

Cross-cultural measurement of eudaimonic well-being

L Schutte
13012584

Thesis submitted for the degree *Doctor Philosophiae* in
Psychology at the Potchefstroom Campus of the North-West
University

Promoter: Prof MP Wissing
Assistant Promoter: Dr SM Ellis

October 2016

Contents

Acknowledgements	ii
Summary	iii
Opsomming	vi
Preface	ix
Letter of Permission 1	x
Letter of Permission 2	xi
Section 1: Introduction	1
Section 2: Manuscript 1. Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa, Australia, and New Zealand	14
2.1. Guidelines for authors: <i>Health and Quality of Life Outcomes</i>	15
2.2. Manuscript: Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa, Australia, and New Zealand	19
Section 3: Manuscript 2. Clarifying the factor structure of the Mental Health Continuum Short Form in three languages: A bifactor model approach	66
3.1. Guidelines for authors: <i>Social Indicators Research</i>	67
3.2. Manuscript: Clarifying the factor structure of the Mental Health Continuum Short Form in three languages: A bifactor model approach	81
Section 4: Manuscript 3. Problematic factorial validity of three language versions of the Basic Psychological Needs Scale (BPNS): Why and what are the implications?	122
4.1. Guidelines for authors: <i>Journal of Personality Assessment</i>	123
4.2. Manuscript	127
Section 5: Conclusions, Implications, and Recommendations	163
Complete Reference List	178

Acknowledgements

First and foremost I thank my Heavenly Father for His unceasing love and grace. Thank you for giving purpose and direction to my life because I know of none of what I do is for myself, but for the greater good and the glory of your Name.

I wish to express my sincerest gratitude to the following persons and institutions:

- To my promotor, Prof. Marié Wissing. It is almost impossible to express in words what an immense privilege it was to do this work under your guidance. Thank you for the wealth of knowledge, absolute dedication, and wisdom with which you walked this road with me. You are a mentor to me on so many levels and in endless ways. Thank you, in particular, for being such an example of what it means to be “functioning well”.
- To my co-promotor, Dr. Suria Ellis. Thank you for your encouragement, inputs, and time throughout this project.
- To my dearest husband, WD Schutte. Thank you for the endless ways in which you supported and motivated me on this road. Thank you for helping me to keep perspective, for supporting me in so many practical ways, and for always being willing to listen and encourage. Thank you for the deeply meaningful experience of sharing life with you.
- To my lovely little Elfriede. Your existence in my life, even though you were not out in the big world yet when I completed this thesis, fills me with absolute wonder and the deepest gratitude. I can’t wait for you to become part of our lives.
- To my family – my parents, Willem and Elfriede, and my brothers and sister, Carel, Anneke, and Hanjo. Thank you for your care, support and sincere interest throughout this project. Thank you for being examples of what it means to be functioning well and serving the greater good.
- To my family-in-law, for your support and encouragement throughout this process.
- To our friends and broader family for sharing life with me, also during this project.
- To my colleagues at AUTHeR, for such a positive and stimulating work environment.
- To Martinus Postma who assisted me with the language editing of this project.
- To the participants in these projects, who shared their time and lives with us.
- To the National Research Foundation (NRF) of South Africa by funding for the FORT3 research project and for a doctoral scholarship. Opinions expressed and conclusions arrived are those of the authors and are not necessarily to be attributed to the NRF. Also to the North-West University, for funding in the form of a doctoral scholarship.

Summary

Cross-cultural measurement of eudaimonic well-being

Keywords: cross-cultural measurement, eudaimonic well-being, modern psychometric techniques, Rasch modelling, meaning in life, mental health continuum, basic psychological needs, positive psychology, language, translation

The overall aim of the present study was to further the understanding and measurement of eudaimonic well-being (EWB) or “functioning well” across different demographic and cultural groups by implementing modern psychometric techniques. This was done by exploring in three manuscripts the cross-cultural psychometric properties of measurement instruments that operationalise three prominent theories associated with EWB.

The first manuscript applied Rasch modelling to explore the psychometric properties of the Meaning in Life Questionnaire (MLQ) among adults from South Africa, Australia, and New Zealand ($N = 601$). The findings revealed that using less response categories and removal of the reversed-phrased item in the scale may improve the scale’s functioning. The Presence subscale exhibited differential item functioning on several items for the country variable. It was found that the Presence subscale was insensitive at high levels of presence of meaning while the majority of the respondents fell in that range. Reasons for, as well as the far-reaching implications of this finding, were contemplated.

The second manuscript explored the suitability of applying a bifactor model to English, Afrikaans, and Setswana versions of the Mental Health Continuum Short Form (MHC-SF) among South African university students ($N = 1060$). A bifactor model consistently displayed superior fit, and it was shown that the scale total score can be used as a

reliable indication of overall positive mental health, but that it is not appropriate to interpret subscale scores. The findings highlighted the lack of focus on contribution towards a greater good beyond the self in the present scale, indicating the necessity of more attention to this important aspect of functioning well in future research. The scale exhibited configural, partial metric, and partial scalar invariance across the three samples and the average scale total score did not differ practically significantly across the groups.

The third manuscript explored the factorial validity of English, Afrikaans, and Setswana versions of the Basic Psychological Needs Scale (BPNS) among South African university students ($N = 1056$). After incorporating a negative-worded method effect and removing several problematic items, the fit of the intended three-factor model was good for the Afrikaans version, marginal for the English version, and poor for the Setswana version of the scale. The resulting factors' reliabilities were low. Configural, metric, and partial scalar invariance were established between the English and Afrikaans versions. These findings not only highlighted problems with the particular scale, but also raised questions about the universality assumption of the basic psychological needs theory. The study also emphasised the linguistic and cultural complexities involved in working within diverse and multicultural communities.

Overall, significant conclusions regarding the theory and measurement of EWB crystallised. The study showed that caution should be applied when transferring theories and their operationalisations from one context or culture to another and that measurement within diverse multicultural communities poses certain complexities and challenges. In addition, it was found that the nuances of positive traits and experiences at high levels may not be well understood. The findings suggested that a strong general factor of well-being exists, while, at the same time, unique subdimensions can be distinguished and interpreted, given that they are measured comprehensively. General pointers that will enhance the construction of

measurement instruments of EWB in diverse contexts flowed from the study. For example, the crucial importance of proper attention to the substantive phase of scale development where the construct is clearly conceptualised along the full spectrum of the latent trait continuum was highlighted, as well as the importance of using reversed-phrased items in an appropriate manner. This study furthered the science of EWB on theoretical and measurement levels, which can, in turn, be applied to develop culture-sensitive interventions that truly enhance the lives of people.

Opsomming

Kruiskulturele meting van eudaimoniese welstand

Sleutelwoorde: kruiskulturele meting, eudaimoniese welstand, moderne psigometriese tegnieke, Raschmodellering, lewensbetekenis, kontinuum van mentale gesondheid, basiese psigologiese behoeftes, positiewe psigologie, taal, vertaling

Die oorhoofse doelstelling van hierdie studie was om die begrip en meting van eudaimoniese welstand (EWS) of “gesonde funksionering” oor verskillende demografiese en kulturele groepe heen te bevorder deur die toepassing van moderne psigometriese tegnieke. Dit is gedoen deur die verkenning van kruiskulturele psigometriese eienskappe van meetinstrumente wat prominente teorieë geassosieer met EWS operasionaliseer soos in drie manuskripte uiteengesit.

In die eerste manuskrip is Raschmodellering toegepas om die psigometriese eienskappe van die Meaning in Life Questionnaire (MLQ) onder volwassenes uit Suid-Afrika, Australië en Nieu Seeland ($N = 601$) te ondersoek. Die bevindings het daarop gedui dat die skaal se funksionering verbeter kan word deur minder responskategorieë te gebruik, en deur die omgekeerde item in die skaal te verwyder. By die “Presence”-subskaal is differensiële itemfunksionering op ’n paar items vir die land-veranderlike geïdentifiseer. Daar is gevind dat die “Presence”-subskaal nie sensitief was vir hoë vlakke van die teenwoordigheid van lewensbetekenis nie, terwyl die meeste van die respondente hoë vlakke van hierdie konstruk gemanifesteer het. Redes vir, asook die verreikende implikasies van hierdie bevinding is uiteengesit.

In die tweede manuskrip is die toepaslikheid ondersoek van die passing van 'n bifaktormodel vir Engelse, Afrikaanse en Setswana-weergawes van die Mental Health Continuum Short Form (MHC-SF) ten opsigte van Suid-Afrikaanse studente. 'n Bifaktormodel het deurgaans die beste passing getoon en dit is aangedui dat die skaal se totaalstelling as 'n betroubare aanduiding van algemene positiewe mentale gesondheid gebruik kan word, maar dat dit nie toepaslik is om subskaaltellings te interpreteer nie. Die bevindings het die gebrek aan fokus op 'n groter belang buite die self in die huidige skaal uitgelig, wat daarop dui dat meer aandag in toekomstige navorsing aan hierdie belangrike aspek van goeie funksionering geskenk moet word. Die skaal het konfigurele, gedeeltelike metriese en gedeeltelike skalaarinvariansie oor die drie steekproewe heen getoon en die gemiddelde skaal-totaaltellings het nie prakties betekenisvol oor die groepe heen verskil nie.

Die derde manuskrip het die faktorale geldigheid van Engelse, Afrikaanse en Setswana-weergawes van die Basic Psychological Needs Scale (BPNS) onder Suid-Afrikaanse universiteitstudente ($N = 1056$) verken. Nadat 'n negatief-bewoorde metode-effek in ag geneem is, en verskeie problematiese items verwyder is, was die passing van die gepostuleerde driefaktormodel goed vir die Afrikaanse weergawe, marginaal vir die Engelse weergawe en swak vir die Setswana-weergawe van die skaal. Die faktore wat verkry is, het lae betroubaarheidsindekse gehad. Die Engelse en Afrikaanse weergawes van die skaal het konfigurele, metriese en gedeeltelike skalaarinvariansie getoon. Hierdie bevindings het nie alleen probleme met die spesifieke skaal uitgelig nie, maar dit het ook vrae laat ontstaan oor die universaliteitsaannames van die teorie van basiese psigologiese behoeftes. Die studie het ook die linguistiese en kulturele kompleksiteite uitgelig wat betrokke is wanneer daar binne diverse en multikulturele kontekste gewerk word.

In geheel het belangrike gevolgtrekkings oor die teorie en meting van EWS gekristaliseer. Die studie het getoon dat die oordrag van teorieë en hul operasionaliserings

van een konteks of kultuur na 'n ander met omsigtigheid gedoen moet word en dat meting binne diverse multikulturele kontekste bepaalde kompleksiteite en uitdagings meebring. Dit is verder bevind dat die nuanses van hoë vlakke van positiewe trekke en ervarings moontlik nie goed verstaan word nie. Die bevindings het daarop gedui dat 'n sterk algemene faktor van welstand bestaan, terwyl daar terselfdertyd unieke subdimensies onderskei en geïnterpreteer kan word, gegewe dat die subdimensies omvattend gemeet word. Algemene riglyne wat die konstruksie van meetinstrumente van EWS in diverse kontekste sal bevorder, het uit die studie voortgevloei. Byvoorbeeld, die kardinale belang van behoorlike aandag aan die substantiewe fase van skaalontwikkeling, waar die konstruk duidelik oor die volle spektrum van die latente veranderlike kontinuum gekonseptualiseer word, is uitgelig, asook die belangrikheid daarvan om omgekeerde items op 'n toepaslike wyse te gebruik. Hierdie studie kan meewerk aan die bevordering van die wetenskap van EWS op teoretiese en metingsvlakke, wat weer toegepas kan word om kultuursensitiewe intervensies te ontwikkel wat werklik die lewens van mense verbeter.

Preface

This thesis is submitted in accordance with rule A.5, and specifically in article format as described in rules A.5.4.1.2., A.5.4.2.7, A.5.4.2.8, and A.5.4.2.9 of the North-West University.

The three manuscripts comprising this thesis have been submitted to Health and Quality of Life Outcomes (Manuscript 1), Social Indicators Research (Manuscript 2), and Journal of Personality Assessment (Manuscript 3). Manuscript 1 was successfully published in 2016.

For each manuscript, the guidelines to the authors provided by the target journal are followed and the referencing style is used in accordance with the target journal's requirements. The referencing style and editorial approach for the rest of the thesis, including the Introduction and Conclusions, Recommendations, and Implications sections are in line with the Publication Manual of the American Psychological Association (6th Edition).

For the purposes of this thesis, the page numbering of the thesis as a whole is consecutive. However, for submission purposes, each individual manuscript was numbered starting from page 1.

Please find attached the letters of permission by the co-authors authorising the use of these manuscripts for the purposes of submission for a Ph.D. degree, as well as an outline of the contribution of each co-author in each manuscript.

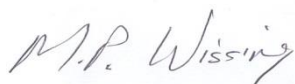
Letter of Permission 1

Permission to submit article for degree purposes

Permission is hereby granted by the co-authors that the following article may be submitted by Lusilda Schutte for the purpose of obtaining a PhD degree in Psychology:

Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa, Australia, and New Zealand

Contributions of authors: Marié P. Wissing (MPW) and Dianne A. Vella-Brodrick (DAVB) participated in the design and planning of the study. MPW, DAVB and Paul E. Jose (PEJ) took responsibility for the data gathering and capturing. Lusilda Schutte (LS) performed and interpreted the statistical analyses, drafted the manuscript, incorporated all suggestions from the co-authors into the manuscript, prepared the final manuscript for submission, and served as the corresponding author during the submission and publication process. MPW provided continuous feedback regarding the draft document. Suria M. Ellis (SME), DAVB, and PEJ revised the manuscript critically for important intellectual content. All authors read and approved the final manuscript.



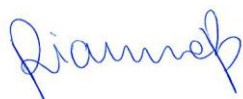
Prof. M. P. Wissing (Promotor)



Dr. S. M. Ellis (Co-promotor)



Prof. P. E. Jose



Prof. D. A. Vella-Brodrick

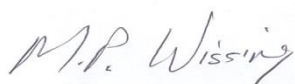
Letter of Permission 2

Permission to submit manuscripts for degree purposes

Permission is hereby granted by the co-authors that the following manuscripts may be submitted by Lusilda Schutte for the purpose of obtaining a PhD degree in Psychology:

- *Clarifying the factor structure of the Mental Health Continuum Short Form in three languages: A bifactor model approach*
- *Problematic factorial validity of three language versions of the Basic Psychological Needs Scale (BPNS): Why and what are the implications?*

Contributions of authors: Marié P. Wissing (MPW) developed, obtained ethical approval, and acquired funding for the broad FORT3-project. The present two manuscripts form part of a sub-study of the FORT3-project. Lusilda Schutte (LS) and MPW participated in the design and planning of the sub-study and in coordinating and conducting the translation of the measurement instruments, data gathering, and data capturing. LS performed and interpreted the statistical analyses, drafted the manuscripts, incorporated suggestions from the co-authors into the manuscript, prepared the final manuscripts for submission, and served as the corresponding author during the submission and publication process. MPW provided continuous feedback regarding the draft documents. Suria M. Ellis (SME) revised the manuscripts critically for important intellectual content.



Prof. M. P. Wissing (Promotor)



Dr. S. M. Ellis (Co-promotor)

Section 1: Introduction

Section 1: Introduction

The focus of the present study was to study the measurement of eudaimonic well-being (EWB) across countries and cultures using modern psychometric techniques. In this process, insight would not only be gained on the applicability of the selected measurement instruments in various contexts, but also on the theories underlying these measures and their transferability to different groups. In addition, the implications of the findings in this study would go beyond the specific instruments and theories, as the use of modern psychometric techniques in the present study would allow for fresh views on the theory and measurement of EWB in general.

Eudaimonic Well-being (EWB)

Interest in exploring the nature of and improving psychosocial well-being receives widespread attention in contemporary literature. Two main approaches to understanding well-being are often distinguished, namely the hedonic perspective, which considers “feeling good” (Diener, 1984; Kahneman, Diener, & Schwarz, 1999) and the eudaimonic view, which focuses on “functioning well” (Deci & Ryan, 2008a; Huta & Ryan, 2010; Waterman, 1993; Waterman et al., 2010). While the distinction between hedonic and eudaimonic well-being is sometimes debated in the literature (cf., Biswas-Diener, Kashdan, & King, 2009; Kashdan, Biswas-Diener, & King, 2008), many researchers agree that, although the two perspectives are interrelated, they are distinct, and each viewpoint contributes to a more comprehensive understanding of mental well-being (Baumeister, Vohs, Aaker, & Garbinsky, 2013; Delle Fave, Brdar, Wissing, & Vella-Brodrick, 2013; Henderson & Knight, 2012; Huta & Ryan, 2010; Joshanloo, 2015; Vittersø & Sørholt, 2011; Waterman, Schwartz, & Conti, 2008). In the past, the well-being research arena was dominated by studies of hedonic well-being, but lately research on eudaimonic well-being came to the fore.

Philosophically, eudaimonia evolved to have different meanings attached to the notion (Waterman, 2013). Different angles within eudaimonic philosophy have been translated into diverse psychological theories of EWB. For example, Waterman's work on eudaimonic identity theory and personal expressiveness (Waterman, 1990, 1993, 2008; Waterman & Schwartz, 2013), Sheldon's self-concordance theory (Sheldon, 2002, 2013), Ryff's theory of psychological well-being (Ryff, 1989b, 2013; Ryff & Singer, 2008), Keyes's (1998) theory of social well-being, and the self-determination theory (Ryan, Curren, & Deci, 2013; Ryan & Deci, 2000b; Ryan, Huta, & Deci, 2008) are associated with EWB. The work of the virtue ethicists, who assert that excellence is not about one's possessions or status, but about the kind of person one is as expressed in virtues such as courage, honesty, and justice, also falls within the domain of eudaimonic research (Fowers, 2005, 2012; Richardson, 2012; Slife, 2012). Furthermore, research about meaning in life is related to EWB (Frankl, 1963; Morgan & Farsides, 2009; Ryff, 1989b; Ryff & Singer, 2008; Steger, 2009; Steger, Frazier, Oishi, & Kaler, 2006; Steger, Shin, Shim, & Fitch-Martin, 2013; Wong, 2011). Recently, the relational nature of meaning was emphasised (Delle Fave & Soosai-Nathan, 2014; Wissing, 2014), where interconnectedness was not only seen as a prominent source of meaning, but actually as "the core essence of meaning itself" (Delle Fave & Soosai-Nathan, 2014, p. 33).

The endeavour to study EWB depends inter alia on the development of high quality measurement instruments that accurately and comprehensively operationalise the relevant theoretical understanding of the construct, and the development of such measures will, in turn, contribute to growth in the theoretical understanding of the construct. It was already acknowledged by Cronbach and Meehl (1955) that the development of an assessment measure and the understanding of the underlying phenomenon go hand-in-hand – the process of measurement and the scientific questions it attempts to answer interact continuously (DeVellis, 2012).

In the present study, three theories that fall within the domain of EWB and their operationalisations were relevant. First, the Meaning in Life Questionnaire (Steger et al., 2006) operationalises Steger's (2009) model of meaning in life. This model differentiates between presence of meaning, which involves "the extent to which people comprehend, make sense of, or see significance in their lives, accompanied by the degree to which they perceive themselves to have a purpose, mission, or overarching aim in life" (Steger, 2009, p. 682) and search for meaning, which refers to "the strength, intensity, and activity of people's desire and efforts to establish and/or augment their understanding of the meaning, significance, and purpose of their lives" (Steger, Kashdan, Sullivan, & Lorentz, 2008, p. 200). Second, the Mental Health Continuum Short Form (Keyes, 2006, 2009; Keyes et al., 2008) is an operationalisation of a theory of overall positive mental health that includes both hedonic well-being (labelled emotional well-being) in terms of positive affect and satisfaction with life, and eudaimonic well-being in terms of Ryff's six dimensions of psychological or personal well-being (1989b), and Keyes's five dimensions of social well-being (1998) in its conceptualisation. Although the theory defines hedonic well-being to be one of the dimensions of well-being, the emphasis in this theory is on a eudaimonic understanding of what it means to be mentally healthy. This is also evident from the ratio of the items in the Mental Health Continuum Short Form, where three items measure hedonic well-being, while 11 items target eudaimonic well-being. In the light of the prominence of EWB in the theory and scale, and the fact that hedonic and eudaimonic well-being are interrelated (Delle Fave et al., 2013; Waterman, 2008), we deemed it appropriate to include the theory and scale in the present study which focuses on EWB. Third, the Basic Psychological Needs Scale (Gagné, 2003) operationalises basic psychological needs theory, a prominent sub-theory of self-determination theory (Deci & Ryan, 2000, 2008b). This theory asserts that the satisfaction of the three basic psychological needs of autonomy, competence, and relatedness is essential for

the actualisation of human potential, and the frustration thereof the source of psychological ill-being and dysfunction (Deci & Ryan, 2000). In this theory, autonomy concerns an experience of volition and choice in behaviour regulation, competence refers to an experience of efficacy concerning one's internal and external environment, and relatedness concerns a sense of being connected to and cared for (Ryan et al., 2008).

The development of all of the above instruments was based on data from the USA. Although several cross-cultural studies have been conducted on some of the scales, a need exists for the application of new frontline statistical procedures to obtain insight into the functioning of the measures and theories, also cross-culturally. In addition, for some of the scales there is a lack of validation studies based on non-Western data. In order for science to grow in its understanding of EWB in a diverse global society, further research to explore the psychometric qualities of measures of EWB across different cultures and contexts, also using modern psychometric techniques, is imperative.

Cross-cultural Measurement and Scale Validation

Many different definitions for culture exist. Triandis (2007) identified three aspects that are considered characteristic of culture, namely that "culture emerges in adaptive interactions between humans and environments", that "culture consists of shared elements", and that "culture is transmitted across time periods and generations" (pp. 64-65). In the past, cross-cultural literature placed much emphasis on differences between individualistic and collectivistic cultures, but, according to Cohen (2009, 2010), many other forms of culture exist, for example, religion, socioeconomic status, and region within a country.

Culture is widely considered a fundamental variable to be integrated in theory and empirical research on human behaviour (Matsumoto & Yoo, 2006). Although cross-cultural comparisons have been conducted on various positive psychology constructs in quantitative, qualitative, and mixed methods studies, there is still much space for exploration and growth

in comprehension (Knoop & Delle Fave, 2013). This is especially the case for studies from a EWB perspective, since much of the cross-cultural explorations to date was done on hedonic well-being constructs (Diener, 2009; Veenhoven, 2010, 2012). One avenue of development involves validly assessing and comparing levels of EWB across different cultural contexts. This involves establishing cross-cultural equivalence of the construct and the measurement instruments designed to assess it – a vital step to ensure valid inferences in studies across cultures (Oishi, 2010; Van de Vijver & Leung, 2011).

Bias is defined as the presence of nuisance factors and occurs when aspects of the measurement instrument do not have the same meaning within and across cultures (Poortinga, 1989). Van de Vijver (2011) and Van de Vijver and Leung (1997, 2011) distinguished three types of bias. Construct bias occurs if the definitions of the construct overlap only partially across different cultures, often because of differential appropriateness of behaviours associated with the construct. Method bias occurs, for example, if the samples from the different cultural groups are not comparable, when there are differences in the administration conditions, when questionnaire instructions differ, due to communication problems between respondents and interviewers, when the different groups have differential familiarity with the stimulus materials or response procedures, or due to differential response styles (e.g., acquiescence) in the various groups. Item bias refers to differential item functioning and occurs due to incongruities at item level, which causes respondents on the same level of the underlying construct to have different mean scores because of different cultural origins. Item bias can, for example, arise due to poor item translations, ambiguities in the original item, or inappropriateness of the item content in some cultures.

Bias has an effect on the comparability of measurement outcomes. Equivalence can be understood as the level of comparability of scores and hierarchically nested types of equivalence are often distinguished (Poortinga, 1989; Van de Vijver, 2011; Van de Vijver &

Leung, 2011). Construct inequivalence implies that the construct lacks shared meaning, which impedes any cross-cultural comparisons. This can, for example, occur when psychological constructs are culture-specific, such as culture-bound syndromes. Examples include Amok in Asian countries like Indonesia and Malaysia, which involves a brief period of persecutory thoughts and aggressive behaviour succeeding an insult, not afterwards recollected by the individual (Van de Vijver & Leung, 2011), as well as Brain Fog Syndrome in African countries, including South Africa, which is a condition experienced by high school or university students (often during periods of intensive reading prior to examinations) where they experience cognitive impairments in concentration, memory and comprehension, as well as somatic symptoms like head and neck pain and burning, blurred vision, and pain and tears in the eyes (Aina & Morakinyo, 2011; Peltzer, 2002; Peltzer, Cherian, & Cherian, 1998). An instrument shows structural equivalence if it measures the same construct(s) in all the groups involved. In operational terms this implies that the instrument must have the same factor structure in all groups. Functional equivalence, a specific type of structural equivalence, involves that the nomological network of the scale is identical across the different groups. In other words, the instrument has similar patterns of convergent and discriminant validity in the various groups. Instruments demonstrate metric equivalence if their measurement scales have the same units of measurement, but different origins. Interval- or ratio-level scale scores are assumed. Any source of bias, such as social desirability or stimulus familiarity, will move the scores of the different groups differentially, but the relative scores of the individuals within a group will not be influenced. Scalar or full score equivalence assumes identical interval or ratio scales and the same origin for all groups in the study. Only if the instrument shows scalar equivalence for the different groups, can cross-cultural comparisons on the mean scores be made (Poortinga, 1989; Van de Vijver, 2011; Van de Vijver & Leung, 2011).

In cross-cultural research, the aim is firstly to establish measurement equivalence for respondents from all groups involved in the study, so that future studies can compare the occurrence, determinants and consequences of the latent factor scores (Van de Schoot, Lugtig, & Hox, 2012). Measurement equivalence can be dealt with by a priori procedures (applied before data collection), as well as a posteriori procedures (applied after data collection). These procedures are complementary and the validity of the conclusions drawn from the cross-cultural study can be maximised by combining these approaches. A priori procedures include the qualitative evaluation of the relevance and representativeness of the content of the measurement instrument within each cultural group where it will be applied, formulation of the items to maximise the appropriateness of the items for all groups involved, careful planning of the sampling approach, thorough training of test administrators, and development of an extensive manual and administration protocol (Fischer & Fontaine, 2011; Van de Vijver & Leung, 2011). A posteriori procedures deal particularly with the assessment of structural equivalence and the detection of item bias. Methods to investigate structural invariance include multidimensional scaling, principal component analysis, exploratory factor analysis, and confirmatory factor analysis (cf., Fischer & Fontaine, 2011). To detect item bias one can use, for example, the delta plot method, standardisation, the Mantel-Haenszel method, item response theory, and logistic regression (cf., Sireci, 2011).

This study will address the limited extent of cross-cultural studies of EWB in existing literature by exploring the psychometric properties of widely used EWB scales cross-culturally using a posteriori procedures from a modern psychometric perspective.

