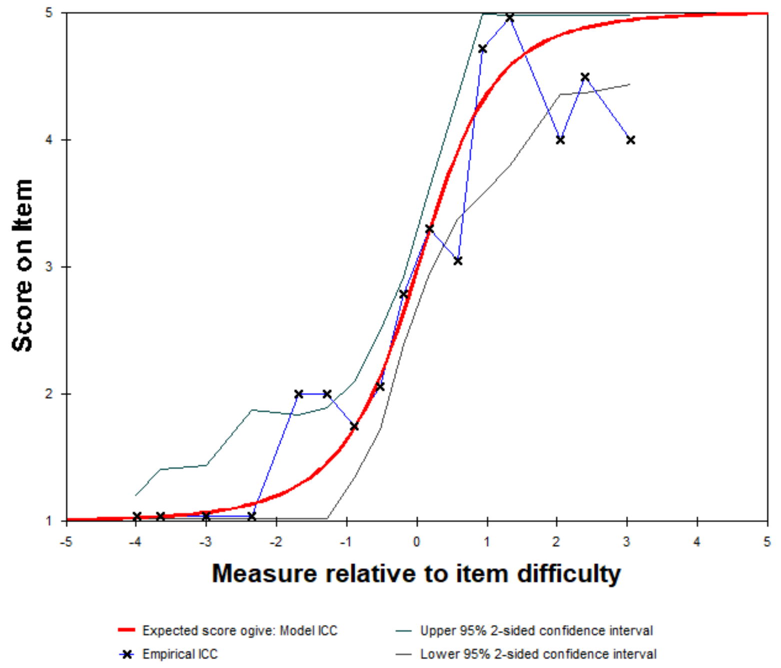
S22 Expected and Empirical ICC



Note. Graph directly imported from *Winsteps*.

Figure 5.9

S26 Expected and Empirical ICCs

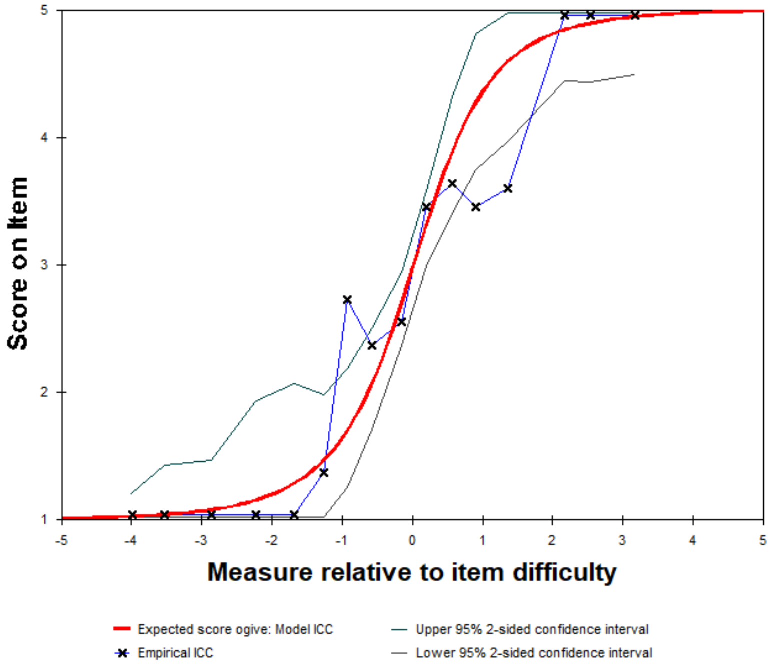


Note. Graph directly imported from *Winsteps*.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.10

S12 Expected and Empirical ICC



Note. Graph directly imported from *Winsteps*.

5.7.5 Enterprising Scale

Three items misfit on the Enterprising scale, namely item E17 ("Negotiate salaries on behalf of people"; outfit MnSq=1.61; infit MnSq=1.36), item E27 ("Teach people how to improve their managerial skills"; outfit MnSq=1.56), and item E12 ("Manage a hotel"; outfit MnSq=1.49; infit MnSq=1.36). Item E17 had an average item difficulty with an item measure of 0.00 (Appendix E). Item E27 was the third easiest item to endorse with an item measure of −0.36, and item E12 was the easiest item to endorse with an item measure of −0.40 (Appendix E). The misfit in item E17 was due to the unexpected responses of eight persons to the item. Twelve test-takers responded unexpectedly to item E27 and 13 people responded unexpectedly to item E12.

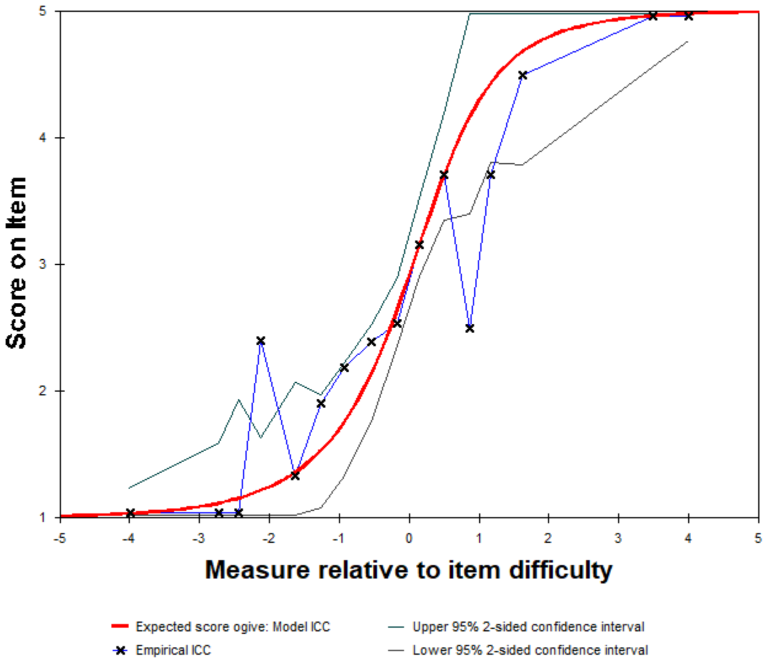
The Item Characteristic Curve of item E17 showed that the item misfits for low performers and moderately high performers (Figure 5.11). Item E27 only functioned properly

Stellenbosch University https://scholar.sun.ac.za

for persons within −1 to 1 logits (Figure 5.12), and item E12 misfits for moderately low performers (Figure 5.13).

Figure 5.11

E17 Expected and Empirical ICCs

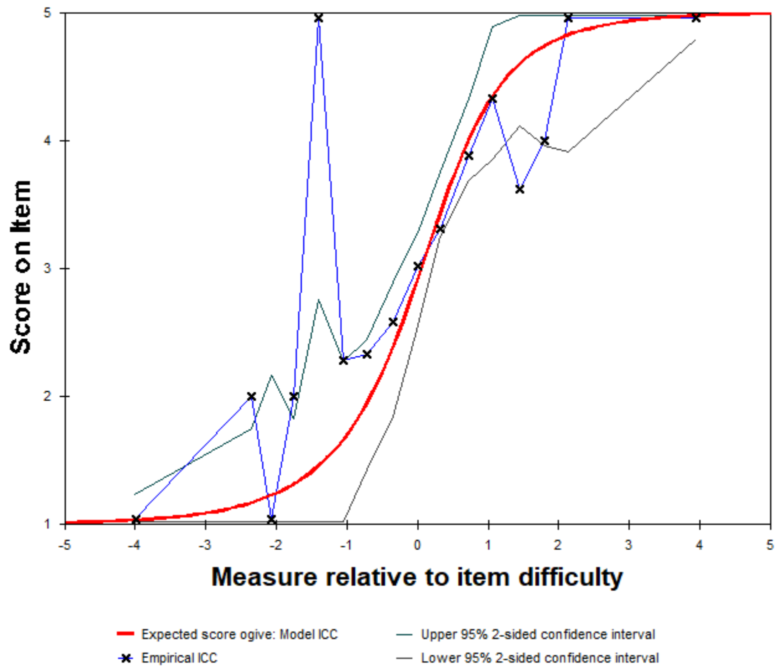


Note. Graph directly imported from *Winsteps*.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.12

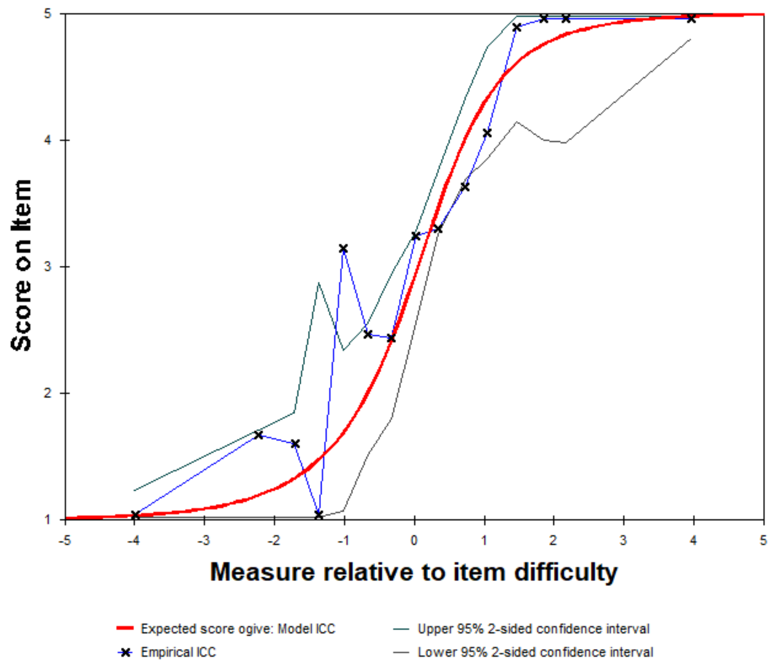
E27 Expected and Empirical ICCs



Note. Graph directly imported from *Winsteps*.

Figure 5.13

E12 Expected and Empirical ICCs



Note. Graph directly imported from *Winsteps*.

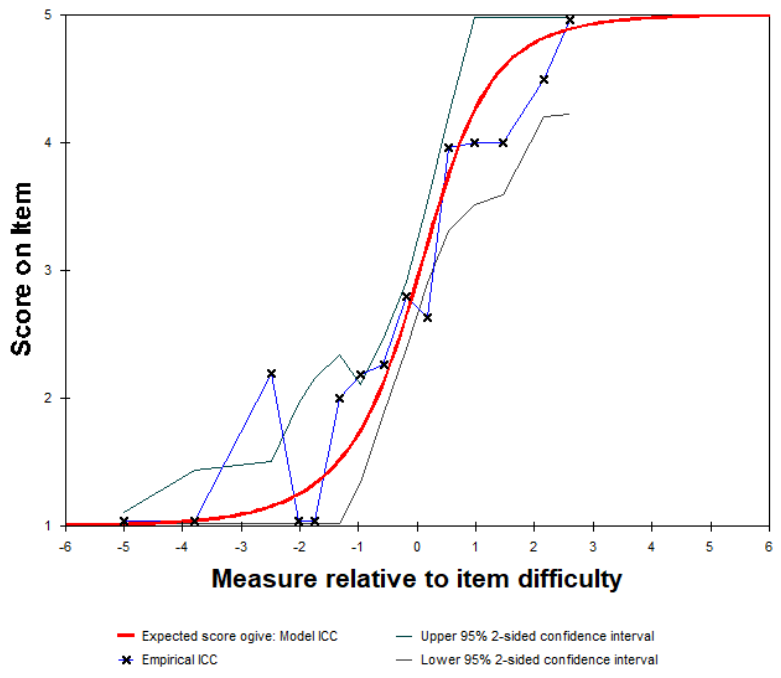
Stellenbosch University https://scholar.sun.ac.za

5.7.6 Conventional Scale

Two items showed item fit values above the cut-off of 1.30, namely item C14 ("Maintain a database"; outfit MnSq=1.51) and item C15 ("Maintain import and export documents"; infit MnSq=1.31). Item C14 had a moderate endorsability with an item measure of −0.05, and item C15 had a moderately low endorsability of −0.09 (Appendix E). The misfit in item C14 was due to the unexpected responses of nine test-takers, while eight persons responded unexpectedly to item C15. The Item Characteristic Curve of item C14 showed that the item misfits for low performers (Figure 5.14). The Item Characteristic Curve of item C15 showed no misfit (Figure 5.15). Due to this, as well as the outfit MnSq indicating no misfit, and the infit MnSq only indicating slight misfit, the misfit of Item C15 was ignored.

Figure 5.14

C14 Expected and Empirical ICCs

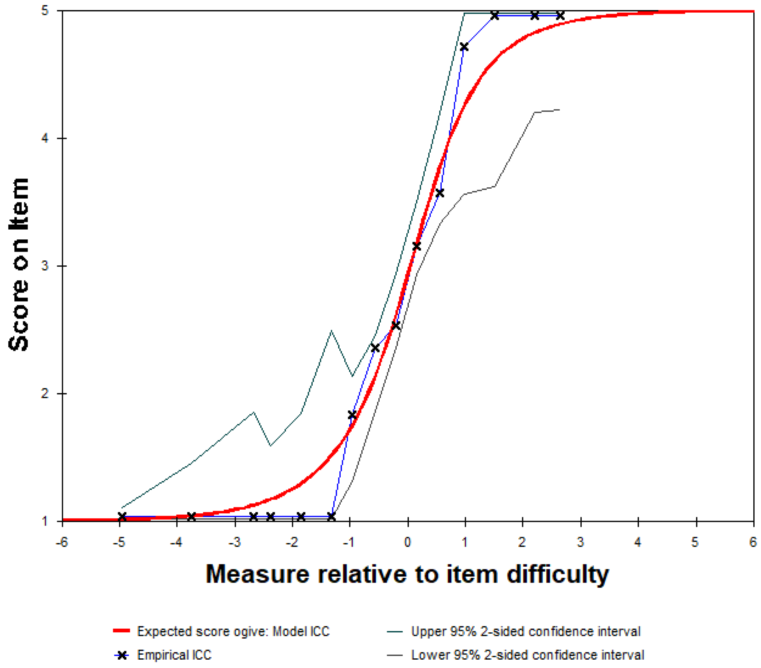


Note. Graph directly imported from *Winsteps*.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.15

C15 Expected and Empirical ICCs



Note. Graph directly imported from *Winsteps.*

5.8 Person Fit

A large number of test-takers displayed both underfit and overfit on each of the six SACII-X scales. On the Realistic Scale, 64 people displayed underfit, with infit and/or outfit MnSq values larger than 1.30, and 74 people displayed overfit with infit and/or outfit MnSq values less than .70. On the Investigative Scale, 72 people displayed underfit, and 74 people displayed overfit. On the Artistic Scale, 77 people displayed underfit, and 72 displayed overfit. The Social scale had 65 underfitting persons, and 68 overfitting persons. On the Enterprising Scale, 60 persons underfit and 72 persons overfit. The Conventional Scale had 80 underfitting persons, and 74 overfitting persons.

It was decided not to remove any misfitting persons prior to the Rasch analysis and the analysis of item fit. This analytical decision was based on a number of reasons. Firstly, the secondary data were cleaned by the primary investigator of the original study, and persons with biased response patterns were removed. Secondly, visual inspection of the raw person data did not reveal any apparent response sets and concerning response patterns in the data. It

Stellenbosch University https://scholar.sun.ac.za

could therefore not be concluded with certainty whether persons purposefully provided invalid responses, or whether inconsistent or poor responding was due to true inherent qualities of the person, or a misunderstanding of the items due to, for example, language- and/or cultural factors.

5.9 Person and Item Targeting

On the Wright map, person and item measures are presented on the same logit scale. On the left-most column of the Wright map, the logit scale is presented. Person locations are presented on the left of the vertical line, and item locations are presented on the right of the vertical line. Both person and item locations are presented from high at the top, to low at the bottom (Araydoust et al., 2021). An average is presented on the vertical line for both person locations and item locations, signified with an *M*. One standard deviation from the mean, is denoted by an *S* and two standard deviations from the mean is denoted by a *T*.

5.9.1 Realistic Scale

The Wright Map (Figure 5.16) indicates that the items covered a relatively narrow range of difficulty. Item measures ranged from −0.46 logits to 0.30 logits (Appendix E). Excluding the five persons who scored the minimum measure on this scale, persons in the sample were located in the range of −3.73 to 2.72 logits.

The average person measure of this sample was −0.51. In general, the items were targeted above the average person ability level. There were only two items located at the person mean, and no items located below the person mean. From this, one can deduce that the items will be less successful differentiating between the person abilities of persons falling below the person mean, than for persons falling above the person mean.

Figure 5.16

Wright Map for the Realistic Scale

```
MEASURE    PERSON - MAP - ITEM
                <more><rare>
    3                 +
                   .


    2                 +

                   .

                   . T
    1              #  +
                   .
                 .##
               .#### ST
                ####  S RR13   RR21   RR24   RR25   RR27   RR28
    0       .########  +M RR10   RR11   RR14   RR17   RR18   RR2    RR20   RR23
                         RR26   RR4    RR7    RR9
         .##########  S RR15   RR19   RR8
           .######## MT RR1    RR16
            .#######
                 ###
   -1            .##  +
                #### S
                  .#
                  .#
                   .
   -2              . T+
                   .
                   .
                   .

   -3              .  +



                   .
   -4             .#  +
                <less><freq>
 EACH "#" IS 4: EACH "." IS 1 TO 3
```

Note. Figure directly imported from *Winsteps*.

5.9.2 Investigative Scale

The average person measure in this sample was −.25. Items were targeted at a slightly higher ability level than the average person measure. But since the item mean and person mean are close together, in general, items are well-targeted to the person sample. All the items are located within one standard deviation above the person mean and about a third of a standard deviation below the person mean, where the majority of persons are located (Figure 5.17). Items will therefore be able to distinguish well between the ability levels of most persons. Items will be able to discriminate less well between individuals on the higher end of the trait, and between individuals on the lower end of the trait. With the exclusion of the five persons with the minimum score on the scale, person ability levels ranged from −2.52 to 2.32. No gaps in item locations were observed.

Figure 5.17

Wright Map for the Investigative Scale

```
MEASURE     PERSON - MAP - ITEM
                  <more><rare>
   3                      +



                      .

   2                  .   +
                      .
                      .
                      .
                      .
                      . T
   1                  .   +
                     .#
                      .
                    .## S
                 ######  T I8
                 .#####  S I11    I13    I19    I25    I5     I6     I9
   0       .###########  +M I1     I17    I2     I20    I21    I22    I24
             .######### MS I10    I16    I26
               .#######  T I12    I23    I4
                   .###
                   .###
                   .### S
  -1                  .   +
                     .#
                      #
                     .# T
                      .
                      .
  -2                      +
                      .

                      .


  -3                 .#   +
                  <less><freq>
 EACH "#" IS 4: EACH "." IS 1 TO 3
```

Note. Figure directly imported from *Winsteps*.

5.9.3 Artistic Scale

The mean of persons in this sample was −.06. The average person measure was therefore only very slightly below the average item difficulty. The Wright map visually shows

Stellenbosch University https://scholar.sun.ac.za

that the mean of persons and the mean of items were well-targeted (Figure 18). Person ability levels ranged from −2.50 to 2.12. However, items were all located very close to the person mean (about half a standard deviation above the person mean, and about a third of a standard deviation below the person mean). This pool of items will therefore not be able to distinguish well between persons who lie further away from the person mean. But since quite a few items lie close to the person mean, items will be able to distinguish well between the ability levels of the majority of persons.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.18

Wright Map for the Artistic Scale

```
MEASURE    PERSON - MAP - ITEM
               <more><rare>
    3                +




                  .
    2                +

                  .
                  .
                  .
                  .
    1             . T+
                  .
                .##
             .##### S
             .#####  T A27
           .#######  S A16    A22    A29    A30    A4     A9
    0      ######## M+M A10    A15    A18    A19    A2     A20    A21    A24
                        A26    A28    A6     A8
      #############  S A1     A11    A23    A25    A5     A7
          .########  T A13
               .###
                 ## S
                .##
   -1             #  +
                 .# T
                  .
                  .
                  .
                  .
   -2             .  +


                  .


   -3                +
               <less><freq>
 EACH "#" IS 4: EACH "." IS 1 TO 3
```

Note. Figure directly imported from *Winsteps*.