Measurement from a Modern Psychometric Perspective

Various approaches exist to explore scales' psychometric attributes. Modern psychometric techniques enable us to develop and assess the functioning of measurement instruments with increasing rigour. Techniques that arose as superior alternatives to

traditional classical test theory approaches to scale validation include item response theory and Rasch modelling and structural equation modelling and confirmatory factor analysis. These techniques will be introduced briefly.

Item Response Theory (IRT) and Rasch Modelling

IRT has lately received growing attention and is often considered a modern and superior alternative to classical test theory (DeVellis, 2012). Classical test theory is based on the assumption that an observed score is the true score plus error. In classical test theory, this error term includes all other influences on the observed variable and there is no differentiation between different sources of error, for example, different time points, settings, or items (DeVellis, 2006, 2012). In contrast, IRT methods discriminate more finely among different sources of error, especially regarding features of individual items that may influence their performance (DeVellis, 2012).

IRT is a set of mathematical models that is utilised to describe the probabilistic relationship between a person's response to an item from a measurement scale and his or her level on the latent trait measured by the questionnaire (Reeve & Fayers, 2005). Different IRT models exist, based on the functional form that is specified for the relationship between the underlying latent trait and the item response probability. The Rasch model (Andrich, 1978; Rasch, 1960) is the simplest IRT model, which specifies only one parameter, namely item difficulty. The Rasch model's significant strength is that it evaluates whether the data correspond with the axioms of additive conjoint measurement. This provides a method whereby ordinal data, such as data from a Likert scale, are transformed into continuous, equal interval units (logits), which allows for the summation of the items' raw scores, where the summed raw score is a sufficient statistic (Da Rocha, Chachamovich, De Almeida Fleck, & Tennant, 2013; Hagquist, Bruce, & Gustavsson, 2009; Karabatsos, 2001). Rasch analysis is useful in the development of new measurement instruments and in the evaluation of the

psychometric properties of existing instruments where it can provide insight into the functioning of the response categories, unidimensionality, the scale's targeting, and differential item functioning across demographic groups (Myers, Wolfe, Feltz, & Penfield, 2006; Tennant & Conaghan, 2007). Given the positive characteristics of the Rasch model and the possibilities it creates in terms of data analysis and interpretation, the analytical procedure was applied in the present study.

Structural Equation Modelling (SEM) and Confirmatory Factor Analysis (CFA)

Structural equation modelling (SEM) is a confirmatory approach to statistical analysis where the structural theory of some phenomenon is tested (Byrne, 2012). The aim of SEM is to test hypotheses regarding the mean, variance, and covariance structures of a set of variables that are strongly rooted in theoretically or empirically derived expectations (Bowen & Guo, 2012). SEM is particularly suitable for assessing hypotheses regarding “latent constructs”, which are abstract phenomena typically found in the social sciences, such as emotions, attitudes, cognitions, behaviour patterns, or experiences that are not directly observable or measurable using a single item (Bowen & Guo, 2012). When applied to latent variables, SEM can involve a measurement model that describes how a set of observed variables represents the latent construct that is measured, and a structural model that postulates how the latent variables are related (Bowen & Guo, 2012; Byrne, 2012). Confirmatory factor analysis (CFA) within a SEM framework pertains specifically to the measurement model.

CFA provides an alternative to traditionally used exploratory factor analysis (EFA). Both CFA and EFA aim to establish which sets of observed variables define latent constructs, as is evident in the fact that they share common variance-covariance features (Schumacker & Lomax, 2010). In EFA, the aim is to explore which sets of observed variables group together to form factors. One would typically explore various alternative models in order to find a

model that fits the data well and then seek for theoretical support for the obtained model (Schumacker & Lomax, 2010). To the contrary, the goal of CFA is to statistically test a hypothesised model based on theory or previous empirical findings in order to confirm whether the sample data endorse that model (Schumacker & Lomax, 2010). EFA is typically appropriate for use in the early phases of developing a scale to obtain preliminary insight into an instrument's factor structure and the functioning of the items. CFA is suitable in the later stages of scale development to determine how well the items capture the hypothesised structure of the latent variables and to determine whether the measure is equivalent across different groups or time points (Bowen & Guo, 2012). Recently, exploratory structural equation modelling (ESEM) has been introduced, which integrates the features of SEM-based CFA and EFA (Marsh, Morin, Parker, & Kaur, 2014).

When used at the appropriate point during scale development, SEM-based CFA has some advantages over traditional EFA approaches. For example, when using CFA it is possible to specify that certain items' error terms are correlated, which is not possible with EFA (DeVellis, 2012). In addition, whereas EFA requires the analyst to either allow all the factors to correlate with one another or to all be independent, CFA provides options where some factors can be specified to correlate while others are independent, as postulated by the underlying theory (DeVellis, 2012). Also, CFA provides appropriate weights (the factor loadings) for all items when composite scores are calculated for use in subsequent analyses, which yield more accurate factor scores than scores obtained when item scores are simply averaged (Bowen & Guo, 2012). Yet another advantage is that SEM-based CFA offers opportunities to test more complex models that better represent the nuances of sophisticated theories, such as higher-order models (Byrne, 2012), bifactor models (Reise, 2012; Reise, Moore, & Haviland, 2010), and models where method effects are accounted for (Brown, 2015). In the light of the fact that the theories and their corresponding instruments that were

investigated in the present study are well-researched in other contexts, or using other statistical methods, it was deemed appropriate to apply confirmatory factor analysis in the present study. In addition, it was hypothesised that more complex models that allow for the incorporation of a bifactor approach or a negative-worded method effect, would suit the data best. In order to apply these models, a CFA approach is needed.

The Present Study

The overall aim of the present study was to further the process of understanding and measuring EWB across countries and cultures by exploring the cross-cultural psychometric properties of scales that operationalise three prominent theories that fall within the domain of eudaimonic well-being using modern psychometric approaches. This was done by preparing three manuscripts for target journals that will be presented in accordance with the guidelines provided for the authors of the respective journals. The first study (Section 2, Manuscript 1) applied Rasch modelling to explore the psychometric properties of the Meaning in Life Questionnaire (Steger et al., 2006) among adults from South Africa, Australia, and New Zealand. This manuscript was published by Health and Quality of Life Outcomes (Schutte, Wissing, Ellis, Jose, & Vella-Brodrick, 2016). The second study (Section 3, Manuscript 2) investigated the appropriateness of a bifactor model for three language versions of the Mental Health Continuum Short Form (Keyes, 2006, 2009; Keyes et al., 2008) with analyses done on single- and multi-group levels. This manuscript was submitted to Social Indicators Research. The third study (Section 4, Manuscript 3) explored the psychometric properties of three language versions of the Basic Psychological Needs Scale (Gagné, 2003). This manuscript was submitted to the Journal of Personality Assessment. The thesis concludes with a final section (Section 5) presenting conclusions, implications, and recommendations based on the findings from this study.

This study contributes by: (i) Exploring the psychometric properties of measurement instruments of eudaimonic well-being across different countries and cultures, and thereby providing insight into validity and usefulness of the scales in these contexts and cross-culturally; (ii) Providing new insights into the functioning of commonly used measures of eudaimonic well-being through the application of modern psychometric techniques; and (iii) Contributing towards the further development of theoretical understandings of eudaimonic well-being.

Section 2: Manuscript 1

**Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa,
Australia, and New Zealand**

Published in

Health and Quality of Life Outcomes

2.1. Guidelines for authors

Health and Quality of Life Outcomes

These guidelines were retrieved from:

<https://hqlo.biomedcentral.com/submission-guidelines>

“Before you submit your manuscript, please ensure that you have satisfied all formatting requirements and have all the required supporting information as listed below. If you have any questions about the submission process or preparation of your manuscript, please contact info@biomedcentral.com.

Manuscripts cannot be submitted to more than one journal at a time and cannot be submitted by anyone other than one of the authors. The submitting author takes responsibility for the submission on behalf of all authors. All authors must consent to the submission.

1. Correctly format your manuscript

Health and Quality of Life Outcomes publishes the following article types. Please click on the relevant article type below for instructions on how to format your manuscript.

- [Research article](#)
- [Commentary](#)
- [Letter to the Editor](#)
- [Review](#)
- [Short report](#)
- [Study protocol](#)

Figures

Figures must be uploaded as separate files and not embedded in your manuscript.

See [Preparing figures](#) for more information.

Tables

Smaller tables that are considered integral to the manuscript can be pasted at the end of the manuscript file in A4 portrait or landscape format. Larger tables can be uploaded separately as additional files.

See [Preparing tables](#) for more information.

Additional files

You may provide datasets, tables, movies, or other information as additional files. These will be published along with the article. Do not include files such as patient consent forms,

certificates of language editing, or revised versions of the main manuscript document with tracked changes.

The maximum file size for additional files is 20MB each, and files will be virus-scanned on submission.

See [Preparing additional files](#) for more information.

2. Prepare supporting information

Please make sure you have the following information available before you attempt to submit your manuscript:

Author information

Full names and email addresses of all co-authors on your manuscript.

Cover letter

A cover letter that includes the following information, as well as any additional information requested in the instructions for your specific article type (see main manuscript section above):

- An explanation of why your manuscript should be published in *Health and Quality of Life Outcomes*
- An explanation of any issues relating to journal policies
- A declaration of any potential competing interests
- Confirmation that all authors have approved the manuscript for submission
- Confirmation that the content of the manuscript has not been published, or submitted for publication elsewhere (see our [duplicate publication policy](#))
- If you are submitting a manuscript to a particular special issue, please refer to its specific name in your covering letter

You may suggest potential peer reviewers for your manuscript. If you wish to do so, please provide institutional email addresses where possible, or information which will help the Editor to verify the identity of the reviewer (for example an [ORCID](#) or [Scopus ID](#)). Intentionally falsifying information, for example, suggesting reviewers with a false name or email address, will result in rejection of your manuscript and may lead to further investigation in line with our misconduct policy. See our [editorial policies](#) for guidance on suggesting peer reviewers.

Excluding peer reviewers

During submission you may enter details of anyone who you would prefer not to review your manuscript.

The following information must be included under the Declarations section of your manuscript. You will be asked about this information during submission of your manuscript.

Declarations

Ethics approval and consent to participate

Manuscripts reporting studies involving human participants, human data or human tissue must include a statement on ethics approval and consent. Studies involving animals must include a statement on ethics approval.

See our [editorial policies](#) for more information.

Consent to publish

All manuscripts containing any individual person's data in any form require consent to publish from that person, or in the case of children, their parent or legal guardian.

See our [editorial policies](#) for more information.

Availability of data and materials

For all journals, you must include an 'Availability of data and materials' section within the 'Declarations' of your manuscript detailing where the data supporting your findings can be found. For some journals, deposition of your data is mandatory (see information on individual article types above). For journals where data deposition is not mandatory, BioMed Central strongly encourages all datasets on which the conclusions of the manuscript rely be either deposited in publicly available repositories or presented in the main paper or additional supporting files. If you do not wish to share your data, please state that data will not be shared, and state the reason.

See our [editorial policies](#) and [preparing your manuscript](#) for more information.

Competing interests

All financial and non-financial competing interests must be declared in this section. You will be asked about this during submission of your manuscript.

See our [editorial policies](#) for more information on competing interests.

Funding

All sources of funding for the research reported should be declared.

3. Agree to conditions of submission, BioMed Central's copyright and license agreement and article-processing charge (APC)

Copyright and license agreement

During submission, you will need to accept and confirm the following conditions of submission:

- All authors of the manuscript have read and agreed to its content and are accountable for all aspects of the accuracy and integrity of the manuscript in accordance with [ICMJE criteria](#)
- That the manuscript is original, has not already been published in a journal, and is not currently under consideration by another journal
- That you agree to the terms of the BioMed Central Copyright and License Agreement and Open data policy, which we strongly recommend you read. For authors who are prevented from being copyright holders (for instance where Crown Copyright applies or researchers are US government employees) BioMed Central can accommodate nonstandard copyright lines. If this applies to you, please contact us and provide details of your situation
- For more information see our [Copyright and license policy](#)

Article-processing charge

An article-processing charge (APC) applies for each article accepted for publication in *Health and Quality of Life Outcomes*. For details of the journal's APC see [About Health and Quality of Life Outcomes](#) and for answers to some frequently-asked questions please see our [article-processing charge page](#).

[Submit your manuscript in Editorial Manager](#)

2.2. Manuscript

**Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa,
Australia, and New Zealand**

Lusilda Schutte*^{1a}, Marié P. Wissing^{1b}, Suria M. Ellis^{1c}, Paul E. Jose², & Dianne A. Vella-Brodrick³

*Corresponding author

^{1a}Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa,
E-mail: Lusilda.Schutte@gmail.com

^{1b} Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa,
E-mail: Marie.Wissing@nwu.ac.za

^{1c}Statistical Consultation Services, North-West University, South Africa, E-mail:
Suria.Ellis@nwu.ac.za

²School of Psychology, Victoria University of Wellington, New Zealand, Email:
Paul.Jose@vuw.ac.nz

³Melbourne Graduate School of Education, University of Melbourne, Australia, Email:
dianne.vella-brodrick@unimelb.edu.au

Abstract

Background: Meaning in life is a key indicator of subjective well-being and quality of life. Further developments in understanding and enhancing the construct will depend inter alia on the sound measurement thereof. This study is at the forefront of applying modern psychometric techniques to the Meaning in Life Questionnaire, a scale widely used to assess meaning in life.

Method: The Rasch rating scale model was applied to the Presence and Search subscales of the Meaning in Life Questionnaire using a sample of 601 adults from South Africa, Australia, and New Zealand.

Results: The Presence subscale was insensitive at high levels of presence of meaning while the majority of the respondents fell in that range. Removal of item 9 (“My life has no clear purpose”) and collapsing the response categories indicative of low and medium levels of the latent construct significantly improved the subscale’s targeting and fit to the Rasch model, resulting in a subscale that exhibited differential item functioning on items 1 (“I understand my life’s meaning”), 4 (“My life has a clear sense of purpose”), and 5 (“I have a good sense of what makes my life meaningful”) for country, but none for gender, age group, or education level. The Search subscale yielded disordered category threshold calibrations, but after collapsing some of the response categories representing low and medium levels of the target construct, a subscale that demonstrated good fit to the Rasch model, good targeting, and no differential item functioning resulted.

Conclusions: In terms of this particular scale, adaptation of the rating scale and removal of item 9 is recommended. Country-level parameter estimates may be needed for items that exhibited differential item functioning. The study also has significant implications for the theory, measurement, and practice of meaning in and quality of life in general.

Reasons for and the far-reaching implications of the insensitivity of the Presence subscale for

high levels of presence of meaning on, for example, the correlation between meaning in life and indicators of health are contemplated. Further investigation of the construct's nature and measurement, especially at high levels, is indicated.

Keywords: Rasch modelling, Meaning in Life Questionnaire, eudaimonic well-being, psychometric properties, modern psychometrics, item response theory, positive psychology, quality of life

Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa,
Australia, and New Zealand

Background

Quality of life involves an evaluative judgement of an individual's physical, cognitive, emotional, and social functioning and can be based on subjective (self-report) and/or objective (independent sources of information) indicators [1, 2]. Although quality of life research traditionally focused on situations and factors that undermine or endanger quality of life, recent research has increasingly stressed the importance of incorporating positive constructs, such as subjective well-being, positive emotions, and character virtues and strengths in the conceptualisation and study of quality of life [2, 3]. One of the key constructs that is widely considered an integral part of a life well-lived and quality of life is meaning in life [4-7]. A myriad of studies have explored the relationship between meaning in life and mental well-being, as well as psychopathology [8]. Also, the association between meaning in life and health-related quality of life has been established in multiple studies [9].

In order to study meaning in life and its quality of life concomitants, the construct has to be conceptualised theoretically. Different models are used in the literature to conceptualise this complex phenomenon, for example those of Wong [10], Schnell [11], and Steger [12]. Steger's model differentiates between presence of meaning, which involves "the extent to which people comprehend, make sense of, or see significance in their lives, accompanied by the degree to which they perceive themselves to have a purpose, mission, or overarching aim in life" [12], and search for meaning, which refers to "the strength, intensity, and activity of people's desire and efforts to establish and/or augment their understanding of the meaning, significance, and purpose of their lives" [13].

Theoretically and empirically sound measurement instruments that assess meaning in life are crucial for the rigorous study of the construct, to understand its associations with

psychological well-being and psychopathology, and to assess the impact of interventions targeting meaning in life. Various models of meaning have been operationalized in self-report questionnaires (see [14] for a systematic review of these measures). One such scale that is widely used and recognized for its outstanding psychometric properties [14] is the Meaning in Life Questionnaire (MLQ) [15], which operationalizes Steger's [12] model of meaning in life. Steger et al. [15] showed that the scale, which consists of two subscales corresponding to the theory, namely Presence of Meaning (MLQ-P) and Search for Meaning (MLQ-S), demonstrated sufficient internal consistency and test-retest reliability, as well as structural, convergent, and discriminant validity in three American student samples.

Since the initial development study of the MLQ [15], which utilised only data from American student samples, good psychometric properties of the scale have been shown in a number of other contexts, cultures, and translations. For example, validity and reliability were shown for the English version of the scale among a web-based survey of adults [16], an American sample of people diagnosed with serious mental illnesses in an inpatient setting [17], and in a multi-cultural South African student setting [18]; for the Japanese translation of the scale among a Japanese student sample [19]; for the Spanish translation of the scale among a Spanish student sample [20]; and for the Turkish version of the scale among a combined college student and adult community sample [21].

Even though the MLQ is widely appraised to possess good psychometric properties [14] and the measure has been found to function well across age groups [13] and cross-culturally [13, 21, 22], the scale has, as far as we are aware of, never been evaluated from an item response theory (IRT) perspective. IRT provides a modern and reputedly superior alternative to classical test theory, as it discriminates more finely among different sources of error, especially regarding features of individual items that may influence their performance [23]. The family of IRT models share the assumption that the probability of a respondent

endorsing any particular item is considered to be a function of the respondent's level on the underlying latent variable that is measured and the characteristics of the item [24].

The Rasch model, specifying only one parameter to characterize each item (item difficulty), is the simplest IRT model and was developed by the Danish mathematician, Georg Rasch [25, 26]. Unlike in other IRT models and classical test theory techniques where the intent is to find a model that best fits the data, the Rasch model requires the data to fit the model in order to yield objective measurement [27]. The Rasch model postulates that useful measurement involves a unidimensional construct increasing or decreasing monotonically along an interval scale [28]. Rasch modelling provides a method to transform ordinal data (e.g. data from Likert-type items) into continuous, equal interval units (logits), which allows for the summation of the items' raw scores, where the summed raw score is a sufficient statistic [29, 30]. Rasch analysis can be used in scale development, for example by reviewing the functioning of the response categories, the unidimensionality of the scale, and the targeting of the measure [31]. Moreover, Rasch modelling can be used to investigate differential item functioning (i.e., when different demographic groups responded differentially to an item despite equal levels of the latent construct), thus enhancing the assessment of item-level cross-cultural invariance of measurement scales [32].

The Present Study

In the present study, the Meaning in Life Questionnaire [15] was examined against the assumptions of the Rasch model. This is the first known study where the scale is analysed using an item response theory (in particular, Rasch modelling) approach. By applying the Rasch model, we explored the unidimensionality of each subscale, the functionality of the response categories, and how well the sample was targeted by the scale. We also examined differential item functioning (DIF) of the scale for a range of demographic variables.

Method

Participants

The sample ($N = 601$) consisted of about equal sized groups of adults from South Africa, New Zealand, and Australia, who all completed the original English version of the MLQ as part of a battery of scales used in the international Eudaimonic and Hedonic Happiness Investigation (EHHI) project [33]. Participants were selected to be fluent in English, have at least secondary education, and be between 30 and 60 years of age. The aim was to factorially cross gender, age (three age groups of 30-39 years, 40-49 years, and 50-60 years), and education level (secondary and tertiary education). The socio-demographic profile of the sample is summarised in Table 1.

Measures

Socio-demographic questionnaire. Demographic information of each participant, including country of residence, gender, age group, and education level, was obtained.

Meaning in Life Questionnaire (MLQ) [15]. The MLQ comprises two subscales that was developed to be relatively independent: Presence of Meaning (MLQ-P) and Search for Meaning (MLQ-S) [15]. Responses to 10 statements are provided on a rating scale with response options 1 = *Absolutely Untrue*, 2 = *Mostly Untrue*, 3 = *Somewhat Untrue*, 4 = *Can't Say True or False*, 5 = *Somewhat True*, 6 = *Mostly True*, and 7 = *Absolutely True*. In the original validation study among American students, the scale exhibited good internal consistency and test-retest reliability, as well as structural, convergent, and discriminant validity, with the Cronbach's alpha values of the Presence subscale varying between 0.82 and 0.86 and for the Search subscale between 0.86 and 0.87 [15]. Good internal consistency reliability was found in South African student [18], New Zealand adult [34], and web-based Australian samples [35], with alpha-values of .85, .90, and .88, respectively, for the MLQ-P, and .94, .91, and .92, respectively, for the MLQ-S. .

Procedure and Ethical Considerations

A mixed-methods cross-sectional survey design was used, where participants responded to open-ended questions related to happiness, meaning in life, and goals, and completed a battery of quantitative measurement scales. For the current investigation, only responses to socio-demographic questions and the MLQ were used. In order to avoid the potential complications of missing values and imputation techniques in Rasch analyses, respondents who generated missing values on the MLQ were removed from the sample. This involved 15 participants from South Africa, whose removal was justified by the fact that the original South African sample was larger than the samples from Australia and New Zealand. The sample from New Zealand contained no missing responses, and for the Australian sample six respondents were removed. Ethical approval was obtained from the respective regulatory ethics committees in each country. Participants were recruited by research leaders within each country using poster and newspaper advertisements and the snowball-method. Participants were provided with information on the study prior to voluntary participation.

Data Analysis

Data were analysed using the Rasch rating scale model [25], which assumes that the distances between the thresholds of polytomous items (i.e., the probabilistic midpoints between adjacent response categories) are equal across all items. The Winsteps® 3.81 software [36] was used for all analyses, except for the graphical presentation of the person-item threshold distributions (Figure 2), which was obtained from RUMM2030™ [37]. The MLQ-P and MLQ-S were evaluated separately, since the scale was designed to yield two relatively independent subscales [15]. Since no single aspect of Rasch analysis is definitive in identifying the optimal data-model relationship, multiple tests and graphical representations should be used to examine the characteristics of the items and persons [30]. The following interrelated facets of Rasch analysis should be considered simultaneously to inform decisions.

Person and item separation and reliability. Person separation and reliability indices indicate how well one can discern persons along the measured variable [28] and values larger than 2 and 0.8, respectively, imply that the items are sensitive enough to differentiate two levels of persons according to their level of intensity on the construct (high and low scorers) [38]. Item separation and reliability indices are indicative of the capacity of the instrument to define a unique hierarchy of items along the measured construct [28] and values larger than 3 and 0.9, respectively, suggest that the sample is large enough to confirm the item challenge order (on three levels of item challenge) [38].

Unidimensionality and local independence. According to the Rasch model, useful measurement is obtained when a unidimensional construct is measured by locally independent items [30]. In terms of unidimensionality, item infit or outfit mean square statistics smaller than 0.6 can be indicative of overfit, and values larger than 1.4 of underfit when the rating scale model is used [28]. The point-biserial correlation of an item indicates whether higher scores on the item correspond with higher levels of the underlying construct and positive values are expected [38]. In addition, lack of unidimensionality may exist when the eigenvalue of the first contrast in a Rasch principal components analysis of the residuals (PCA-R) (i.e., the first component after the Rasch component has been removed) is larger than 2.0, and when the variance explained by the Rasch component is small (e.g., < 40%) [38]. Correlations between the residuals of item pairs of around 0.7 are indicative of high local dependence, while correlations around 0.4 are considered to be low [38].

Response category functioning. Rasch analysis enables the researcher to investigate how the respondents used the rating scale so that scale developers can decide on the optimal number and combination of rating scale categories [39, 31]. This task can be accomplished by examining how the data fit the Rasch model after response categories were collapsed. Bond and Fox [28] provided guidelines in this regard, including that the collapse should make

intuitive sense and that the ideal is to create a uniform frequency distribution over the categories with each category containing at least 10 observations. Also, the average measures of the categories and the category threshold estimates should increase monotonically, with the category threshold estimates having steep gradients (at least 1.4 logits, but no more than 5.0 logits) to ensure that each category represents a distinct portion of the latent variable – this can also be investigated graphically by looking at the category probability curves. Lastly, the infit and outfit mean square statistics of each response category should be less than 2.0.

Targeting. Rasch analysis can be used to detect gaps in the continuum of the measured construct by identifying poor targeted items or persons, such as items for which there is an insufficient number of persons with an intensity level comparable to the item challenge^a, or persons for which there is an insufficient number of items with a challenge level comparable to the person's intensity [40]. This goal can be attained by examining the person-item threshold distributions generated by RUMM2030TM, which offers a visual comparison of the distribution of the person intensity levels (top part of the graph) and the item challenge levels (bottom part of the graph) along the latent trait continuum, with the information provided by the items also mapped onto the person distribution.

Differential item functioning. Rasch analysis can assist in identifying differential item functioning (DIF), which occurs when different groups of people within the sample responded in a different way to an item despite equal levels of the construct that was measured. In this study, uniform DIF [31] was investigated for country, gender, age group, and education level. The degree of DIF was assessed by comparing *p*-values from the polytomous version of the Mantel-Haenszel statistic [41, 42] against a Bonferroni-corrected 5% significance level, as well as the DIF Contrast, which is indicative of moderate to large DIF when it is larger than or equal to 0.64 [38].

Results

Results for the Presence Subscale

MLQ-P. Although the MLQ-P yielded person and item separation and reliability indices that were in line with the guidelines and the results from the PCA-R suggested sufficient unidimensionality and local independence of the items (see Table 2), item 9 (“My life has no clear purpose”) showed misfit based on its infit and outfit mean square statistics (see Table 3). Also, response category 1 (*Absolutely untrue*) exhibited a low frequency and misfit based on its outfit mean square statistic (see Table 4). Although the average measures and threshold calibrations increased monotonically as the categories increased, the threshold calibrations were close to each other, indicating that categories 2 (*Mostly untrue*), 3 (*Somewhat untrue*), and 4 (*Can't say true or false*) were the most likely to be endorsed on only a small portion of the latent construct (see Table 4 and Figure 1). From the person-item threshold distribution (Figure 2) it was clear that the person intensity was in general higher than the item challenge, indicating that the scale exhibited poor targeting for persons with high levels of the latent construct. The MLQ-P showed DIF for country on items 1 (“I understand my life’s meaning”), 4 (“My life has a clear sense of purpose”), and 9 (“My life has no clear purpose”), as depicted in Table 6. There was no significant DIF for gender, age group, or education level.

In an attempt to remedy the problems highlighted for MLQ-P, all possible combinations of response category collapses were explored, but none of the collapses resolved the problems with item 9. Therefore the next step was to remove item 9, resulting in a 4-item scale (hereafter labelled MLQ-P-4).

Results for the MLQ-P-4. The person and item separation and reliability indices improved significantly after item 9 was dropped from the scale (see Table 2). The PCA-R yielded results that confirmed satisfactory unidimensionality and local independence (Table 2) and all point-biserial correlations (values ranged between .79 and .85) and item infit and

outfit mean square statistics (Table 3) pointed towards good fit. Although none of the response categories showed misfit based on their infit and outfit mean square statistics, the category probability curve (not shown) and threshold calibrations (see Table 4) still revealed that response categories 2 (*Mostly untrue*), 3 (*Somewhat untrue*), and 4 (*Can't say true or false*) were the most likely to be endorsed over only a small portion of the latent variable, suggesting redundant response categories. Category 1 (*Absolutely untrue*) also still generated a low frequency. The person-item threshold distribution (not displayed) suggested even worse targeting for persons with high levels of the latent construct when compared to the full MLQ-P. The MLQ-P-4 showed DIF for country on items 1 (“I understand my life’s meaning”), 4 (“My life has a clear sense of purpose”), and 5 (“I have a good sense of what makes my life meaningful”) as depicted in Table 6. No significant DIF was found for gender, age group, or education level. In order to address the redundancy of the response categories, the next step was to explore all possible combinations of category collapses.

Results for the MLQ-P-4, response categories collapsed. Based on Rasch model diagnostics, two combinations of category collapses produced superior performance: One where category 1 (*Absolutely untrue*) was collapsed with category 2 (*Mostly untrue*), and category 3 (*Somewhat untrue*) with category 4 (*Can't say true or false*) – hereafter labelled MLQ-P-4 1122345; and one where categories 2, 3, and 4 were collapsed – hereafter labelled MLQ-P-4 1222345. For both, the separation and reliability indices and the results from the PCA-R were in line with the results before collapsing categories (see Table 2). Due to space limitations only the results of the MLQ-P-4 1122345 are displayed in Tables 3, 4, and 5, and Figures 1 and 2. Results for the MLQ-P-4 1222345 were similar, unless indicated in the text. The item infit and outfit mean square statistics (Table 3) and point-biserial correlations (values ranged between .83 and .88 for MLQ-P-4 1122345) indicated that all items fitted the Rasch model well and the response categories showed good fit, with threshold calibrations

increasing monotonically and being sufficiently distanced from each other (see Table 4 and Figure 1). For the MLQ-P-4 1222345, the frequency of category 1 (*Absolutely untrue*) was low, while the MLQ-P-4 1122345 yielded a larger frequency for category 1. Collapsing the categories improved the targeting of the scale considerably (see Figure 2). Both the MLQ-P-4 1122345 and the MLQ-P-4 1222345 showed DIF for country on items 1 (“I understand my life’s meaning”) and 5 (“I have a good sense of what makes my life meaningful”) as shown in Table 6. No significant DIF was found for gender, age group, or education level.

Results for the Search Subscale

MLQ-S. The separation and reliability indices for the MLQ-S were in line with the guidelines, and the results from the PCA-R pointed to sufficient unidimensionality and local independence (see Table 2). Considering the item infit and outfit mean square statistics (Table 3) and the point-biserial correlations (values ranged between .80 and .85), all items fitted the Rasch model well. Although the infit and outfit mean square statistics of the response categories adhered to the guidelines, the threshold calibrations of categories 2 (*Mostly untrue*), 3 (*Somewhat untrue*), and 4 (*Can’t say true or false*) were disordered, pointing towards problematic use of the rating scale (see Table 5), which is also evident in the category probability curve (Figure 1). The person-item threshold distribution (Figure 2) portrayed that the average item challenge was slightly lower than the average person intensity, but from the information curve it is clear that there was substantial information available for the majority of respondents. There was no significant DIF for country, gender, age group, or education level. In an attempt to remedy the disordered threshold calibrations, all possible combinations of response category collapses were explored.

Results for the MLQ-S, response categories collapsed. Based on Rasch model diagnostics, two combinations of category collapses stood out as superior: One where category 2 (*Mostly untrue*) was collapsed with category 3 (*Somewhat untrue*), as well as

category 4 (*Can't say true or false*) with category 5 (*Somewhat true*) – hereafter labelled MLQ-S 1223345; and one where categories 3 (*Somewhat untrue*) and 4 (*Can't say true or false*) were collapsed – hereafter labelled MLQ-S 1233456. Although the item separation dropped slightly after collapsing the categories, the person separation increased and the person and item reliability indices remained unchanged (see Table 2). Results of the PCA-R suggested sufficient unidimensionality and local independence (see Table 2). Based on the item infit and outfit mean square statistics (Table 3) and the point-biserial correlations (values ranged between .82 and .89 for MLQ-S 1223345 and between .82 and .88 for MLQ-S 1233456), all items manifested adequate fit. The problem of disordered category thresholds has been resolved, the distances between the threshold calibrations have improved, and the infit and outfit mean square statistics of the response categories pointed towards satisfactory fit (see Table 5 and Figure 1). The person-item threshold distribution (Figure 2) suggested improved targeting for MLQ-S 1233456, but for MLQ-S 1223345 (not shown) the average item challenge level was found to be more than the average person intensity level, which suggests less optimal targeting. There was no significant DIF for country, gender, age group, or education level.

Discussion

Rasch analyses were conducted to investigate the psychometric properties of the Presence of and Search for Meaning subscales of the Meaning in Life Questionnaire, a scale which measures meaning in life – a key indicator of quality of life, from a Rasch modelling point of view. Data from three different countries were used. For MLQ-P, removal of item 9 (“My life has no clear purpose”) and collapsing response categories indicative of low and medium levels of the latent construct significantly improved the fit to the Rasch model and the targeting of the scale, resulting in a scale that exhibited DIF on items 1 (“I understand my life’s meaning”), 4 (“My life has a clear sense of purpose”), and 5 (“I have a good sense of

what makes my life meaningful”) for country, but no DIF for gender, age group, or education level. The MLQ-S yielded disordered category threshold calibrations, but after collapsing some of the response categories representing low and medium levels of the target construct, a scale that demonstrated good fit to the Rasch model, good targeting, and no DIF resulted. Several specific aspects of the results will now be discussed.

Reversed Item

The first significant finding that warrants discussion is the poor performance of item 9 (“My life has no clear purpose”), the only reversed item in the MLQ-P scale. In a review on misresponse to reversed and negated items, Weijters and Baumgartner [43] advocated for the inclusion of reversed items in measurement scales as it can provide many benefits (e.g., control acquiescence, disrupt careless responding, and promote a broader coverage of the content domain), but stressed that it should be done with caution. A reversed item that is merely the negation of an item in the main direction (in point of fact, item 9 is basically the negation of item 4, “My life has a clear sense of purpose”), does not hold the benefit of broadening the content domain tapped by the instrument, and has the disadvantages inherent in negated items (e.g., accurately assessing level of agreement with statements that contain negation requires considerable cognitive strain) and reversed items (e.g., cross-cultural differences in response styles such as acquiescence). We therefore follow the guidance provided by Weijters and Baumgartner [43], who advised against the use of negated reversals, and consequently we recommend the removal of item 9, which will result in a 4-item Presence of Meaning subscale.

Steger et al. [15] stated that the reversed item was retained in the hope of discouraging automatic response sets. It is our view that this concern is to a large extent already handled by the mixed administration of the Presence and Search subscales. If item 9 is removed, however, the remaining items 4 to 6 will tap presence of meaning and the last three items will

tap search for meaning. To guard against careless responding and response sets, we recommend shuffling the last six items (item 9 excluded) so that the respondent does not respond to three items from the same subscale in sequence.

Number of Response Categories

For both subscales, the response categories indicative of low and medium levels of the latent construct appeared to be redundant and for the search subscale, the category thresholds were disordered. These findings suggest that the respondents were unable to distinguish reliably among the categories, and consequently fewer categories should yield more consistent, reliable scores. Weijters, Cabooter, and Schillewaert [44] suggested that seven response categories may be acceptable for populations who are expected to have high cognitive abilities, verbal skills, or questionnaire experience, such as college students, but that a 5-point scale may be more appropriate for the general population. For future use, we recommend either a 6-point rating scale where the midpoint category 4 = *Can't say true or false* is dropped, or a 5-point scale with categories 1 = *Absolutely untrue*, 2 = *Untrue*, 3 = *Unsure*, 4 = *True*, 5 = *Absolutely true* (the issue of whether to include a midpoint category is much debated in the literature [44, 45]).

Targeting

In the present study, the average level of meaning in life captured by the items was substantially lower than the average level of presence of meaning manifested by persons who completed the scale, suggesting poor targeting. In fact, the scale provided little information for respondents with high levels of presence of meaning while at the same time most of the respondents fell within that range. This could have significant practical implications. Correlations in correlational studies will be largely influenced by the minority of people exhibiting lower levels of presence of meaning as reflected by lower scores on the MLQ-P, while nuances of presence of meaning at the higher end of the continuum will not be captured

well. This can, for example, influence outcomes of studies where the associations between meaning in life and indicators of health and quality of life are studied significantly. In addition, in experimental studies or studies where intervention programs are evaluated, the MLQ-P would probably not detect changes in meaning in life of people on the higher end of the continuum, which involves the majority of people, as the scale is not sensitive to changes at the higher end of the continuum.

Different explanations can be given for the findings regarding the targeting of the MLQ-P. One apparently obvious explanation is that there are not enough items or response options to capture high levels of the presence of meaning continuum and such items or response options should be added. However, given that the questionnaire already allows respondents to rate statements like “I understand my life’s meaning” to be “*absolutely true*”, it is not clear what kind of items or response options can be added to capture even higher levels of presence of meaning in life.

Another possible explanation pertains to the nature of presence of meaning as a construct and its distribution in the general population. The fact that the majority of the respondents endorsed high levels of presence of meaning according to their scores on the MLQ-P could simply tell us that most people indeed experience their lives as basically meaningful: Most respondents’ level of presence of meaning were higher than the levels where the scale had optimal information, merely because there is not much variability at the upper end of the underlying construct continuum. Such an explanation speaks to the findings of Heintzelman and King [46], who conducted a review of research on meaning in life from epidemiological data and studies using the MLQ-P [15] and the Purpose in Life Test [47]. They found that diverse samples rated themselves significantly above the midpoint on self-report measures of meaning in life and concluded that most people experience their lives as “pretty meaningful”. This line of thought can be linked to psychopathology literature where

“quasi-traits” are distinguished. Reise and Waller [48] defines a quasi-trait as “a unipolar construct in which one end of the scale represents severity and the other pole represents its absence (depression versus not depressed)” which “is in contrast to a bipolar construct, where both ends of the scale represent meaningful variation (depression versus happiness)”. In psychopathology research, the existence of quasi-traits with their associated peaked information curves (with the peaks in the range representing severe levels of the trait) has been found in many item response theory applications and often led researchers to conclude that items needed to be added or adapted to provide information at low (less severe) levels of the trait continuum [48]. According to Reise and Waller [48] this reasoning is problematic when working with quasi-traits: If the underlying latent construct is a quasi-trait, such attempts may be futile – it will be difficult (if not impossible) to formulate items that yield information across the continuum of the trait. Similarly we can ask whether it would be possible to develop items designed to capture even higher levels of presence of meaning, or whether we should conclude that the variation of presence of meaning is limited at the higher end of the continuum, although the majority of people attain such high levels.

If we settle with the conclusion that the majority of the population attained maximum levels of presence of meaning, we will inevitably have to re-evaluate the usefulness of, for example, interventions that aim to enhance meaning in life in the general population (most of whom have attained high levels of meaning in life). The question would be what the (large) portion of people with high levels of meaning would gain from interventions that intend to enhance meaning. Accepting that the majority of the population have already attained levels of presence of meaning that do not allow for much improvement may pose further questions. For example, could it be possible that icons of eudaimonic living, such as Mahatma Gandhi, Mother Theresa, or Nelson Mandela, who sacrificed their lives for a greater cause, have experienced levels of meaning in life similar to the majority of people? Or should we rather

conclude that the nuances of presence of meaning at higher levels are just not captured by the current conceptualisation and operationalization of the construct?

Another way to explain the poor targeting of the MLQ-P may be that the subscale applies a rather narrow understanding of meaning in life, with all items paraphrasing the notion of having found a sense of meaning or purpose in life. By repeating the same content using slightly different syntax, the scale actually operates in a similar way to a one or two-item measure, which could contribute to the inability of the scale to differentiate well at the higher end of the continuum. Alternative measures that capture a broader sense of meaning in life, such as the Sources of Meaning and Meaning in Life Questionnaire (SoMe) that operationalises meaningfulness through coherence, significance, direction, and belonging [11], may display better sensitivity.

In addition, one can argue that participants' presence of meaning in life was not really as high as they indicated it to be – social desirability may have augmented their scores artificially. However, presence of meaning in life has been found to be unrelated to scores on measures of social desirability in several studies [49, 15] and, as argued by Heintzelman and King [46], high presence of meaning scores have been found consistently among diverse samples, including anonymous samples where social desirability may not have been a big concern. The high scores could have also been due to a generalisation effect – when asked to respond to items that concern global meaning in life, people may not be sure what meaning actually refers to. They may have a broad understanding of meaning and therefore think that they generally experience meaning. However, if the constituents of meaning are spelled out, they might realise that they don't have as much meaning as they initially thought.

One may also reason that the lack of sensitivity to varying nuances of meaning in life at the higher end of the continuum relates to the fact that the scale relies on self-report and alternative avenues to capture meaning in life should be explored. This approach may be

problematic because meaning in life is, at its heart, a subjective experience. Several studies have argued that self-report is the best way to capture meaning in life [50, 46, 51]. However, obtaining self-report using less structured approaches may add value, for example by using experience sampling methods [52] or qualitative methods.

Differential Item Functioning (DIF)

The data in this study were gathered in three different countries and two gender groups, three age groups, and two levels of education were distinguished. Of all these demographic variables, significant DIF was only detected for items from the Presence subscale based on the country variable. The absence of DIF is the desirable outcome should data from the different demographic groups be combined or compared [53].

The significant country DIF for items from the Presence of Meaning subscale warrants further attention. Before removal of item 9 (“My life has no clear purpose”), the item exhibited DIF for country: Given equal levels of the latent trait, respondents from South Africa tended to respond more strongly towards the extreme *True* response categories than respondents from New Zealand and Australia, and, similarly participants from New Zealand were more inclined to extreme responses in the *True* direction than participants from Australia. After removal of item 9 and before collapsing the response categories, item 1 (“I understand my life’s meaning”) manifested DIF, where Australians found it harder to endorse the item than South Africans given equal levels of the latent construct. After collapsing categories, this finding was extended – respondents from both New Zealand and Australia found it significantly harder to endorse item 1 than respondents from South Africa given equal levels of the construct. Also, before collapsing categories, participants from New Zealand found it harder to endorse item 4 (“My life has a clear sense of purpose”) than participants from South Africa given equal levels of the latent trait. Last, given equal levels of the latent construct, participants from South Africa found it harder to endorse item 5 (“I

have a good sense of what makes my life meaningful”) than respondents from Australia and New Zealand, both before and after collapsing categories. Country-specific parameter estimates may be needed for these items of the Presence subscale, that is, the dataset can be split by country and these items should be calibrated separately for each country [54].

The two items that respondents from Australia and New Zealand found harder to endorse than South Africans given equal levels of the latent construct (i.e., items 1 and 4) refer to comprehending one’s life meaning and having a clear sense of purpose – both can be seen as a global state of grasping one’s life meaning, without referring to the elements that brings meaning to one’s life. South Africa is a developing country and together with the many challenges the country faces come multiple opportunities for individuals to contribute and to have a sense of purpose. This may especially be the case for educated individuals who may feel that they have skills and knowledge that can really make a difference in a country with many challenges (based on the selection criteria of this study all participants had at least secondary education). Australia and New Zealand, on the other hand, are first world countries with a lot more stability and certainty. People from such countries may feel that things “go right” regardless of their contribution which may possibly lead to having a less clear sense of purpose and meaning comprehension. Another possible explanation may be connected with the fact that the specific South African group in this study exhibited a higher frequency of religious practice (mostly Christianity) than the participants from Australia and New Zealand. Religiosity may be associated with a clear sense of purpose and meaning comprehension.

The item that South Africans found harder to endorse than respondents from Australia and New Zealand given equal levels of the latent trait (item 5) refers to an awareness of the constituents of a meaningful life – the elements that make one’s life meaningful. One possibility is to infer that people (in this case, South Africans) who find it easier to agree with items referring to a global comprehension of one’s life’s meaning (items 1 and 4), may not

have such a pressing need to know what the elements are that make their lives feel meaningful – one may argue that they take it for granted or that they spend less time attending to the specific details of why they find their lives meaningful. In contrast, people who find it more challenging to agree with items related to comprehending one's life meaning and having a clear sense of purpose (in this case respondents from Australia and New Zealand), may be more attentive to the things that add life meaning.

For both items 4 and 9, South Africans tended to answer more strongly in the *True* direction when compared to respondents from Australia and New Zealand given equal levels of the latent construct. In other words, South Africans were more inclined to find both the non-reversed, non-negated statement “My life has a clear sense of purpose” (item 4) and the reversed, negated statement “My life has no clear purpose” (item 9) true. This points to a discrepancy which poses questions about the possible influence of response styles involved in responding to the reversed item that could have caused DIF. This finding provides additional support for the deletion of item 9.

Since all aspects of Rasch analysis are interconnected [30], the existence of cross-country DIF on the Presence subscale could have influenced the rest of the findings. Future research should explore whether the findings of this study replicate in more culturally homogeneous samples where DIF is not present.

Limitations and Future Directions

While the study makes important contributions to the body of knowledge about meaning in life and the measurement thereof across three countries, it also possessed several limitations. This study made use of the Rasch model, which is considered to be a one-parameter IRT model that includes only item difficulty as a parameter. Although the Rasch model has very attractive mathematical properties, analysing MLQ data using more complex IRT models will also be of value.

In this study, recommendations regarding the removal of item 9 (“My life has no clear purpose”) and category collapses were made a posteriori based on removing the item from and collapsing categories of data attained using the original full scale. These recommendations should be tested in new datasets gathered with a revised scale.

The fact that the sample in this study comes from three different countries can be seen as a strength in the sense that diversity is reflected in the study of an already well-established scale. In addition, it allowed us to investigate DIF across the three countries. The fact that evidence was found for DIF across the countries, however, points towards the possibility that the scale may function differentially across the different country groups which could have had an influence on the rest of the results. This suggests the need for repetition of the study in more culturally homogeneous groups to investigate whether the findings replicate when cross-country influences are not present.

Another important avenue for future research is the revisiting of presence of meaning in life as a construct, in particular with regards to the higher end of the construct continuum. The content domain of presence of meaning should be explored qualitatively in order to deepen our understanding of the construct, especially at high levels. For example, by investigating lay people’s conceptualisations of meaning in life, we may identify sub-facets of meaning in life which may provide greater variance at the upper end of the continuum.

Conclusions

The rigorous measurement of meaning in life is essential for the study of this key aspect of well-being and quality of life. The present study was the first to apply item response theory, in particular Rasch modelling, to investigate the psychometric properties and item-level equivalence of the MLQ across different demographic variables. The study offered valuable insights into the functioning of the MLQ in groups from South Africa, Australia, and New Zealand and the construct of meaning in life and the measurement thereof in general. In

particular, the MLQ displayed good psychometric potential from a Rasch modelling perspective. However, several directions for revision were highlighted. First, the study pointed out that seven response categories may be too many when measuring meaning in life in the general population, and suggested that five or six response categories may be more appropriate. In addition, the study confirmed the potential problems involved in reversed, negated items, and suggested that this type of item should rather be avoided – removing the reversed phrased item 9 (“My life has no clear purpose”) was indicated. Although no DIF was found for the Search subscale, the Presence subscale displayed significant DIF for the country variable on four of its five items. Hypotheses were articulated to explore possible sources of the DIF, and it was suggested that country-level parameter estimates may be needed for these items. The existence of DIF pointed to the necessity of repeating this study in more mono-cultural settings to investigate whether the findings replicate. Furthermore, it was shown that people with high scores on presence of meaning were not targeted well by this highly commended meaning in life scale, while at the same time most of the respondents fell in that range. Reasons for and the extensive implications of this finding were contemplated and the vital importance of further exploration of the nature of the construct of presence of meaning and the measurement thereof, particularly at high levels, was indicated.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

MPW and DAVB participated in the design and planning of the study. MPW, DAVB and PEJ took responsibility for the data gathering. LS performed and interpreted the statistical analyses, drafted the manuscript, incorporated all suggestions from the co-authors into the manuscript and prepared the final manuscript for submission. MPW provided continuous feedback regarding the draft document. SME, DAVB, and PEJ revised the

manuscript critically for important intellectual content. All authors read and approved the final manuscript. The study forms part of the doctoral thesis of LS.

Acknowledgements

The South African research was financially supported by the National Research Foundation (NRF) of South Africa by funding for the FORT3 research project and for a doctoral scholarship. Opinions expressed and conclusions arrived are those of the authors and are not necessarily to be attributed to the NRF. Funding was also provided in the form of a doctoral scholarship by the North-West University, South Africa. The New Zealand research was financially supported by a Marsden grant from the Royal Society of New Zealand awarded to the fourth author. This study is based on data from the international Eudaimonic-Hedonic Happiness Investigation (EHHI). The authors wish to thank the anonymous reviewers for insightful comments on the manuscript.

Endnotes

a. In the ability testing environment, where Rasch modelling originated, the term *item difficulty* is often used to refer to the level of the latent construct captured by the item, while *person ability* describes the level of the latent trait held by the respondent. For the purpose of the current study where meaning in life is studied, we deemed the terms *item challenge* and *person intensity* more suitable and used them in the manuscript.

References

1. Efklides A, Moraitou D. Introduction: Looking at quality of life and well-being from a positive psychology perspective. In: Efklides A, Moraitou D, editors. A positive psychology perspective on quality of life. New York, NY: Springer; 2013. p. 1-14.
2. Keyes CL, Fredrickson BL, Park N. Positive psychology and the quality of life. In: Land KC, Michalos AC, Sirgy MJ, editors. Handbook of social indicators and quality of life research. New York, NY: Springer; 2012. p. 99-112.

3. Efklides A, Moraitou D, editors. *A positive psychology perspective on quality of life*. New York, NY: Springer; 2013.
4. Frankl VE. *Man's search for meaning: An introduction to logotherapy*. New York, NY: Washington Square Press; 1963.
5. Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*. 1989;57(6):1069-81. doi:10.1037/0022-3514.57.6.1069.
6. Steger MF, Shin JY, Shim Y, Fitch-Martin A. Is meaning in life a flagship indicator of well-being? In: Waterman AS, editor. *The best within us: Positive psychology perspectives on eudaimonia*. Washington, DC: American Psychological Association; 2013. p. 159-82.
7. Baumeister RF, Vohs KD. The pursuit of meaningfulness in life. In: Snyder CR, Lopez SJ, editors. *Handbook of positive psychology*. New York, NY: Oxford University Press; 2002. p. 608-18.
8. Steger MF. Experiencing meaning in life: Optimal functioning at the nexus of well-being, psychopathology, and spirituality. In: Wong PTP, editor. *The human quest for meaning: Theories, research, and applications 2nd ed*. New York, NY: Routledge; 2012. p. 165-84.
9. Park CL. Integrating positive psychology into health-related quality of life research. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*. 2015;24(7):1645-51. doi:10.1007/s11136-014-0889-z.
10. Wong PTP. Meaning therapy: An integrative and positive existential psychotherapy. *Journal of Contemporary Psychotherapy*. 2010;40(2):85-93. doi:10.1007/s10879-009-9132-6.
11. Schnell T. The Sources of Meaning and Meaning in Life Questionnaire (SoMe): Relations to demographics and well-being. *The Journal of Positive Psychology*. 2009;4(6):483-99. doi:10.1080/17439760903271074.

12. Steger MF. Meaning in life. In: Lopez SJ, Snyder CR, Lopez SJ, Snyder CR, editors. Oxford handbook of positive psychology. 2nd ed. Oxford library of psychology. New York, NY: Oxford University Press; 2009. p. 679-87.
13. Steger MF, Kashdan TB, Sullivan BA, Lorentz D. Understanding the search for meaning in life: Personality, cognitive style, and the dynamic between seeking and experiencing meaning. *Journal of Personality*. 2008;76(2):199-228. doi:10.1111/j.1467-6494.2007.00484.x.
14. Brandstätter M, Baumann U, Borasio GD, Fegg MJ. Systematic review of meaning in life assessment instruments. *Psycho-Oncology*. 2012;21(10):1034-52. doi:10.1002/pon.2113.
15. Steger MF, Frazier P, Oishi S, Kaler M. The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*. 2006;53(1):80-93. doi:10.1037/0022-0167.53.1.80.
16. Steger MF, Oishi S, Kashdan TB. Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology*. 2009;4(1):43-52. doi:10.1080/17439760802303127.
17. Schulenberg SE, Strack KM, Buchanan EM. The meaning in life questionnaire: Psychometric properties with individuals with serious mental illness in an inpatient setting. *Journal of Clinical Psychology*. 2011;67(12):1210-9. doi:10.1002/jclp.20841.
18. Temane L, Khumalo IP, Wissing MP. Validation of the Meaning in Life Questionnaire in a South African context. *Journal of Psychology in Africa*. 2014;24(1):81-95. doi:10.1080/14330237.2014.904088.
19. Steger MF, Kawabata Y, Shimai S, Otake K. The meaningful life in Japan and the United States: Levels and correlates of meaning in life. *Journal of Research in Personality*. 2008;42(3):660-78. doi:10.1016/j.jrp.2007.09.003.

20. Steger MF, Frazier PA, Zacchanini JL. Terrorism in two cultures: Stress and growth following September 11 and the Madrid train bombings. *Journal of Loss & Trauma*. 2008;13(6):511-27. doi:10.1080/15325020802173660.
21. Boyraz G, Lightsey OR, Can A. The Turkish version of the Meaning in Life Questionnaire: Assessing the measurement invariance across Turkish and American adult samples. *Journal of Personality Assessment*. 2013;95(4):423-31. doi:10.1080/00223891.2013.765882
22. Khumalo IP, Wissing MP, Schutte L, De Klerk W. Translation and measurement invariance of the Meaning in Life Questionnaire across three languages in South Africa. 2015. Manuscript submitted for publication.
23. DeVellis RF. *Scale development: Theory and applications*. vol Book, Whole. Thousand Oaks, CA: Sage; 2012.
24. Thomas ML. The value of item response theory in clinical assessment: A review. *Assessment*. 2011;18(3):291-307. doi:10.1177/1073191110374797.
25. Andrich D. Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*. 1978;2(4):581-94. doi:0.1080/14330237.2014.904088.
26. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark Danmarks Peadagogiske Institut; 1960.
27. Andrich D. Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care*. 2004;42(1):1-16.
28. Bond TG, Fox CM. *Applying the Rasch model: Fundamental measurement in the human sciences* 2nd ed. New York, NY: Routledge; 2007.