5.9.4 Social Scale

The average person measure in this sample was 0.5. As can also be visually seen on the Wright map, the person mean and item mean were thus well targeted (Figure 5.19). Items are located from −.37 logits to .34 logits. Excluding the three persons with the minimum measure on the scale, person ability level ranged from −.31 logits to 3.29 logits. Items will be able to distinguish well between the ability level of the majority of persons, but not between persons falling outside two standard deviations from the person mean.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.19

Wright Map for the Social Scale

```
MEASURE     PERSON - MAP - ITEM
                <more><rare>
    4                   +



                     .
    3                   +
                     .

                     .
    2                   +
                       T
                     .
                     .
    1                .##  +
                    .### S
                .#######  T
               .########  S S16    S2     S22    S25    S26    S28
    0 .############ M+M S12    S15    S17    S18    S20    S27    S3     S8
                   .####  S S10    S11    S23    S24    S4     S5     S6     S7
                   #####  T
                      ## S
   -1                 ##  +
                     .
                     . T

   -2                .  +



                     .
   -3                   +

                     .


   -4                .  +
                <less><freq>
  EACH "#" IS 5: EACH "." IS 1 TO 4
```

Note. Figure directly imported from *Winsteps*.

5.9.5 Enterprising Scale

The average person measure in this sample was −.17. As can also be seen on the Wright map, the item mean and person mean were very close together. This shows that, in general, the item pool is well-targeted to the person sample. Excluding the extreme persons in this sample, person ability ranged from −2.72 to 3.48, while item difficulty ranged from −40 to .53 logits. All the items were located within one standard deviation above the person mean and almost half a standard deviation below the person mean, where many persons were also

Stellenbosch University https://scholar.sun.ac.za

located. Items will thus be able to distinguish well between the ability levels of the majority of test-takers. Items will not be able to distinguish well between persons located beyond these boundaries.

Figure 5.20

Wright Map for the Enterprising Scale

```
MEASURE    PERSON - MAP - ITEM
                  <more><rare>
    4                   .  +

                        .

    3                      +




    2                      +
                        .

                        .
                      .# T
    1                   .  +
                        #
                     .## S  E22
                   #####  T E3
                   #####  S E13    E18    E2     E20    E21    E28    E9
    0         .#########  +M E1     E16    E17    E23    E25    E5     E7     E8
            .########### MS E10    E14    E15    E19    E24
                   .####  T E11    E12    E27
                     .##
                      .# S
   -1                 ##  +
                      .#
                        #
                        . T
                        .
   -2                   .  +
                        .
                        .

                        .
   -3                   .  +
                  <less><freq>
 EACH "#" IS 5: EACH "." IS 1 TO 4
```

Note. Figure directly imported from *Winsteps*.

Stellenbosch University https://scholar.sun.ac.za

5.9.6 Conventional Scale

Appendix E shows that the average person mean was −.34. Thus, in general, items were targeted slightly above the average person ability, as can also be seen on the Wright Map (Figure 5.21). Excluding five extreme persons, test-takers' ability levels ranged from −3.85 to 2.56. Item endorsability ranged from −40 to .37. Items therefore covered a narrow range of endorsability, compared to person ability level. Hence, items will be more effective at differentiating between persons close to the person mean than differentiating between persons further away from the person mean.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.21

Wright Map for the Conventional Scale

```
MEASURE     PERSON - MAP - ITEM
                <more><rare>
   3                 +
                   
                   
                  .
                  
                  .
   2                 +
                   
                   
                  .
                  .
                  . T
   1              .  +
                 .#
                 ##
               .### ST
           .#######    C28    C3     C5
           .#######  S C11    C12    C16    C18    C20    C24    C25    C29
                        C9
   0    .##########  +M C1      C14
      .########### MS C15    C17    C19    C2     C21    C27    C7     C8
       ###########     C22    C26    C4
          .########  T
             .#####
             .#####
  -1           #### S+
                 .#
                  .
                  .
                    T
                  .
  -2                 +
                  .
                   
                 .#
                   
                  .
  -3                 +
                   
                   
                   
                   
                  .
  -4             .#  +
                <less><freq>
 EACH "#" IS 3: EACH "." IS 1 TO 2
```

Note. Figure directly imported from *Winsteps*.5.10 Differential Item Functioning

Prior to the DIF analysis, one person was deleted from the analysis, as this person did not indicate their gender. The DIF Contrast value only highlighted one item of the SACII-X as displaying uniform DIF. Several items across the SACII-X showed indications of non-uniform DIF.

5.10.1 Realistic Scale

Table 5.6 presents the results of the uniform DIF analysis. The DIF Contrast value indicated that item R10 ("Fix a car engine") showed evidence of slight uniform DIF. Item

Stellenbosch University https://scholar.sun.ac.za

R10 had a DIF Contrast value of .47 (*p*=.0001) and was easier to endorse for males than for females. Visual inspection of Item Characteristic Curve indicated non-uniform DIF in items RR17 ("Monitor electricity use at a power plant"; Figure 5.22) and RR19 ("Operate heavy machinery"; Figure 5.23). On item RR17, low scoring males found the item easier to endorse than low scoring females, whereas low scoring females found item RR19 easier to endorse than low scoring males.

Table 5.6

Realistic Scale DIF

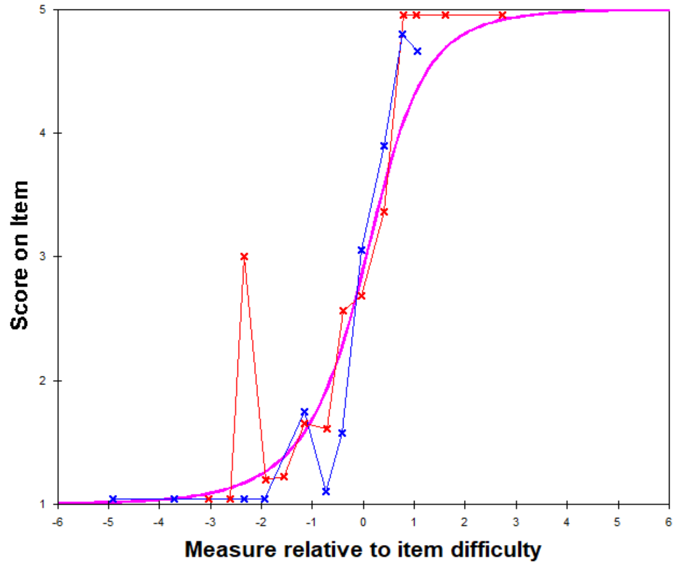
			Rasch-Welch			Mantel	
Item	DIF Contrast	Joint *S.E.*	*t*	*d.f.*	Prob.	Chi-square	Prob.
RR1	-.09	.11	-.83	210	.4078	.6858	.4076
RR10	-.47	.12	-4.00	233	**.0001**	9.2713	**.0023**
RR11	-.15	.11	-1.36	223	.1764	1.3436	.2464
RR13	.22	.12	1.83	218	.0686	4.0433	.0443
RR14	-.09	.12	-.79	225	.4308	.6121	.4340
RR15	.12	.11	1.07	218	.2846	.1123	.7376
RR16	-.02	.11	-.19	210	.8462	.7470	.3874
RR17	.06	.11	.48	221	.6311	1.0525	.3049
RR18	.40	.11	3.52	211	.0005	4.8802	.0272
RR19	.22	.11	2.00	214	.0465	1.1098	.2921
RR2	.12	.12	1.08	219	.2813	2.0958	.1477
RR20	.13	.12	1.10	220	.2723	.1775	.6735
RR21	.00	.12	.00	228	1.000	2.0417	.1530
RR23	-.10	.11	-.90	222	.3677	1.6499	.1990
RR24	.00	.12	.00	228	1.000	.7465	.3876
RR25	-.14	.12	-1.11	233	.2675	.0889	.7656
RR26	-.31	.12	-2.66	233	.0084	5.5805	.0182
RR27	-.18	.12	-1.51	231	.1313	.1295	.7189
RR28	-.10	.12	-.85	231	.3986	.0613	.8045
RR4	-.12	.12	-1.03	226	.3028	1.3116	.2521
RR7	.14	.11	1.20	219	.2301	.0399	.8417
RR8	.15	.11	1.31	215	.1925	4.2172	.0400
RR9	.19	.12	1.67	217	.0971	1.0009	.3171

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.22

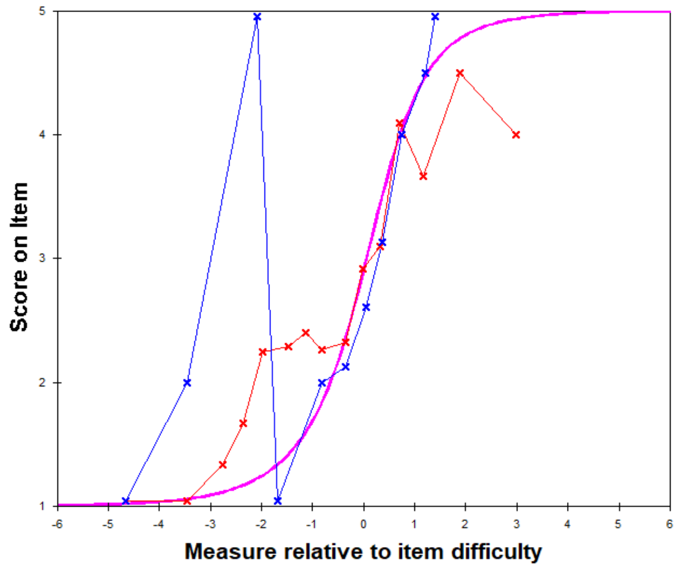
RR17 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.23

RR19 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC, and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

5.10.2 Investigative Scale

Table 5.7 presents the results of the uniform DIF analysis. Based on the DIF Contrast, no items showed uniform DIF on the Investigative Scale. Non-uniform DIF was indicated in four items. Low performing males found item I9 ("Research laser technology"; Figure 5.24) easier to endorse than low performing females. Item I12 ("Research ways to improve space flight"; Figure 5.25) favoured low performing females over low performing males. And lastly, high scoring males found items I21 ("Study plants in the environment"; Figure 5.26) and I26 ("Use new ways of removing diseases from crops"; Figure 5.27) harder to endorse than high performing females.

Stellenbosch University https://scholar.sun.ac.za

Table 5.7

Investigative Scale DIF

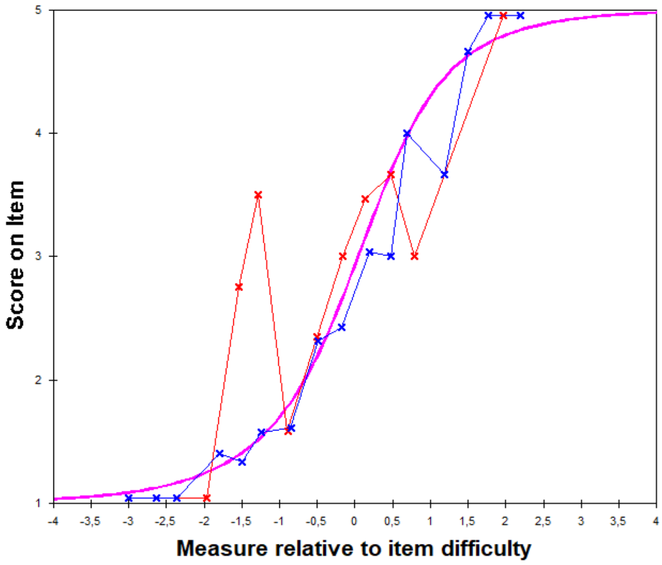
Item	DIF Contrast	Joint *S.E.*	Rasch-Welch *t*	*d.f.*	Prob.	Mantel Chi-square	Prob.
I1	.00	.11	.00	207	1.000	.3178	.5729
I10	-.14	.11	-1.32	205	.1896	.7040	.4014
I11	.07	.11	.60	208	.5511	.3904	.5321
I12	.00	.11	.00	206	1.000	.2926	.5886
I13	.20	.11	1.72	212	.0873	4.4626	.0346
I16	-.41	.11	-3.69	201	.0003	15.2128	.0001
I17	-.06	.11	-.52	205	.6043	.2761	.5992
I19	.22	.11	1.87	214	.0628	6.1563	.0131
I2	-.26	.11	-2.32	201	.0211	5.0429	.0247
I20	.24	.11	2.11	211	.0357	.4330	.5105
I21	.16	.11	1.41	208	.1614	2.2139	.1368
I22	-.06	.11	-.57	205	.5698	1.1173	.2905
I23	-.21	.11	-1.91	207	.0580	5.1707	.0230
I24	.15	.11	1.36	208	.1742	.9036	.3418
I25	.07	.11	.59	208	.5587	.3566	.5504
I26	-.09	.11	-.85	205	.3941	.2598	.6103
I4	-.23	.11	-2.11	206	.0359	2.2779	.1312
I5	.00	.11	.00	207	1.000	.0000	.9974
I6	-.06	.11	-.57	204	.5700	.5836	.4449
I8	.18	.12	1.57	214	.1190	3.2809	.0701
I9	.31	.11	2.74	214	.0066	11.5390	.0007

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.24

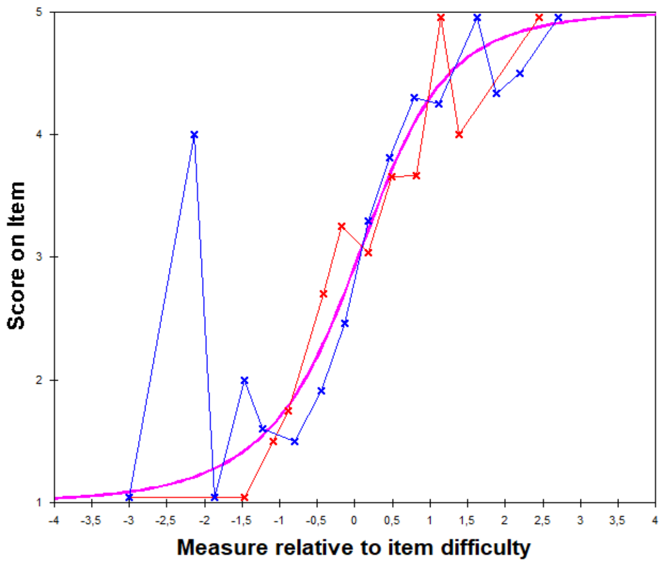
I9 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.25

I12 Non-Uniform DIF ICCs for Males and Females

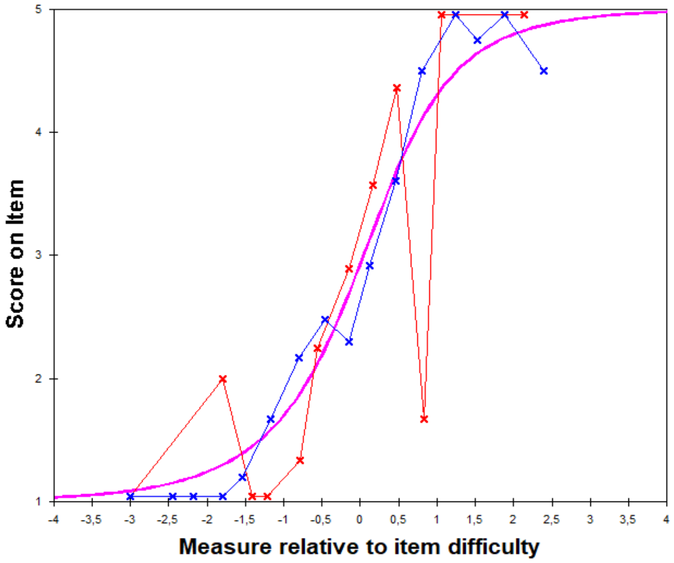


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.26

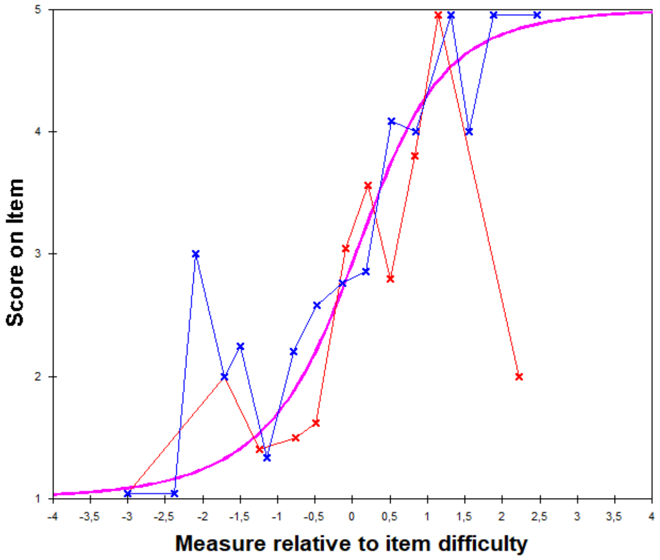
I21 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.27

I26 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

5.10.3 Artistic Scale

Table 5.8 presents the results of the uniform DIF analysis. Based on the DIF Contrast, no items showed uniform DIF on the Artistic scale, but four items showed evidence of non-uniform DIF. Visual inspection of the Item Characteristic Curves indicated that items A7 ("Go to theatre productions"; Figure 5.28) and A11 ("Play a musical instrument"; Figure 5.29) were easier to endorse for low performing females than for low performing males, item A23 ("Write a musical score"; Figure 5.30) was easier to endorse for moderately high performing males than for moderately high performing females, and item A27 ("Write jingles for advertisements"; Figure 5.31) was easier to endorse for average scoring males than for average scoring females.