29. da Rocha NS, Chachamovich E, de Almeida Fleck MP, Tennant A. An introduction to Rasch analysis for Psychiatric practice and research. *Journal of Psychiatric Research*. 2013;47(2):141-8. doi:10.1016/j.jpsychires.2012.09.014.
30. Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies*. 2009;46(3):380-93. doi:10.1016/j.ijnurstu.2008.10.007.
31. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism: Arthritis Care & Research*. 2007;57(8):1358-62. doi:10.1002/art.23108.
32. Myers ND, Wolfe EW, Feltz DL, Penfield RD. Identifying Differential Item Functioning of Rating Scale Items With the Rasch Model: An Introduction and an Application. *Measurement in Physical Education & Exercise Science*. 2006;10(4):215-40. doi:10.1207/s15327841mpee1004_1.
33. Delle Fave A, Brdar I, Freire T, Vella-Brodrick D, Wissing MP. The eudaimonic and hedonic components of happiness: Qualitative and quantitative findings. *Social Indicators Research*. 2011;100(2):185-207. doi:10.1007/s11205-010-9632-5.
34. Grouden ME, Jose PE. How do sources of meaning in life vary according to demographic factors? *New Zealand Journal of Psychology*. 2014;43(3):29-38.
35. Cohen K, Cairns D. Is searching for meaning in life associated with reduced subjective well-being? Confirmation and possible moderators. *Journal of Happiness Studies*. 2012;13(2):313-31. doi:10.1007/s10902-011-9265-7.
36. Linacre JM. *Winsteps (Version 3.81.0)*. Beaverton, Oregon: Winsteps.com; 2014.
37. Andrich D, Sheridan B, Luo G. *Rumm 2030 for Windows (Version 5.52)*. Perth, Australia: Rumm Laboratory; 2014.

38. Linacre JM. A user's guide to Winsteps Ministep Rasch-model computer programs: Program manual 3.81.0. Beaverton, Oregon: Winsteps.com; 2014.
39. Fox CM, Jones JA. Uses of Rasch modeling in counseling psychology research. *Journal of Counseling Psychology*. 1998;45(1):30. doi:0022-0167/98/53.00.
40. Chao RC-L, Green KE. Rasch analysis of the Outcome Questionnaire with African Americans. *Psychological Assessment*. 2013;25(2):568-82. doi:10.1037/a0032083.
41. Mantel N. Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association* 1963;58(303):690-700.
42. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal Of The National Cancer Institute*. 1959;22(4):719-48.
43. Weijters B, Baumgartner H. Misresponse to Reversed and Negated Items in Surveys: A Review. *Journal of Marketing Research (JMR)*. 2012;49(5):737-47. doi:10.1509/jmr.11.0368
44. Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*. 2010;27(3):236-47. doi:10.1016/j.ijresmar.2010.02.004.
45. Nowlis SM, Kahn BE, Dhar R. Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research*. 2002;29(3):319-34. doi:10.1086/344431
46. Heintzelman SJ, King LA. Life is pretty meaningful. *The American Psychologist*. 2014;69(6):561-74. doi:10.1037/a0035049.
47. Crumbaugh JC, Maholick LT. An experimental study in existentialism: The psychometric approach to Frankl's concept of noogenic neurosis. *Journal of Clinical Psychology*. 1964;20(2):200-7. doi:10.1002/1097-4679(196404)20:2<200::AID-JCLP2270200203>3.0.CO;2-U

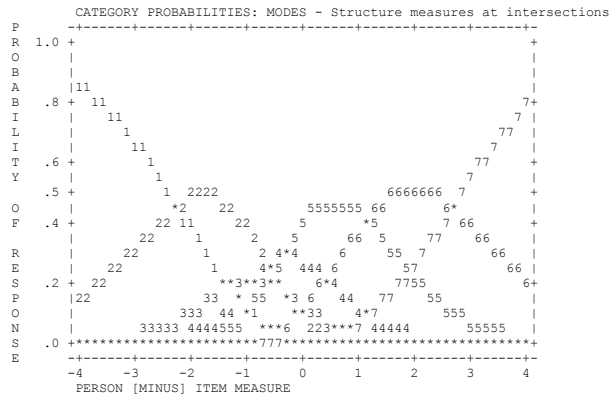
48. Reise SP, Waller NG. Item response theory and clinical measurement. *Annual Review of Clinical Psychology*. 2009;5:27-48. doi:10.1146/annurev.clinpsy.032408.153553.
49. Liu S-S, Gan Y-Q. Reliability and validity of the Chinese version of the Meaning in Life Questionnaire. *Chinese Mental Health Journal*. 2010;24(6):478-82.
50. Heintzelman SJ, King LA. On knowing more than we can tell: Intuitive processes and the experience of meaning. *Journal of Positive Psychology*. 2013;8(6):471-82. doi:10.1080/17439760.2013.830758.
51. Kashdan T, Biswas-Diener R, King L. Reconsidering happiness: The costs of distinguishing between hedonics and eudaimonia. *The Journal of Positive Psychology*. 2008;3(4):219-33. doi:10.1080/17439760802303044.
52. Trull TJ, Ebner-Priemer UW. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*. 2009;21(4):457-62. doi:10.1037/a0017653.
53. Sireci SG. Evaluating test and survey items for bias across languages and cultures. In: Matsumoto D, Van de Vijver FJR, editors. *Cross-cultural research methods in psychology. Culture and psychology*. New York, NY: Cambridge University Press; 2011. p. 216-43.
54. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*. 2007;46(1):1-18. doi:10.1348/014466506X96931.

Figure 1.

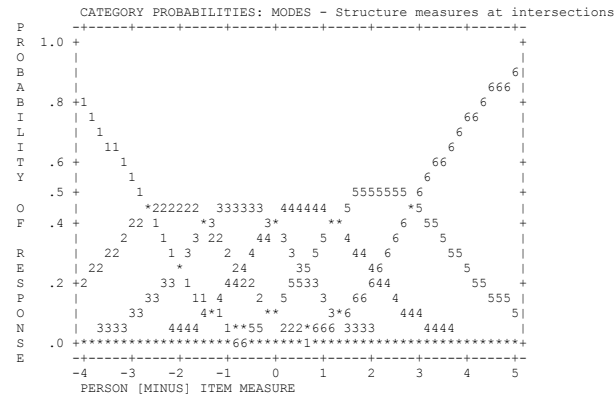
Title: Category probability curves of the Meaning in Life Questionnaire (MLQ).

Legend: MLQ-P = Original Presence of Meaning subscale of the MLQ; MLQ-P-4 1122345 = MLQ-P-4 with response categories 1 and 2 collapsed and categories 3 and 4 collapsed. MLQ-S = Original Search for Meaning subscale of the MLQ; MLQ-S 1233456 = MLQ-S with response categories 3 and 4 collapsed.

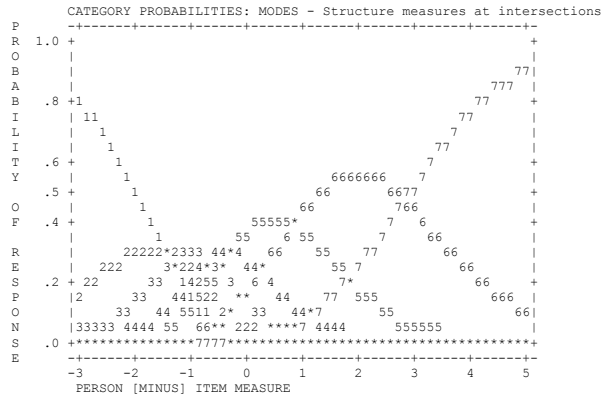
MLQ-P



MLQ-P-4 1122345



MLQ-S



MLQ-S 1233456

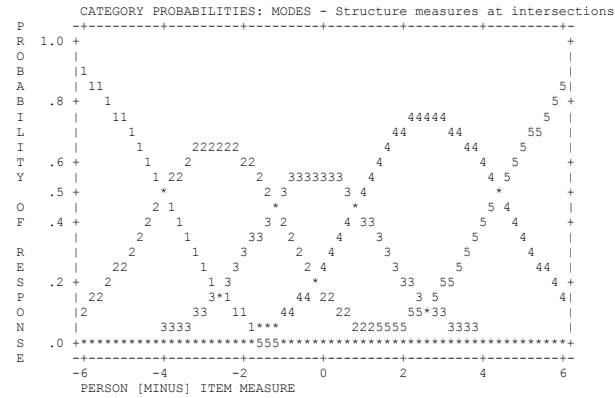
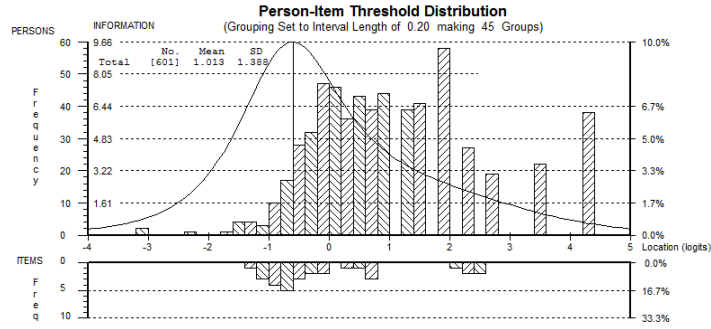


Figure 2.

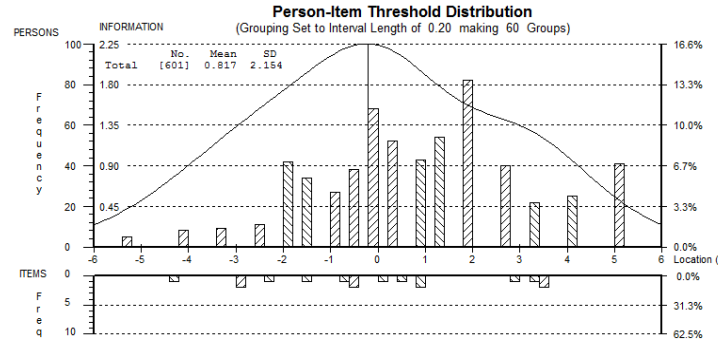
Title: Person-item threshold distributions of the Meaning in Life Questionnaire (MLQ).

Legend: MLQ-P = Original Presence of Meaning subscale of the MLQ; MLQ-P-4 1122345 = MLQ-P-4 with response categories 1 and 2 collapsed and categories 3 and 4 collapsed. MLQ-S = Original Search for Meaning subscale of the MLQ; MLQ-S 1233456 = MLQ-S with response categories 3 and 4 collapsed. Person-item threshold distributions were obtained from RUMM2030™.

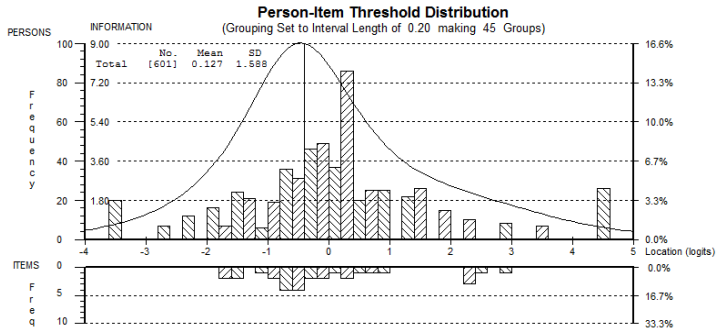
MLQ-P



MLQ-P-4 1122345



MLQ-S



MLQ-S 1233456

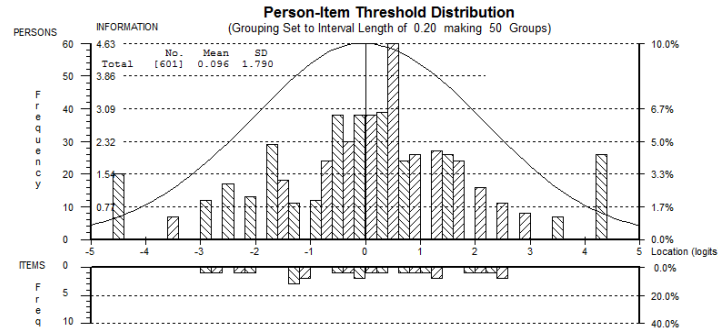


Table 1

Demographic Profile of the Sample

	South Africa	New Zealand	Australia	Total
Gender				
Male	101	107	79	287
Female	115	108	91	314
Age				
<i>M (SD)</i>	44.11 (8.53)	44.45 (8.85)	44.62 (8.84)	44.38 (8.72)
30-39	77	71	53	201
40-49	71	72	58	201
50-60	68	72	59	199
Education level				
Secondary	106	106	68	280
Tertiary	106	109	102	317
Missing	4	0	0	4
Total	216	215	170	601

Table 2.

Separation, Reliability, Fit, and Dimensionality by Subscale and Analysis

Analysis	Separation		Reliability		Infit MNSQ		Outfit MNSQ		Dimensionality and local independence		
	Person	Item	Person	Item	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Eigenvalue of 1 st contrast	% Variance explained	Max residual <i>r</i>
MLQ-P	2.00	3.84	.80	.94	1.01	0.57	1.01	0.59	1.7	59.3	.04
MLQ-P-4	2.42	5.81	.85	.97	0.99	0.16	0.96	0.14	1.6	69.7	None
MLQ-P-4 1122345	2.41	5.91	.85	.97	0.99	0.14	0.98	0.14	1.6	69.1	None
MLQ-P-4 1222345	2.43	5.71	.86	.97	0.99	0.14	0.96	0.14	1.4	69.0	None
MLQ-S	2.45	5.90	.86	.97	0.99	0.17	0.97	0.17	1.6	68.7	None
MLQ-S 1223345	2.50	5.59	.86	.97	0.99	0.17	0.99	0.17	1.6	63.5	None
MLQ-S 1233456	2.52	5.74	.86	.97	0.99	0.16	0.98	0.99	1.6	66.5	None
<i>Ideal values</i>	>2	>3	>0.8	>0.9	<1.4	<i>Small</i>	<1.4	<i>Small</i>	<2	>40%	<0.2

Note. MNSQ = mean square statistic; Eigenvalue of 1st contrast = eigenvalue of the 1st contrast in the Rasch principal components analysis of the residuals; % Variance explained = % variance explained by the Rasch component in the Rasch principal components analysis of the residuals; Max residual *r* = maximum positive residual correlation; MLQ-P = original Presence of Meaning subscale of the Meaning in Life

Questionnaire; MLQ-P-4 = MLQ-P with item 9 removed; MLQ-P-4 1122345 = MLQ-P-4 with response categories 1 and 2 collapsed and categories 3 and 4 collapsed; MLQ-P-4 1222345 = MLQ-P-4 with response categories 2, 3, and 4 collapsed; MLQ-S = original Search for Meaning subscale of the Meaning in Life Questionnaire; MLQ-S 1223345 = MLQ-S with response categories 2 and 3 collapsed and categories 4 and 5 collapsed; MLQ-S 1233456 = MLQ-S with response categories 3 and 4 collapsed.

Table 3

Meaning in Life Questionnaire: Item Measures, Standard Errors, and Fit Statistics

	δ	SE	Infit	Outfit		δ	SE	Infit	Outfit		δ	SE	Infit	Outfit
	MLQ-P					MLQ-P-4					MLQ-P-4 1122345			
1. I understand my life's meaning.	0.09	0.04	0.92	0.90		0.10	0.06	1.23	1.17		0.09	0.07	1.21	1.20
4. My life has a clear sense of purpose.	0.17	0.04	0.60	0.62		0.23	0.06	0.79	0.80		0.31	0.07	0.83	0.82
5. I have a good sense of what makes my life meaningful.	-0.34	0.05	0.71	0.66		-0.61	0.06	1.04	0.97		-0.75	0.07	0.96	0.94
6. I have discovered a satisfying life purpose.	0.19	0.04	0.69	0.70		0.27	0.06	0.92	0.88		0.36	0.07	0.94	0.94
9. My life has no clear purpose. ^a	-0.12	0.05	2.13	2.18										
	MLQ-S					MLQ-S 1223345					MLQ-S 1233456			
2. I am looking for something that makes my life feel meaningful.	-0.30	0.04	1.14	1.11		-0.44	0.07	1.09	1.09		-0.36	0.05	1.11	1.10
3. I am always looking to find my life's purpose.	-0.06	0.04	0.96	0.93		-0.12	0.07	0.96	0.96		-0.06	0.05	0.99	0.98
7. I am always searching for something that makes my life feel significant.	-0.20	0.04	1.05	1.03		-0.29	0.07	1.02	1.04		-0.24	0.05	1.02	1.04

8. I am seeking a purpose or mission for my life.	0.10	0.04	0.67	0.65		0.18	0.07	0.69	0.69		0.12	0.05	0.69	0.69
10. I am searching for meaning in my life.	0.46	0.04	1.12	1.12		0.67	0.07	1.19	1.17		0.54	0.05	1.13	1.11
<i>Ideal values</i>			>0.6	>0.6				>0.6	>0.6				>0.6	>0.6
			<1.4	<1.4				<1.4	<1.4				<1.4	<1.4

Note. MLQ = Meaning in Life Questionnaire; MLQ-P = original Presence of Meaning subscale of the MLQ; MLQ-P-4 = MLQ-P with item 9 removed; MLQ-P-4 1122345 = MLQ-P-4 with response categories 1 and 2 collapsed and categories 3 and 4 collapsed; MLQ-P-4 1222345 = MLQ-P-4 with response categories 2, 3, and 4 collapsed; MLQ-S = original Search for Meaning subscale of the MLQ; MLQ-S 1223345 = MLQ-S with response categories 2 and 3 collapsed and categories 4 and 5 collapsed; MLQ-S 1233456 = MLQ-S with response categories 3 and 4 collapsed. δ = average item challenge; SE = standard error of the item challenge; Infit = infit mean square statistic; Oufit = outfit mean square statistic.

^aThe original item 9 was reversed in these analyses.

Table 4

Meaning in Life Questionnaire – Presence Subscale: Comparisons of the Rating Scale

Categories

	Observed count	Average measure	Infit	Outfit	Threshold calibration
MLQ-P					
1 – <i>Absolutely untrue</i>	79	-1.02	1.58	2.44	None
2 – <i>Mostly untrue</i>	111	-0.65	1.26	1.54	-1.47
3 – <i>Somewhat untrue</i>	208	-0.46	0.82	0.81	-1.19
4 – <i>Can't say true or false</i>	352	-0.07	0.67	0.66	-0.67
5 – <i>Somewhat true</i>	700	0.64	0.67	0.60	-0.31
6 – <i>Mostly true</i>	919	1.71	0.87	0.89	0.83
7 – <i>Absolutely true</i>	636	2.55	1.39	1.19	2.82
MLQ-P-4					
1 – <i>Absolutely untrue</i>	56	-2.23	1.29	1.48	None
2 – <i>Mostly untrue</i>	73	-1.72	0.87	0.83	-2.59
3 – <i>Somewhat untrue</i>	152	-1.09	0.95	1.01	-2.09
4 – <i>Can't say true or false</i>	288	-0.34	0.87	0.93	-1.32
5 – <i>Somewhat true</i>	633	0.83	0.74	0.74	-0.55
6 – <i>Mostly true</i>	810	2.85	0.96	0.91	1.49
7 – <i>Absolutely true</i>	392	4.56	1.75	1.21	5.05
MLQ-P-4 1122345					
1 (Original categories 1 and 2 collapsed)	129	-3.40	1.16	1.13	None
2 (Original categories 3 and 4 collapsed)	440	-1.82	0.94	1.00	-4.00

3 (Original category 5)	633	0.02	0.75	0.73	-1.22
4 (Original category 6)	810	2.20	0.92	0.94	0.78
5 (Original category 7)	392	3.98	1.44	1.27	4.43
<i>Ideal values</i>			<2	<2	

Note. MLQ-P = original Presence of Meaning subscale of the Meaning in Life Questionnaire;

MLQ-P-4 = MLQ-P with item 9 removed; MLQ-P-4 1122345 = MLQ-P-4 with response

categories 1 and 2 collapsed and categories 3 and 4 collapsed; MLQ-P-4 1222345 = MLQ-P-

4 with response categories 2, 3, and 4 collapsed; Infit = infit mean square statistic; Oufit =

oufit mean square statistic.

Table 5

Meaning in Life Questionnaire – Search Subscale: Comparisons of the Rating Scale

Categories

	Observed count	Average measure	Infit	Outfit	Threshold calibration
MLQ-S					
1 – <i>Absolutely untrue</i>	335	-1.99	1.53	1.52	None
2 – <i>Mostly untrue</i>	406	-1.45	0.72	0.77	-2.23
3 – <i>Somewhat true</i>	254	-0.55	0.83	0.84	-0.46
4 – <i>Can't say true or false</i>	408	-0.07	0.75	0.67	-0.81
5 – <i>Somewhat true</i>	731	0.51	0.81	0.84	-0.38
6 – <i>Mostly true</i>	530	1.38	0.98	0.99	1.17
7 – <i>Absolutely true</i>	341	2.14	1.70	1.30	2.70
MLQ-S 1223345					
1 (Original category 1)	335	-3.11	1.43	1.39	None
2 (Original categories 2 and 3 collapsed)	660	-1.87	0.79	0.77	-3.55
3 (Original categories 4 and 5 collapsed)	1139	0.21	0.74	0.75	-1.27
4 (Original category 6)	530	1.72	0.91	0.91	1.64
5 (Original category 7)	341	2.68	1.37	1.43	3.18
MLQ-S 1233456					
1 (Original category 1)	335	-2.47	1.52	1.46	None
2 (Original category 2)	406	-1.80	0.69	0.71	-2.65
3 (Original categories 3 and 4 collapsed)	662	-0.39	0.78	0.76	-1.45
4 (Original category 5)	731	0.53	0.76	0.79	-0.06

5 (Original category 6)	530	1.52	0.96	0.99	1.25
6 (Original category 7)	341	2.35	1.55	1.43	2.90
<i>Ideal values</i>			<2	<2	

Note. MLQ-S = original Search for Meaning subscale of the Meaning in Life Questionnaire;

MLQ-S 1223345 = MLQ-S with response categories 2 and 3 collapsed and categories 4 and 5

collapsed; MLQ-S 1233456 = MLQ-S with response categories 3 and 4 collapsed. Infit =

infit mean square statistic; Oufit = outfit mean square statistic.

Note. MLQ-P = original Presence of Meaning subscale of the Meaning in Life Questionnaire; MLQ-P-4 = MLQ-P with item 9 removed; MLQ-P-4 1122345 = MLQ-P-4 with response categories 1 and 2 collapsed and categories 3 and 4 collapsed; Bonferroni α = Bonferroni-corrected significance level; DIF measure = item challenge for the particular country; DIF S.E. = standard error of item challenge for the particular country; Contrast = if the DIF contrast (i.e., the difference between the two countries' DIF measures) was larger than or equal to 0.64, the countries are specified in this column; MH = if the *p*-value of the Mantel-Haenszel test was smaller than Bonferroni α , the countries are specified in this column; AU = Australian sample; SA = South African sample; NZ = sample from New Zealand. In columns DIF Contrast and MH, $x > y$ implies that respondents from country *x* found it significantly harder to endorse the item than respondents from country *y* given equal levels of presence of meaning.

^aThe original item 9 was reversed in these analyses.

Section 3: Manuscript 2

Clarifying the factor structure of the Mental Health Continuum Short Form in three languages: A bifactor model approach

Submitted to

Social Indicators Research

3.1. Guidelines for authors

Social Indicators Research

These guidelines were retrieved from:

<http://www.springer.com/social+sciences/journal/11205>

“General remarks

Language

We appreciate any efforts that you make to ensure that the language is corrected before submission. This will greatly improve the legibility of your paper if English is not your first language.

Offprints and Page Charges

The author will receive, without charge, 25 copies of the article. Additional offprints (both hard copies and PDF files) may be ordered by the corresponding author.

No page charges are levied on authors or their institutions.

Manuscript Submission

Manuscript Submission

Submission of a manuscript implies: that the work described has not been published before; that it is not under consideration for publication anywhere else; that its publication has been approved by all co-authors, if any, as well as by the responsible authorities – tacitly or explicitly – at the institute where the work has been carried out. The publisher will not be held legally responsible should there be any claims for compensation.

Permissions

Authors wishing to include figures, tables, or text passages that have already been published elsewhere are required to obtain permission from the copyright owner(s) for both the print and online format and to include evidence that such permission has been granted when submitting their papers. Any material received without such evidence will be assumed to originate from the authors.

Online Submission

Please follow the hyperlink “Submit online” on the right and upload all of your manuscript files following the instructions given on the screen.

Title page

Title Page

The title page should include:

- The name(s) of the author(s)
- A concise and informative title
- The affiliation(s) and address(es) of the author(s)
- The e-mail address, telephone and fax numbers of the corresponding author

Abstract

Please provide an abstract of 150 to 250 words. The abstract should not contain any undefined abbreviations or unspecified references.

Keywords

Please provide 4 to 6 keywords which can be used for indexing purposes.

Text

Text Formatting

Manuscripts should be submitted in Word.

- Use a normal, plain font (e.g., 10-point Times Roman) for text.
- Use italics for emphasis.
- Use the automatic page numbering function to number the pages.
- Do not use field functions.
- Use tab stops or other commands for indents, not the space bar.
- Use the table function, not spreadsheets, to make tables.
- Use the equation editor or MathType for equations.
- Save your file in docx format (Word 2007 or higher) or doc format (older Word versions).

Manuscripts with mathematical content can also be submitted in LaTeX.

- [LaTeX macro package \(zip, 182 kB\)](#)

Headings

Please use the decimal system of headings with no more than three levels.

Abbreviations

Abbreviations should be defined at first mention and used consistently thereafter.

Footnotes

Footnotes can be used to give additional information, which may include the citation of a reference included in the reference list. They should not consist solely of a reference citation,

and they should never include the bibliographic details of a reference. They should also not contain any figures or tables.

Footnotes to the text are numbered consecutively; those to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data).

Footnotes to the title or the authors of the article are not given reference symbols.

Always use footnotes instead of endnotes.

Acknowledgments

Acknowledgments of people, grants, funds, etc. should be placed in a separate section on the title page. The names of funding organizations should be written in full.

References

Citation

Cite references in the text by name and year in parentheses. Some examples:

- Negotiation research spans many disciplines (Thompson 1990).
- This result was later contradicted by Becker and Seligman (1996).
- This effect has been widely studied (Abbott 1991; Barakat et al. 1995; Kelso and Smith 1998; Medvec et al. 1999).

Reference list

The list of references should only include works that are cited in the text and that have been published or accepted for publication. Personal communications and unpublished works should only be mentioned in the text. Do not use footnotes or endnotes as a substitute for a reference list.

Reference list entries should be alphabetized by the last names of the first author of each work.

- Journal article

Harris, M., Karper, E., Stacks, G., Hoffman, D., DeNiro, R., Cruz, P., et al. (2001). Writing labs and the Hollywood connection. *Journal of Film Writing*, 44(3), 213–245.
- Article by DOI

Slifka, M. K., & Whitton, J. L. (2000) Clinical implications of dysregulated cytokine production. *Journal of Molecular Medicine*, doi:10.1007/s001090000086
- Book

Calfee, R. C., & Valencia, R. R. (1991). *APA guide to preparing manuscripts for journal publication*. Washington, DC: American Psychological Association.
- Book chapter

O'Neil, J. M., & Egan, J. (1992). Men's and women's gender role journeys: Metaphor for healing, transition, and transformation. In B. R. Wainrib (Ed.), *Gender issues across the life cycle* (pp. 107–123). New York: Springer.

- Online document

Abou-Allaban, Y., Dell, M. L., Greenberg, W., Lomax, J., Peteet, J., Torres, M., & Cowell, V. (2006). Religious/spiritual commitments and psychiatric practice. Resource document. American Psychiatric Association. http://www.psych.org/edu/other_res/lib_archives/archives/200604.pdf. Accessed 25 June 2007.

Journal names and book titles should be italicized.

For authors using EndNote, Springer provides an output style that supports the formatting of in-text citations and reference list.

- [EndNote style \(zip, 3 kB\)](#)

Tables

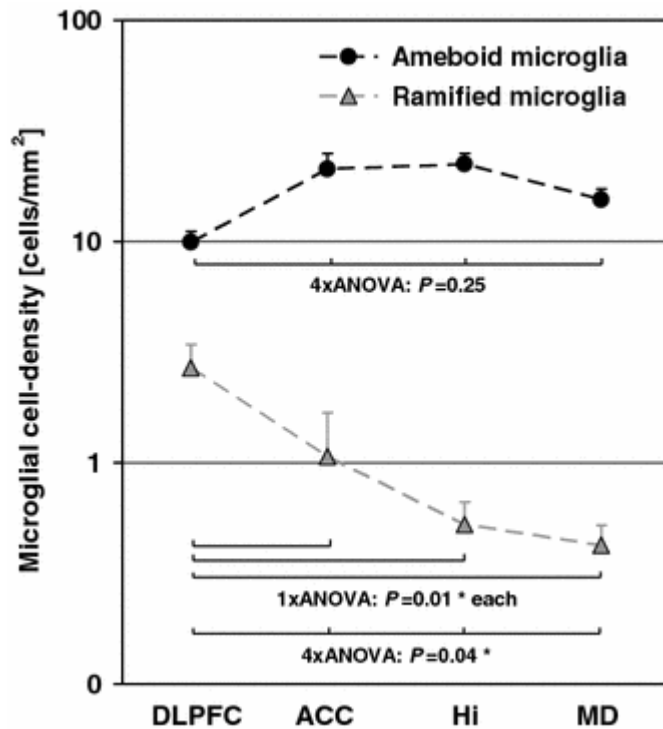
- All tables are to be numbered using Arabic numerals.
- Tables should always be cited in text in consecutive numerical order.
- For each table, please supply a table caption (title) explaining the components of the table.
- Identify any previously published material by giving the original source in the form of a reference at the end of the table caption.
- Footnotes to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data) and included beneath the table body.

Artwork and Illustrations Guidelines

Electronic Figure Submission

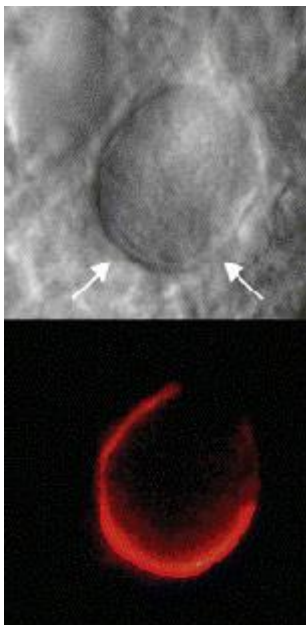
- Supply all figures electronically.
- Indicate what graphics program was used to create the artwork.
- For vector graphics, the preferred format is EPS; for halftones, please use TIFF format. MSOffice files are also acceptable.
- Vector graphics containing fonts must have the fonts embedded in the files.
- Name your figure files with "Fig" and the figure number, e.g., Fig1.eps.

Line Art



- Definition: Black and white graphic with no shading.
- Do not use faint lines and/or lettering and check that all lines and lettering within the figures are legible at final size.
- All lines should be at least 0.1 mm (0.3 pt) wide.
- Scanned line drawings and line drawings in bitmap format should have a minimum resolution of 1200 dpi.
- Vector graphics containing fonts must have the fonts embedded in the files.

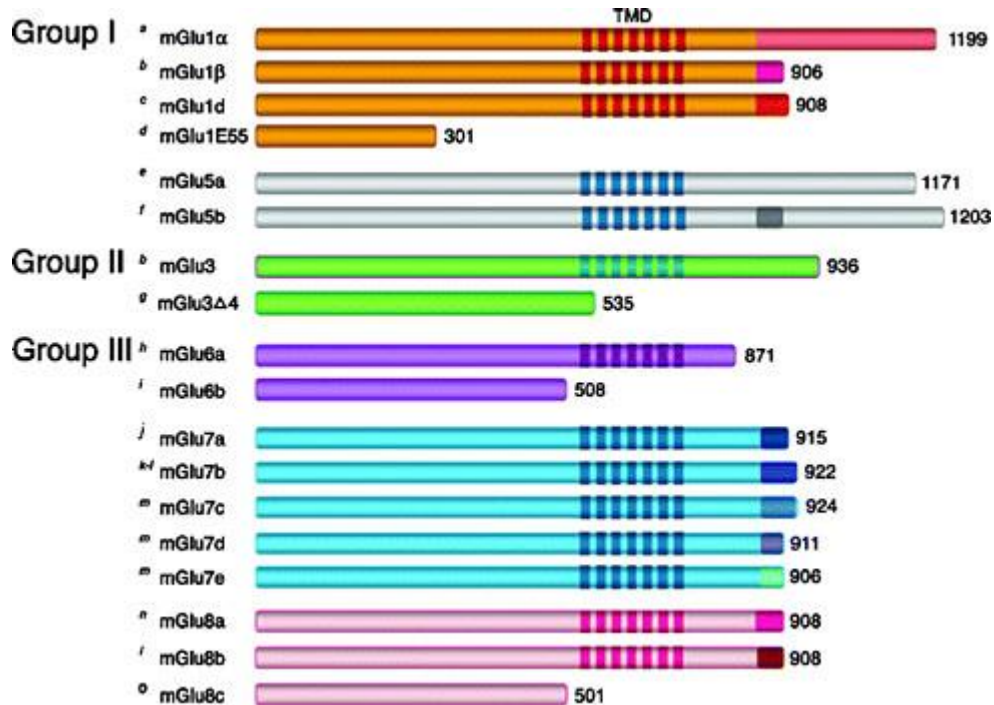
Halftone Art



- Definition: Photographs, drawings, or paintings with fine shading, etc.

- If any magnification is used in the photographs, indicate this by using scale bars within the figures themselves.
- Halftones should have a minimum resolution of 300 dpi.

Combination Art



- Definition: a combination of halftone and line art, e.g., halftones containing line drawing, extensive lettering, color diagrams, etc.
- Combination artwork should have a minimum resolution of 600 dpi.

Color Art

- Color art is free of charge for online publication.
- If black and white will be shown in the print version, make sure that the main information will still be visible. Many colors are not distinguishable from one another when converted to black and white. A simple way to check this is to make a xerographic copy to see if the necessary distinctions between the different colors are still apparent.
- If the figures will be printed in black and white, do not refer to color in the captions.
- Color illustrations should be submitted as RGB (8 bits per channel).

Figure Lettering

- To add lettering, it is best to use Helvetica or Arial (sans serif fonts).
- Keep lettering consistently sized throughout your final-sized artwork, usually about 2–3 mm (8–12 pt).
- Variance of type size within an illustration should be minimal, e.g., do not use 8-pt type on an axis and 20-pt type for the axis label.
- Avoid effects such as shading, outline letters, etc.
- Do not include titles or captions within your illustrations.

Figure Numbering

- All figures are to be numbered using Arabic numerals.
- Figures should always be cited in text in consecutive numerical order.
- Figure parts should be denoted by lowercase letters (a, b, c, etc.).
- If an appendix appears in your article and it contains one or more figures, continue the consecutive numbering of the main text. Do not number the appendix figures, "A1, A2, A3, etc." Figures in online appendices (Electronic Supplementary Material) should, however, be numbered separately.

Figure Captions

- Each figure should have a concise caption describing accurately what the figure depicts. Include the captions in the text file of the manuscript, not in the figure file.
- Figure captions begin with the term Fig. in bold type, followed by the figure number, also in bold type.
- No punctuation is to be included after the number, nor is any punctuation to be placed at the end of the caption.
- Identify all elements found in the figure in the figure caption; and use boxes, circles, etc., as coordinate points in graphs.
- Identify previously published material by giving the original source in the form of a reference citation at the end of the figure caption.

Figure Placement and Size

- Figures should be submitted separately from the text, if possible.
- When preparing your figures, size figures to fit in the column width.
- For most journals the figures should be 39 mm, 84 mm, 129 mm, or 174 mm wide and not higher than 234 mm.
- For books and book-sized journals, the figures should be 80 mm or 122 mm wide and not higher than 198 mm.

Permissions

If you include figures that have already been published elsewhere, you must obtain permission from the copyright owner(s) for both the print and online format. Please be aware that some publishers do not grant electronic rights for free and that Springer will not be able to refund any costs that may have occurred to receive these permissions. In such cases, material from other sources should be used.

Accessibility

In order to give people of all abilities and disabilities access to the content of your figures, please make sure that

- All figures have descriptive captions (blind users could then use a text-to-speech software or a text-to-Braille hardware)
- Patterns are used instead of or in addition to colors for conveying information (colorblind users would then be able to distinguish the visual elements)

- Any figure lettering has a contrast ratio of at least 4.5:1

Electronic Supplementary Material

Springer accepts electronic multimedia files (animations, movies, audio, etc.) and other supplementary files to be published online along with an article or a book chapter. This feature can add dimension to the author's article, as certain information cannot be printed or is more convenient in electronic form.

Submission

- Supply all supplementary material in standard file formats.
- Please include in each file the following information: article title, journal name, author names; affiliation and e-mail address of the corresponding author.
- To accommodate user downloads, please keep in mind that larger-sized files may require very long download times and that some users may experience other problems during downloading.

Audio, Video, and Animations

- Aspect ratio: 16:9 or 4:3
- Maximum file size: 25 GB
- Minimum video duration: 1 sec
- Supported file formats: avi, wmv, mp4, mov, m2p, mp2, mpg, mpeg, flv, mxf, mts, m4v, 3gp

Text and Presentations

- Submit your material in PDF format; .doc or .ppt files are not suitable for long-term viability.
- A collection of figures may also be combined in a PDF file.

Spreadsheets

- Spreadsheets should be converted to PDF if no interaction with the data is intended.
- If the readers should be encouraged to make their own calculations, spreadsheets should be submitted as .xls files (MS Excel).

Specialized Formats

- Specialized format such as .pdb (chemical), .wrl (VRML), .nb (Mathematica notebook), and .tex can also be supplied.

Collecting Multiple Files

- It is possible to collect multiple files in a .zip or .gz file.

Numbering

- If supplying any supplementary material, the text must make specific mention of the material as a citation, similar to that of figures and tables.
- Refer to the supplementary files as “Online Resource”, e.g., "... as shown in the animation (Online Resource 3)", "... additional data are given in Online Resource 4”.
- Name the files consecutively, e.g. “ESM_3.mpg”, “ESM_4.pdf”.

Captions

- For each supplementary material, please supply a concise caption describing the content of the file.

Processing of supplementary files

- Electronic supplementary material will be published as received from the author without any conversion, editing, or reformatting.

Accessibility

In order to give people of all abilities and disabilities access to the content of your supplementary files, please make sure that

- The manuscript contains a descriptive caption for each supplementary material
- Video files do not contain anything that flashes more than three times per second (so that users prone to seizures caused by such effects are not put at risk)

Does Springer provide English language support?

Manuscripts that are accepted for publication will be checked by our copyeditors for spelling and formal style. This may not be sufficient if English is not your native language and substantial editing would be required. In that case, you may want to have your manuscript edited by a native speaker prior to submission. A clear and concise language will help editors and reviewers concentrate on the scientific content of your paper and thus smooth the peer review process.

The following editing service provides language editing for scientific articles in all areas
Springer publishes in:

- [Edanz English editing for scientists](#)

Use of an editing service is neither a requirement nor a guarantee of acceptance for publication.

Please contact the editing service directly to make arrangements for editing and payment.

For Authors from China

文章在投稿前进行专业的语言润色将对作者的投稿进程有所帮助。作者可自愿选择使用Springer推荐的编辑服务，使用与否并不作为判断文章是否被录用的依据。提高文章的语言质量将有助于审稿人理解文章的内容，通过对学术内容的判断来决定文章的

取舍，而不会因为语言问题导致直接退稿。作者需自行联系 Springer 推荐的编辑服务公司，协商编辑事宜。

- [理文编辑](#)

For Authors from Japan

ジャーナルに論文を投稿する前に、ネイティブ・スピーカーによる英文校閲を希望されている方には、Edanz社をご紹介します。サービス内容、料金および申込方法など、日本語による詳しい説明はエダンズグループジャパン株式会社の下記サイトをご覧ください。

- [エダンズグループジャパン](#)

For Authors from Korea

영어 논문 투고에 앞서 원어민에게 영문 교정을 받고자 하시는 분들께 Edanz 회사를 소개해 드립니다. 서비스 내용, 가격 및 신청 방법 등에 대한 자세한 사항은 저희 Edanz Editing Global 웹사이트를 참조해 주시면 감사하겠습니다.

- [Edanz Editing Global](#)

Ethical Responsibilities of Authors

This journal is committed to upholding the integrity of the scientific record. As a member of the Committee on Publication Ethics (COPE) the journal will follow the COPE guidelines on how to deal with potential acts of misconduct.

Authors should refrain from misrepresenting research results which could damage the trust in the journal, the professionalism of scientific authorship, and ultimately the entire scientific endeavour. Maintaining integrity of the research and its presentation can be achieved by following the rules of good scientific practice, which include:

- The manuscript has not been submitted to more than one journal for simultaneous consideration.
- The manuscript has not been published previously (partly or in full), unless the new work concerns an expansion of previous work (please provide transparency on the re-use of material to avoid the hint of text-recycling (“self-plagiarism”).
- A single study is not split up into several parts to increase the quantity of submissions and submitted to various journals or to one journal over time (e.g. “salami-publishing”).
- No data have been fabricated or manipulated (including images) to support your conclusions
- No data, text, or theories by others are presented as if they were the author’s own (“plagiarism”). Proper acknowledgements to other works must be given (this includes material that is closely copied (near verbatim), summarized and/or paraphrased), quotation marks are used for verbatim copying of material, and permissions are secured for material that is copyrighted.

Important note: the journal may use software to screen for plagiarism.

- Consent to submit has been received explicitly from all co-authors, as well as from the responsible authorities - tacitly or explicitly - at the institute/organization where the work has been carried out, **before** the work is submitted.
- Authors whose names appear on the submission have contributed sufficiently to the scientific work and therefore share collective responsibility and accountability for the results.

In addition:

- Changes of authorship or in the order of authors are not accepted **after** acceptance of a manuscript.
- Requesting to add or delete authors at revision stage, proof stage, or after publication is a serious matter and may be considered when justifiably warranted. Justification for changes in authorship must be compelling and may be considered only after receipt of written approval from all authors and a convincing, detailed explanation about the role/deletion of the new/deleted author. In case of changes at revision stage, a letter must accompany the revised manuscript. In case of changes after acceptance or publication, the request and documentation must be sent via the Publisher to the Editor-in-Chief. In all cases, further documentation may be required to support your request. The decision on accepting the change rests with the Editor-in-Chief of the journal and may be turned down. Therefore authors are strongly advised to ensure the correct author group, corresponding author, and order of authors at submission.
- Upon request authors should be prepared to send relevant documentation or data in order to verify the validity of the results. This could be in the form of raw data, samples, records, etc.

If there is a suspicion of misconduct, the journal will carry out an investigation following the COPE guidelines. If, after investigation, the allegation seems to raise valid concerns, the accused author will be contacted and given an opportunity to address the issue. If misconduct has been established beyond reasonable doubt, this may result in the Editor-in-Chief's implementation of the following measures, including, but not limited to:

- If the article is still under consideration, it may be rejected and returned to the author.
- If the article has already been published online, depending on the nature and severity of the infraction, either an erratum will be placed with the article or in severe cases complete retraction of the article will occur. The reason must be given in the published erratum or retraction note.
- The author's institution may be informed.

Compliance with Ethical Standards

To ensure objectivity and transparency in research and to ensure that accepted principles of ethical and professional conduct have been followed, authors should include information regarding sources of funding, potential conflicts of interest (financial or non-financial), informed consent if the research involved human participants, and a statement on welfare of animals if the research involved animals.

Authors should include the following statements (if applicable) in a separate section entitled "Compliance with Ethical Standards" when submitting a paper:

- Disclosure of potential conflicts of interest
- Research involving Human Participants and/or Animals
- Informed consent

Please note that standards could vary slightly per journal dependent on their peer review policies (i.e. single or double blind peer review) as well as per journal subject discipline. Before submitting your article check the instructions following this section carefully. The corresponding author should be prepared to collect documentation of compliance with ethical standards and send if requested during peer review or after publication. The Editors reserve the right to reject manuscripts that do not comply with the above-mentioned guidelines. The author will be held responsible for false statements or failure to fulfill the above-mentioned guidelines.

Disclosure of potential conflicts of interest

Authors must disclose all relationships or interests that could influence or bias the work. Although an author may not feel there are conflicts, disclosure of relationships and interests affords a more transparent process, leading to an accurate and objective assessment of the work. Awareness of real or perceived conflicts of interests is a perspective to which the readers are entitled and is not meant to imply that a financial relationship with an organization that sponsored the research or compensation for consultancy work is inappropriate. Examples of potential conflicts of interests **that are directly or indirectly related to the research** may include but are not limited to the following:

- Research grants from funding agencies (please give the research funder and the grant number)
- Honoraria for speaking at symposia
- Financial support for attending symposia
- Financial support for educational programs
- Employment or consultation
- Support from a project sponsor
- Position on advisory board or board of directors or other type of management relationships
- Multiple affiliations
- Financial relationships, for example equity ownership or investment interest
- Intellectual property rights (e.g. patents, copyrights and royalties from such rights)
- Holdings of spouse and/or children that may have financial interest in the work

In addition, interests that go beyond financial interests and compensation (non-financial interests) that may be important to readers should be disclosed. These may include but are not limited to personal relationships or competing interests directly or indirectly tied to this research, or professional interests or personal beliefs that may influence your research. The corresponding author collects the conflict of interest disclosure forms from all authors. In author collaborations where formal agreements for representation allow it, it is sufficient for the corresponding author to sign the disclosure form on behalf of all authors. Examples of forms can be found

- [here:](#)

The corresponding author will include a summary statement **on the title page that is separate from their manuscript**, that reflects what is recorded in the potential conflict of interest disclosure form(s).

See below examples of disclosures:

Funding: This study was funded by X (grant number X).

Conflict of Interest: Author A has received research grants from Company A. Author B has received a speaker honorarium from Company X and owns stock in Company Y. Author C is a member of committee Z.

If no conflict exists, the authors should state:

Conflict of Interest: The authors declare that they have no conflict of interest.

After acceptance

Upon acceptance of your article you will receive a link to the special Author Query Application at Springer's web page where you can sign the Copyright Transfer Statement online and indicate whether you wish to order OpenChoice, offprints, or printing of figures in color.

Once the Author Query Application has been completed, your article will be processed and you will receive the proofs.

Open Choice

In addition to the normal publication process (whereby an article is submitted to the journal and access to that article is granted to customers who have purchased a subscription), Springer provides an alternative publishing option: Springer Open Choice. A Springer Open Choice article receives all the benefits of a regular subscription-based article, but in addition is made available publicly through Creative Commons Attribution-NonCommercial 4.0 International License.

- [Springer Open Choice](#)

Copyright transfer

Authors will be asked to transfer copyright of the article to the Publisher (or grant the Publisher exclusive publication and dissemination rights). This will ensure the widest possible protection and dissemination of information under copyright laws.

Open Choice articles do not require transfer of copyright as the copyright remains with the author. In opting for open access, the author(s) agree to publish the article under the Creative Commons Attribution License.

- [Creative Commons Attribution-NonCommercial 4.0 International License](#)

Offprints

Offprints can be ordered by the corresponding author.

Color illustrations

Online publication of color illustrations is free of charge. For color in the print version, authors will be expected to make a contribution towards the extra costs.

Proof reading

The purpose of the proof is to check for typesetting or conversion errors and the completeness and accuracy of the text, tables and figures. Substantial changes in content, e.g., new results, corrected values, title and authorship, are not allowed without the approval of the Editor.

After online publication, further changes can only be made in the form of an Erratum, which will be hyperlinked to the article.

Online First

The article will be published online after receipt of the corrected proofs. This is the official first publication citable with the DOI. After release of the printed version, the paper can also be cited by issue and page numbers.”

3.2. Manuscript

Clarifying the Factor Structure of the Mental Health Continuum Short Form in Three
Languages: A Bifactor Model Approach

Lusilda Schutte, Marié P. Wissing and Suria M. Ellis

Author note

Lusilda Schutte, Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa; Marié P. Wissing, Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa; Suria M. Ellis, Unit for Business Mathematics and Informatics, North-West University, South Africa.

Correspondence concerning this article should be addressed to Lusilda Schutte, Africa Unit for Trans-disciplinary Health Research, North-West University, Potchefstroom Campus, South Africa, 2531. Email: Lusilda.Schutte@nwu.ac.za. Tel: +27 18 299 1104.

Acknowledgements

The research was financially supported by the National Research Foundation (NRF) of South Africa by funding for the FORT3 research project granted to the second author (grant numbers 85649 and 91557), and for a doctoral scholarship granted to the first author (grant number 84376). Opinions expressed and conclusions arrived are those of the authors and are not necessarily to be attributed to the NRF. Funding was also provided in the form of a doctoral scholarship to the first author by the North-West University, South Africa. The authors wish to thank I. P. Khumalo and L. Temane for their assistance with the translation of the questionnaire and gathering of the data. The study forms part of the doctoral thesis of the first author.

Conflict of interest

The authors declare that they have no conflict of interest.

Abstract

The Mental Health Continuum Short Form was developed to measure and diagnose overall positive mental health, a key indicator of quality of life. The aim of the present study was to explore the psychometric properties and measurement invariance of three language versions of the scale when administered in diverse cultural groups, in particular whether a bifactor model exhibited superior fit in comparison with other commonly used models. Confirmatory factor analysis was applied to analyse data from South African students ($N = 1060$) who completed either the English ($n = 324$), Afrikaans ($n = 478$), or Setswana ($n = 258$) version of the scale. A bifactor model consistently displayed superior fit, and it was shown that the scale total score can be used as a reliable indication of overall positive mental health, but that it is not appropriate to interpret and calculate subscale scores. The 'social contribution' item appeared to target a different aspect (contribution towards a greater good) than the other items (all self-orientated) and removal of the item significantly improved the model fit. This finding highlighted the lack of focus on contribution towards a greater good beyond the self in the present scale and, possibly, well-being literature in general, indicating the necessity of more attention to this important aspect of functioning well in future research. The scale exhibited configural, partial metric, and partial scalar invariance across the three samples and the average scale total score did not differ practically significantly across the groups.

Keywords: positive mental health, cross-cultural equivalence, scale validation, structural validity, translation of measurement instruments

Clarifying the Factor Structure of the Mental Health Continuum Short Form in Three Languages: A Bifactor Model Approach

The assessment of mental health should cover the whole continuum from pathology to functioning well in order to be comprehensive and holistic. However, traditionally psychological health was often considered to merely be the absence of psychological disease, disorder, or malfunctioning (Keyes 2002). In reaction to this mainly one-sided focus, Keyes (2002) proposed that mental health and mental illness run along two related, but distinct, continua. He suggested that the absence of mental illness is not equal to the presence of mental health and the aim of mental health programmes should not only be to reduce mental illness, but also to promote and maintain positive mental health (also called flourishing, Keyes, 2007). This two-continua model of mental health was empirically supported in multiple studies (e.g., Keyes 2006; Keyes et al. 2008; Lamers et al. 2011). A recent upsurge of research on positive mental health addressed the previous neglect of this continuum.

In the conceptualisation of positive mental health, two main streams of thought can be distinguished, namely hedonic and eudaimonic perspectives on well-being. The hedonic tradition focuses on satisfaction with life, the presence of positive affect, and the absence of negative effect, and is often associated with happiness and feeling good (cf., Diener 1984; Kahneman et al. 1999; Keyes 2006). Eudaimonic well-being, on the other hand, focuses on positive functioning in life that results from developing capacities and abilities that enable one to become a more fully functioning person and citizen (Keyes 2006). Different psychological theories are associated with the eudaimonic tradition. One prominent example is Ryff's (1989b, 1989a) psychological well-being theory which distinguishes six dimensions of psychological well-being, namely self-acceptance, positive relations with others, autonomy, environmental mastery, purpose in life, and personal growth. Another prominent theory that relates to eudaimonic well-being is Keyes's (1998) theory of social well-being,

that defines social well-being as the assessment of one's functioning and circumstance in society and proposes that social well-being comprises five dimensions, namely social integration, social contribution, social coherence, social actualization, and social acceptance. Aspects of both hedonic and eudaimonic well-being showed negative correlations with psychological problems, such as depression, insomnia, and anxiety (cf., Vázquez et al. 2015; Steger 2012; Waterman et al. 2010).

One of the important scientific endeavours that will enhance the understanding and promotion of positive mental health involves the operationalisation of relevant constructs in the form of measurement scales. Measurement instruments can either target specific aspects of positive mental health, or assess the broader construct. One of the most widely used instruments that attempts to measure the broad construct of overall positive mental health is Keyes's Mental Health Continuum (Keyes et al. 2008; Keyes 2002). This scale measures both hedonic well-being or feeling good (labelled emotional well-being) in terms of positive affect and satisfaction with life, and eudaimonic well-being or functioning well in terms of Ryff's six dimensions of psychological or personal well-being (1989b), and Keyes's five dimensions of social well-being (1998). Initially, a long form of the scale was developed (Keyes 2002), that measures hedonic (emotional) well-being with six items of positive affect and one item of life satisfaction, psychological well-being using Ryff's (1989b) measure where each of the six dimensions of psychological well-being is measured by three items, and social well-being using Keyes's (1998) measure where each of the five aspects of social well-being is assessed by three items. Later, a short form of the scale (MHC-SF) was constructed (Keyes 2006; Keyes et al. 2008; Keyes 2009) that uses two items of positive affect and one satisfaction with life item to measure hedonic (emotional) well-being, and one item representing each of the six dimensions of psychological well-being together with one item representing each aspect of social well-being to measure eudaimonic well-being. Keyes et al.