Stellenbosch University https://scholar.sun.ac.za

Table 5.8

Artistic Scale DIF

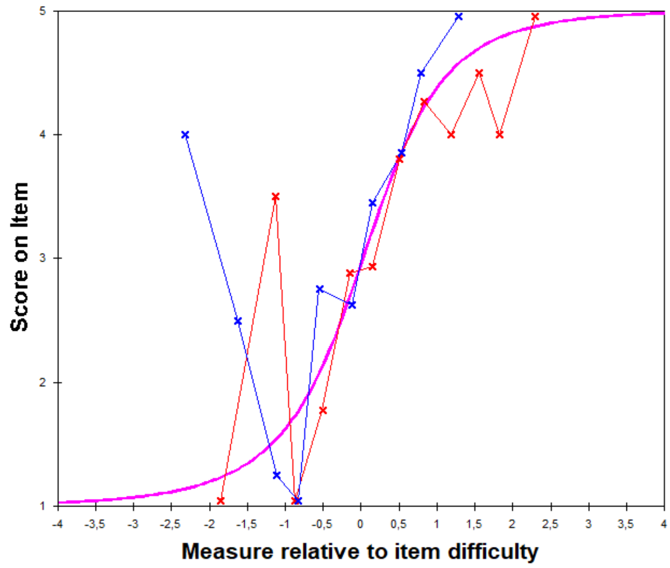
Item	DIF Contrast	Rasch-Welch				Mantel	
		Joint *S.E.*	*t*	*d.f.*	Prob.	Chi-square	Prob.
A1	-.20	.10	-1.90	208	.0582	2.0419	.1530
A10	.00	.10	.00	209	1.000	.1018	.7497
A11	.00	.10	.00	209	1.000	.4475	.5035
A13	.30	.10	2.88	213	.0043	6.6158	.0101
A15	.00	.10	.00	208	1.000	.1857	.6665
A16	.00	.11	.00	208	1.000	.1153	.7342
A18	-.15	.10	-1.49	210	.1371	7.9612	.0048
A19	.03	.10	.30	209	.7617	.3356	.5624
A2	-.12	.10	-1.19	210	.2367	1.2437	.2648
A20	-.13	.10	-1.27	210	.2050	.7003	.4027
A21	.02	.10	.23	208	.8221	.1206	.7284
A22	.07	.11	.70	207	.4857	.0433	.8351
A23	-.03	.10	-.30	209	.7620	.0604	.8058
A24	-.14	.10	-1.34	210	.1818	1.3530	.2448
A25	.00	.10	.00	209	1.000	.0263	.8711
A26	.10	.10	1.00	208	.3190	2.4056	.1209
A27	.12	.11	1.09	204	.2771	.4024	.5258
A28	.15	.10	1.48	208	.1392	1.1478	.2840
A29	.13	.11	1.22	206	.2239	.3184	.5726
A30	.00	.11	.00	208	1.000	.1262	.7224
A4	.10	.11	.92	206	.3590	.9537	.3288
A5	.16	.10	1.57	210	.1180	1.9289	.1649
A6	-.10	.10	-.93	209	.3523	.2823	.5952
A7	-.16	.10	-1.49	207	.1379	1.1794	.2775
A8	-.03	.10	-.30	209	.7677	.0699	.7915
A9	-.20	.11	-1.88	214	.0618	1.9016	.1679

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.28

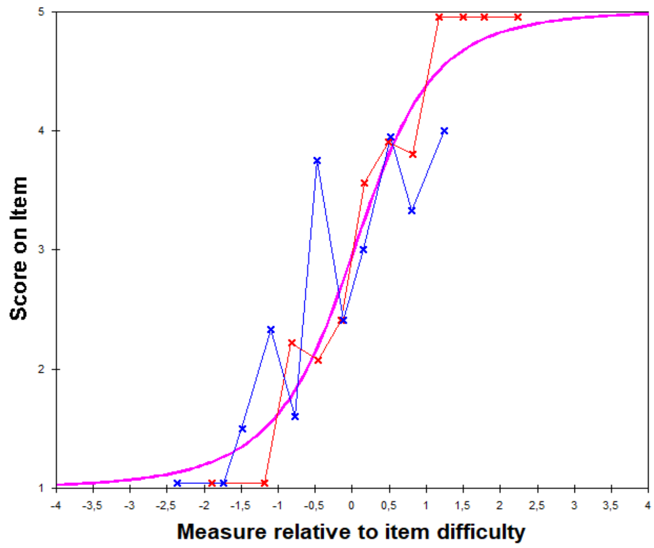
A7 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.29

A11 Non-Uniform DIF ICCs for Males and Females

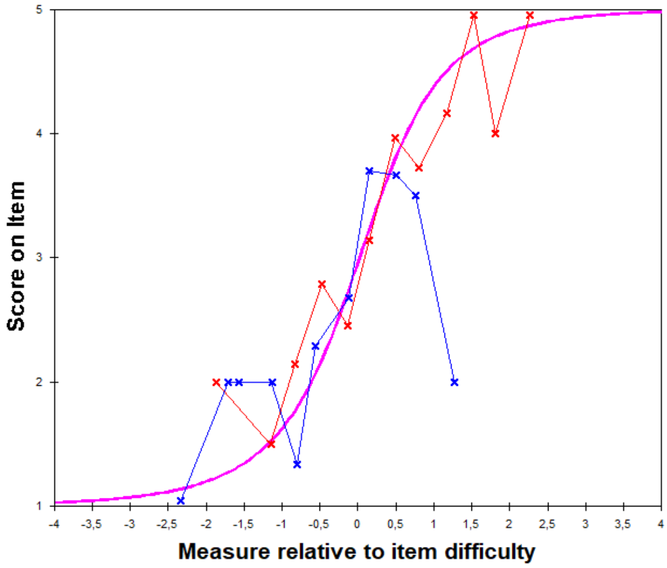


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC

Stellenbosch University https://scholar.sun.ac.za

Figure 5.30

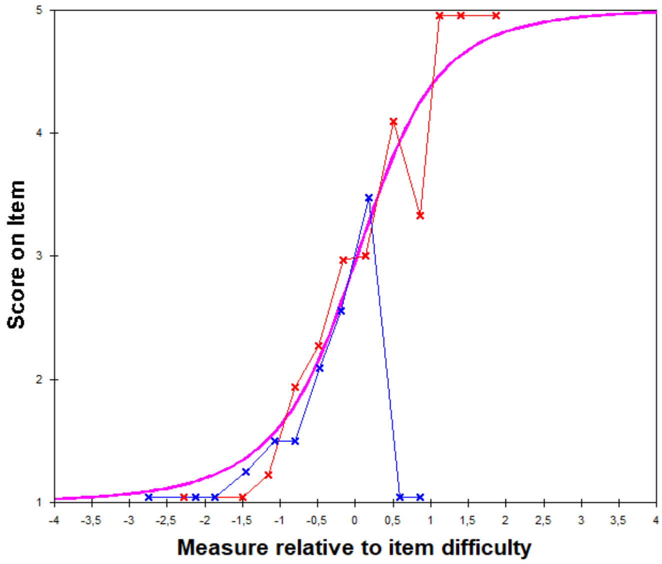
A23 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.31

A27 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

5.10.4 Social Scale

Table 5.9 presents the results of the uniform DIF analysis. Based on the DIF Contrast, no items showed uniform DIF on the Social scale. Four items showed evidence of non-uniform DIF. Items S11 (“Look after people who are sick”; Figure 5.32) and S28 (“Work with elderly people”; Figure 5.33) were easier to endorse for high performing males than for high performing females, item S12 (“Model clothes”; Figure 5.34) was easier to endorse for moderately high performing males than for moderately high performing females, and item S22 (“Teach at a school”; Figure 5.35) was easier to endorse for high performing females than for high performing males.

Stellenbosch University https://scholar.sun.ac.za

Table 5.9

Social Scale DIF

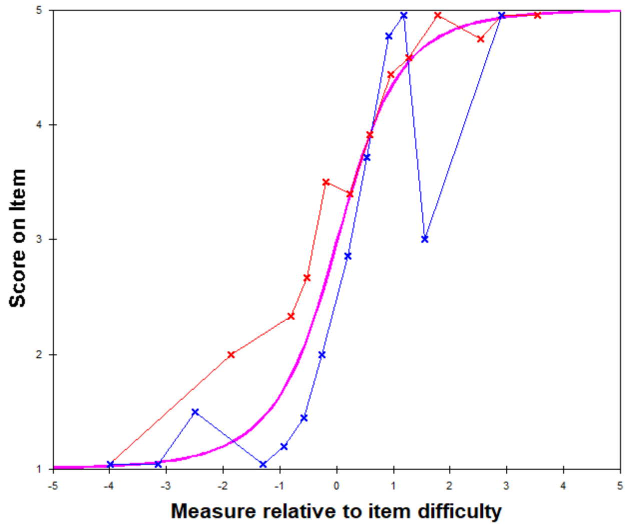
Item	DIF Contrast	Joint *S.E.*	Rasch-Welch			Mantel	
			t	*d.f.*	Prob.	Chi-square	Prob.
S10	.07	.11	.63	213	.5322	.0027	.9586
S11	.35	.11	3.12	218	.0020	3.1386	.0765
S12	-.02	.11	-.19	205	.8498	.8940	.3444
S15	.00	.11	.00	208	1.000	.0635	.8010
S16	-.31	.11	-2.85	209	.0049	3.8181	.0507
S17	-.18	.11	-1.68	206	.0945	1.0150	.3137
S18	-.09	.11	-.84	206	.4023	.0032	.9547
S2	.02	.11	.23	204	.8202	.3861	.5343
S20	.00	.11	.00	207	1.000	.0115	.9145
S22	-.22	.11	-2.06	207	.0403	1.2981	.2546
S23	.03	.11	.27	210	.7853	.9266	.3357
S24	.25	.11	2.28	213	.0235	4.4485	.0349
S25	-.11	.11	-1.00	205	.3175	1.7238	.1892
S26	-.21	.11	-1.91	207	.0570	1.8427	.1746
S27	.00	.11	.00	207	1.000	1.2417	.2651
S28	-.20	.11	-1.85	207	.0655	2.3204	.1277
S3	.12	.11	1.14	205	.2557	1.6895	.1937
S4	.24	.11	2.18	212	.0303	2.5811	.1081
S5	.07	.11	.59	214	.5571	.0896	.7647
S6	.13	.11	1.15	215	.2503	.1096	.7406
S7	.00	.11	.00	210	1.000	.0759	.7830
S8	.07	.11	.62	207	.5332	.3330	.5639

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.32

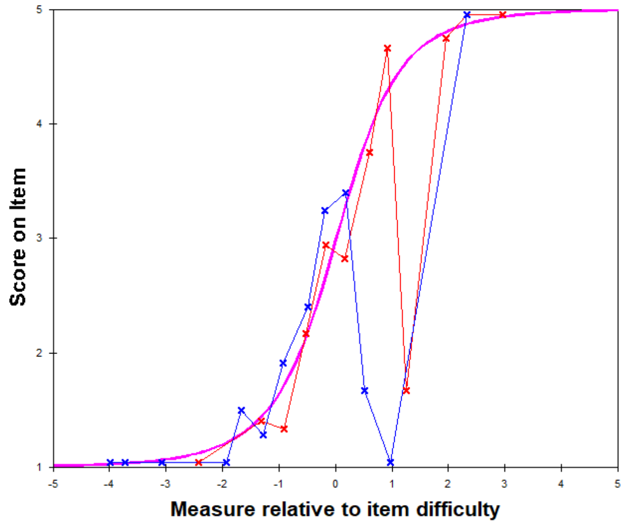
S11 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.33

S28 Non-Uniform DIF ICCs for Males and Females

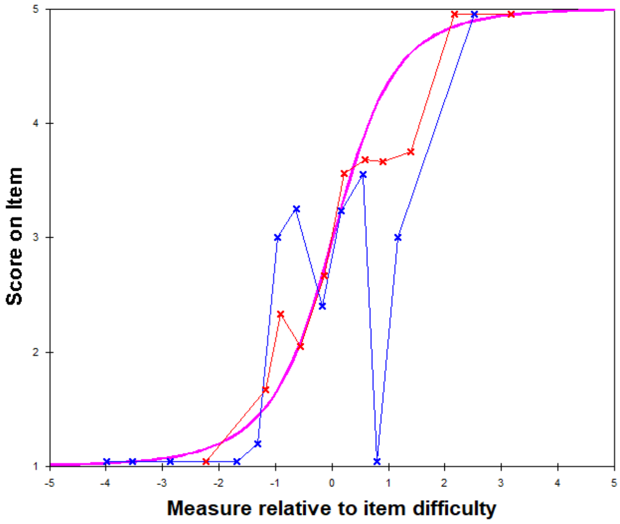


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.34

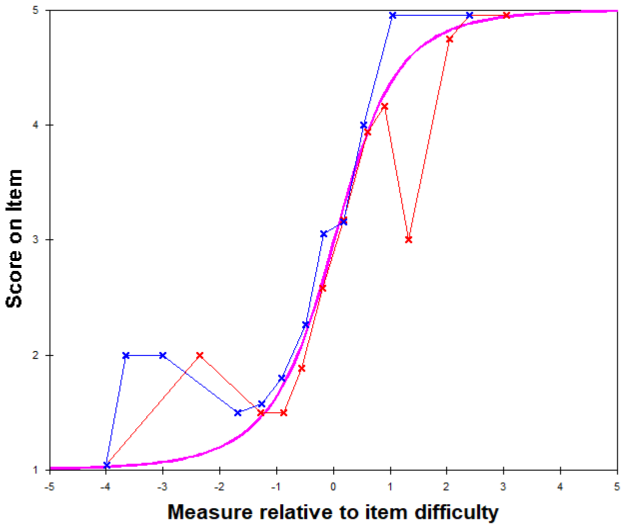
S12 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.35

S22 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

5.10.5 Enterprising Scale

Table 5.10 present the results of the uniform DIF analysis. Based on the DIF Contrast values, the Enterprising scale showed no evidence of uniform DIF. Non-uniform DIF was indicated in four items. Item E17 (“Negotiate salaries on behalf of people”; Figure 5.36) was easier to endorse for low performing females than for low performing males, item E20 (“Sell goods at an auction”; Figure 5.37) was easier to endorse for moderately high performing males than moderately high performing females, item E27 (“Teach people how to improve their managerial skills”; Figure 5.38) was easier to endorse for low performing males than for low performing females, and item E3 (“Assist customers to buy products”; Figure 39) was easier to endorse for moderately high performing females than moderately high performing males.

Stellenbosch University https://scholar.sun.ac.za

Table 5.10

Enterprising Scale DIF

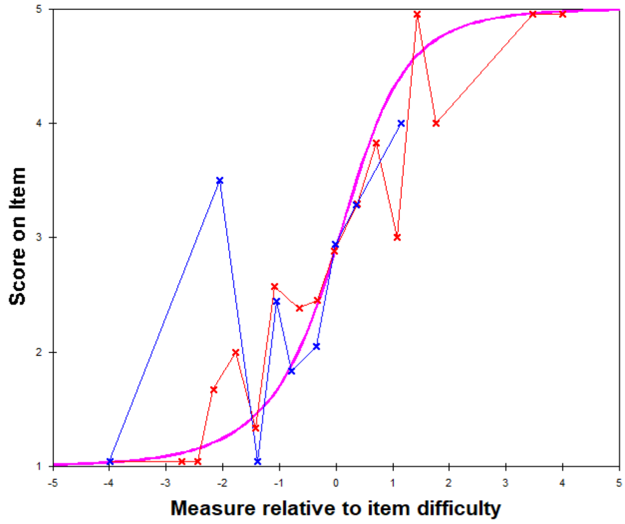
Item	DIF Contrast	Joint *S.E.*	Rasch-Welch			Mantel	
			t	*d.f.*	Prob.	Chi-square	Prob.
E1	-.15	.11	-1.36	213	.1746	1.7021	.1920
E10	.00	.11	.00	211	1.000	.1907	.6623
E11	-.31	.11	-2.74	204	.0068	3.6707	.0554
E12	-.13	.11	-1.17	207	.2425	1.6877	.1939
E13	-.07	.11	-.67	212	.5056	.9093	.3403
E14	-.05	.11	-.50	211	.6194	.2605	.6097
E15	.07	.11	.66	210	.5109	.0334	.8549
E16	.08	.11	.71	209	.4782	.2665	.6057
E17	.00	.11	.00	211	1.000	.0010	.9752
E18	.00	.12	.00	210	1.000	.0008	.9772
E19	.18	.11	1.66	210	.0977	.6622	.4158
E2	.21	.12	1.83	202	.0686	2.3928	.1219
E20	.14	.11	1.25	206	.2131	1.0253	.3113
E21	-.11	.11	-.98	213	.3300	.0556	.8136
E22	-.28	.12	-2.32	222	.0213	.4194	.5173
E23	-.02	.11	-.19	211	.8476	.0713	.7895
E24	.07	.11	.61	211	.5406	.2922	.5888
E25	.06	.11	.52	210	.6068	.8096	.3682
E27	.12	.11	1.09	212	.2784	.0169	.8965
E28	.14	.11	1.27	206	.2063	.0917	.7621
E3	.02	.12	.18	208	.8590	.2848	.5936
E5	-.09	.11	-.78	212	.4335	.5499	.4584
E7	.16	.11	1.41	209	.1591	.9662	.3256
E8	-.10	.11	-.87	212	.3837	.0237	.8777
E9	.07	.11	.60	208	.5472	.5107	.4748

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.36

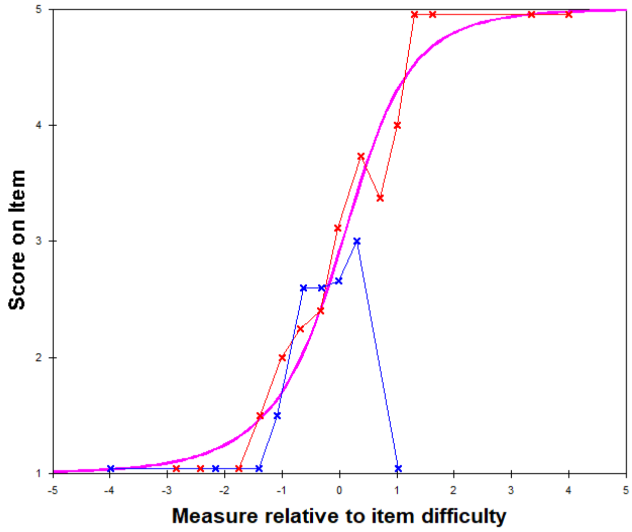
E17 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.37

E20 Non-Uniform DIF ICCs for Males and Females

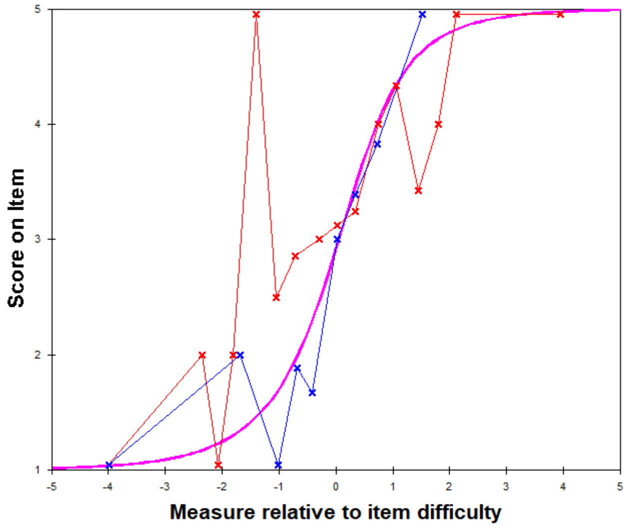


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.38

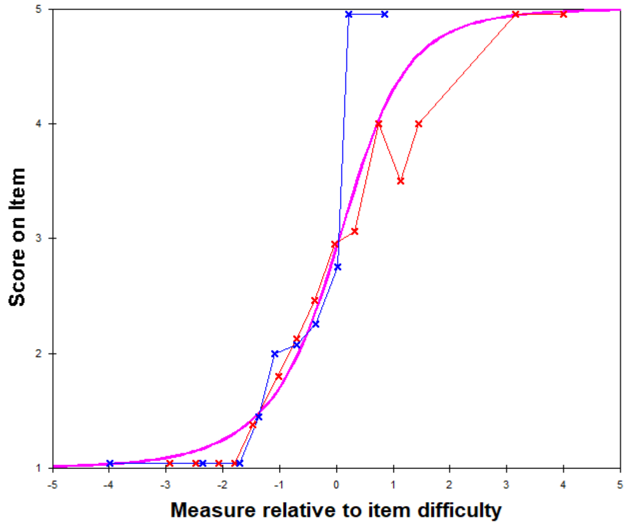
E27 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.39

E3 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

5.10.6 Conventional Scale

Table 5.11 presents the results of the uniform DIF analysis. The Conventional scale showed no evidence of uniform DIF based on the DIF Contrast values, but non-uniform DIF was indicated in seven items. Items C16 ("Maintain records of income received from clients";

Stellenbosch University https://scholar.sun.ac.za

Figure 5.40), C18 (“Monitor credit scores”; Figure 5.41), C25 (“Organise transportation of products from a factory”; Figure 5.42) and C4 (“Calculate wages and salaries of employees”; Figure 5.43) were easier to endorse for low performing females than low performing males. Item C21 (“Monitor the accounts of a company”; Figure 5.44) was easier to endorse for low performing females than low performing males, and easier to endorse for high performing females than high performing males. Item C26 (“Prepare contracts for a company”; Figure 5.45) was easier to endorse for low performing males than for low performing females. Item C29 (“Take merchandise orders”; Figure 5.46) was easier to endorse for average performing males than for average performing females.