(2008) suggested that the MHC-SF can be used to reach a diagnosis of mental health, where high scores on at least one hedonic well-being item and six eudaimonic well-being items result in a diagnosis of flourishing, a low score on at least one hedonic well-being item and six eudaimonic well-being items lead to a diagnosis of languishing, and individuals who do not fit the criteria for flourishing or languishing are diagnosed as being moderately mentally healthy.

The validity and reliability of the MHC-SF have been studied in various contexts, cultures and translations. Noticeably, the factorial structure of the scale has been a contentious issue in these studies. Most of the authors compared the fit of a three-factor first-order model (distinguishing between emotional, social, and psychological well-being), with the fit of a single-factor first-order model (specifying only overall positive mental health as a factor) and a two-factor first-order model (distinguishing between hedonic and eudaimonic well-being), and found the three-factor model to be superior. Examples of such studies include Keyes (2006) who administered a 12-item English version of the scale, where two dimensions of psychological well-being (self-acceptance and purpose in life) were not included, to an American adolescent sample; Keyes et al. (2008) who administered the Setswana version of the scale to a Setswana-speaking South African community sample using a structured interview format; Lamers et al. (2011), who administered the Dutch version of the MHC-SF to a representative sample of Dutch adults in an online study; Karaś, Ciecuch and Keyes (2014) who administered the Polish adaptation of the MHC-SF to a diverse Polish sample; and Lim (2014) who administered the Korean version of the scale to a South Korean adolescent sample. Although all of these studies concluded that the three-factor first-order structure adequately fitted the data, fit indexes were mostly indicative of only marginally acceptable fit. Reliability indexes of the subscales were also not always sufficient. Petrillo, Capone, Caso, and Keyes (2015) explored how well a three-factor second-order structure, as

well as a three-factor first-order model fitted the data, and fit indexes were, once again, indicative of only fair fit. In summary, although multiple validation studies claimed support for the three-factor first- and/or second-order models, the fit was in matter of fact not convincingly good.

The lack of good fit of the three-factor first- and second-order models was only recently addressed by Jovanovic (2015) and de Bruin and du Plessis (2015) who independently showed that an orthogonal bifactor model outperformed the other models for the MHC-SF in Serbian student and adult, and South African student samples, respectively. A bifactor model stipulates that the covariance among a set of items can be accounted for by a general factor that reflects the common item response variance among all the items in a measurement instrument and specific factors that reflect the common item response variance among subsets of items, such as the items comprising subscales. The general factor denotes the broad target dimension that the scale intends to measure, while the specific factors denote the conceptually constricted sub-dimensions and represent the factors that may explain common variance not accounted for by the general factor. The bifactor model is particularly appropriate for multidimensional scales that measure a general broad construct while specifying multiple conceptually distinct sub-constructs, such as the MHC-SF that measures overall positive mental health (general factor), with emotional, social, and psychological well-being as sub-dimensions (specific factors). In the orthogonal bifactor model it is assumed that all factors are orthogonal (Reise et al. 2010; Reise 2012). One of the recommendations of both Jovanovic and de Bruin and du Plessis (2015) was that studies are needed to explore whether their findings of the good fit of the bifactor model replicate in other samples.

When a scale is valid and reliable in one context, it does not automatically mean that the scale will also have good psychometric properties in another context or translation. For

example, the typical behaviours that represent the target construct may differ across different contexts (construct bias), response styles may differ for different groups (a form of method bias), or items may be poorly translated (item bias, Van de Vijver and Leung 2011).

Therefore, scales need to be validated in every context and translation that it is applied. In addition, evidence of validity of the scale in different contexts does not imply that the scores can just be compared across the contexts. Before this can be done, the measurement invariance of the scale across the contexts must be established. Measurement invariance involves that people across different groups interpret the items, as well as the underlying construct, in the same way. Once measurement invariance was established, the means, causes, and consequences of the latent factor scores can be compared across the groups (Van de Schoot et al. 2012).

The Present Study

In the light of the need for the replication of the findings of Jovanovic (2015) and de Bruin and du Plessis (2015) that the bifactor model fitted the MHC-SF considerably better than previously used models, as well as the need for cross-cultural adaptation and the assessment of measurement invariance of measurement scales, the objectives of the present study were as follows: (1) to examine the validity of the English, Afrikaans, and Setswana versions of the MHC-SF among South African students; (2) to examine whether recent findings on the bifactor structure of the MHC-SF replicate in these three groups; (3) to explore the cross-group invariance of the scale; and (4) if scalar invariance was found, to compare the average well-being scores across the three groups.

Method

Design and Participants

A cross-sectional survey design was implemented in this study. Participants ($N = 1060$) were students from three campuses of a South African university. English ($n = 324$),

Afrikaans ($n = 478$), and Setswana ($n = 258$) versions of the research battery were administered to students fluent in each language, respectively. Participants were asked to complete the scales in their native language, should it be Afrikaans, English, or Setswana (main languages in the area of data gathering), or in the language of choice if none of these was their native language. The demographic profile of the participants is shown in Table 1.

Since the fluency of the participants in the language of the test battery can play such a significant role in the validity of the results, the issue of participants' competence in the language of the scale administration warrants some elaboration. For the Afrikaans version, the majority of the participants (99.2%) indicated that Afrikaans was their native language. For the English version, only 18.2% of the participants indicated that English was their native language, while 6.8% selected Afrikaans, 18.8% selected Setswana, and 54.9% chose "Other", which was probably one of the other eight official South African languages. From the participants who completed the Setswana version of the research battery, 66.7% indicated Setswana as their native language, 0.8% selected Afrikaans, 22.1% English, and 9.3% "Other". Since the language of tuition at the university where the data was gathered was English (two campuses) or mainly Afrikaans (one campus), we reasoned that all participants should be fluent in either English or Afrikaans, based on the assumption that students should be fluent in the language in which they receive tertiary education. Since all participants had the choice to complete the battery in English, Afrikaans, or Setswana, no participant was obliged to complete the battery in a language that was not either their home language or the language of tuition at the tertiary institution where they studied. We therefore considered it reasonable to assume sufficient fluency in the language in which the participant chose to complete the questionnaires. In the remainder of this paper, the terms "English sample", "Afrikaans sample", and "Setswana sample" will be used to refer to the groups that completed the English, Afrikaans, and Setswana scales, respectively.

Measures

Participants completed a research battery containing various scales of psychological wellbeing and illness. The scales that are relevant to this paper will now be introduced.

Socio-demographic questionnaire. A questionnaire on relevant socio-demographic information was administered.

Mental Health Continuum Short Form (MHC-SF) (Keyes 2006; Keyes et al. 2008; Keyes 2009). The MHC-SF is a 14-item measure of positive overall mental health that consists of three items of emotional well-being (EWB), five items of social well-being (SWB), and six items of psychological well-being (PWB). Respondents report how frequently each of the 14 statements occurred in the past month using a 6-point Likert scale ranging from 0 (*never*) to 5 (*almost every day*).

Procedure

Preparation of scales. A research committee approach was used to translate the MHC-SF from English to Afrikaans and Setswana (Van de Vijver and Leung 1997). The scale was translated into each of the target languages by one translator and back-translated into English by another translator. The original and back-translated English versions were then compared by research committees comprising native Afrikaans and Setswana speaking academics, respectively, who were also fluent in English (Van de Vijver and Leung 1997; Brislin 1973). Preliminary evaluation of the translated scales involved the administration thereof to small pilot samples from the Afrikaans and Setswana speaking populations, respectively, to determine whether the items were comprehensible and culturally appropriate.

Data gathering. Authorities from the data gathering sites were approached to get permission and make practical arrangements regarding data gathering. Volunteers who were willing to participate in the study were recruited by their lecturers during class periods. Any

student who was willing to participate was asked to provide written informed consent, where after they completed the research battery in the classroom or at home.

Data Analysis

Confirmatory factor analysis was implemented using Mplus 7.31 (Muthén and Muthén 1998-2014). As the data showed slight deviations from normality (skewness-values ranged from -1.32 to 0.25 for the English group, from -1.07 to 0.15 for the Afrikaans group, and from -1.33 to -0.11 for the Setswana group; kurtosis-values ranged from -1.02 to 1.64 for the English group, from -0.85 to 1.78 for the Afrikaans group, and from -0.82 to 1.85 for the Setswana group), the robust maximum likelihood (MLR) estimator was applied to accommodate any deviations from normality. The models were parameterized by fixing the variances of the latent factors to one (Van de Schoot et al. 2012). Full information maximum likelihood estimation was applied to deal with missing data.

Prior to assessing measurement invariance across the three samples we established a best-fitting baseline model for each group. This involved comparing the fit of different models, as suggested by previous research and the theory underlying the scale, for each group. The following models were tested: a single-factor first-order model, where all items load on one underlying factor of overall positive mental health (Model 1); a two-factor first-order model, where a hedonic well-being factor (comprising items 1-3 of the EWB subscale) and a eudaimonic well-being factor (comprising items 4-8 of the SWB subscale and items 9-14 of the PWB subscale) are modelled and allowed to correlate (Model 2); a three-factor first-order model, where the three subscales form three correlated factors (Model 3); a three-factor second-order model, where the items load on the three subscales to form three first order factors which, in turn, load on overall mental health as a second order factor (Model 4); and a bifactor model, where the items of the subscales load on the three dimensions of well-

being, respectively, as well as on the general mental health factor (Model 5). Graphical representations of these models are depicted in Figure 1.

In order to determine how well each model fitted the data, we used model fit indexes from different classes based on guidelines provided by Byrne (2012). The comparative fit index (CFI) and Tucker-Lewis fit index (TLI) are reported from the class of incremental fit indexes. Earlier, a cut-off value of 0.90 was suggested, with values larger than the cut-off value being indicative of good model fit. More recently, the cut-off value was revised to be close to 0.95. The root mean square error of approximation (RMSEA) with its 90% confidence interval and standardized root mean square residual (SRMR) are reported from the absolute fit indexes class. For the RMSEA, values smaller than 0.05 indicate good fit and values of 0.08 or less suggest reasonable fit. Small values of the SRMR (e.g., less than 0.05) indicate good model fit. In addition, we report the chi-square test statistic with its number of degrees of freedom and corresponding *p*-value for the sake of completeness (small *p*-values for this test would suggest that the hypothesized model does not fit the data well), but due to its sensitivity to sample size, this result is not used for decision making. If the model fit indexes suggested that the model did not adequately fit the data, we were guided by large modification index (MI) values to highlight potential revisions of the model that may lead to improved model fit, given that these revisions were also theoretically justifiable.

Some additional indexes of unidimensionality and reliability were considered for the bifactor model. First, the explained common variance (ECV) attributable to the general factor (MH) and the specific factors (EWB, SWB, and PWB) were considered as an indication of the unidimensionality of the scale, and therefore the appropriateness of a total scale score (Bentler 2009). The ECV of the general factor is calculated as the variance explained by the general factor divided by the variance explained by the general plus the specific factors. In parallel, the ECV of a specific factor is defined as the variance explained by the specific

factor divided by the variance explained by the general plus specific factors (Reise 2012). Second, coefficient omega hierarchical was calculated as an indicator of reliability. This index is superior to Cronbach's alpha when the data are multidimensional (Reise et al. 2010). The total score omega hierarchical coefficient estimates the proportion of item response variance that is due to the general factor, while the omega hierarchical coefficient for a specific factor estimates the proportion of item response variance that is due to the specific factor after accounting for the general factor. Coefficient omega hierarchical indicates whether it is reasonable to calculate total scale and subscale scores (Reise 2012).

Once a sufficiently fitting baseline model was established for each group, we assessed the measurement invariance across the three groups. In this study we were interested in examining three increasingly stringent levels of measurement invariance, as described by Byrne and van de Vijver (2010). First, we examined configural invariance, which requires the number of factors and the configuration of their factor loadings (i.e., the pattern of freely estimated and fixed parameters) to be the similar across the groups, but the parameters are not constricted to be equal. The configural model should fit the data well, because it becomes the baseline against which subsequent models (all nested within the configural model) are compared. Second, we examined metric invariance, which involves that the factor loadings across the groups are constrained to be equal. If there is evidence that particular factor loadings are inequivalent across the groups (as indicated by large modification indexes), these loadings may be allowed free estimation to obtain partial metric equivalence. Third, we examined scalar invariance, which involves that the factor loadings and item intercepts are constrained to be equal across the groups. Once again, partial scalar invariance can be attained by allowing intercepts that show evidence of inequivalence (based on large modification indexes) to be estimated freely, while constricting the other intercepts to be equal across the groups. Once scalar invariance has been established, the latent factor means

can be compared across the groups (Van de Vijver and Leung 2011). In order to determine whether each next level of invariance holds, the likelihood ratio test (difference in chi-square between nested models) is often used. However, this test is largely dependent on sample size and invariance will mostly be rejected if the sample sizes are large (Cheung and Rensvold 2002). Cheung and Rensvold (2002) and Chen (2007) suggested that the difference in the CFI- and RMSEA-values of the nested tests (ΔCFI and ΔRMSEA , respectively) should rather be used, where $|\Delta\text{CFI}|$ smaller than .01 and $|\Delta\text{RMSEA}|$ smaller than 0.015 are indicative of measurement invariance. In addition, the Akaike's information criterion (AIC) will be reported, where smaller values indicate better fit (Byrne 2012).

Ethical Considerations

Data for this study were gathered as part of the FORT 3 research-project of Wissing (2008/2012), titled "The prevalence of levels of psychosocial health: dynamics and relationships with biomarkers of (ill) health in South African social context". The project was approved by the Ethics Committee of the North-West University, with project number NWU 00002-07-A2. All participants gave written informed consent prior to participating in the study and participation was voluntary. Data gathering and capturing were done anonymously.

Results

Fitting the Baseline Models

Table 2 presents the results of the baseline models, where the MHC-SF was analysed using CFA on each of the groups separately. The single-factor and two-factor first-order models (Model 1 and 2, respectively) displayed poor fit to the data, with the two-factor model performing slightly better than the single-factor model. The first- and second-order three-factor models (Model 3 and 4, respectively) fitted the data better, but the fit was still insufficient for the English sample and marginal for the Setswana sample. The inter-factor correlations of Model 3 were large for the English, Afrikaans, and Setswana samples (.61,

.65, and .59, respectively, for EWB with SWB; .79, .77, and .67, respectively, for EWB with PWB; and .66, .65, and .61, respectively, for SWB with PWB), suggesting the presence of a general mental health factor which points towards the potential appropriateness of a second-order or bifactor model. The bifactor model (Model 5) performed better than any of the other models for the English and Afrikaans samples, attaining marginal fit for the English sample and good fit for the Afrikaans sample¹. For the Setswana sample, the fit of the bifactor model was comparable with the fit of the first- and second-order three-factor models (i.e., marginal).

Since the fit of the bifactor model was still only marginal for the English and Setswana groups, we examined the modification indexes provided for these models. For both the English and Setswana groups, the highest modification index suggested that allowing the residuals of item 4 (social contribution: “In the past month, how often did you feel that you had something important to contribute to society?”) and item 5 (social integration: “In the past month, how often did you feel that you belonged to a community [like a social group, your neighbourhood, or city]?”) to correlate will improve the model fit. In addition, the R^2 – value of item 4 was the smallest of all items’ R^2 – values for the English and Setswana groups (.165 and .235, respectively). We therefore explored two avenues of improving the model: allowing the residuals of items 4 and 5 to correlate (Model 6); and removing item 4 (Model 7). Although the Afrikaans sample’s model fit was adequate for the bifactor model and the problems with items 4 and 5 were only evident for the English and Setswana samples, we

¹ For the Afrikaans sample, the residual variance of item 1 was negative for Models 5, 6, and 7, which resulted in a residual covariance (theta) matrix that was not positive definitive. A similar problem occurred for the English sample where Model 6 resulted in a negative residual variance for item 13, yielding a non-positive definite residual covariance (theta) matrix. We addressed these problems by fixing the residual variances of these items to zero in the respective models (F. Chen et al. 2001).

also explored what the impact would be on the model fit if we would apply the same changes to the Afrikaans sample. The results of these analyses are shown in Table 2.

For the English group, Models 6 and 7 showed a significant improvement over Model 5 and the fit of Models 6 and 7 were relatively similar (with Model 6 outperforming Model 7 in terms of some of the fit indexes and vice versa). For the Afrikaans sample, the fit of Model 6 was similar to the fit of Model 5, but Model 7 had a consistently better fit on all fit indexes when compared to all other models. For the Setswana sample, the fit of Models 6 and 7 were substantially better than the fit of Model 5, and Model 7 outperformed Model 6 slightly on all fit indexes. Based on the consistently similar or superior performance of Model 7 when compared to all other models, this model was used in the measurement invariance tests.

The Bifactor Model

The factor loadings of Model 7 are shown in Table 3 for all samples. Note that the factor loadings of the general factor (MH) on the items were significant for all items across all groups. The factor loadings of the specific factors (EWB, SWB, and PWB) were significant for the Afrikaans group (except for one SWB item, namely item 5 [social integration]). For the English and Setswana groups, the factor loadings of the EWB and SWB factors were significant for all items, but the factor loadings of the PWB factor were significant for only two items in the English group (items 12 [personal growth] and 14 [purpose in life]) and no items in the Setswana group. Most of the items from the EWB and PWB factors had higher loadings on the general factor (except for item 1 [happiness] in all samples, item 3 [satisfaction] in the Setswana sample, and item 12 [personal growth] in the English and Setswana samples) than on the specific factors. In contrast, most items from the SWB factor had higher loadings on the specific factor (except for item 5 [social integration] which loaded the highest on the general factor across all samples).

The general mental health factor explained 66.3%, 63.9%, and 64.6% of the common variance in the English, Afrikaans, and Setswana groups, respectively, while the EWB factor accounted for 10.1%, 9.1%, and 13.5% of the common variance, the SWB factor for 15.5%, 13.8%, and 15.6% of the common variance, and the PWB factor for 8.2%, 13.2%, and 6.4% of the common variance in the English, Afrikaans, and Setswana groups, respectively.

Clearly, the general factor accounted for a significantly larger portion of the variance of the scale than the three specific factors of positive mental health.

The omega hierarchical coefficients of the general mental health factor were 0.79, 0.78, and 0.78 for the English, Afrikaans, and Setswana samples, respectively, showing that there exists a strong and reliable general well-being factor that influences the variation across all 13 items (recall that item 4 was removed in Model 7 which was used in these analyses). For the specific factors, the omega hierarchical coefficients for the EWB subscale were .27, .18, and .40, for the SWB subscale .39, .35, and .41, and for the PWB subscale .09, .25, and .00 for the English, Afrikaans, and Setswana samples, respectively, suggesting that only a small proportion of the subscale item response variance is attributable to the specific components. This implies that the degree to which the MHC-SF subscales reliably measure the specific variances of EWB, SWB, and PWB is low. Based on standard convention which requires reliability estimates to be above .70 (Nunnally 1978), the total scale had an adequate reliability, while the subscales did not measure the specific variances of EWB, SWB, and PWB reliably.

Measurement Invariance

We assessed the measurement invariance of Model 7. The findings for each step in the invariance testing process are summarised in Table 4 and will now be presented.

Configural invariance. We started our invariance testing by evaluating configural invariance for Model 7 (Invariance Model 1 in Table 4). The fit indices indicated that the configural model adequately fitted the data.

Metric invariance. Large $|\Delta\text{CFI}|$ and $|\Delta\text{RMSEA}|$ values indicated that Model 7 did not show sufficient metric invariance (Invariance Model 2A). We examined the modification indexes and, based on these, allowed the factor loading of the general mental health factor on item 14 (purpose in life) to be estimated freely in all groups (Invariance Model 2B). This model then yielded sufficient fit as indicated by a small $|\Delta\text{CFI}|$ and $|\Delta\text{RMSEA}|$ values indicating partial metric invariance.

Scalar invariance. Next, we assessed the partial scalar invariance of Model 7 (Invariance Model 3A), where the factor loading of the general mental health factor on item 14 was still allowed to be estimated freely in all groups because this loading was shown to be inequivalent in the metric invariance step. The $|\Delta\text{CFI}|$ and $|\Delta\text{RMSEA}|$ values indicated that this model did not exhibit invariance. One by one we allowed the intercepts of items 2 (interest in life), 10 (environmental mastery), 1 (happiness), and 12 (personal growth) to be estimated freely in all groups as suggested by the modification indexes. This resulted in a final model (Invariance Model 3B) which showed sufficient partial scalar invariance as indicated by small $|\Delta\text{CFI}|$ and $|\Delta\text{RMSEA}|$ values².

² We also assessed the measurement invariance for Model 6. Configural invariance was attained ($\text{CFI} = .955$; $\text{RMSEA} = 0.051$), but to attain metric invariance the following factor loadings had to be allowed free estimation: the general mental health factor's loadings on items 14, 12, and 10 (in this order) and the social well-being factor's loading on item 4. Partial metric invariance was then attained ($\Delta\text{CFI} = -.010$ and $\Delta\text{RMSEA} = 0$ when compared with the configural model). To attain partial scalar invariance, the intercepts of items 5, 6, 3,

Comparison of Latent Factor Means

The English group was set as the reference group (all factor means were fixed to zero for this group) for the purpose of evaluating the differences between the latent factor mean scores across the groups. The results are shown in Table 5. The Afrikaans group had a statistically significantly higher average score on the general mental health factor than the English and Setswana groups. The English group, in turn, had a statistically significantly higher factor mean than the Setswana group at a 5% significance level, but not at a 1% significance level. For the social and psychological well-being residual factors, the Afrikaans group's scores were statistically significantly lower than the English and Setswana groups' scores. For emotional well-being, the Afrikaans group scored lower than both the English and Setswana groups, but the difference was only statistically significant in comparison with the Setswana group. The English group had statistically significantly lower scores on all residual scales than the Setswana group. None of these differences were practically significant (i.e., close to 0.5 or higher, Cohen, 1988), except for the difference between the English versus Setswana and Afrikaans versus Setswana scores on SWB.

Discussion

The objectives of the present study were to (1) examine the validity of the English, Afrikaans, and Setswana versions of the MHC-SF among South African students; (2) examine whether recent findings on the bifactor structure of the MHC-SF replicate in these three groups; (3) explore the cross-group invariance of the scale; and (4) if scalar invariance was found, compare the mean well-being scores across the three groups. We found that a bifactor model with item 4 removed outperformed all other models in this study across the three groups. In this model, the general mental health factor explained by far the greatest part

10, 11, 4, and 12 (in this order) had to be allowed free estimation in addition, resulting in $\Delta\text{CFI} = -.009$ and $\Delta\text{RMSEA} = 0.002$ when compared with the partial metric model.

of the common variance in the item responses and exhibited sufficient reliability, while the specific factors (subscales) accounted for only a small portion of the response variance and attained low reliability scores. This implies that the scale total score can be used as a reliable indicator of general mental health, but caution should be applied when interpreting and using subscale scores in subsequent analyses. Partial scalar invariance was found, which allowed for the comparison of factor mean scores. The Afrikaans group's average score on the general mental health factor was statistically significantly higher than the English and Setswana groups' scores, while the English group's scores were statistically significantly higher than the Setswana group's scores. Both the Afrikaans and English groups had statistically significantly lower scores than the Setswana group on all residual scales, while the Afrikaans group's scores were statistically significantly lower than the English group on the social and psychological well-being subscales. Considering the effect sizes, these differences were not significant in practice, except for the English and Afrikaans scores that were practically significantly lower than the Setswana scores on the social well-being residual scale. The implications of these findings will now be discussed.

Baseline Models

Bifactor model. Similar to Jovanovic (2015) and de Bruin and du Plessis (2015), results of the bifactor model in all three groups of the present study strongly supported the presence of a general factor of mental health within the items of the scale. This was evident in the fact that a large proportion of the item response variance was explained by the general factor, the high coefficient omega hierarchical score for this factor, and the significant factor loadings of the general factor on all items. Although the existence of specific factors was supported by the remarkably better fit of the bifactor model in comparison with the single-factor model, the ability of the subscales to reliably measure the specific factors was countered by the small proportion of the item response variance explained by the specific

factors, the low coefficient omega hierarchical scores for these factors and the small factor loadings of some items on the specific factors. These findings suggest that calculating and interpreting an overall mental health score based on the sum of all the MHC-SF items (item 4 excluded) is justified, but the calculation and interpretation of subscale scores are not empirically supported.

The confirmation of the existence of the general and specific factors of mental health is in line with previous studies that assessed a bifactor and/or hierarchical structure for well-being. For example, F. Chen, Jing, Hayes, and Lee (2013) confirmed a bifactor structure of well-being and found that both a general well-being factor and specific factors of psychological (eudaimonic) and subjective (hedonic) well-being exist in own right; Wissing and Temane (2008) found that well-being has a hierarchical structure, with general psychological well-being as a higher-order factor across different cultural groups, and lower-order factors which exhibited varying nuances in relatively collectivistic versus individualistic contexts; Gallagher, Lopez, and Preacher (2009) confirmed the hierarchical structure of well-being by showing that well-being has a general higher-order factor with hedonic, eudaimonic and social well-being as correlated but distinct subcomponents; and Joshanloo (2015) found that hedonic and eudaimonic aspects of well-being are correlated but distinct using exploratory structural equation modelling. In line with previous research, the present study confirmed the multidimensional nature of mental health, while suggesting a strong general mental health factor.

Despite the confirmation of a general well-being factor with an underlying multidimensional structure, the variance explained by the specific factors and the reliabilities of the subscales were very low in this study. This corresponds with the findings of Jovanovic (2015) and de Bruin and du Plessis (2015) when they applied the bifactor model to the MHC-SF and implies that the subscale scores cannot be trusted and should not be used, for

example, in regression analyses. This is in contrast with F. Chen et al. (2013) who, in their study of the suitability of a bifactor structure for well-being, found that the subdimensions of well-being were clearly distinct when their overlap with the general mental health factor was partialled out. The main difference between the study of F. Chen et al. (2013) and the studies using the MHC-SF, is that F. Chen and colleagues used separate and more comprehensive measurement instruments to assess the specific factors of mental health, while the MHC-SF relies on very brief subscales (three to six items each) to assess the specific factors, where the subcomponents of the specific factors are mostly measured by a single item. We conclude that, although short scales like the MHC-SF may yield usable and reliable scores for the general mental health factor, more comprehensive scales are needed to reliably measure the subdimensions of the general construct.

Comparison of fit of the baseline models across the groups. It is noticeable that the English group performed worse than the other two groups on almost all baseline models that we explored in this study. In addition, the Afrikaans group's fit consistently outperformed the other groups' fit. This is against what one would expect, given that the English version was the original and well-validated scale, while the Setswana and Afrikaans versions were both translations of the original scale – which one may expect to lead to more issues. There are at least two possible ways to understand these findings. First, in the Afrikaans group the great majority of participants indicated Afrikaans as native language (99.2%), in the Setswana group the native language Setswana speakers were still the majority but less than in the Afrikaans group (66.7%), but in the English group the native language English speakers were by far the minority (18.8%). Although we argued that the participants should all be fluent in the language in which they completed the assessment, it seems as if native versus non-native language scale administration may have played a role in the resulting psychometric

properties. This finding highlights the potential danger of not administering scales in participants' native language.

A second way to understand the finding pertains to the cultural heritage of the groups. Afrikaans is a language that comes from Dutch and is often spoken by South Africans with a strong Western heritage. The Setswana and English samples mostly indicated African languages as native language, suggesting a strong African heritage. The fact that the Afrikaans scale displayed better psychometric properties than the English and Setswana versions may point towards the notion that theories and their operationalisations developed in Western contexts do not necessarily merely transfer to other, non-Western contexts (cf., Henrich et al. 2010). This highlights the necessity of also incorporating an emic approach in scale adaptation, where the phenomena are studied in a bottom-up manner from the perspective of members of the specific cultural group. In particular, qualitative research is needed to understand what positive mental health looks like in an African context.

Problematic items. For the MHC-SF, allowing items 4 (social contribution: "In the past month, how often did you feel that you had something important to contribute to society") and 5 (social integration: "In the past month, how often did you feel that you belonged to a community [like a social group, your neighbourhood, or city]") to covary improved the fit of the English and Setswana scales significantly. That was not the case for the Afrikaans version. On face value, items 4 and 5 seem to be different from items 6 (social actualisation: "In the past month, how often did you feel that our society is becoming a better place for people"), 7 (social acceptance: "In the past month, how often did you feel that people are basically good"), and 8 (social coherence: "In the past month, how often did you feel that the way our society works makes sense to you") which form the rest of the subscale. Whereas items 4 and 5 have a personal tone to them (personally contributing to society and personally belonging to a community), items 6 to 8 involve general opinions about society

and people in general. Allowing items 4 (social contribution) and 5 (social integration) to correlate based on empirical and substantive evidence is one way to deal with these items in future. Note that the social contribution and social integration subscales also exhibited issues in studies where the mental health continuum was administered using comprehensive scales. Joshanloo (2015) conducted exploratory structural equation modelling on the long form of the MHC-SF and reported that the social integration subscale cross-loaded on the psychological and social well-being factors, while the social contribution factor did not load strongly on social well-being, but had a large loading on psychological well-being. Also, social contribution has been found to load on psychological well-being in addition to social well-being in Spanish immigrant samples when a longer form of the MHC was administered (Bobowik et al. 2015). Clearly, the conceptualization and measurement of social contribution and social integration requires more attention in future research.

Apart from the implied residual correlation between items 4 and 5, item 4 attained a low R^2 -value for both the English and Setswana groups. Removal of item 4 significantly improved the fit in all groups. Upon examination of the item, we noted that item 4 is the only item in the entire scale that extends the notion of well-being from a self-orientated satisfaction with oneself and the world around one towards a contribution to a greater good. More specifically, all items except item 4 is about what one feels (items 1-3), what one thinks about oneself (items 9, 10, and 12-14), or what the world has to offer one (items 5-8 and 11). Item 4 is the only item that explicitly focuses on looking beyond oneself to what one can offer the world. The fact that the item stood alone in its contribution orientation may have caused it not to fit well with the other items and, for the purpose of model fit, we removed the item in our analyses. The poor fit of item 4 alerted us to the lack of explicit focus on contribution towards others and the greater good in the conceptualisation of mental health that was operationalised in this scale. This is particularly significant since 11 of the 14 items

in this scale aim to measure eudaimonic well-being, which is all about functioning well.

Reflecting on prominent theories on eudaimonic well-being, one may realise that the focus is very often on the self and what one gains from the world as is evident in, for example, Waterman's eudaimonic identity theory and his work on personal expressiveness (Waterman 2008; Waterman and Schwartz 2013), the self-determination theory (Ryan and Deci 2000; Ryan et al. 2008; Ryan et al. 2013) and theories of meaning in life (Steger et al. 2006; Wong 2011). Although the interpersonal component features in many of these theories, it still tends to be self-orientated. While the intrapsychic component of eudaimonic well-being is undoubtedly important, a greater emphasis on the interconnectedness aspect of functioning well and, specifically, the expression thereof in contribution towards a greater good, may enhance our measurement of well-being in general and eudaimonic well-being in particular. Recent perspectives that address these aspects of well-being is the relationality-meaning model of Wissing (2014), the altruism model of Delle Fave and Soosai-Nathan (2014), and the sustainable well-being model of Helne and Hirvilammi (2015).

Measurement Invariance

In this study, sufficient configural invariance, and partial metric and scalar invariance was found. This implies that not all factor loadings and intercepts were invariant across the groups and some needed to be freed before comparing the mean scores.

Comparison of mean scores. Since the reliabilities of the residual scales were so low, we do not interpret the comparison of the subscale scores for the groups at hand. The comparison of the average overall scale scores can, however, be interpreted. Although the Afrikaans group had statistically significantly higher scores than the English group on the general factor, who, in turn, had statistically higher scores than the Setswana group, it is well-known that statistical significance is greatly influenced by sample size and when sample sizes are large statistically significant differences are often found. We therefore base our

interpretation on the effect sizes, which give an indication of the practical significance of the differences. The overall mental health of the participants who completed the English, Afrikaans, and Setswana versions of the scale were not practically significantly different, suggesting that cultural heritage as reflected in language does not seem to have a large influence on the level of mental health experienced by participants. This is very significant in a country with a history of political inequalities. In this study, the majority of the Afrikaans sample may typically come from previously advantaged groups who may still enjoy the heritage of good opportunities (supported by the standard of living levels depicted in Table 1), but, at the same time, currently experience political and economic marginalisation as a consequence of transformation actions. The majority of the English and Setswana groups may, on the other hand, typically come from previously disadvantaged groups with a heritage of fewer opportunities, but currently experience multiple opportunities as a result of the present political climate and transformation practices in the country. One may have expected that cultural heritage and the current political climate should play a significant role in the mental health of citizens from such a politicised country, but the contrary seems to be true. The fact that all participants in this group have the opportunity to study at a university may have a larger influence than their cultural heritage. Although culturally diverse, the group is quite homogeneous in terms of educational background (all should have at least secondary education), age group, life phase, marital status, and future possibilities. Future research should investigate whether other factors, such as socio-economic status, age, life phase, and so forth are associated with differences on overall well-being.

Conclusions

In order for the study of mental health to be comprehensive and holistic, measurement instruments should not only assess mental illness, but also positive mental health. The MHC-SF is an instrument that was developed to address the traditional lack of focus on positive

mental health in research. In the present study, the psychometric properties of three language versions of the MHC-SF were explored. A bifactor model showed superior fit. The general mental health factor explained the majority of the common variance and demonstrated good reliability, indicating that the scale total score can be calculated and used in subsequent analyses. The specific factors of mental health, on the other hand, explained only a small proportion of the common variance and displayed low reliability, suggesting that it is not appropriate to interpret and calculate the subscale scores. It seems as if more comprehensive measurement of the specific factors is needed to yield reliable scores. The social contribution item of the scale was shown to be problematic and removal of the item significantly improved the model fit. The item seems to target a different aspect (contribution towards a greater good) than the other items (all self-orientated) and for that reason did not fit well with the rest of the scale. This important aspect of eudaimonic well-being that involves contribution towards a greater good beyond the self is underrepresented in the MHC-SF and, possibly, eudaimonic well-being literature in general, suggesting that more attention to this aspect is recommended for future research. The scale exhibited configural, partial metric, and partial scalar invariance across samples that completed the English, Afrikaans, and Setswana versions of the scale and the average scale total score did not differ practically significantly across the groups.

Limitations and Recommendations

Although this study makes important contributions to the understanding of the structure and measurement of mental health, it is not without limitations. The conclusions regarding the fit of the bifactor model in the present contexts and the extent to which the scale total score and subscale scores are reliable and interpretable, are based on the administration of a very short instrument of overall mental health, namely the MHC-SF. In order to establish whether scores for the specific factors could actually be calculated and

interpreted in these contexts, and to confirm the bifactor structure of mental health in non-Western contexts, the study deserves replication using more comprehensive measurement of the specific factors of mental health. Although the present study had the strength that participants were culturally diverse and came from understudied cultural groups, the samples were homogeneous in the sense that all participants were university students. This limits the extent to which the results can be generalised and future studies should investigate whether the findings replicate in other contexts. In this study, the measurement scale and underlying theory come from a Western context. Although the application thereof in the present context may yield important information to help us grow in our understanding of well-being in various contexts, an emic approach to understanding the well-being construct in the present context will add further value.

References

- Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, 74(1), 137-143, doi:10.1007/S11336-008-9100-1
- Bobowik, M., Basabe, N., & Páez, D. (2015). The bright side of migration: Hedonic, psychological, and social well-being in immigrants in Spain. *Social Science Research*, 51, 189-204, doi:10.1016/j.ssresearch.2014.09.011
- Brislin, R. W. (1973). *Cross-cultural research methods*. New York, NY: John Wiley.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Byrne, B. M., & Van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, 10(2), 107-132, doi:10.1080/15305051003637306
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*, 29(4), 468-508, doi:10.1177/0049124101029004003
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464-504.
- Chen, F. F., Jing, Y., Hayes, A., & Lee, J. M. (2013). Two concepts or two approaches? A bifactor analysis of psychological and subjective well-being. *Journal of Happiness Studies*, 14(3), 1033-1068, doi:10.1007/s10902-012-9367-x
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255, doi:10.1207/S15328007SEM0902_5

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum, NJ: Hillsdale.
- De Bruin, G. P., & Du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum-short form (MHC-SF). *Psychological Reports*, 116(2), 438-446, doi:10.2466/03.02.PR0.116k20w6
- Delle Fave, A., & Soosai-Nathan, L. (2014). Meaning as inter-connectedness: Theoretical perspectives and empirical evidence. *Journal of Psychology in Africa*, 24(1), 53-68, doi:10.1080/14330237.2014.904090
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575.
- Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality*, 77(4), 1025-1050, doi:10.1111/j.1467-6494.2009.00573.x
- Helne, T., & Hirvilammi, T. (2015). Wellbeing and sustainability: A relational approach. *Sustainable Development*, 23(3), 167-175, doi:10.1002/sd.1581
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83, doi:10.1017/S0140525X0999152X
- Joshanloo, M. (2015). Revisiting the empirical distinction between hedonic and eudaimonic aspects of well-being using exploratory structural equation modeling. *Journal of Happiness Studies*, 1-14, doi:10.1007/s10902-015-9683-z
- Jovanović, V. (2015). Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. *Personality and Individual Differences*, 75, 154-159, doi:10.1016/j.paid.2014.11.026
- Kahneman, D., Diener, E., & Schwarz, N. (1999). *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage Foundation.

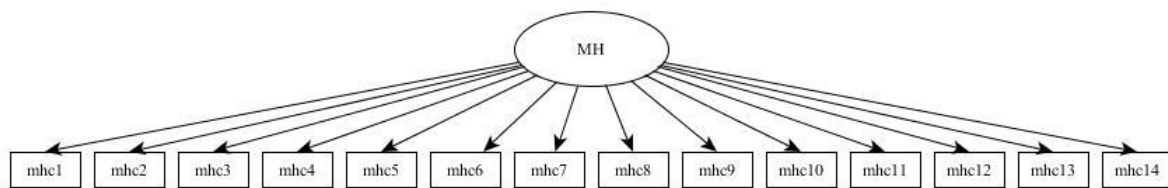
- Karaś, D., Cieciuch, J., & Keyes, C. L. M. (2014). The Polish adaptation of the Mental Health Continuum-Short Form (MHC-SF). *Personality and Individual Differences*, 69, 104-109, doi:10.1016/j.paid.2014.05.011
- Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, 61(2), 121-140.
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health & Social Behavior*, 43(2), 207-222.
- Keyes, C. L. M. (2006). The subjective well-being of America's youth: Toward a comprehensive assessment. *Adolescent and Family Health*, 4(1), 3-11, doi:10.1037/0002-9432.76.3.395
- Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. *American Psychologist*, 62(2), 95-108, doi:10.1037/0003-066X.62.2.95
- Brief description of the Mental Health Continuum Short Form (MHC-SF) (2009). <https://www.aacu.org/sites/default/files/MHC-SFEnglish.pdf>. Accessed 31 October 2015.
- Keyes, C. L. M., Wissing, M. P., Potgieter, J. P., Temane, Q. M., Kruger, A., & Van Rooy, S. (2008). Evaluation of the Mental Health Continuum–Short Form (MHC–SF) in Setswana-speaking South Africans. *Clinical Psychology & Psychotherapy*, 15(3), 181-192, doi:10.1002/cpp.572
- Lamers, S. M. A., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. M. (2011). Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF). *Journal of Clinical Psychology*, 67(1), 99-110, doi:10.1002/jclp.20741

- Lim, Y.-J. (2014). Psychometric characteristics of the Korean Mental Health Continuum-Short Form in an adolescent sample. *Journal of Psychoeducational Assessment, 32*(4), 356-364, doi:10.1177/0734282913511431
- Muthén, L. K., & Muthén, B. O. (1998-2014). *Mplus (Version 7.31)*. Los Angeles, CA: Muthén and Muthén.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Petrillo, G., Capone, V., Caso, D., & Keyes, C. L. M. (2015). The Mental Health Continuum-Short Form (MHC-SF) as a measure of well-being in the Italian context. *Social Indicators Research, 121*(1), 291-312, doi:10.1007/s11205-014-0629-3
- Reise, S. P. (2012). Invited Paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696, doi:10.1080/00273171.2012.715555
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544-559, doi:10.1080/00223891.2010.496477
- Ryan, R. M., Curren, R. R., & Deci, E. L. (2013). What humans need: Flourishing in Aristotelian philosophy and self-determination theory. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 57-75). Washington, DC: American Psychological Association.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78, doi:10.1037/110003-066X.55.1.68

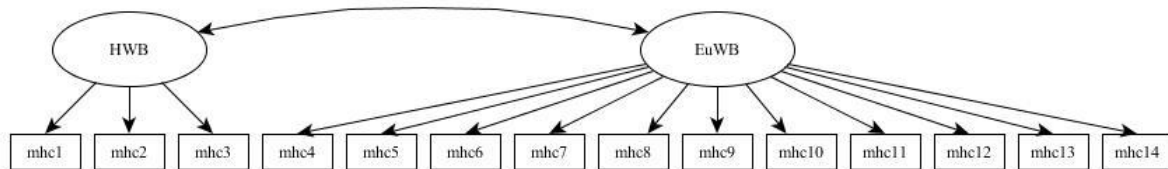
- Ryan, R. M., Huta, V., & Deci, E. L. (2008). Living well: A self-determination theory perspective on eudaimonia. *Journal of Happiness Studies*, 9(1), 139-170, doi:10.1007/s10902-006-9023-4
- Ryff, C. D. (1989a). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful ageing. *International Journal of Behavioral Development*, 12(1), 35-55, doi:10.1177/016502548901200102
- Ryff, C. D. (1989b). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069-1081, doi:10.1037/0022-3514.57.6.1069
- Steger, M. F. (2012). Experiencing meaning in life: Optimal functioning at the nexus of well-being, psychopathology, and spirituality. In P. T. P. Wong (Ed.), *The human quest for meaning: Theories, research, and applications* (2nd ed., pp. 165-184). New York, NY: Routledge.
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53(1), 80-93, doi:10.1037/0022-0167.53.1.80
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492, doi:10.1080/17405629.2012.686740
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto, & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17-45). New York, NY: Cambridge University Press.

- Vázquez, C., Rahona, J. J., Gómez, D., Caballero, F. F., & Hervás, G. (2015). A national representative study of the relative impact of physical and psychological problems on life satisfaction. *Journal of Happiness Studies*, 16(1), 135-148, doi:10.1007/s10902-014-9501-z
- Waterman, A. S. (2008). Reconsidering happiness: A eudaimonist's perspective. *The Journal of Positive Psychology*, 3(4), 234-252, doi:10.1080/17439760802303002
- Waterman, A. S., & Schwartz, S. J. (2013). Eudaimonic identity theory. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 99-118). Washington, DC: American Psychological Association.
- Waterman, A. S., Schwartz, S. J., Zamboanga, B. L., Ravert, R. D., Williams, M. K., Agocha, V. B., et al. (2010). The Questionnaire for Eudaimonic Well-Being: Psychometric properties, demographic comparisons, and evidence of validity. *The Journal of Positive Psychology*, 5(1), 41-61, doi:10.1080/17439760903435208
- Wissing, M. P. (2008/2012). The prevalence of levels of psychosocial health: Dynamics and relationships with biomarkers of (ill) health in South African social contexts (FORT3) (Research project with ethical approval registered at the North-West University).
- Wissing, M. P. (2014). Meaning and relational well-being: A reflection on the state of the art and a way forward. *Journal of Psychology in Africa*, 24(1), 183-192, doi:10.1080/14330237.2014.904100
- Wissing, M. P., & Temane, Q. M. (2008). The structure of psychological well-being in cultural context: Towards a hierarchical model of psychological health. *Journal of Psychology in Africa*, 18(1), 45-55, doi:10.1080/14330237.2008.10820170
- Wong, P. T. P. (2011). Positive psychology 2.0: Towards a balanced interactive model of the good life. *Canadian Psychology*, 52(2), 69-81, doi:10.1037/a0022511

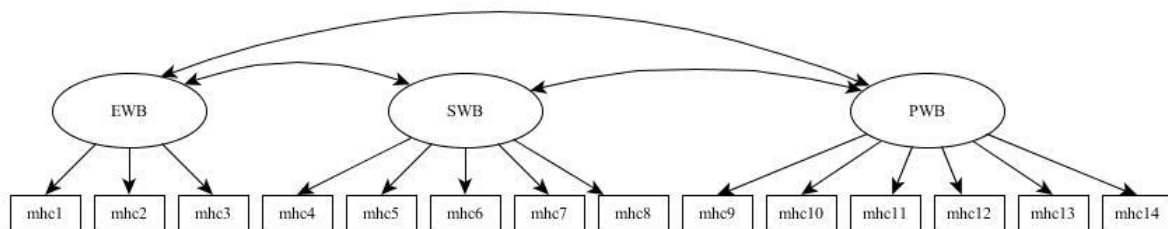
Model 1: Single-factor first-order model



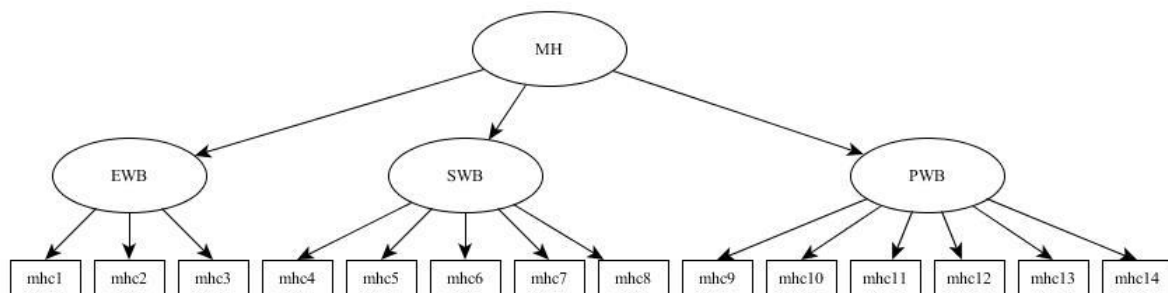
Model 2: Two-factor first-order model



Model 3: Three-factor first-order model



Model 4: Three-factor second-order model



Model 5: Bifactor model

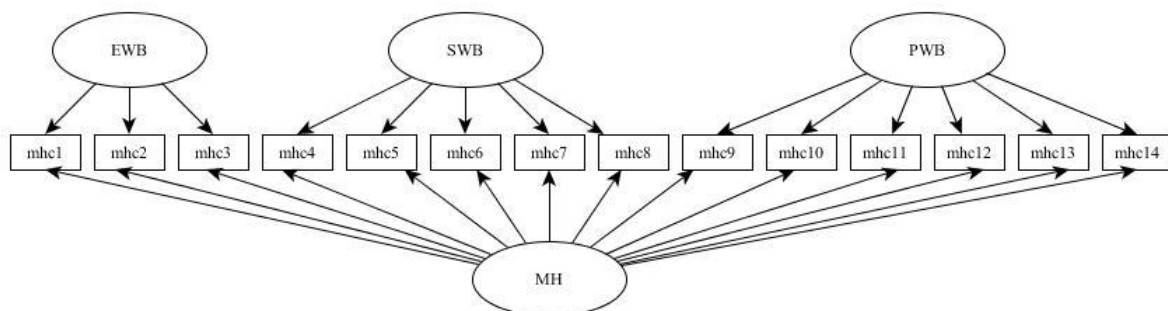


Figure 1. Competing models for the baseline structure of the MHC-SF. MH = mental health; HWB = hedonic well-being; EuWB = eudaimonic well-being; EWB = emotional well-being; SWB = social well-being; PWB = psychological well-being.

Table 1

Demographic profile of the participants

Demographic variable	All (<i>N</i> = 1060)	Eng (<i>n</i> = 324)	Afr (<i>n</i> = 478)	Set (<i>n</i> = 258)
Gender				
Male	336 (31.7%)	80 (24.7%)	171 (35.8%)	85 (32.9%)
Female	721 (68.0%)	241 (74.4%)	307 (64.2%)	173 (67.1%)
Age				
<i>M</i> (<i>SD</i>)	20.60 (3.91)	21.03 (4.10)	19.78 (3.14)	21.61 (4.60)
Range	17 – 67	18 – 54	18 – 67	17 – 46
Native language				
English	118 (11.1%)	59 (18.2%)	2 (0.4%)	57 (22.1%)
Afrikaans	498 (47.0%)	22 (6.8%)	474 (99.2%)	2 (0.8%)
Setswana	233 (22.0%)	61 (18.8%)	2 (0.4%)	172 (66.7%)
Other	204 (19.2%)	178 (54.9%)	0 (0%)	24 (9.3%)
Standard of living				
Below average	77 (7.3%)	30 (9.3%)	7 (1.5%)	40 (15.5%)
Average	786 (74.2%)	256 (79.0%)	327 (68.4%)	203 (78.7%)
Above average	192 (18.1%)	36 (11.1%)	143 (29.9%)	13 (5.0%)

Note. All = Total sample; Eng = English sample; Afr = Afrikaans sample; Set = Setswana sample; If percentages do not add to 100%, the outstanding participants did not respond to the particular demographic item.

Table 2

Confirmatory factor analysis fit statistics of the baseline models for each sample

Model	χ^2	<i>df</i>	<i>p</i>	CFI	TLI	RMSEA	90% CI	SRMR
Eng (<i>n</i> = 324)								
Model 1	317.17	77	<.001	.790	0.752	0.098	[0.087, 0.109]	0.072
Model 2	278.04	76	<.001	.824	0.789	0.091	[0.079, 0.102]	0.069
Model 3	206.11	74	<.001	.885	0.858	0.074	[0.062, 0.086]	0.059
Model 4	208.65	75	<.001	.883	0.858	0.074	[0.062, 0.086]	0.063
Model 5	167.35	63	<.001	.909	0.868	0.072	[0.058, 0.085]	0.044
Model 6 ^a	124.43	63	<.001	.946	0.923	0.055	[0.041, 0.069]	0.042
Model 7	110.49	52	<.001	.945	0.918	0.059	[0.044, 0.074]	0.038
Afr (<i>n</i> = 476)								
Model 1	456.82	77	<.001	.812	0.778	0.102	[0.093, 0.111]	0.069
Model 2	379.36	76	<.001	.850	0.821	0.092	[0.083, 0.101]	0.065
Model 3	224.89	74	<.001	.925	0.908	0.065	[0.056, 0.075]	0.054
Model 4	228.63	75	<.001	.924	0.908	0.066	[0.056, 0.075]	0.055
Model 5 ^a	151.64	64	<.001	.957	0.938	0.054	[0.043, 0.065]	0.035
Model 6 ^a	151.89	63	<.001	.956	0.937	0.054	[0.043, 0.066]	0.035
Model 7 ^a	107.64	53	<.001	.971	0.957	0.047	[0.034, 0.059]	0.030
Set (<i>n</i> = 258)								
Model 1	222.77	77	<.001	.800	0.764	0.086	[0.073, 0.099]	0.077
Model 2	182.34	76	<.001	.854	0.825	0.074	[0.060, 0.087]	0.070
Model 3	113.10	74	.002	.946	0.934	0.045	[0.027, 0.061]	0.051
Model 4	113.31	75	.003	.947	0.936	0.044	[0.026, 0.061]	0.051
Model 5	100.88	63	.002	.948	0.925	0.048	[0.030, 0.065]	0.052

Model 6	86.88	62	.020	.966	0.950	0.039	[0.016, 0.058]	0.049
Model 7	72.62	52	.031	.969	0.953	0.039	[0.012, 0.059]	0.045

Note. Eng = English sample; Afr = Afrikaans sample; Set = Setswana sample; χ^2 = chi-square test statistic; *df* = degrees of freedom; *p* = probability value; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval of the RMSEA; SRMR = standardized root mean square residual. Model 1 = single-factor first-order model; Model 2 = two-factor first-order model; Model 3 = three-factor first-order model; Model 4 = three-factor second-order model; Model 5 = bifactor model; Model 6 = bifactor model where the residual terms of items 4 and 5 were allowed to correlate; Model 7 = bifactor model with item 4 removed.

^aSee footnote 1.

Table 3

Factor loadings of Model 7 (bifactor model with item 4 deleted) for the MHC-SF in the three groups

Item	Eng (<i>n</i> = 324)		Afr (<i>n</i> = 476)		Set (<i>n</i> = 258)	
	MH	SF	MH	SF	MH	SF
1. Positive affect, EWB	.46**	.60**	.67**	.74**	.49**	.56**
2. Positive affect, EWB	.67**	.26**	.71**	.13*	.50**	.48**
3. Satisfaction with life, EWB	.66**	.42**	.71**	.19**	.46**	.50**
5. Social integration, SWB	.46**	.23**	.52**	.11	.41**	.30**
6. Social actualisation, SWB	.42**	.59**	.45**	.57**	.41**	.46**
7. Social acceptance, SWB	.46**	.52**	.49**	.54**	.37**	.54**
8. Social coherence, SWB	.42**	.52**	.46**	.55**	.43**	.56**
9. Self-acceptance, PWB	.63**	.13	.64**	.25**	.70**	.17
10. Environmental mastery, PWB	.67**	.05	.50**	.39**	.68**	.15
11. Positive relations with others, PWB	.68**	-.13	.58**	.20**	.68**	.22
12. Personal growth, PWB	.31**	.43**	.43**	.49**	.53**	-.12
13. Autonomy, PWB	.59**	.33	.53**	.49**	.70**	-.49
14. Purpose in life, PWB	.66**	.41**	.71**	.40**	.56**	-.15

Note. MHC-SF = Mental Health Continuum Short Form; Eng = English sample; Afr =

Afrikaans sample; Set = Setswana sample; MH = general mental health; SF = specific factor;

EWB = emotional well-being; SWB = social well-being; PWB = psychological well-being.