Stellenbosch University https://scholar.sun.ac.za

Table 5.11

Conventional Scale DIF

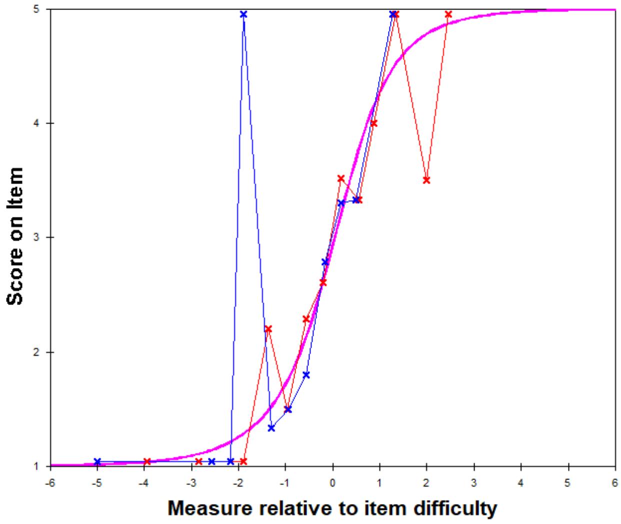
			Rasch-Welch			Mantel	
Item	DIF Contrast	Joint *S.E.*	*t*	*d.f.*	Prob.	Chi-square	Prob.
C1	.07	.11	.61	209	.5437	.0001	.9935
C11	-.09	.12	-.73	214	.4671	1.6063	.2050
C12	.02	.12	.19	209	.8492	.1083	.7420
C14	.14	.11	1.28	208	.2031	.3444	.5573
C15	-.08	.11	-.67	212	.5031	.2789	.5974
C16	.07	.12	.57	209	.5691	.7294	.3931
C17	.20	.11	1.78	209	.0760	3.6080	.0575
C18	.17	.12	1.43	205	.1538	5.7216	.0168
C19	.00	.11	.00	210	1.000	.0052	.9427
C2	.06	.11	.56	210	.5758	.0648	.7990
C20	.00	.12	.00	211	1.000	.5777	.4472
C21	-.14	.11	-1.24	211	.2175	1.2185	.2696
C22	-.05	.11	-.46	210	.6452	.0119	.9133
C24	.14	.12	1.23	206	.2211	.4242	.5148
C25	.00	.12	.00	210	1.000	.0805	.7766
C26	-.10	.11	-.94	209	.3494	1.0909	.2963
C27	.00	.11	.00	211	1.000	.0015	.9687
C28	-.32	.12	-2.59	225	.0101	3.4113	.0648
C29	.10	.12	.82	207	.4111	.0002	.9875
C3	-.17	.12	-1.43	217	.1537	1.8477	.1740
C4	-.22	.11	-2.00	207	.0468	.3730	.5414
C5	.13	.12	1.07	204	.2876	.6727	.4121
C7	-.06	.11	-.51	210	.6076	.0036	.9522
C8	.17	.11	1.49	208	.1373	2.5764	.1085
C9	-.08	.12	-.71	213	.4758	3.3369	.0677

Note. Prob. = probability

Stellenbosch University https://scholar.sun.ac.za

Figure 5.40

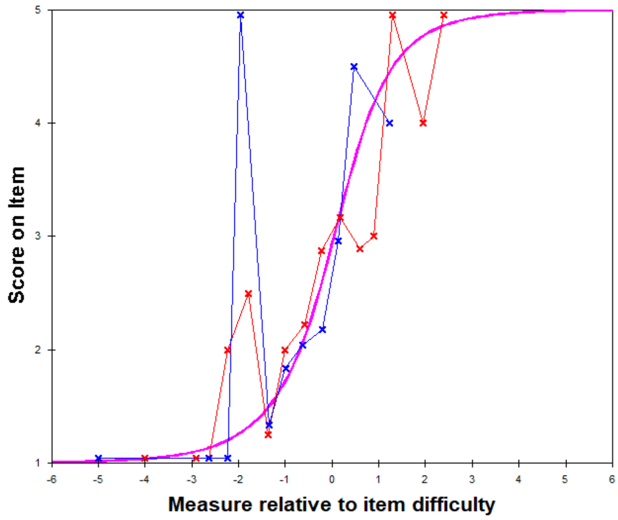
C16 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.41

C18 Non-Uniform DIF ICCs for Males and Females

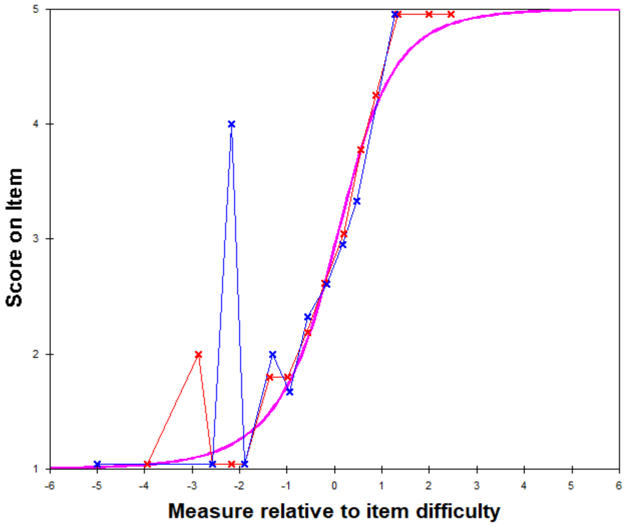


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.42

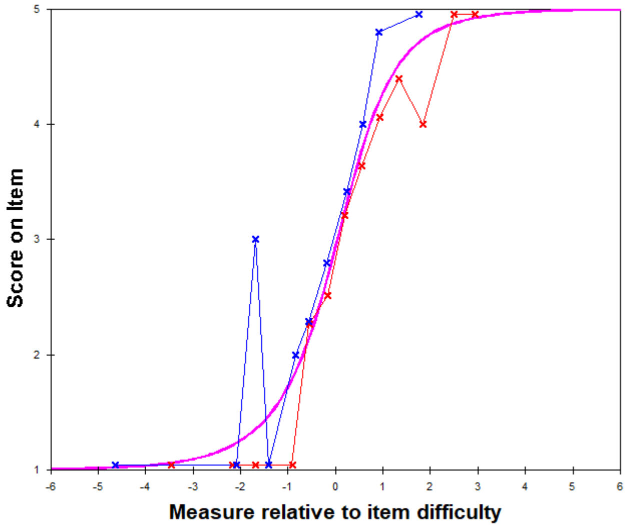
C25 Non-Uniform ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.43

C4 Non-Uniform DIF ICCs for Males and Females

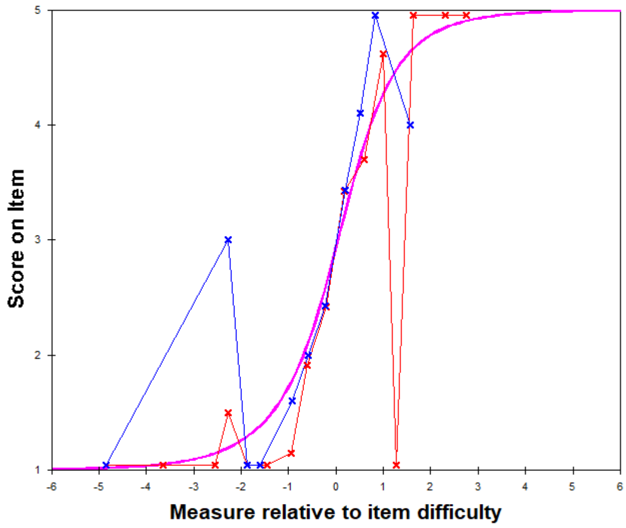


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.44

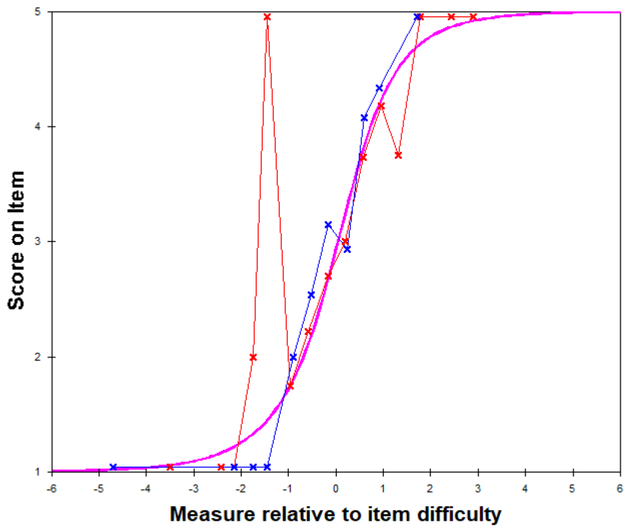
C21 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Figure 5.45

C26 Non-Uniform DIF ICCs for Males and Females

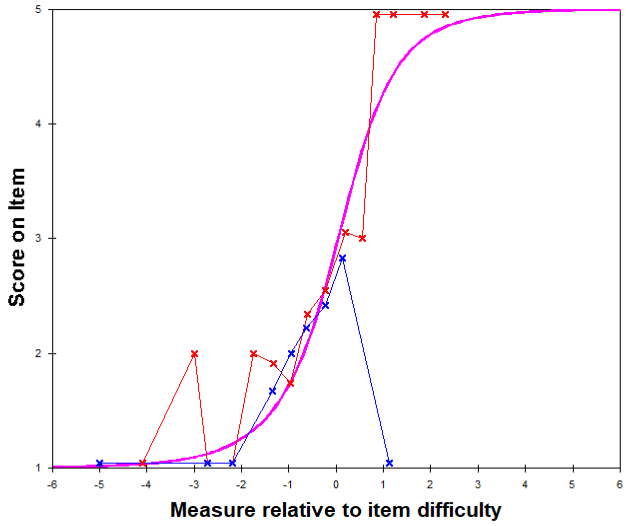


Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.46

C29 Non-Uniform DIF ICCs for Males and Females



Note. The pink ogive represents the model ICC. The red line represents the male ICC and the blue line represents the female ICC.

5.11 Item Removal

5.11.1 Realistic Scale

The two misfitting items on the Realistic scale were removed iteratively. First, item RR19 (“Operate heavy machinery”) was removed. The item’s removal resulted in a decrease of .01 logits in both the person separation reliability and the item separation reliability, with the new person separation reliability equal to .86, and the new item separation reliability equal to .91. With the removal of the item, item RR20 (“Plan the layout of cables”) still misfit with an outfit MnSq of 1.71.

Item RR19 was subsequently included in the analysis, and item RR20 removed. After the removal of RR20, item RR19 still misfit with an outfit MnSq of 1.99. Therefore, both misfitting items were removed. With the removal of both items, the person separation reliability decreased with 0.02 logits to .85, while the item separation reliability stayed the same (.92). However, after the removal of these two misfitting items, RR17 (“Monitor electricity use at a power plant”) emerged as misfitting with an outfit MnSq of 1.32. The Item

Stellenbosch University https://scholar.sun.ac.za

Characteristic Curve of the item showed that the item performed poorly for persons with a lower standing on the Realistic vocational personality type. Item RR17 was removed from the analysis, together with items RR19 and RR20. After the removal of all three items, no more items showed misfit. With all three these items removed, the person separation index decreased with 0.16 logits to 2.37 and the item separation index increased with 0.16 logits to 3.48. The person separation reliability decreased with .02 logits to .85, and the item separation reliability stayed the same at .92.

It is important to note that RR19 (“Operate heavy machinery”) was one of only five items at and below one standard deviation from the item mean. The removal of item RR19 therefore decreases the number of items that can differentiate between lower ability persons on the Realistic scale. Item RR20 (“Plan the layout of cables”) was one of nine items at the item mean, and the removal of this item will not have a big effect on the variability of item measures. RR17 (“Monitor electricity use at a power plant”) was also located at the item mean, and its removal will not have a large effect on the item location variability of the Realistic scale.

5.11.2 Investigative Scale

The removal of the only misfitting item on the Investigative scale, I26 (“Use new ways of removing diseases from crops”), led to no change in the person separation reliability, and improved the item separation reliability slightly from .90 to .91 logits. After the removal of I26, no more items misfit. On the Wright map (Figure 5.17), item I26 was one of six items below the item mean.

5.11.3 Artistic Scale

As unidimensionality and local independence analyses of the Artistic scale showed that three items (A1, A2, and A6) possibly load onto a secondary dimension, and that these same items formed three locally dependent item pairs, the items were removed iteratively to

Stellenbosch University https://scholar.sun.ac.za

determine which items had the smallest influence on the reliability of the scale. Item A1 ("Act in a play") was removed first. With item A1 removed, the person separation reliability decreased by .01 logits to .86, while there was no change in the item separation reliability. The unexplained variance in the first contrast decreased to 2.49.

Item A1 was then placed back into the analysis, and item A2 ("Audition for a play") was removed. After the removal of item A2, the person reliability stayed the same, and the item separation reliability increased to .86. Item A7 ("Go to theatre productions") still misfit with an outfit MnSq of 1.48. The unexplained variance in the first contrast decreased to 2.53. Items A1 ("Act in a play") and A6 ("Direct a play") were still locally dependent with a standardised residual correlation of .41.

Hereafter, item A6 ("Direct a play"), another item locally dependent on A1 and on A2, was removed from the scale to determine its influence on the reliability of the scale. Item A6's removal resulted in a decrease in the person separation reliability to .86, while the item separation reliability stayed the same. The unexplained variance in the first contrast decreased to 2.55. Items A1 ("Act in a play") and A2 ("Audition for a play") remained locally dependent.

Therefore, out of the three items, item A2 ("Audition for a play") contributes the least to the reliability of the scale and is the first candidate for removal. Items A1 ("Act in a play") and A6 ("Direct a play") contributes equally to the scale's reliability. Item A1 ("Act in a play"), however, is one of six items below the item mean, while item A6 ("Direct a play") is one of 12 items at the item mean. Since item A6 ("Direct a play") has the same item location as 11 other items of the Artistic scale, item A6 is a candidate for removal.

Subsequently, items A2 ("Audition for a play") and A6 ("Direct a play") were both removed from the scale. With both these items removed, the person separation reliability only decreased slightly by .01 logits to .86 and the item separation reliability increased by .01

Stellenbosch University https://scholar.sun.ac.za

logits to .86. The unexplained variance in the first component decreased to 2.19. The smallest disattenuated correlation was equal to .50, and items A7 (“Go to theatre productions”), A1 (“Act in a play”) and A24 (“Write a play”) clustered together on the scree-plot. No more item pairs were locally dependent.

The removal of none of the above items (A1, A2, and A6) eliminated the misfit in item A7 (“Go to theatre productions”). Item A7 also showed evidence of non-uniform DIF in the DIF analysis. Subsequently, items A2, A6, and A7 were removed from the scale. The person separation reliability decreased by .01 logits to .87, and the item separation reliability increased by .01 logits to .86. The unexplained variance in the first component decreased to 2.01, and the smallest disattenuated correlation was equal to .68. It was therefore concluded that the scale is unidimensional enough. No more item pairs were locally dependent. The removal of the three items caused item A1 to misfit with an outfit MnSq of 1.53. Despite this, no further items were removed from the analysis.

5.11.4 Social Scale

The three misfitting items on the Social scale were also removed iteratively, starting with the worst-fitting item, item S22 (“Teach at a school”). As discussed previously, item S22 also showed evidence of non-uniform DIF. Upon the removal of item S22, person separation reliability decreased with .01 logits to .86, and the item reliability stayed the same. Item S12 (“Model clothes”; outfit MnSq=1.42) and S26 (“Teach people how to do public speaking”; outfit MnSq=1.40) still misfit.

The removal of the other two misfitting items separately, item S26 (“Teach people how to do public speaking”) and item S12 (“Model clothes”), did not improve the misfit in the other items. Hence, items S22 (“Teach at a school”) and S26 (“Teach people how to do public speaking”) were removed simultaneously. The person separation reliability decreased with .01 logits to .86, and the item separation reliability stayed the same. Item S12 (“Model clothes”)

Stellenbosch University https://scholar.sun.ac.za

still misfitted (outfit MnSq=1.45) and S25 ("Teach people how to cook a meal") also emerged as misfitting (outfit MnSq=1.34). Since the outfit MnSq of item S25 ("Teach people how to cook a meal") was close to the cut-off value of 1.30, the misfit in this item was ignored.

Lastly, all three misfitting items were removed. The person separation reliability decreased by .02 logits to .85, while the item separation reliability stayed the same. Three items emerged as misfitting, namely S25 ("Teach people how to cook a meal"; outfit MnSq=1.38), S28 ("Work with elderly people"; outfit MnSq=1.34), and S18 ("Show people how to dress more fashionably"; outfit MnSq=1.32). The removal of these new misfitting items, caused more items to misfit. It was thus decided to stop the removal of items.

5.11.5 Enterprising Scale

First, locally dependent items were removed iteratively, to determine which item contributes the least to the reliability of the Enterprising scale. Item E11 ("Manage a business") was removed first. With E11 removed, there was no change in either of the item- and person separation reliabilities. The unexplained variance in the first component decreased slightly to 2.19, and no more items showed local dependence. Together with the previously misfitting items, item E22 ("Sell properties") also showed misfit (outfit MnSq=1.32). Since the outfit MnSq of item E22 was close to 1.30, the misfit in the item was ignored. Item E11 ("Manage a business") was one of only five items below the person mean. Removing this item therefore decreases the variability in item difficulties. Item E24 ("Start a business") is also one of the five items below the person mean, but slightly closer to the person mean than item E11.

With item E11 ("Manage a business") placed back in the analysis, and item E24 ("Start a business") removed, the person separation reliability and item separation reliability stayed the same. The unexplained variance decreased very slightly to 2.25, and no more items showed local dependence. Therefore, the change caused in the separation- and reliability

Stellenbosch University https://scholar.sun.ac.za

indices by the removal of each locally dependent item, was very similar. The only difference was that the removal of item E11 (“Manage a business”), resulted in a 0.02 decrease in the item separation index, while the removal of item E24 (“Start a business”) resulted in a 0.19 size decrease in the item separation index. Based on this index, item E24 (“Start a business”) would be a better item for removal than item E11 (“Manage a business”).

The removal of the misfitting items, E17 (“Negotiate salaries on behalf of people”), E27 (“Teach people how to improve their managerial skills”), and E12 (“Manage a hotel”) separately, did not eliminate the misfit in the other items. Items E17 (“Negotiate salaries on behalf of people”), and E27 (“Teach people how to improve their managerial skills”) also showed evidence of non-uniform DIF in the DIF analysis. All three misfitting items were subsequently removed. This resulted in no change in the person separation reliability, while the item separation reliability decreased with .01 logits to .92. Items E11 (“Manage a business”) and E24 (“Start a business”) were still locally dependent (standardised residual correlation=.31). Items E14 (“Mentor staff at a company”; infit MnSq=1.33; outfit MnSq=1.52), E22 (“Sell properties”; outfit MnSq=1.49), and E11 (“Manage a business”; outfit MnSq=1.35) emerged as misfitting.

Item E17 (“Negotiate salaries on behalf of people”) was one of eight items at the item mean and removing this item can therefore be equated to removing a redundant item. Items E27 (“Teach people how to improve their managerial skills”) and E12 (“Manage a hotel”) were two of only three items located two standard deviations from the person mean. Removing these items will therefore decrease the already limited number of items that can distinguish between lower performers.

Lastly, the chosen locally dependent item, E11 (“Manage a business”), was removed with the three misfitting items, E17 (“Negotiate salaries on behalf of people”), E27 (“Teach people how to improve their managerial skills”), and E12 (“Manage a hotel”). The person

Stellenbosch University https://scholar.sun.ac.za

separation reliability remained the same, and the item separation reliability decreased to .91. The unexplained variance in the first contrast decreased to 1.86 and no items were locally dependent. Items E22 ("Sell properties"; outfit MnSq=1.67) and E14 ("Mentor staff at a company"; infit MnSq=1.36, outfit MnSq=1.66) showed misfit. No more items were removed.

5.11.6 Conventional Scale

First, each locally dependent item was removed iteratively to investigate its impact on the reliability- and separation indices. First, item C27 ("Prepare financial records for a company") was removed. This removal decreased the person separation reliability by .01 logits to .86 and increased the item separation reliability with .01 logits to .94. The unexplained variance in the first contrast increased slightly to 2.01. No more items showed local dependence, but item C14 still misfitted ("Maintain a database"; outfit MnSq=1.46).

Hereafter, item C27 ("Prepare financial records for a company") was placed back in the analysis and item C8 ("Control accounting records"), the other locally dependent item in the pair, was removed. The person separation reliability decreased to .86, and the item separation reliability increased to .94. The removal of the two items therefore had the same impact on the reliability of the scale. The unexplained variance in the first contrast increased to 2.05, and no more items showed local dependence. Item C14 ("Maintain a database") still misfit.

One of these items can be removed as the local independence in these items, as well as their same location on the Wright map, implies that one of these items may be redundant and not contributing useful information above the other. As both items have the same effect on the reliability of the scale, each item was removed iteratively with item C14 ("Maintain a database"), the misfitting item.

Stellenbosch University https://scholar.sun.ac.za

With both items C27 (“Prepare financial records for a company”) and C14 (“Maintain a database”) removed, the person separation reliability decreased by .01 logits to .86, and the item separation reliability increased by .01 logits to .94. The removal of items C8 (“Control accounting records”) and C27 (“Prepare financial records for a company”) led to the same change in the reliability indices. Therefore, either of the two locally dependent items can be removed from the scale with the same effect on the scale.

5.12 Chapter Summary

In this chapter, the results of the Rasch Rating Scale Model analyses were provided. Overall, relatively good fit of the data to the Rasch model was found. The separation and reliability indices were satisfactory across all scales. While the majority of the scales were adequately unidimensional, some deviations from unidimensionality were observed on the Artistic scale. Despite the good overall fit, 12 items showed misfit across the SACII-X scales. One item of the Realistic Scale showed evidence of uniform DIF, while 25 items showed signs of non-uniform DIF. Lastly, the misfitting SACII-X items were removed iteratively to determine the influence of their removal on model fit. Overall, the reliability and validity of the SACII-X were supported. Some problematic items were identified that can potentially be adapted or removed from the SACII-X. In the ensuing chapter, a discussion of the results is provided.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 6

DISCUSSION AND CONCLUSION

6.1 Introduction

In this chapter, the results of the study are discussed and interpreted. First, the aims, objectives and hypotheses of the study are presented. A discussion of the results for each objective follows. This is followed by the implications of the findings for theory, research, and practice. Finally, limitations of the study, and future recommendations are discussed.

6.2 Aims and Objectives

This study aimed to investigate the psychometric properties of the SACII-X, identify improvements to the SACII-X, and investigate differential item functioning across gender groups. To achieve these aims, the study had three objectives, namely:

- To determine the reliability and validity of the SACII-X items, through the analysis of the adherence of the data to the Rasch model assumptions and the fit of the data to the Rasch model.
- To investigate the item functioning of the SACII-X using Rasch analysis to identify problematic items that should be adapted or removed.
- To investigate differential item functioning in the SACII-X, across gender groups.

It was hypothesised that firstly, the data will fit the Rasch model, hence providing support for the reliability and validity of the SACII-X, secondly, that the Rasch model will identify improvements to items across all six scales of the SACII-X, and thirdly, that some items will show differential item functioning across gender groups

6.3 Objective 1

The first objective was to determine the reliability and the validity of the SACII-X through the analysis of the adherence of the data to the Rasch model assumptions and the fit

Stellenbosch University https://scholar.sun.ac.za

of the data to the Rasch model. Specifically, Rasch separation and reliability indices, global model fit, response category functioning, dimensionality and local independence, and item fit were examined. It was hypothesised that the Rasch model will provide support for the reliability and the validity of the SACII-X.

Overall, the data showed good fit to the Rasch model across all scales with mean standardised residual correlations equal to or close to 0, and satisfactory infit and outfit MnSq values. Morgan (2014) similarly reported overall good model fit of the final SACII-E to the Rasch model.

The separation- and reliability indices for all scales were satisfactory. Person and item separation indices were above 2 for all the SACII-X scales, indicating that the test can separate persons into a minimum of two levels of standing on the latent trait, and that test-takers were able to separate the SACII-X items into two levels of endorsability. Both person and item separation reliability indices were above .80 across all six SACII-X scales. This demonstrates stability in the estimation of person measures across equivalent sets of items, and stability in the estimation of item measures across different samples (Bond & Fox, 2015; Wright & Masters, 1982).

These reliability values are slightly lower than those reported for the final version of the SACII-E in Morgan (2014). Morgan (2014) reported Cronbach alpha and/or Guttman's Lambda 6 reliability coefficients of ≥.92 across all six scales and across gender. The Rasch item separation reliability coefficients in Morgan (2014) were ≥.90, and the Rasch person separation reliability coefficients were ≥.84 across the six scales. Similar to Morgan's (2014) study, the lowest item separation reliability coefficient across the six scales, was also observed on the Investigative scale in the present study. However, the SACII-E version used by Morgan in the reliability analysis included 169 items, which is more than the number of items in the SACII-X used in this study. This could have increased the reliability coefficient

Stellenbosch University https://scholar.sun.ac.za

obtained by Morgan (2014). The present study's reliability coefficients are further similar to the Cronbach's alpha and Guttman's Lambda 6 coefficients of ≥.88 across the six SACII-X scales for the 266 isiXhosa home language-speaking participants composing the sample in the current study, found by Rabie (2017).

The analysis of response category functioning showed relatively satisfactory functioning of the response categories across the six scales. Average person measures increased monotonically with each successive response category, indicating that persons with higher standings on the latent trait are more likely to endorse higher categories and vice versa. The infit and outfit MnSq values of all categories across all scales were also satisfactory. However, Rasch-Andrich threshold disordering was present across all six scales. Category 4 (*Agree)* had a consistent lower endorsement than Category 5 (*Strongly Agree)* across all six scales, resulting in a disordering of the Rasch-Andrich thresholds between Categories 3 and 4 (*Somewhat Disagree/Somewhat Agree* and *Agree*), and 4 and 5 (*Agree* and *Strongly Agree*). Despite this, Category 4 (*Agree*) was still endorsed by more than 10% of the sample, indicating that the category is a useful response option. Due to this, and the monotonic increase in average person measures with response categories, the rating scale was not collapsed and rescored.

Morgan (2014) also found Rasch-Andrich threshold disordering in the SACII-E during the development of the SACII-E. In Morgan's (2014) study, however, the *Neutral/Uncertain* category had a lower endorsement and never became a modal category for a portion of the latent trait, across any of the six SACII-E scales. Upon Morgan's collapsing and recoding of the categories into four response categories, the fit of the data to the Rasch model improved. Morgan (2014) suggested that a five-point Likert scale may not be the most optimal rating scale for the SACII-X, and that a four-point Likert scale might be a more optimal rating scale.

Stellenbosch University https://scholar.sun.ac.za

However, similarly to Morgan's (2014) findings, apart from the threshold disordering, the rating scale functioned well.

Dimensionality analysis found that the majority of the SACII-X scales demonstrated only slight deviations from unidimensionality. In addition, the vast majority of items of the SACII-X did not show local dependence between item pairs. The Artistic scale, however, did show a practically significant deviation from unidimensionality. Based on these findings, five of the six SACII-X scales can be assumed to be unidimensional for all practical purposes. The findings on the possible multidimensionality of the Artistic scale, calls for further research and refinement of the scale. These findings mostly echo those in Morgan (2014). Morgan's (2014) dimensionality analysis of the SACII-E scales with the Rasch model demonstrated that there were only slight deviations from unidimensionality across the six SACII-X scales. Morgan (2014) concluded that for practical purposes, the SACII-E scales can be assumed to be unidimensional.

The overwhelming majority of the SACII-X items showed good fit to the Rasch model. A total of 12 items showed misfit to the model, distributed as follows: (R) 2; (I) 1; (A) 1; (S) 3; (E) 3; and (C) 2. Specifically, all the items showed underfit to the model with infit and/or outfit MnSq values above 1.30. These results compare favourably to those reported by Morgan (2014) for the final SACII-E, and by Rabie (2017) for the SACII-X. On the final SACII-E a total of 68 items demonstrated misfit. The distribution of misfitting items across the scales were as follows: (R) 18; (I) 8; (A) 10; (S) 10; (E) 5; and (C) 11. However, in contrast to the present study where all the misfitting items demonstrated underfit, 28 items of the total number of misfitting items in Morgan (2014), demonstrated overfit. Rabie (2017) found a total of 34 poor functioning items on the SACII, using CTT methodology, and 31 poor functioning items through the investigation of angular locations and communality estimates. A more comprehensive discussion of item fit is presented in the subsequent section.

Stellenbosch University https://scholar.sun.ac.za

The above-mentioned results indicate that, overall, the data showed good fit to the Rasch model. The good fit provides support for the reliability and validity of the SACII-X. The first objective was thus met, and the first hypothesis was supported.

6.4 Objective 2

The second objective was to investigate the item functioning of the SACII-X using Rasch analysis to identify problematic items that should be adapted or removed. This was achieved through the analysis of point-measure correlations, item measures, unidimensionality, local independence, item fit, and uniform and non-uniform DIF. It was hypothesised that the Rasch measurement model will identify improvements to items in the Enterprising and Conventional subscales of the SACII-X.

Point-measure correlations of all items across all scales were positive, indicating that each item correlates positively to the latent trait being measured (Aryadoust et al., 2011). The range of items on each scale was relatively narrow. Morgan (2014) reported a wider range of item measures across all scales of the SACII-E, than the present study. However, there were multiple items around the person mean of each scale of the SACII-X. Hence, the SACII-X scales will be able to distinguish well between the latent trait levels of the majority of test-takers. The SACII-X will be less able to distinguish between persons with higher and lower standings on the latent trait.

As mentioned in the previous section, only the Artistic scale showed a significant deviation from unidimensionality, with three items on the scale contrasting strongly against the other scale items and demonstrating local dependence between item pairs. These items were item A1 (“Act in a play”), A2 (“Audition for a play”), and A6 (“Direct a play”). With the removal of items A2 and A6, Rasch analysis supported the unidimensionality of the scale and the local independence of the items. The removal of these two items, further led to negligible changes to the item and person separation reliability indices. In addition, these two

Stellenbosch University https://scholar.sun.ac.za

items are each one of 12 items at the item mean and person mean calculated by the Rasch analysis. Considering these results, the two items are redundant items not contributing useful information over and above the other items. Items A2 ("Audition for a play") and A6 ("Direct a play") are therefore strong candidates for removal from the SACII-X.

The Enterprising scale showed local dependence between two items, namely E11 ("Manage a business") and E24 ("Start a business"). The removal of either item E11 or E24 had a similar influence on the reliability- and separation indices of the Enterprising scale. However, the removal of item E11 resulted in a larger decrease in the item separation index than the removal of item E24. Furthermore, item E24 is one of five items located 1 S*D* from the item mean. These findings indicate that item E24 ("Start a business") may be redundant, and a candidate for removal from the SACII-X.

On the Conventional scale, items C27 ("Prepare financial records for a company") and C8 ("Control accounting records") were locally dependent. Both items were one of 11 items located 1 *SD* below the item mean. The removal of both of these items had the same effect on the reliability- and separation indices of the Conventional scale. Given this information, a conclusive decision cannot be made regarding which item out of the locally dependent pair is a better candidate for removal.

The majority of the SACII-X items showed good fit to the Rasch model and can be assumed to be functioning well. A total of 12 items showed misfit to the model - two items on the Realistic scale, one item on the Investigative scale, one item on the Artistic scale, three items on the Social scale, three items on the Enterprising scale, and two items on the Conventional scale. Specifically, all the items showed underfit to the model with infit and/or outfit MnSq values above 1.30.

On the Realistic Scale, items RR19 ("Operate heavy machinery") and RR20 ("Plan the layout of cables") did not function well. On the Investigative Scale, only item I26 ("Use new

Stellenbosch University https://scholar.sun.ac.za

ways of removing diseases from crops") showed misfit to the scale, and on the Artistic scale, item A7 ("Go to theatre productions") misfit. Three items misfit on the Social scale, namely item S22 ("Teach at a school"), S26 ("Teach people how to do public speaking"), and S12 ("Model clothes"). The Enterprising scale contained three misfitting items, namely item E17 ("Negotiate salaries on behalf of people"), E27 ("Teach people how to improve their managerial skills"), and E12 ("Manage a hotel"). Item C14 ("Maintain a database") misfit on the Conventional scale. Plausible reasons for the poor functioning of the items could be that the items are vague and/or ambiguous. As a result, the Grade 9 learners in this sample may not fully grasp the meaning of these items.

Some of these items were also previously indicated as poor functioning through IRT and CTT analyses by Morgan (2014), and CTT analysis, PCA, arc-tangent transformations, and angular location and communality investigations by Rabie (2017). Findings from Morgan (2014), Rabie (2017), and the present study were considered in the identification of items for removal from the SACII-X.

Rabie (2017) found item RR19 ("Operate heavy machinery") to be a poor functioning item in CTT analysis and PCA. Furthermore, item RR19 is one of three items located 1 *SD* below the item mean. In the final SACII-E questionnaire developed by Morgan (2014), item RR20 ("Plan the layout of cables") demonstrated overfit to the Rasch model. Item RR20 is further one of 12 items located at the item mean. This item location indicates that item RR20 is potentially redundant. Considering the poor functioning of the items highlighted using different methods and across different samples, and the item locations of the items, both items RR19 ("Operate heavy machinery") and RR20 ("Plan the layout of cables") are strong candidates for removal from the SACII-X.

Item I26 ("Use new ways of removing diseases from crops") was the only item identified as misfitting on the Investigative Scale. Rabie's (2017) CTT analysis, arc-tangent

Stellenbosch University https://scholar.sun.ac.za

transformation, and angular location and communality estimates investigation also highlighted this item as problematic in the SACII-X. Item I26 is one of three items located 1 *SD* from the item mean in the Rasch model, and one of three items at the person mean of the Rasch model. Within this context, item I26 ("Use new ways of removing diseases from crops") is a strong candidate for removal from the SACII-X.

Items S22 ("Teach at a school"), S26 ("Teach people how to do public speaking"), and S12 ("Model clothes") were identified as misfitting on the Social scale. Neither items S22 nor S26 has previously been indicated as misfitting items on the SACII-X. Item S12, however, has been found to underfit the Rasch model by Morgan (2014), and function poorly in CTT analysis and PCA by Rabie (2017). Item S12 is further one of eight items located at the item- and person mean of the Rasch model. Item S12 ("Model clothes") is therefore a prime candidate for removal from the SACII-X. Items S22 and S26 are both one of six items located 1 *SD* above the item- and person mean. Due to these items not previously being indicated as poor functioning, a conclusive decision cannot be made from the results of the present analysis to remove these items from the SACII-X.

Item E17 ("Negotiate salaries on behalf of people") was also indicated as overfitting in the SACII-E pilot questionnaire (Morgan, 2014). This item is furthermore one of eight items located at the item mean of the scale. The item is therefore a candidate for removal from the SACII-X. Items E27 ("Teach people how to improve their managerial skills") and E12 ("Manage a hotel") was indicated as problematic items in Rabie (2017). Both items are one of three items located 2 *SD* from the item mean. The removal of both items will therefore decrease the limited number of items able to distinguish between lower performers. But since these items have been indicated as poor functioning across multiple analysis methods and different samples, they are good candidates for adaptation, or removal from the SACII-X.

Stellenbosch University https://scholar.sun.ac.za

Rabie (2017) found item C14 (“Maintain a database”) to be problematic in CTT analysis. Item C14 is one of only two items located just above the item mean, and the removal of this item will thus reduce the limited number of items at this level of item endorsability. The item is a candidate for adaptation or removal, but the functioning of the item should be investigated in more samples before a conclusive decision is made.

The second objective was therefore met, and the second hypothesis was supported. The analysis of item functioning identified a relatively equivalent number of poor functioning items across all six SACII-X scales.

6.5 Objective 3

The third objective was to determine differential item functioning across gender across the SACII-X scales. It was hypothesised that indications of DIF will be present in some SACII-X items.

Across the SACII-X scales, one item on the Realistic scale demonstrated uniform DIF and 24 items across the SACII-X demonstrated non-uniform DIF. The distribution of the non-uniform DIF items across the six scales were as follows: (R) 2; (I) 4; (A) 3; (S) 4; (E) 4; and (C) 7. Item R10 (“Fix a car engine”) displayed slight uniform DIF on the Realistic scale, favouring males over females. Non-uniform DIF was found in items RR17 (“Monitor electricity use at a power plant”), RR19 (“Operate heavy machinery”), I9 (“Research laser technology”), I12 (“Research ways to improve space flight”), I21 (“Study plants in the environment”), I26 (“Use new ways of removing diseases from crops”), A7 (“Go to theatre productions”), A11 (“Play a musical instrument”), A23 (“Write a musical score”), A27 (“Write jingles for advertisements”), S11 (“Look after people who are sick”), S28 (“Work with elderly people”), S12 (“Model clothes”), S22 (“Teach at a school”), E17 (“Negotiate salaries on behalf of people”), E20 (“Sell goods at an auction”), E27 (“Teach people how to improve their managerial skills”), E3 (“Assist customers to buy products”), C16 (“Maintain

Stellenbosch University https://scholar.sun.ac.za

records of income received from clients"), C18 ("Monitor credit scores"), C25 ("Organise transportation of products from a factory"), C4 ("Calculate wages and salaries of employees"), C21 ("Monitor the accounts of a company"), C26 ("Prepare contracts for a company"), and C29 ("Take merchandise orders"). The total number of items displaying uniform DIF is larger than findings by Mintram et al. (2019), the only study investigating DIF across gender in the SACII items. Mintram et al. (2019) identified 18 items showing DIF, using a two-way ANOVA analysis.

It is generally accepted that sex-typing of careers influences career interest development during childhood and adolescence (Einarsdóttir & Rounds, 2009; Weisgram et al. 2010). According to Weisgram et al. (2010), subscales formed from Holland's six vocational personality types, include both careers typically conceived as masculine, and careers typically perceived as feminine. This could provide an explanation as to why males and females with the same standing on one of the vocational personality types in the SACII-X, differentially endorsed some of the items on the scales, depending on whether the item is viewed as a masculine activity or a feminine activity. However, not all items showing non-uniform DIF behaved in this expected way. It is possible that other unmodelled variables other than gender, contributed to the differential item functioning found in some items. Due to the small sample size in this study, the results of the DIF analysis should be interpreted with caution. Nevertheless, seven of the items on which non-uniform DIF was found, was also indicated as misfitting in the fit analysis.

The third objective was therefore met, and the third hypothesis was supported. The analysis of differential item functioning identified a small number of items across the SACII-X scales showing indications of differential item functioning.

Stellenbosch University https://scholar.sun.ac.za

6.6 Implications for Theory, Research, and Practice

This study has corroborated previous research supporting the SACII-X as a potentially reliable and valid measure of career interests that can fairly be used among isiXhosa-speaking secondary school learners. It has shown that assessment in the home language of test-takers is a fair and valid option for interest assessment in South Africa. The study has also shown that IRT models such as the Rasch model can be applied to career interest assessments to investigate the psychometric properties and item functioning of the assessments.

The study has further contributed to the information available on the SACII-X as an instrument under development. Problematic items that are candidates for potential removal or adaptation were identified. This can contribute to the future refinement of the SACII-X. The removal of these items could shorten the SACII-X without significant losses in the reliability of the scale. Shortening the SACII-X could reduce the burden of response on test-takers.

The poor functioning items found in this study can potentially be improved by revising the items to better represent the activities for an adolescent sample. For example, *Operate heavy machinery* (item RR19) on the Realistic Scale, can be revised to *Drive a truck or forklift*, to make the meaning of the item activity more clear for a secondary school sample.

With regards to the use of the SACII-X in a career guidance setting, even though research on the SACII-X has shown that the instrument is potentially reliable and valid among adolescent samples, two points are worth considering. Firstly, the functioning of the SACII-X has only been studied in one sample group. Consequently, the item functioning, and reliability and validity of the SACII-X cannot yet be extrapolated to other samples and other age groups. Secondly, clear areas for refinement of the instrument have been identified in this study, as well as in Rabie (2017). Before fully incorporating the SACII-X into the career counselling field, the poor functioning items and potentially gender biased items found in this study and Rabie (2017) should be further investigated, and their poor functioning addressed.

Stellenbosch University https://scholar.sun.ac.za

6.7 Limitations and Recommendations

Similar to Morgan's (2014) study, this study found that the rating scale of the SACII-X could be reduced to four response categories. Future research could focus on the rating scale functioning of the SACII-X and investigate whether collapsing the rating scale to four response options, is warranted.

No misfitting persons were removed prior to the Rasch analysis and the analysis of item fit. The removal of misfitting persons could possibly have resulted in better item fit. However, this analytical decision was made as the secondary data was previously cleaned by the primary investigator of the original study, removing persons with biased response patterns. In addition, visual inspection of the raw data did not reveal noticeable response sets. Hence, it could not unequivocally be concluded whether persons purposefully provided inaccurate responses, or whether inconsistent responding was due to true inherent qualities of the person, or a misunderstanding of the items due to, for example, language- and/or cultural factors.

Due to the small sample size of the study, results of the DIF analysis should be interpreted with caution. Parameter estimation in smaller samples is more prone to error than in larger samples (Davis et al., 2014). As a result, it may be difficult to distinguish true group differences from differences obtained due to chance. Future studies can investigate the reasons for the potential differential item functioning of the items indicated in this study.

Furthermore, the functioning of the SACII-X has only been investigated in an adolescent sample with the majority of test-takers aged 14 to 15 years old. Even though this is a strength in the sense that it focuses specifically on the item functioning and test reliability in the intended population, it may lead to sample-dependent results that do not generalise outside of the current sample. Future studies could focus on the functioning of the SACII-X in different age groups.

Stellenbosch University https://scholar.sun.ac.za

This study has highlighted some items as problematic, that have not previously been shown to be problematic. The functioning of these items should be investigated among more samples to determine whether they truly function poorly. To ameliorate the poor functioning of these items, the SACII-X can also be administered with a glossary simplifying the technical language used in the test, to ensure participants' understanding of the test items and to make the items more appropriate for an adolescent sample group. Table 5.12 provides examples of how such glossary entries may look.

Table 5.12

Glossary Entries for the SACII-X Items

Item number	isiXhosa item	English item	Explanation
45	Ukuya kwimidlalo yeqonga	Go to theatre productions	Watch a live performance of a musical or a play
16	Ukusebenzisa imitshini emikhulu	Operate heavy machinery	Drive heavy machinery such as trucks or forklifts

Future studies could also compare item responses of the same sample groups across the SACII-E and SACII-X. The item functioning of the two versions of the SACII can then be compared to investigate the equivalence of the two test versions.

6.8 Conclusion

Language is a critical moderator of assessment performance, and language issues in psychometric assessment is a threat to fair and ethical testing practices. This study set out to investigate the reliability and validity of the SACII-X as a career interest measure in an indigenous South-African language. It further aimed to investigate the item functioning of the SACII-X using the Rasch measurement model to inform the refinement of the SACII-X. The findings have supported the SACII-X as a test than can potentially be reliably and validly

Stellenbosch University https://scholar.sun.ac.za

applied to isiXhosa home language-speaking secondary school learners. The Rasch measurement model has further identified a number of items that functions poorly in the SACII-X. Previous research findings on the SACII-E and SACII-X, together with the findings of the present study, were used to inform recommendations regarding the adaptation or removal of the identified items in the SACII-X. The findings of this study should be used to further refine the SACII-X for future application in diverse settings.

Stellenbosch University https://scholar.sun.ac.za

References

Abrahams, F., & Mauer, K. F. (1999a). The comparability of the constructs of the 16PF in the South African context. *Journal of Industrial Psychology, 25*(1), 53-59. https://sajip.co.za/index.php/sajip/article/view/679

Abrahams, F., & Mauer, K. F. (1999b). Qualitative and statistical impacts of home language on responses to the items of the Sixteen Personality Factor questionnaire (16PF) in South Africa. *South African Journal of Psychology, 29*(2), 76-86. https://doi.org/10.1177/008124639902900204

Aiken, L. R., & Groth-Marnat, G. (2006). Psychological testing and assessment (12th ed.). *Annual Review of Clinical Psychology, 7*, 383-409.

Albien, A. (2013). *Grade 12 Kayamandi adolescents' career influences using the systems theory framework of career development* (Master's thesis). Sunscholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Albien, A. J. (2018). A mixed-methods analysis of black adolescents' vocational identity status and career adaptability competencies in a South African township (Doctoral dissertation). Sunscholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Albien, A. J., & Naidoo, A. V. (2016). Social career influences of Xhosa adolescents elicited using the Systems Theory Framework in a peri-urban South African township. *South African Journal of Higher Education, 30*(3), 111-137. https://doi.org/10.20853/30-3-668

Allen, R. (2000a). Vocation. In R. Allen (Ed.), *The New Penguin English Dictionary.* The Penguin Press, Penguin Books Ltd.

Allen, R. (2000b). Career. In R. Allen (Ed.), *The New Penguin English Dictionary.* The Penguin Press, Penguin Books Ltd.

Stellenbosch University https://scholar.sun.ac.za

Allen, R. (2000c). Stable. In R. Allen (Ed.), *The New Penguin English Dictionary.* The Penguin Press, Penguin Books Ltd.

American Psychological Association. (2017). *Ethical Principles of Psychologists and Code of Conduct.* https://www.apa.org/ethics/code

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573. https://doi.org/ 10.1007/BF02293814

Armstrong, P. I., Hubert, L., & Rounds, J. (2003). Circular unidimensional scaling: A new look at group differences in interest structure. *Journal of Counselling Psychology, 50*(3), 297-308. https://doi.org/10.1037/0022-0167.50.3.297

Arnold, J. (2004). The congruence problem in John Holland's theory of vocational decisions. *Journal of Occupational and Organisational Psychology, 77*(1), 95-113. https://doi.org/10.1348/096317904322915937

Aryadoust, V., Mehran, P., & Alizadeh, M. (2016). Validating a computer-assisted language learning attitude instrument used in Iranian EFL context: An evidence-based approach. *Computer Assisted Language Learning, 29*(3), 561-595. https://doi.org/10.1080/09588221.2014.1000931

Aryadoust, V., Ng, L. Y., & Sayama, H. (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing, 38*(1), 6-40. https://doi.org/10.1177/0265532220927487

Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-analysis of the relationship between the Five-Factor Model of personality and Holland's occupational types. *Personnel Psychology, 56*(1), 45-74. https://doi.org/10.1111/j.1744-6570.2003.tb00143.x

Bedell, B., Van Eeden, R., & Van Staden, E. (1999). Culture as moderator variable in psychological test performance. *South African Journal of Industrial Psychology, 25*(3). https://doi.org/10.4102/sajip.v25i3.681

Stellenbosch University https://scholar.sun.ac.za

Berg, A., Buys, M., Schaap, P., & Olckers, C. (2004). The comparability of the construct validity of Schepers' Locus of Control Inventory for first and second language respondents. *South African Journal of Industrial Psychology, 30*(3), 87-96. https://doi.org/10.4102/sajip.v30i3.165

Bless, C., Higson-Smith, C., & Sithole, S. L. (2013). The types of research. In C. Bless, C. Higson-Smith, & S .L. Sithole (Eds.), *Fundamentals of Social Research Methods - An African Perspective* (pp. 56-70). Juta and Company Ltd.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental Measurement in the Human Sciences* (3rd ed.). Lawrence Erlbaum Associates.

Boone, W. J. (2016). Rasch analysis for instrument development: Why, when, and how? *CBE Life Sciences Education, 15*(4). https://doi.org/10.4102/sajip.v30i3.165

Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education, 4*(1). https://doi.org/10.1080/2331186X.2017.1416898

Boone, W., & Rogan, J. (2005). Rigour in quantitative analysis: The promise of Rasch analysis techniques. *African Journal of Research in Mathematics, Science and Technology Education, 9*(1), 25-38. https://doi.org/ 10.1080/10288457.2005.10740574

Bryman, A. (2016). *Social Research Methods.* Oxford University Press.

Byrne, B. M., Oakland, T., Leong, F. T. L., Van de Vijver, F. J. R., Hambleton, R. K., Cheung, F. M., & Bartram, D. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training in Education in Professional Psychology, 3*(2), 94-105. https://psycnet.apa.org/doi/10.1037/a0014516

Stellenbosch University https://scholar.sun.ac.za

Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review, 14*(3), 333–349. https://doi.org/10.5465/amr.1989.4279063

Chatman, J. A. (1991). Matching people and organizations: Selection and socialization in public accounting firms. *Administrative Science Quarterly, 36*, 459–484. https://doi.org/10.5465/ambpp.1989.4980837

Chen, D. P., & Simpson, P. A. (2015). Does personality matter? Applying Holland's typology to analyze students' self-selection into science, technology, engineering, and mathematics majors. *Journal of Higher Education, 86*(5), 725-750. https://doi.org/10.4102/sajip.v30i3.165

Chou, Y. T., & Wang, W. C. (2010). Checking dimensionality in item response models with principal component analysis on standardised residuals. *Educational and Psychological Measurement, 70*(5), 717-731. https://doi.org/10.1177/0013164410379322

Coetzer, W. J., & Rothmann, S. (2007). A psychometric evaluation of measures of affective well-being in an insurance company. *South African Journal of Industrial Psychology, 33*(2), 7-15. https://sajip.co.za/index.php/sajip/article/download/372/365

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-344. https://doi.org/10.1007/BF02310555

Darcy, M. U. A. (2005). Examination of the structure of Irish students' vocational interests and competence perceptions. *Journal of Vocational Behaviour, 67*(2), 321-333. https://doi.org/10.1016/j.jvb.2004.08.007

Darcy, M. U. A., & Tracey, T. J. G. (2007). Circumplex structure of Holland's RIASEC interests across gender and time. *Journal of Counseling Psychology, 54*(1), 17-31. https://doi.org/ 10.1037/0022-0167.54.1.17

Stellenbosch University https://scholar.sun.ac.za

Davis, G., Pecar, B., Santana, L., & Burke, A. (2014). Introduction to Statistics. In G. Davis, B. Pecar, L. Santana, and A. Burke (Eds.), *Statistics for the Social Sciences* (pp. 1-10). Oxford University Press Southern Africa (Pty) Ltd.

Dawis, R. V. (2005). The Minnesota Theory of Work Adjustment. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: Putting theory and research to work* (pp. 3–23). John Wiley & Sons, Inc.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment: An individual-differences model and its applications*. University of Minnesota press.

De Bruin, G. P., & Buchner, M. (2010). Factor and item response theory analysis of the Protean and Boundaryless Career Attitude scales. *South African Journal of Industrial Psychology, 36*(2). https://doi.org/10.4102/sajip v36i2.932

De Bruin, G. P., & De Bruin, K. (2017). Career assessment. In G. B. Stead & M. B. Watson (Eds.), *Career Psychology in the South African Context* (3rd ed., pp. 185-193). Van Schaik Publishers.

De Bruin, G. P., & De Bruin, K. (2018). Career counselling assessment. In C. Foxcroft & G. Roodt (Eds.), I*ntroduction to Psychological Assessment in the South African Context* (5th ed., pp. 243-254). Oxford University Press Southern Africa (Pty) Ltd.

De Bruin, G. P., & Taylor, N. (2005). Development of the Sources of Work Stress Inventory. *South African Journal of Psychology, 35*(4), 748-765. https://core.ac.uk/download/pdf/43590458.pdf

De Bruin, G. P., Hill, C., Henn, C. M., & Muller, K. (2013). Dimensionality of the UWES-17: An item response modelling analysis. *South African Journal of Industrial Psychology, 39*(2). https://doi.org/10.4102/sajip.v39i2.1148

Stellenbosch University https://scholar.sun.ac.za

De Bruin, K., & De Bruin, G. P. (2011). Development of the Learner Self-Directedness in the Workplace scale. *South African Journal of Industrial Psychology, 37*(1). https://doi.org/ 10.4102/sajip.v37i1.926

De Kock, F., & Foxcroft, C. (2018). Cross-cultural test adaptation, translation, and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to Psychological Assessment in the South African Context* (pp. 100-122). Oxford University Press.

De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to Psychological Assessment in the South African context*, (4th ed., pp. 83-106). Oxford University Press Southern Africa (Pty) Ltd.

Desai, Z. (2016). Learning through the medium of English in multilingual South Africa: Enabling or disabling learners from low income contexts. *Comparative Education, 52*(3), 343-358. https://doi.org/10.4102/sajip.v25i3.681

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology,* 7, 383-409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Du Plessis, G. A., & De Bruin, G. P. (2015). Using Rasch modelling to examine the International Personality Item Pool (IPIP) Values in Action (VIA) measure of character strengths. *Journal of Psychology in Africa, 25*(6), 512-521. https://doi.org/ 10.4102/sajip.v39i2.1148

Du Toit, R., & De Bruin, G. P. (2002). The structural validity of Holland's R-I-A-S-E-C model of vocational personality types for young Black South African men and women. *Journal of Career Assessment, 10*(1), 62-77. https://doi.org/10.1177/1069072702010001004

Stellenbosch University https://scholar.sun.ac.za

Du Toit, R., Prinsloo, W. B. J., Gevers, J., & Harilall, R. (1993). *Manual for the South African Vocational Interest Inventory (SAVII).* Human Sciences Research Council. Pretoria.

Dube-Addae, T. (2019). Narrative identities: The voices of the unemployed youth in a low-income community in South Africa (Unpublished Doctoral dissertation). Sunscholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Edwards, M. C., Houts, C. R., & Cai, L. (2018). A diagnostic procedure to detect departures from local independence in item response theory models. *Psychological Methods, 23*(1), 138-149. https://doi.apa.org/doi/10.1037/met0000121

Einarsdóttir, S., Rounds, J., & Su, R. (2010). Holland in Iceland revisited: An emic approach to evaluating U.S. vocational interest models. *Journal of Counseling Psychology, 57*(3), 361-367. https://doi.org/10.1037/a0019685

Einarsdóttir, S., Rounds, J., Ægisdottir, S., & Gerstein, L. H. (2002). The structure of vocational interests in Iceland: Examining Holland's and Gati's RIASEC models. *European Journal of Psychological Assessment, 18*(1), 85-95. https://doi.org/10.1027//1015-5759.18.1.85

Elosua, P. (2007). Assessing vocational interests in the Basque country using paired comparison design. *Journal of Vocational Behaviour, 71*, 135-147. https://doi.org/10.1016/j.jvb.2007.04.001

Finchilescu, G., & Tredoux, C. (2010). The changing landscape of intergroup relations in South Africa. *Journal of Social Issues, 66*(2), 223-236. http://www.psychology.uct.ac.za/sites/default/files/image_tool/images/117/CGTredouxdocs/Finchilescu%20&%20Tredoux,%202010,%20JSI.pdf

Fouad, N. A. (2007). Work and vocational psychology: Theory, research, and applications. *Annual Review of Psychology, 58*, 543-564. https://doi.org/10.1146/annurev.psych.58.110405.085713

Stellenbosch University https://scholar.sun.ac.za

Foxcroft, C. D. (2004). Planning a psychological test in the multicultural South-African context. *South African Journal of Industrial Psychology, 30*(4), 8-15. https://journals.co.za/docserver/fulltext/psyc/30/4/psyc_v30_n4_a2.pdf?expires=1558039049&id=id&accname=57926&checksum=4DBA7B1D5E19515052F65D536E4CD5CD

Foxcroft, C. D. (2011). Ethical issues related to psychological testing in Africa: What I have learned (so far). *Online Readings in Psychology and Culture, 2*(7), 1-19. https://doi.org/10.9707/2307-0919.1022

Foxcroft, C. D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAIS-III. *South African Journal of Industrial Psychology, 32*(4), 97-102. https://journals.co.za/content/psyc/32/4/EJC89105

Foxcroft, C. D., & Grieve, K. W. (2018). Interpreting and reporting assessment results. In C. Foxcroft & G. Roodt (Eds.), *Introduction to Psychological Assessment in the South African Context* (5th ed., 301-312). Oxford University Press Southern Africa (Pty) Ltd.

Foxcroft, C., & Roodt, G. (2013). An overview of assessment: Definition and scope. In C. Foxcroft & G. Roodt (Eds.), *Introduction to Psychological Assessment in the South African context* (4th ed., pp. 3-27). Oxford University Press Southern Africa (Pty) Ltd.

Foxcroft, C., Paterson, H., Le Roux, N., & Herbst, D. (2004). *Psychological assessment in South Africa: A needs analysis – The test use patterns and needs of psychological assessment practitioners* (Final report, July). http://hdl.handle.net/20.500.11910/7498

Foxcroft, C., Roodt, G., & Abrahams, F. (2018). Psychological assessment: A brief retrospective overview. In C. Foxcroft & G. Roodt (Eds.), *Introduction to Psychological Assessment in the South African context* (4th ed., pp. 12-36). Oxford University Press Southern Africa (Pty) Ltd.

Stellenbosch University https://scholar.sun.ac.za

French B. F. (2014) Test Bias. In A.C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research*. Springer. https://doi.org/10.1007/978-94-007-0753-5_2998

Furnham, A. (2001). Vocational preference and P-O fit: Reflections on Holland's theory of vocational choice. *Applied Psychology, 50*(1), 5-29. https://doi.org/10.1111/1464-0597.00046

Gevers, J., du Toit, R., & Harilall, R. (1997). SDS manual. Pretoria, South Africa: Human Sciences Research Council.

Glosenberg, A., Tracey, T. J. G., Behrend, T. S., Blustein, D. L., & Foster, L. L. (2019). Person-vocation fit across the world of work: Evaluating the generalizability of the circular model of vocational interests and Social Cognitive Career theory across 74 countries. *Journal of Vocational Behaviour, 112*, 92-108. https://doi.org/10.1016/j.jvb.2019.01.002

Gottfredson, G. D., & Johnstun, M. L. (2009). John Holland's contributions: A theory-ridden approach to career assistance. *Career Development Quarterly, 58*(2), 99-107. https://doi.org/10.1002/j.2161-0045.2009.tb00050.x

Gregory, R. J. (2007). *Psychological Testing: History, Principles, and Applications* (5th ed.). Pearson Education, Inc.

Grobler, S., & De Beer, M. (2015). Psychometric evaluation of the Basic Traits Inventory in the multilingual South African environment. *Journal of Psychology in Africa, 25*(1), 50-55. https://doi.org/10.1080/14330237.2014.997033

Gupta, S., Tracey, T. J. G., & Gore, P. A. (2008). Structural examination of RIASEC scales in high school students: Variation across ethnicity and method. *Journal of Vocational Behaviour, 72*, 1-13. https://doi.org/10.1016/j.jvb.2007.10.013

Stellenbosch University https://scholar.sun.ac.za

Harrington, T., & Long, J. (2013). The history of interest inventories and career assessments in career counselling. *The Career Development Quarterly, 61*, 83-92. https://doi.org/10.1002/j.2161-0045.2013.00039.x

Health Professions Act 1974 (Act No. 56 of 1974).

Health Professions Council of South Africa. (2017). Health Professions Act 1974 (Act No. 56 of 1974) List of Classified and Certified Psychological Tests. *Government Gazette, 41100*, 279. https://www.hpcsa.co.za/Uploads/PSB_2019/List_of_Classified_tests_Board_Notice_155_of_2017.pdfhttps://www.hpcsa.co.za/Uploads/PSB_2019/List_of_Classified_tests_Board_Notice_155_of_2017.pdf

Hedrih, V. (2008). Structure of vocational interests in Serbia: Evaluation of the Spherical model. *Journal of Vocational Behaviour, 73*, 13-23. https://doi.org/10.1016/j.jvb.2007.12.004

Hedrih, V., Šverko, I., & Pedovic, I. (2018). Structure of vocational interests in Macedonia and Croatia – Evaluation of the Spherical model. *FACTA UNIVERSITATIS: Philosophy, Sociology, Psychology and History, 17*(1), 19-36. https://doi.org/10.22190/fupsph1801019h

Hirschi, A., & Läge, D. (2007). Holland's secondary constructs of vocational interests and career choice readiness of secondary students. *Journal of Individual Differences, 28*(4), 205-218. https://doi.org/10.1027/1614-0001.28.4.205

Holland, J. L. (1985). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments* (2nd ed.). Prentice-Hall.

Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments* (3rd ed.). Psychological Assessment Resources.

Stellenbosch University https://scholar.sun.ac.za

International Test Commission. (2001). International guidelines for test use. *International Journal of Testing, 1*(2), 93-114. https://www.intestcom.org/

International Test Commission. (2005). *International Guidelines on Computer-Based and Internet Delivered Testing.* https://www.intestcom.org/

Jafari, P., Bagheri, Z., Ayatollahi, S. M. T., & Soltani, Z. (2012). Using Rasch Rating Scale Model to reassess the psychometric properties of the Persian Version of the pedsQl TM4.0 Generic Core scales in school. *Health and Quality of Life Outcomes, 10*(27). https://doi.org/10.1186/1477-7525-10-27

Joseph, L., & Van Lill, B. (2008). Investigating subscale differences among race and language groups on the Occupational Personality Profile. *South African Journal of Psychology, 38*(3), 501-514. https://doi.org/10.1177/008124630803800305

Kantamneni, N., & Fouad, N. (2011). Structure of vocational interests for diverse groups on the 2005 Strong Interest Inventory. *Journal of Vocational Behaviour, 78*(2), 193-201. https://doi.org/10.1016/j.jvb.2010.06.003

Kean, J., Bisson, E. F., Brodke, D. S., Biber, J., & Gross, P. H. (2017). An introduction to item response theory and Rasch analysis: Application using the Eating Assessment Tool (EAT-10). *Brain Impairment, 19*(1), 91-102. https://doi.org/10.1017/BrImp.2017.31

Koch, E. (2007). Die evaluering van 'n eentalige toelatinsgtoets wat vir toelating tot hoër onderwys in 'n veeltalige konteks gebruik word. *South African Journal of Industrial Psychology, 33*(1), 90-101. https://doi.org/10.4102/sajip.v33i1.255

Laher, S., & Cockcroft, K. (2013). Contextualising psychological assessment in South Africa. In S. Laher & K. Cockcroft (Eds.), *Psychological Assessment in South Africa: Research and Applications* (pp. 1-14). Wits University Press.

Stellenbosch University https://scholar.sun.ac.za

https://library.oapen.org/bitstream/handle/20.500.12657/25725/9781776143580_OpenAccessPDF.pdf

Laher, S., & Cockroft, K. (2014). Psychological assessment in post-Apartheid South Africa: The way forward. *South African Journal of Psychology, 44*(3), 303-314. https://doi.org/10.1177/0081246314533634

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six interests and Big Five personality factors. *Journal of Vocational Behaviour, 61*, 217-239. https://doi.org/10.1006/jvbe.2001.1854

Law, K., Wong, C., & Leong, F. (2001). The cultural validity of Holland's model and its implications for human resource management: The case of Hong Kong. *International Journal of Human Resource Management, 12*(3), 484-496. https://doi.org/10.1080/09585190010026257

Leung, S. A., & Hou, Z. (2005). The structure of vocational interests among Chinese students. *Journal Of Career Development, 32*(1), 74-90. https://doi.org/10.1177/0894845305277036

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions, 7*(4), 328. https://www.rasch.org/rmt/rmt74m.htm

Linacre, J. M. (2002a). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*(2), 878. https://www.rasch.org/rmt/rmt162f.htm

Linacre, J. M. (2002b). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.424.2811&rep=rep1&type=pdf

Stellenbosch University https://scholar.sun.ac.za

Linacre, J. M. (2009). Local independence and residual covariance: A study of Olympic figure skating ratings. *Journal of Applied Measurement, 10*(2), 157–169. https://www.winsteps.com/a/Linacre-figure-skating.pdf

Linacre, J. M. (2021a). *Winsteps* (Version 5.0.0) [Computer software]. Winsteps. https://www.winsteps.com/index.htm

Linacre, J. M. (2021b). *A User's Guide to Winsteps Ministep Rasch-Model Computer Program.* https://www.winsteps.com/manuals.htm

Long, L., Adams, R. S., & Tracey, T. J. G. (2005). Generalizability of interest structure to China: Application of the Personal Globe Inventory. *Journal of Vocational Behaviour, 66*(1), 66-80. https://doi.org/10.1016/j.jvb.2003.12.004

Long, L., Watanabe, N., & Tracey, T. J. G. (2006). Structure of interests in Japan: Application of the Personal Globe Inventory occupational scales. *Measurement and Evaluation in Counseling and Development, 38*(4), 222-235. https://doi.org/10.1080/07481756.2006.11909784

Low, K. S. D., & Rounds, J. (2007). Interest change and continuity from early adolescence to middle adulthood. *International Journal for Educational and Vocational Guidance, 7*(1), 23-36. https://doi.org/10.1007/s10775-006-9110-4

Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin, 131*(5), 713-737. https://doi.org/ 10.1037/0033-2909.131.5.713

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–174.

Stellenbosch University https://scholar.sun.ac.za

McDonald, E., & Van Eeden, R. (2014). The impact of home language on the understanding of the vocabulary used in the South African version of the Sixteen Personality Factor Questionnaire Fifth Edition. *South African Journal of Psychology, 44*(2), 228-242. https://doi.org/10.1177/0081246314522366

McIlveen, P. (2007). A test for theoretical integration: Systems Theory Framework and dialogical self. *Australian Journal of Career Development*, *16*(3), 31-37. http://eprints.usq.edu.au/3398/1/McIlveen_2007_AJCD_STF_Authorversion.pdf

Meiring, D., Van de Vijver, A. J. R., Rothmann, S., & Barrick, M. R. (2005). Construct, item, and method bias of cognitive and personality tests in South Africa. *South African Journal of Industrial Psychology, 31*(1), 1-8. https://doi.org/10.4102/sajip.v31i1.182

Meiring, D., Van de Vijver, F. J. R., & Rothmann, S. (2006). Bias in an adapted version of the 15FQ+ in South Africa. *South African Journal of Psychology, 36*(2), 340-356. https://doi.org/10.1177%2F008124630603600208

Mintram, K., Morgan, B., & De Bruin, G. P. (2019). An investigation of gender differences in Holland's circumplex model of vocational personality types in South Africa. *International Journal for Educational and Vocational Guidance, 20*, 293-310. https://link-springer-com.ez.sun.ac.za/article/10.1007/s10775-019-09404-4

Morgan, B. (2014). *Toward the Development and Evaluation of the South African Career Interest Inventory* (Unpublished Doctorate dissertation). University of Johannesburg, South Africa.

Morgan, B., & De Bruin, G. P. (2018). Structural validity of Holland's circumplex model of vocational personality types in Africa. *Journal of Career Assessment, 26*(2), 275-290. https://doi.org/10.1177/1069072717692747

Stellenbosch University https://scholar.sun.ac.za

Morgan, B., & De Bruin, G. P. (2019). Development and evaluation of a short RIASEC interest inventory. *International Journal for Educational and Vocational Guidance, 19*(3), 437-454. https://doi.org/10.1007/s10775-019-09387-2

Morgan, B., De Bruin, G. P., & De Bruin, K. (2015a). Constructing Holland's hexagon in South Africa: Development and initial validation of the South African Career Interest Inventory. *Journal of Career Assessment, 23*(3), 493-511. https://doi.org/10.1177/1069072714547615

Morgan, B., De Bruin, G. P., & De Bruin, K. (2015b). Gender differences in Holland's circular/circumplex interest structure as measured by the South African Career Interest Inventory. *South African Journal of Psychology, 45*(3), 349-360. https://doi.org/10.1177/0081246315572514

Nagy, G., Trautwein, U., & Lüdtke, O. (2010). The structure of vocational interests in Germany: Different methodologies, different conclusions. *Journal of Vocational Behaviour, 76*(2), 153-169. https://doi.org/10.1016/j.jvb.2007.07.002

Naidoo, A. V., Pretorius, T. B., & Nicholas, L. (2017). The emergence of career psychology in South Africa: A socio-historical perspective. In G. B. Stead & M. B. Watson (Eds.), *Career Psychology in the South African Context* (3rd ed., pp.1-20). Van Schaik Publishers.

Nauta, M. M. (2010). The development, evolution, and status of Holland's theory of vocational personalities: Reflections and future directions for counseling psychology. *Journal of Counseling Psychology, 57*(1), 11-22. https://doi.org/10.1037/a0018213

Nauta, M. M. (2012). Are RIASEC interests traits? Evidence based on self-other agreement. *Journal of Career Assessment, 20*(4), 426-439. https://doi.org/10.1177/1069072712448895

Stellenbosch University https://scholar.sun.ac.za

Nauta, M. M., & Khan, J. H. (2007). Identity status, consistency, and differentiation of interests, and career decision self-efficacy. *Journal of Career Assessment, 15*(1), 55-65. https://doi.org/10.1177/1069072705283786

Nell, V. (2000). C*ross-cultural Neuropsychological Assessment: Theory and Practice.* Lawrence Erlbaum Associates, Inc.

O*NET. (2021). *Browse by O*NET Data.* https://www.onetonline.org/find/descriptor/browse/Interests/

Panayides, P., Robinson, C., & Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal, 36*(5), 611-626. https://doi.org/10.1080/01411920903018182

Posel, D., & Casale, D. (2011). Language proficiency and language policy in South Africa: Findings from new data. *International Journal of Educational Development, 31*(5), 449-457. https://doi.org/10.1016/j.ijedudev.2010.09.003

Program Committee of the Institute for Objective Measurement. (2000). *Definition of Objective Measurement.* https://www.rasch.org/define.htm

Psychological Society of South Africa. (2007). *South African Professional Conduct Guidelines in Psychology 2007*. https://www.psyssa.com/wp-content/uploads/2016/12/SOUTH-AFRICAN-PROFESSIONAL-CONDUCT-GUIDELINES-IN-PSYCHOLOGY-2007-PsySSA_updated_01-12-2016pdf.pdf

Rabie, S. (2017). *A translation and psychometric investigation of the South African Career Interest Inventory across gender and race among secondary school learners* (Doctoral Dissertation). Stellenbosch University. Sunscholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Stellenbosch University https://scholar.sun.ac.za

Rabie, S., & Naidoo, A. V. (2019). Validating the adaptation of the first career measure in isiXhosa: The South African Career Interest Inventory-isiXhosa version. *South African Journal of Psychology, 49*(1), 109-121. https://doi.org/10.1177/0081246318772419

Rabie, S., Visser, M., Naidoo, A.V., van den Berg, F., & Morgan, B. (2021). Beyond the individual: A group-based career development intervention implemented in resource-constrained schools in South Africa. *Journal for Specialists in Group Work, 46*(1), 48-61. DOI: 10.1080/01933922.2020.1856252

Rasch, G. (1960). *Studies in Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment tests*. Nielsen & Lydiche.

Rasch, G. (1980). *Probabilistic Models for Some Intelligence and Attainment Tests.* The University of Chicago Press.

Republic of South Africa. (1998, October 19). Employment Equity Act, 1998. *Government Gazette, 400* (19370). https://www.uwc.ac.za/SO/HR/Documents/EE%20Act%2055%20of%201998.pdf

Retief, L., Potgieter, M., & Lutz, M. (2013). The usefulness of the Rasch model for the refinement of Likert scale questionnaires. *African Journal of Research in Mathematics, Science, and Technology Education, 17*(1-2). https://doi.org/10.1080/10288457.2013.828407

Rock, D. (2011). The location shall be called Kaya Mandi: A history of Kaya Mandi (Master’s thesis). SunScholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Stellenbosch University https://scholar.sun.ac.za

Rounds, J. B., & Tracey, T. J. (1990). From Trait-and Factor to Person-Environment Fit counselling: Theory and process. In W. B. Walsh & S. H. Osipow (Eds.), *Career Counselling: Contemporary Topics in Vocational Psychology* (pp.1-44). Lawrence Erlbaum.

Rounds, J., & Tracey, T. T. (1996). Cross-cultural structural equivalence of RIASEC models and measures. *Journal of Counseling Psychology, 43*(3), 310-329. https://doi.org/10.1037/0022-0167.43.3.310

Schaap, P., Buys, M. A., & Olckers, C. (2003). The construct validity of Schepers' Locus of Control Inventory for Black and White tertiary students. *South African Journal of Industrial Psychology, 29*(1), 32-43. https://doi.org/10.4102/sajip.v29i1.90

Schermer, J. A. (2016). Vocational interests. In V. Zeigler-Hill & T. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences.* https://link.springer.com/referenceworkentry/10.1007%2F978-3-319-28099-8_801-1

Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*(3), 437-453. https://doi.org/10.1111/j.1744-6570.1987.tb00609.x

Schoeman, A., De Beer, M., & Visser, D. (2008). The relationship between learning potential, English language proficiency, and work-related training test results. *South African Journal of Labour Relations, 32*(1), 47-62. http://uir.unisa.ac.za/bitstream/handle/10500/5318/labour_v32_n1_a4%5B1%5D.pdf?sequence=1&isAllowed=y

Sheldon, K. M., Holliday, G., Titova, L., & Benson, C. (2020). Comparing Holland and self-determination theory measures of career preference as predictors of career choice. *Journal of Career Assessment, 28*(1), 28-42. https://doi.org/10.1177/1069072718823003

Stellenbosch University https://scholar.sun.ac.za

Society for Industrial and Organisational Psychology. (2005). *Guidelines for the Validation and Use of Assessment Procedures for the Workplace*.

Society for Industrial and Organisational Psychology. (2006). *Code of Practice for Psychological and Other Similar Assessment in the Workplace.*

Spokane, A. R., Moya, M., & Faris, E. (2017). Vocational personalities in work environments: Perspectives on Holland's theory. In G. B. Stead & M.B. Watson (Eds.), *Career Psychology in the South African context* (3rd ed., pp. 49-66). Van Schaik Publishers.

Statistics South Africa. (2011a). *Census – 2011.* (StatsSA Publication No. P0301.4). https://www.statssa.gov.za/publications/P03014/P030142011.pdf

Statistics South Africa. (2011b). *Khayamandi.* http://www.statssa.gov.za/?page_id=4286&id=106

Statistics South Africa. (2018). *General Household Survey – 2018* (StatsSA Publication No. P0318). http://www.statssa.gov.za/publications/P0318/P03182018.pdf

Stead, G. B., & Subich, L. M. (2017). Career counselling practice. In G. B. Stead & M. B. Watson (Eds.), *Career Psychology in the South African context* (3rd ed., pp. 119-135). Van Schaik Publishers.

Stellenbosch Municipality. (2008). *Local Economic Development Strategy* (Main Report). https://stellenbosch.gov.za/documents/all-categories/departments/office-of-the-municipal-manager/93-local-economic-development-strategy/file

Stoll, G., Rieger, S., Nagengast, B., Trautwein, U., & Rounds, J. (2021). Stability and change in vocational interests after graduation from high school: A six-wave longitudinal study. *Journal of Personality and Social Psychology: Personality Processes and Individual Differences, 120*(4), 1091-1116. https://doi.apa.org/doi/10.1037/pspp0000359

Stellenbosch University https://scholar.sun.ac.za

Su, R., Murdock, C., & Rounds, J. (2015). Person-Environment Fit. In P. J. Hartung, M. L. Savickas, & W. B. Walsh (Eds.), *APA Handbook of Career Interventions* (pp. 81-98). American Psychological Association.

Sung, Y., Cheng, Y., & Wu, J. (2016). Constructing a situation-based career interest assessment for junior high school students and examining their interest structure. *Journal of Career Assessment, 24*(2), 347-365. https://doi.org/10.1177/1069072715580419

Šverko, I. (2008). Spherical model of interests in Croatia. *Journal of Vocational Behaviour, 72*(1), 14-42. https://doi.org/ 10.1016/j.jvb.2007.10.001

Šverko, I., & Babarović, T. (2006). The validity of Holland's theory in Croatia. *Journal of Career Assessment, 14*(4), 490-507. https://doi.org/10.1177/1069072706288940

Swanepoel, I., & Krüger, C. (2011). Revisiting validity in cross-cultural psychometric test development: A systems-informed shift towards qualitative research designs. *South African Journal of Psychiatry, 17*(1), 10-15. https://doi.org/10.4102/sajpsychiatry.v17i1.250

Tak, J. (2004). Structure of vocational interests for Korean college students. *Journal of Career Assessment, 12*(3), 298-311. https://doi.org/10.1177/1069072703261555

Tang, M. (2001). Investigation of the structure of vocational interests of Chinese college students. *Journal of Career Assessment, 9*(4), 365-379. https://doi.org/10.1177%2F106907270100900404

Tang, M. (2009). Examining the application of Holland's theory of vocational interests and choices of Chinese college students. *Journal of Career Assessment, 17*(1), 86-98. https://doi.org/10.1177/1069072708325743

Stellenbosch University https://scholar.sun.ac.za

Teye-Kwadjo, E. (2021). The Job-Search Self-Efficacy (JSSE) Scale: An item response theory investigation. *International Journal of Applied Positive Psychology, 6*, 301-314. https://doi.org/10.1007/s41042-021-00050-2

Teye-Kwadjo, E., & De Bruin, G. (2021). Rasch analysis of the Proactive Personality Scale. *Psychological Reports, 0*(0), 1-21. https://doi.org/10.1177/00332941211028110

Tien, H. S. (2009). Cultural encountering: The applicability of Holland's typology in Taiwan. *Asian Journal of Counselling, 16*(2), 193-226. https://www.hkier.cuhk.edu.hk/journal/document/AJC/ajc_v16n2_193-226.pdf

Toms, E. (2015). *A community assessment identifying support organisations in Kayamandi, Stellenbosch* (Master's thesis). Stellenbosch University. Sunscholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/60

Tracey, T. J. G., Wille, B., Durr II, M. R., & De Fruyt, F. (2014). An enhanced examination of Holland's consistency and differentiation hypotheses. *Journal of Vocational Behaviour, 84*, 237-247. https://www.academia.edu/download/45314047/An_enhanced_examination_of_Hollands_con20160503-23234-o2t8u1.pdf

Tracey, T. J. G., & Robbins, S. B. (2005). Stability of interests across ethnicity and gender: A longitudinal examination of Grades 8 through 12. *Journal of Vocational Behaviour, 67*, 335-364. https://doi.org/10.1016/j.jvb.2004.11.003

Tracey, T. J. G., & Rounds, J. (1996). The spherical representation of vocational interests. *Journal of Vocational Behaviour, 48*(1), 3-41. https://doi.org/10.1006/jvbe.1996.0002

Tracey, T. J. G., Robbins, S. B., & Hofsess, C. D. (2005). Stability and change in interests: A longitudinal study of adolescents from Grades 8 through 12. *Journal of Vocational Behaviour, 66*, 1-25. https://doi.org/10.1016/j.jvb.2003.11.002

Stellenbosch University https://scholar.sun.ac.za

Tracey, T. J., & Rounds, J. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin, 113*(2), 229-246. https://psycnet-apa-org.ez.sun.ac.za/doi/10.1037/0033-2909.113.2.229

Tsotetsi, I. N. (2012). *The relationship between personality and vocational interests* (Master's Dissertation). University of Johannesburg. https://www.proquest.com/openview/e9d5780633dfdb8e45c865417a1b8b81/1?pq-origsite=gscholar&cbl=2026366&diss=y

Van de Vijver, A. J. R., & Rothmann, S. (2004). Assessment in multicultural groups: The South African case. *South African Journal of Industrial Psychology. 30*(4), 1-7. https://doi.org/10.1111/j.1464-0597.2004.00169.x

Van Eeden, R., & Mantsha, T. R. (2007). Theoretical and methodological considerations in the translation of the 16PF5 into an African language. *South African Journal of Psychology, 37*(1), 62-81. https://journals.sagepub.com/doi/pdf/10.1177/008124630703700105

Vardarlı, B., Özyürek, R., Wilkins-Yel, K. G., & Tracey, T. J. G. (2017). Examining the structure of vocational interests in Turkey in the context of the Personal Globe. *International Journal for Educational and Vocational Guidance, 17*(3), 347-359. https://doi.org/10.1007/s10775-016-9338-6

Visser, D., & Viviers, R. (2010). Construct equivalence of the OPQ32n for Black and White people in South Africa. *South African Journal of Industrial Psychology, 36*(1), 1-11. https://doi.org/10.4102/sajip.v36i1.748

Wallis, T., & Birt, M. (2003). A comparison of native and non-native English-speaking groups' understanding of the vocabulary contained within the 16PF (SA92). *South African Journal of Psychology, 33*(3), 182-190. https://doi.org/10.1177/008124630303300307

Stellenbosch University https://scholar.sun.ac.za

Watson, M. B., & Stead, G. B. (2017). An overview of career theory. In M. B. Watson & G. B. Stead (Eds.), *Career Psychology in the South African Context* (3rd ed., pp. 21-47). Van Schaik Publishers.

Wang, S., Wiernik, B. M., & Morgan, B. (n.d.). *Gender measurement bias trait differences in vocational interest assessment* [Presented paper]. Society for Industrial and Organisational Psychology.

Weisgram, E., Bigler, R., & Liben, L. (2010). Gender, values, and occupational interests among children, adolescents, and adults. *Child Development, 81*(3), 778-796. https://doi.org/10.1111/j.1467-8624.2010.01433.x

Wiernik, B. M. (2016). *The nomological network of classic and contemporary career preferences* (Doctoral Dissertation). Minneapolis: University of Minnesota.

Wilkins, K. G., Ramkissoon, M., & Tracey, T. J. G. (2013). Structure of interest in a Caribbean sample: Application of the Personal Globe Inventory. *Journal of Vocational Behaviour, 83*(3), 367-372. https://doi.org/10.1016/j.jvb.2013.06.005

Wong, C., & Wong, P. (2006). Validation of Wong's Career Interest Assessment Questionnaire and Holland's revised hexagonal model of occupational interest in four Chinese societies. *Journal of Career Development, 32*(4), 378-393. https://doi.org/10.1177/0894845305284765

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370. *https://www.rasch.org/rmt/rmt83b.htm*

Wright, B. D., & Stone, M. H. (1979). *Best Test Design.* MESA Press. https://www.winsteps.com/index.htm

Wright, B., & Masters, G. (1982). *Rating Scale Analysis.* MESA PRESS. https://www.winsteps.com/index.htm

Stellenbosch University https://scholar.sun.ac.za

Yang, W., Stokes, G. S., & Hu, C. H. (2005). Cross-cultural validation of Holland's interest structure in Chinese populations. *Journal of Vocational Behaviour, 67*(3), 379-396. https://doi.org/10.1016/j.jvb.2004.08.003

Zhang, Y., Kube, E., Wang, Y., & Tracey, T. J. G. (2013). Vocational Interests in China: An evaluation of the Personal Globe Inventory-Short. *Journal of Vocational Behaviour, 83*, 99-105. https://doi.org/ 10.1016/j.jvb.2013.03.009

Stellenbosch University https://scholar.sun.ac.za

Appendix A: Research Ethics Committee Approval Letter

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

NOTICE OF APPROVAL

REC: Social, Behavioural and Education Research (SBER) - Initial Application Form

3 September 2021

Project number: 23176

Project Title: A Psychometric Investigation of the South African Career Interest Inventory among Secondary School Learners using the Rasch Measurement Model

Dear Ms N Jansen van Vuuren

Co-investigators:

Your REC: Social, Behavioural and Education Research (SBER) - Initial Application Form submitted on 01/08/2021 16:17 was reviewed and approved by the REC: Social, Behavioural and Education Research (REC: SBE).

Please note below expiration date of this approved submission:

Ethics approval period:

Protocol approval date (Humanities)	Protocol expiration date (Humanities)
3 September 2021	2 September 2024

GENERAL REC COMMENTS PERTAINING TO THIS PROJECT:

INVESTIGATOR RESPONSIBILITIES

Please take note of the General Investigator Responsibilities attached to this letter. You may commence with your research after complying fully with these guidelines.

If the researcher deviates in any way from the proposal approved by the REC: SBE, the researcher must notify the REC of these changes.

Please use your SU project number (23176) on any documents or correspondence with the REC concerning your project.

Please note that the REC has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

Stellenbosch University https://scholar.sun.ac.za

CONTINUATION OF PROJECTS AFTER REC APPROVAL PERIOD

You are required to submit a progress report to the REC: SBE before the approval period has expired if a continuation of ethics approval is required. The Committee will then consider the continuation of the project for a further year (if necessary).

Once you have completed your research, you are required to submit a final report to the REC: SBE for review.

Included Documents:

Document Type	File Name	Date	Version
Research Protocol/Proposal	20069138_ Research Proposal Final	19/07/2021	1
Co-investigator CV	STEPHAN RABIE_CV_2021	19/07/2021	1
Default	Stephan Rabie REC approval letter	19/07/2021	1
Default	DESC Review report 1_Jansen Van Vuuren	19/07/2021	1
Default	DESC Review Report 2_Jansen van Vuuren	19/07/2021	1
Default	DESC Review_Masters student feedback report form	19/07/2021	1
Proof of permission	Permission_Access to SACIIX	27/07/2021	1
Budget	MA Thesis Budget	28/07/2021	1

If you have any questions or need further help, please contact the REC office at cgraham@sun.ac.za.

Sincerely,

Clarissa Graham

REC Coordinator: Research Ethics Committee: Social, Behavioral and Education Research

National Health Research Ethics Committee (NHREC) registration number: REC-050411-032.
The Research Ethics Committee: Social, Behavioural and Education Research complies with the SA National Health Act No.61 2003 as it pertains to health research. In addition, this committee abides by the ethical norms and principles for research established by the Declaration of Helsinki (2013) and the Department of Health Guidelines for Ethical Research: Principles Structures and Processes (2nd Ed.) 2015. *Annually a number of projects may be selected randomly for an external audit.*

Stellenbosch University https://scholar.sun.ac.za

Appendix B: Rabie (2017) Research Ethics Committee Approval Letter

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Approved with Stipulations

Response to Modifications- (New Application)

Il-Mar-2016
Rabie, Stephan S

Proposal SU-HSD-000640

Title: Investigating gender and ethnic differences on the South African Career Interest Inventory among high school students

Dear Mr. Stephan Rabie,

Your **Response to Modifications** - (*New Application*) received on **01-Mar-2016**, was reviewed by members of the **Research Ethics Committee: Human Research (Humanities)** via Expedited review procedures on **10-Mar-2016.**

Please note the following information about your approved research proposal:

Proposal Approval Period: **11-Mar-2016 -10-Mar-2017**

The following stipulations are relevant to the approval of your project and must be adhered to:

The researcher may proceed with the envisaged research provided that the following stipulations, relevant to the approval of your project are adhered to or addressed.

1) **The REC acknowledges the changes made to the invitation letter, informed assent form and informed consent form. The user friendliness of these documents can, however, still be improved by removing some of the academic jargon. For example: do the learners or their parents really need to know that the current study is a "quantitative" study; that they need to complete a "demographic" questionnaire; and that the scale will be a 5-point "Likert-type" scale? Consider rephrasing "vocational" interests into words the readers will understand.**

2) **In the researcher's response letter, he mentions that "The SACII will be administered by a psychometrist (independent practice) who received training in the administration of career assessment and is registered with the Health Professions Council of South Africa." Is this person the researcher or will someone else visit to schools to administer the questionnaires? If the latter, the principals of the participating schools should be made aware of this fact.**

Stellenbosch University https://scholar.sun.ac.za

3) **The word "ethnic group" is still used in the biographical questionnaire. Please rephrase this or use an alternative. If an alternative cannot be found, the concept should be defined (grounded in existing literature).**

Please provide a letter of response to all the points raised IN ADDITION to HIGHLIGHTING or using the TRACK CHANGES function to indicate ALL the corrections/amendments of ALL DOCUMENTS clearly in order to allow rapid scrutiny and appraisal.

Please take note of the general Investigator Responsibilities attached to this letter. You may commence with your research after complying fully with these guidelines.

Please remember to use your **proposal number (SU-HSD-000640)** on any documents or correspondence with the REC concerning your research proposal.

Please note that the REC has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

Also note that a progress report should be submitted to the Committee before the approval period has expired if a continuation is required. The Committee will then consider the continuation of the project for a further year (if necessary).
This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki and the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health). Annually a number of projects may be selected randomly for an external audit. National Health Research Ethics Committee (NHREC) registration number REC-050411-032.

We wish you the best as you conduct your research.
If you have any questions or need further help, please contact the REC office at 218089183.

Sincerely,

Clarissa Graham
REC Coordinator
Research Ethics Committee: Human Research (Humanities)

Stellenbosch University https://scholar.sun.ac.za

Appendix C: Rabie (2017) WCED Approval Letter

Western Cape Government
Education

Audrey.wyngaard@westerncape.gov.za
tel: +27 021 467 9272
Fax: 0865902282
Private Bag x9114, Cape Town, 8000
wced.wcape.gov.za

REFERENCE: 20150930-3811
ENQUIRIES: Dr A T Wyngaard

Mr Stephan Rabie
69 Skaamrosie Street
Proteavalley
Bellville
7530

Dear Mr Stephan Rabie

RESEARCH PROPOSAL: INVESTIGATING GENDER AND ETHNIC DIFFERENCES ON THE SOUTH AFRICAN CAREER INTEREST INVENTORY AMONG HIGH SCHOOL STUDENTS

Your application to conduct the above-mentioned research in schools in the Western Cape has been approved subject to the following conditions:

1. Principals, educators and learners are under no obligation to assist you in your investigation.
2. Principals, educators, learners and schools should not be identifiable in any way from the results of the investigation.
3. You make all the arrangements concerning your investigation.
4. Educators' programmes are not to be interrupted.
5. The Study is to be conducted from **18 August 2015 till 30 September 2016.**
6. No research can be conducted during the fourth term as schools are preparing and finalizing syllabi for examinations (October to December).
7. Should you wish to extend the period of your survey, please contact Dr A.T Wyngaard at the contact numbers above quoting the reference number?
8. A photocopy of this letter is submitted to the principal where the intended research is to be conducted.
9. Your research will be limited to the list of schools as forwarded to the Western Cape Education Department.
10. A brief summary of the content, findings and recommendations is provided to the Director: Research Services.
11. The Department receives a copy of the completed report/dissertation/thesis addressed to:

    The Director: Research Services
    Western Cape Education Department
    Private Bag X9114
    CAPE TOWN
    8000

We wish you success in your research.

Kind regards.
Signed: Dr Audrey T Wyngaard
Directorate: Research
DATE: 30 September 2015

Stellenbosch University https://scholar.sun.ac.za

Appendix D: Permission to Access SACII-X Data

University of Cape Town

Groote Schuur Hospital

Western Cape Government
Health

Department of Psychiatry and Mental Health

Professor **Dan J Stein, FRCPC, PhD, DPhil**
Head: Department of Psychiatry and Mental Health, UCT & GSH
Email: dan.stein@uct.ac.za
UCT Dept of Psychiatry and Mental Health, Groote Schuur Hospital (J-2), Anzio Rd,
Observatory 7925, Cape Town, South Africa
Phone: +27 21 4042137 Fax: +27 21 4488158
URL: http://www.health.uct.ac.za/departments/psychiatry/about/

20/09/2021

RE: Request for access to SACII-X data

Dear Nicola,

I write this letter to grant you permission to access the isiXhosa sample of my PhD data. In total, there 266 datapoints that will be shared with you. All the data will be deidentified to ensure anonymity and confidentiality. The data will be shared with you via encrypted electronic transfer. Please note, access to these data is granted to only you and your statisticians.

All the best with your study.

Sincerely,

Dr Stephan Rabie

Senior Research Officer

HIV Mental Health Research Unit

Phone: +27 21 650 4658

Email: stephan.rabie@uct.ac.za

Stellenbosch University https://scholar.sun.ac.za

Appendix E: Item Measures and Point-Measure Correlations of the SACII-X Items

Item	Measure	Model S.E.	Point-Measure Correlation	Expected Point-Measure Correlation
Realistic				
RR21	.30	.06	.48	.49
RR24	.30	.06	.49	.49
RR25	.27	.06	.55	.49
RR28	.24	.06	.55	.50
RR13	.20	.06	.48	.50
RR27	.13	.06	.55	.51
RR20	.08	.06	.43	.52
RR26	.08	.06	.55	.52
RR9	.04	.06	.51	.52
RR14	.04	.06	.57	.52
RR2	.03	.06	.53	.52
RR4	.02	.06	.59	.52
RR10	.01	.06	.54	.53
RR7	.00	.06	.53	.53
RR17	-.02	.06	.54	.53
RR18	-.04	.06	.52	.53
RR11	-.08	.06	.58	.54
RR23	-.08	.06	.58	.54
RR15	-.11	.06	.50	.54
RR8	-.26	.05	.55	.56
RR19	-.26	.05	.41	.56
RR1	-.45	.05	.58	.58
RR16	-.46	.05	.53	.59
Investigative				
I8	.33	.06	.47	.48
I19	.23	.06	.53	.49
I13	.18	.06	.52	.50
I25	.16	.06	.50	.50
I11	.13	.06	.53	.50
I5	.12	.06	.54	.50
I9	.12	.06	.48	.50
I6	.10	.05	.50	.50
I20	.08	.05	.53	.50
I22	.07	.05	.46	.50
I17	.01	.05	.49	.51
I2	-.01	.05	.52	.51
I24	-.02	.05	.52	.51
I1	-.03	.05	.50	.51
I21	-.06	.05	.52	.52
I16	-.12	.05	.52	.52

Stellenbosch University https://scholar.sun.ac.za

Item	Measure	Model S.E.	Point-Measure Correlation	Expected Point-Measure Correlation
I26	-.14	.05	.45	.52
I10	-.19	.05	.51	.53
I4	-.26	.05	.52	.53
I23	-.32	.05	.56	.54
I12	-.38	.05	.53	.54
		Artistic		
A27	.26	.05	.52	.50
A16	.22	.05	.53	.50
A9	.20	.05	.46	.50
A4	.15	.05	.48	.50
A30	.13	.05	.60	.50
A29	.11	.05	.51	.50
A22	.11	.05	.60	.50
A15	.07	.05	.55	.51
A8	.06	.05	.49	.51
A2	.03	.05	.46	.51
A21	.02	.05	.51	.51
A24	.02	.05	.48	.51
A18	.01	.05	.58	.51
A20	.00	.05	.49	.51
A10	-.02	.05	.40	.51
A28	-.03	.05	.58	.51
A26	-.04	.05	.57	.51
A19	-.05	.05	.52	.51
A6	-.07	.05	.50	.51
A11	-.12	.05	.49	.51
A1	-.14	.05	.45	.51
A23	-.15	.05	.43	.51
A5	-.16	.05	.52	.51
A7	-.18	.05	.43	.51
A25	-.20	.05	.56	.51
A13	-.25	.05	.52	.51
		Social		
S16	.34	.05	.49	.52
S28	.33	.05	.48	.52
S22	.25	.05	.45	.53
S26	.25	.05	.45	.53
S25	.22	.05	.45	.53
S2	.16	.05	.57	.53
S12	.12	.05	.44	.53
S17	.11	.05	.51	.53
S3	.06	.05	.55	.54
S18	.03	.05	.50	.54

Stellenbosch University https://scholar.sun.ac.za

Item	Measure	Model S.E.	Point-Measure Correlation	Expected Point-Measure Correlation
S8	.03	.05	.53	.54
S27	.01	.05	.58	.54
S20	-.02	.05	.60	.54
S15	-.05	.05	.60	.54
S4	-.13	.05	.61	.55
S23	-.14	.05	.61	.55
S24	-.15	.05	.54	.55
S7	-.22	.05	.57	.55
S11	-.26	.05	.52	.55
S10	-.29	.05	.58	.55
S6	-.30	.05	.61	.55
S5	-.37	.05	.58	.56
Enterprising				
E22	.53	.06	.39	.49
E3	.31	.06	.51	.50
E18	.25	.06	.57	.50
E2	.24	.06	.50	.50
E21	.19	.06	.57	.51
E9	.13	.06	.53	.51
E20	.12	.05	.50	.51
E28	.12	.05	.57	.51
E13	.11	.05	.56	.51
E5	.04	.05	.56	.51
E8	.04	.05	.57	.51
E16	.03	.05	.49	.51
E23	.01	.05	.56	.51
E17	.00	.05	.39	.51
E1	-.01	.05	.57	.52
E25	-.06	.05	.53	.52
E7	-.08	.05	.53	.52
E10	-.11	.05	.59	.52
E15	-.13	.05	.55	.52
E19	-.16	.05	.52	.52
E14	-.18	.05	.47	.52
E24	-.24	.05	.51	.53
E27	-.36	.05	.42	.53
E11	-.39	.05	.49	.53
E12	-.40	.05	.42	.53
Conventional				
C28	.39	.06	.47	.45
C5	.33	.06	.47	.46
C3	.30	.06	.47	.46
C29	.24	.06	.43	.47

Stellenbosch University https://scholar.sun.ac.za

Item	Measure	Model S.E.	Point-Measure Correlation	Expected Point-Measure Correlation
C12	.22	.06	.44	.47
C11	.19	.06	.52	.47
C20	.17	.06	.45	.47
C18	.15	.06	.41	.48
C9	.13	.06	.49	.48
C25	.10	.06	.44	.48
C16	.10	.06	.49	.48
C24	.09	.06	.51	.48
C1	.08	.06	.50	.48
C14	-.05	.05	.41	.50
C27	-.08	.05	.56	.50
C15	-.09	.05	.46	.50
C8	-.12	.05	.53	.50
C2	-.17	.05	.55	.51
C21	-.20	.05	.53	.51
C17	-.23	.05	.56	.51
C19	-.23	.05	.55	.51
C7	-.24	.05	.50	.52
C22	-.34	.05	.50	.53
C26	-.34	.05	.49	.53
C4	-.40	.05	.52	.53

Note. S.E.=Standard Error