Boldface print indicates the factor (general or specific) where the factor loading was the highest.

Table 4

Measurement invariance of Model 7 (bifactor model with item 4 deleted)

Model	χ^2	<i>df</i>	<i>p</i>	CFI	RMSEA	AIC	Model comparison	$\Delta\chi^2$	<i>df</i>	<i>p</i>	Δ CFI	Δ RMSEA
Inv model 1	289.18	157	<.001	.963	0.049	39442.44						
Inv model 2A	385.39	201	<.001	.948	0.051	39487.57	2A vs 1	93.87	44	<.001	-.015	0.002
Inv model 2B	352.58	199	<.001	.957	0.047	39451.69	2B vs 1	65.55	42	.012	-.006	-0.002
Inv model 3A	553.15	217	<.001	.906	0.066	39642.66	3A vs 2B	256.13	18	<.001	-.051	0.019
Inv model 3B	390.99	209	<.001	.949	0.050	39478.10	3B vs 2B	38.31	10	<.001	-.008	0.003

Note. χ^2 = chi-square test statistic; *df* = degrees of freedom; *p* = probability value; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike’s information criterion; $\Delta\chi^2$ = likelihood ratio test (difference in chi-square between nested models); Δ CFI = difference in CFI between nested models; Δ RMSEA = difference in RMSEA between nested models. Inv Model 1 = configural invariance model; Inv Model 2A = metric invariance model; Inv Model 2B = partial metric invariance model where the factor loading of general mental health on item 14 was allowed free estimation in all groups; Inv Model 3A = partial scalar invariance model where the factor loading of general mental health on item 14 was still allowed free estimation in all groups; Inv Model 3B = partial scalar invariance model where the factor loading of general mental health in item 14 as well as the intercepts of items 2, 10, 1, and 12 were allowed free estimation in all groups.

Table 5

Comparison of the latent factor mean scores across the three samples

Scale	<i>M (SD)</i>			<i>z</i> -score, <i>p</i> -value, and <i>d</i> -value								
	Eng (<i>n</i> = 324)	Afr (<i>n</i> = 476)	Set (<i>n</i> = 258)	Eng vs Afr			Eng vs Set			Afr vs Set		
				<i>z</i>	<i>p</i>	<i>d</i>	<i>z</i>	<i>p</i>	<i>d</i>	<i>z</i>	<i>p</i>	<i>d</i>
MH	0 (0)	0.72 (0.21)	-0.63 (0.25)	3.43	.001	0.25	- 2.56	.011	-0.21	4.94	<.001	0.38
EWB	0 (0)	-1.09 (0.65)	2.18 (0.80)	-1.68	.092	-0.12	2.73	.006	0.23	-3.20	.001	-0.25
SWB	0 (0)	-0.73 (0.20)	1.39 (0.22)	-3.76	<.001	-0.27	6.18	<.001	0.52	-7.89	<.001	-0.61
PWB	0 (0)	-2.50 (0.64)	1.66 (0.56)	-3.90	<.001	-0.28	2.98	.003	0.25	-4.46	<.001	-0.34

Note. *M* = Mean factor score relative to the English sample; *SD* = Standard error of the mean score; *z* = *z*-score for the test comparing the means of two samples; *p* = probability value; *d* = Cohen's *d*-value; MH = general mental health, EWB = emotional well-being, SWB = social well-being, PWB = psychological well-being. Eng = English sample; Afr = Afrikaans sample; Set = Setswana sample.

Section 4: Manuscript 3

Problematic factorial validity of three language versions of the Basic Psychological Needs Scale (BPNS): Why and what are the implications?

Submitted to

Journal of Personality Assessment

4.1. Guidelines for authors

Journal of Personality Assessment

These guidelines were retrieved from:

<http://www.tandfonline.com/action/authorSubmission?journalCode=hjpa20&page=instructions#.Vthh2-Y6gg4>

“Thank you for choosing to submit your paper to us. These instructions will ensure we have everything required so your paper can move through peer review, production and publication smoothly. Please take the time to read them and follow the instructions as closely as possible.



Should you have any queries, please visit our [Author Services website](#) or contact us at authorqueries@tandf.co.uk.

SCHOLARONE MANUSCRIPTS™

This journal uses ScholarOne Manuscripts (previously Manuscript Central) to peer review manuscript submissions. Please read the [guide for ScholarOne authors](#) before making a submission. Complete guidelines for preparing and submitting your manuscript to this journal are provided below.

Aims and Scope: The *Journal of Personality Assessment* (*JPA*) seeks to publish papers on the methods and processes related to the psychological assessment of personality, both normative and pathological. While test and measurement development manuscripts will be considered for publication, the journal is most interested in papers that advance the clinical utility and ecological validity of personality assessment in clinical and nonclinical populations. These papers could include studies on enhancing the psychometric properties of a known measure, revision or modification of an existing instrument to enhance its usefulness, or the methods and mechanisms that expand the functionality of personality assessment in clinical practice.

JPA regularly publishes three sections that focus upon some of its broad areas of interest: Statistical Developments and Applications, Clinical Case Applications, and Personality Assessment in the Diagnostic Manuals. Papers devoted to test construction, methods, processes, and applicability of personality assessment in health care settings are especially desired. *JPA* also is interested in articles addressing under-studied areas and promising but under-utilized methods. These include (a) systematic reviews or meta-analyses that summarize a body of evidence, (b) the effective integration of nomothetic empirical findings with the idiographic requirements of practice in which the assessor reasons through test and extra-test information to make individualized judgments and provide assessment feedback, (c) the practical value of the clinical assessment process for individuals receiving services and/or those who refer them for evaluation, and (d) an evaluation of cultural, ethnic, and other sociodemographic factors that affect assessment. Finally, *JPA* will periodically publish special sections or a special issue related to a topic on personality assessment.

JPA invites comments and product reviews. Each comment should express a substantive opinion on an issue germane to personality assessment, including articles from recent *JPA* issues; comments on *JPA* articles should be submitted within two months of the target article's publication. The Book, Software and Test Review section publishes brief, paragraph-length descriptions of new personality assessment products as well as detailed reviews written by an expert in the field. Reviews should evaluate the strengths and limitations of books, tests, or software that are relevant to personality assessment practice or research. Comments and reviews typically will be limited to 1,000 words, and an abstract is not required.

JPA is the official journal of the Society for Personality Assessment.

Please note that *Journal of Personality Assessment* uses [CrossCheck™](#) software to screen papers for unoriginal material. By submitting your paper to *Journal of Personality Assessment* you are agreeing to any necessary originality checks your paper may have to undergo during the peer review and production processes.

Submission of Manuscripts: *Journal of Personality Assessment* receives all manuscript submissions electronically via their ScholarOne Manuscripts website located at: <http://mc.manuscriptcentral.com/JPersAssess>. ScholarOne Manuscripts allows for rapid submission of original and revised manuscripts, as well as facilitating the review process and internal communication between authors, editors and reviewers via a web-based platform. For ScholarOne Manuscripts technical support, you may contact them by e-mail or phone support via <http://scholarone.com/services/support/>. Manuscripts should not exceed 35 pages in length. Any paper longer than this will be returned to the author without review. Authors who would like to submit a longer manuscript should contact the Editor directly. If you have any other requests please contact the journal at jpa_office@emich.edu.

Publishers may send new book, tests, or software review materials to Virginia M. Brabender, Institute for Graduate Clinical Psychology, Widener University, One University Place, Chester, PA 19013, USA, vmbrabender@widener.edu (phone: 610-499-1208).

Each manuscript must be accompanied by a statement that it has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere. Authors are required to sign an agreement for the transfer of copyright to the publisher. As an author, you are required to secure permission if you want to reproduce any figure, table, or extract from the text of another source. This applies to direct reproduction as well as "derivative reproduction" (where you have created a new figure or table which derives substantially from a copyrighted source). All accepted manuscripts, artwork, and photographs become the property of the publisher.

All parts of the manuscript should be typewritten and double-spaced, with margins of at least one inch on all sides. Manuscript pages should be consecutively numbered throughout the paper. Manuscripts should not exceed 35 pages in length. Any paper longer than this will be returned to the author without review. Authors who would like to submit a longer manuscript should contact the Editor directly. Authors should also supply a shortened version of the title suitable for the running head, not exceeding 50 character spaces. Each article should be summarized in an abstract of not more than 200 words. The abstract should avoid abbreviations, diagrams, and reference to the text.

All authors should qualify for authorship by having made a substantial contribution to the conception, design, analyses, or interpretation of data, and by participating in writing or revising the manuscript. If portions of a manuscript were presented at a conference, the Author Notes should provide specifics. All submissions, but especially case reports, protect anonymity by avoiding or disguising information that could potentially identify a client. The cover letter for a case submission should clearly explain how issues of privacy and confidentiality were addressed. Manuscripts are sent anonymously to reviewers and should be ready for blind review when submitted. Thus, place Author Notes on the title page and do not intentionally reveal author identity in other ways.

Manuscripts are to be prepared according to the latest edition of the *Publication Manual of the American Psychological Association* (<http://www.apastyle.org/>). Authors of research manuscripts should incorporate the recommendations in “Statistical Methods in Psychology Journals: Guidelines and Explanations” by Wilkinson and the APA Task Force on Statistical Inference (1999, *American Psychologist*, 54, 594-604; or www.apa.org/journals/amp/amp548594.html) and must report standard effect size measures (e.g., r , Cohen’s d) for all statistical results. Authors are also encouraged to compute and report effect sizes when describing specific findings from previous studies. Researchers studying the diagnostic accuracy of tests are strongly encouraged to follow the Standards for Reporting Diagnostic Accuracy (STARD) guidelines. Resources for following the STARD guidelines and for computing effect sizes are available at www.stard-statement.org/. For more useful diagnostic and statistical guidelines, resources, and templates, visit www.tandf.co.uk/journals/authors/hjpa/.

Publishing Ethics

The Editor and Taylor & Francis Group are committed to the highest academic, professional, legal, and ethical standards in publishing work in this journal. To this end, we have adopted a [set of guidelines](#), to which all submitting authors are expected to adhere, to assure integrity and ethical publishing for authors, reviewers, and editors.

Taylor & Francis is a member of the Committee of Publications Ethics (COPE). COPE aims to provide a forum for publishers and editors of scientific journals to discuss issues relating to the integrity of their work, including conflicts of interest, falsification and fabrication of data, plagiarism, unethical experimentation, inadequate subject consent, and authorship disputes. For more information on COPE please visit <http://publicationethics.org>.

References: Cite in the text by author and date (Smith, 1983). Prepare reference list in accordance with the APA Publication Manual, latest ed. Examples:

Journal: Brierley, B., Medford, N., Shaw, P., & David, A. S. (2007). Emotional memory for words: Separating content and context. *Cognition and Emotion*, 21, 495–521. doi: [10.1080/02699930600684963](https://doi.org/10.1080/02699930600684963)

Book: Smith, E. R., & Mackie, D. M. (2007). *Social psychology* (3rd ed.). Hove, England: Psychology Press/Taylor & Francis (UK).

Contribution to a Book: Tanner, W. P., Jr., & Swets, J. A. (2000). A decision-making theory

of visual detection. In S. Yantis (Ed.), *Visual perception: Essential readings* (pp. 48–55). New York, NY: Psychology Press.

Illustrations: Illustrations submitted (line drawings, halftones, photos, photomicrographs, etc.) should be clean digital files and should follow these guidelines:

- 300 dpi or higher
- Sized to fit on a journal page
- Submitted on separate pages of the manuscript file, not embedded in a page of regular text

Color Reproduction: Color art will be reproduced in color in the online publication at no additional cost to the author. Color illustrations will also be considered for print publication; however, the author will be required to bear the full cost involved in color art reproduction. Please note the color reprints can only be ordered if print reproduction costs are paid. Print Rates: \$900 for the first page of color; \$450 per page for the next three pages of color. A custom quote will be provided for articles with more than four pages of color. Art not supplied at a minimum of 300 dpi will not be considered for print.

Tables and Figures: Tables and figures (illustrations) should not be embedded in the text, but should be included as separate pages at the end of the manuscript file. A short, descriptive title should appear above each table with a clear legend and any footnotes suitably identified below. All units must be included. Figures should be completely labeled, taking into account necessary size reduction. Captions should be typed, double-spaced, on a separate page at the end of the manuscript file.

Proofs and Reprints: Page proofs are sent to the designated author using Taylor & Francis' Central Article Tracking System (CATS). They must be carefully checked and returned within 48 hours of receipt. Authors for whom we receive a valid e-mail address will be provided an opportunity to purchase reprints of individual articles, or copies of the complete print issue. These authors will also be given complimentary access to their final article on *Taylor & Francis Online* .

Open Access: Taylor & Francis Open Select provides authors or their research sponsors and funders with the option of paying a publishing fee and thereby making an article fully and permanently available for free online access – *open access* – immediately on publication to anyone, anywhere, at any time. This option is made available once an article has been accepted in peer review. [Full details of our Open Access programme](#) .

Search Engine Optimization: Search Engine Optimization (SEO) is a means of making your article more visible to anyone who might be looking for it. Please consult our guide [here](#).”

4.2. Manuscript

Problematic Factorial Validity of Three Language Versions of the Basic Psychological Needs

Scale (BPNS): Why and What are the Implications?

Lusilda Schutte, Marié P. Wissing and Suria M. Ellis

North-West University, South Africa

Author note

Lusilda Schutte, Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa; Marié P. Wissing, Africa Unit for Trans-disciplinary Health Research, North-West University, South Africa; Suria M. Ellis, Unit for Business Mathematics and Informatics, North-West University, South Africa.

Correspondence concerning this article should be addressed to Lusilda Schutte, Africa Unit for Trans-disciplinary Health Research, North-West University, Potchefstroom Campus, South Africa, 2531. Email: Lusilda.Schutte@nwu.ac.za. Tel: +27 18 299 1104.

The research was financially supported by the National Research Foundation (NRF) of South Africa by funding for the FORT3 research project granted to the second author and for a doctoral scholarship granted to the first author. Opinions expressed and conclusions arrived are those of the authors and are not necessarily to be attributed to the NRF. Funding was also provided in the form of a doctoral scholarship to the first author by the North-West University, South Africa. The authors wish to thank I. P. Khumalo, L. Temane, and L. Ndimba for their assistance with their assistance at different stages of the research. The study forms part of the doctoral thesis of the first author.

Abstract

Self-determination theory is a macrotheory of human motivation that describes fundamental matters such as personality development, goals and aspirations, and self-regulation. Basic psychological needs theory, a subtheory of self-determination theory, postulates that the needs for autonomy, competence, and relatedness are universal and the satisfaction thereof essential for human functioning. Despite the theory's strong universality claim, almost no studies tested the assumption on the African continent. The present study addressed this by exploring the factorial validity of English, Afrikaans, and Setswana versions of the Basic Psychological Needs Scale ($N = 1056$). After incorporating a negative-worded method effect and removing several problematic items, the fit of the intended three-factor model was good for the Afrikaans version, marginal for the English version, and poor for the Setswana version. The resulting factors' reliabilities were low. Configural, metric, and partial scalar invariance were established between the English and Afrikaans versions. These findings not only highlighted problems with the particular scale, but also raised questions about the universality assumption of basic psychological needs theory and/or assumptions about denotations or manifestations of main constructs in various cultural contexts. The study indicated the conceptual and linguistic complexities involved in assessment across diverse and multicultural contexts.

Keywords: Basic Psychological Needs Scale (BPNS), scale validation, measurement invariance, cross-cultural assessment, self-determination theory, positive psychology, South Africa

Problematic Factorial Validity of Three Language Versions of the Basic Psychological Needs Scale (BPNS): Why and What are the Implications?

The question of what drives and motivates human beings has interested psychology since its earliest days. This matter has been addressed from multiple angles and perspectives, for example by the study of people's drives, goals, aspirations, purposes, and needs. One of the prominent theories that connects many of these constructs is self-determination theory (SDT), which can be considered a macrotheory of human motivation that describes fundamental matters such as personality development, goals and aspirations, self-regulation, basic psychological needs, and the impact of social environments on motivation, behaviour, and well-being (Deci & Ryan, 2008b). SDT works from the premise that people are naturally internally motivated, curious, interested, and vital (Deci & Ryan, 2008a). In order to realise this proactive nature, the three basic psychological needs of autonomy, relatedness, and competence have to be satisfied – they are considered essential for human functioning (Deci & Ryan, 2008a). Basic psychological needs theory or subtheory is considered one of the corner stones of SDT (Deci & Ryan, 2000).

Basic Psychological Needs Theory

In basic psychological needs theory, autonomy concerns an experience of volition and choice in behaviour regulation, competence refers to an experience of efficacy concerning one's internal and external environment, and relatedness concerns a sense of being connected to and cared for by others (Ryan, Huta, & Deci, 2008). The theory postulates that social environments that satisfy individuals' three basic psychological needs support their natural growth tendency and will facilitate optimal motivation which will result in psychological well-being (Vansteenkiste & Ryan, 2013). On the other hand, social milieus that frustrate or thwart the basic psychological needs activate less optimal types of motivation. This results in psychological ill-being and may manifest in seeking unhealthy need substitutes as expressed

in the striving for extrinsic goals, compensatory behaviours, or rigid behavioural patterns (Vansteenkiste & Ryan, 2013).

A large number of studies have shown that the satisfaction of the three basic psychological needs is associated with health and well-being. For example, Ng et al. (2012) showed in a meta-analysis that basic psychological need satisfaction was positively associated with beneficial mental and physical health outcomes and Milyavskaya, Philippe, and Koestner (2013) found that basic psychological need satisfaction at three levels of experience (general, domain-specific, and episodic) correlated positively with well-being. In addition, need frustration or thwarting has been related to psychological ill-being in multiple studies, for example, need thwarting predicted negative affect when being physically active (Gunnell, Crocker, Wilson, Mack, & Zumbo, 2013) and low levels of basic psychological need fulfilment were associated with burnout in elite rugby players (Hodge, Lonsdale, & Ng, 2008). See Vansteenkiste and Ryan (2013) for a review of correlates of basic psychological need satisfaction and frustration.

One of the central premises of basic psychological needs theory is that the needs are universal: Every human being, regardless of culture, life stage, or context, needs these elements to function and flourish, just like plants need water and sunlight to grow and thrive (Deci & Ryan, 2000). This will be the case regardless of whether the individual considers the needs as important or not (Ryan & Deci, 2000a). Several cross-country studies have found support for the positive association between need satisfaction and well-being and/or between need frustration and ill-being across different cultures (e.g., B. Chen, Vansteenkiste, et al., 2015; Church et al., 2012). However, since need satisfaction is proposed to be facilitated by the integration and internalisation of culturally endorsed values and behaviours, the expression thereof may differ for different cultural climates (Ryan & Deci, 2000b). In other words, while the benefits related to need satisfaction are claimed to be universal, the ways in

which the needs are met and the paths taken to experience satisfaction of the needs may differ from culture to culture (B. Chen, Vansteenkiste, et al., 2015).

Measurement of Basic Psychological Needs

In order to test the premises of SDT and basic psychological needs theory, good measurement of the constructs is essential. Furthermore, to study claims such as the universality of the needs, the measurement instruments should be cross-culturally valid and invariant. Several domain-specific measurement instruments of basic psychological needs have been developed, for example for the work context (e.g., Van den Broeck, Vansteenskiste, De Witte, Soenens, & Lens, 2010), for physical exercise (e.g., Vlachopoulos & Michailidou, 2006), and for educational settings (e.g., Longo, Gunz, Curtis, & Farsides, 2016). Although these measures are appropriate in studies where the theory is applied in a specific context, domain-general measures of psychological need satisfaction are needed when the constructs are studied from a context-free perspective. For many years, the Basic Psychological Needs Scale (BPNS, Gagné, 2003) has been commonly used to assess context-free basic psychological need satisfaction. More recently, the Balanced Measure of Psychological Needs (BMPN, Sheldon & Hilpert, 2012) and the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS, B. Chen, Vansteenkiste, et al., 2015) have been developed for this purpose. In this study, the focus will fall on the BPNS.

The BPNS was first used by (Gagné, 2003), who modified the domain-specific Basic Need Satisfaction at Work Scale (Deci et al., 2001; Ilardi, Leone, Kasser, & Ryan, 1993; Kasser, Davey, & Ryan, 1992). For example, the item “I really like the people I work with” was modified to read “I really like the people I interact with”¹. The BPNS has been applied in numerous studies after Gagné first used the instrument, sometimes obtaining subscale scores (e.g., Kashdan, Mishra, Breen, & Froh, 2009), sometimes a total scale score (e.g., Philippe, Koestner, Beaulieu-Pelletier, & Lecours, 2011), and sometimes both (e.g., Meyer, Enström,

Harstveit, Bowles, & Beevers, 2007). The first proper validation study of the scale has been conducted by Johnston and Finney (2010) among American students. These authors found that only after removal of five problematic items and when incorporating a negative-worded method effect did the scale show sufficient structural validity. An exploration of the shortened scale's external validity confirmed the distinctiveness of the three subscales.

Sheldon and Hilpert (2012) reported that, among American students, both the 21-item version (original scale) and 16-item version (obtained by Johnston and Finney, 2010) of the scale displayed only marginal fit, with the 16-item measure not performing better than the 21-item version. Both Johnston and Finney (2010) and Sheldon and Hilpert (2012) pointed out that findings needed confirmation in replication studies in other contexts.

Basic Psychological Needs in an African Context

Even though self-determination theory claims that the three basic psychological needs and their links with psychological well- and ill-being are universal (c.f., B. Chen, Vansteenkiste, et al., 2015; Chirkov, Ryan, & Sheldon, 2011; Church et al., 2012) very few studies investigated this theory in an African context. In addition, as far as we are aware, no in-depth validation studies of measurement instruments of basic psychological needs have been done based on data from Africa. Examples of studies that measured domain-general basic psychological needs include Van Zyl and Rothmann (2012) who applied the BPNS to a sample of South African students, Roman et al. (2015) who administered the Balanced Measure of Psychological Needs (Sheldon & Hilpert, 2012) to a group of South African adolescents, and B. Chen, Van Assche, Vansteenkiste, Soenens, and Beyers (2015) who applied the Basic Psychological Need Satisfaction and Frustration Scale (B. Chen, Vansteenkiste, et al., 2015) to a sample of South African students. In more specific contexts, Rothmann, Diedericks, and Swart (2013) applied the Work-related Basic Need Satisfaction Scale (Van den Broeck et al., 2010) to a sample of employees of South African agricultural

companies and Muller and Louw (2004) measured the satisfaction of basic psychological needs in a learning context among a sample of South African university students. Apart from Cronbach's alpha-values, none of these studies gave any indication of the scales' validity or reliability in the new context. In addition, in some of the studies the Cronbach's alpha-values were low and/or items had to be removed to obtain sufficient alpha-values. Clearly there is a gap in the literature when it comes to the validation of measurement instruments for basic psychological needs theory and the underlying assumptions of this theory in an African context. Addressing this gap is imperative not only for the development of psychological theories that are nuanced and sensitive for cultural variation, but also for the development of interventions and applications that are culturally sensitive, appropriate, and effective.

The Present Study

The aim of the present study was to address the gap in the literature by assessing the psychometric properties of three language versions of the BPNS using data from South Africa. We hypothesised that a three-factor model consisting of the factors Autonomy, Competence, and Relatedness with incorporation of a negative-worded method effect would best fit the data and that removal of some items was expected to improve the functioning of the scale. In addition, we aimed to evaluate the measurement invariance of the scale among the three groups and expected invariance after removal of problematic items. This study is the first to thoroughly explore the validity of a measure of basic psychological needs in an African context and will in that way contribute significantly to the body of knowledge about basic psychological needs theory and its applicability across the world.

Method

Design and Participants

The study implemented a cross-sectional survey design with students ($N = 1056$) from the three campuses of a university in South Africa. Respondents who completed less

than 50% of the BPNS were removed from the dataset prior to data analysis. The research battery was administered in three of the main languages in the areas of data gathering, namely English ($n = 322$, 75% female, $M_{\text{age}} = 21.04$, $SD_{\text{age}} = 4.11$), Afrikaans ($n = 478$, 64% female, $M_{\text{age}} = 19.78$, $SD_{\text{age}} = 3.14$), and Setswana ($n = 256$, 68% female, $M_{\text{age}} = 21.61$, $SD_{\text{age}} = 4.60$). From the group who completed the English scale, 18% indicated English as their native language, 18% Setswana, 7% Afrikaans, and 55% "Other". 99% of the group who completed the Afrikaans scale indicated that Afrikaans was their native language. From the group who completed the Setswana scale, 22% indicated English as their native language, 67% Setswana, 1% Afrikaans, and 9% "Other". "Other" probably indicated another African language. Participants were requested to complete the questionnaires in their native language or, if they did not speak one of the three languages natively, in their language of choice. Since either English or Afrikaans was the language of tuition at the respective campuses, it was assumed that participants who did not speak one of the three languages at home would select the language in which they receive tertiary education, making them sufficiently fluent in the language of the instruments. The groups who completed the English, Afrikaans, and Setswana versions of the research battery will be referred to as the English, Afrikaans, and Setswana samples, respectively.

Measures

Multiple scales were included in the research battery. Instruments relevant to this study will now be briefly presented.

Socio-demographic questionnaire. Socio-demographic information was obtained from the participants, for example gender, age, and native language.

Basic Psychological Needs Scale (BPNS, Gagné, 2003). This 21-item scale consists of three subscales, where respondents rate the extent to which their needs for autonomy (7 items), relatedness (6 items), and competence (8 items) are met on a scale from 1 (*not true at*

all) to 7 (*very true*). Nine of the items are reversed-phrased. Subscale scores can be calculated to get an indication of the extent to which each individual need is satisfied, while the total score can be used as a general index of need satisfaction. Gagné (2003) obtained Cronbach's alpha scores of .69, .86 and .71 for the three subscales, respectively. Johnston and Finney (2010) conducted a validation study of the scale and found that a reduced 16-item three-factor model displayed adequate structural and external validity.

Procedure

The BPNS was translated from English to Afrikaans and Setswana using a research committee approach (Van de Vijver & Leung, 1997). One translator translated the scale into the target language and another translator back-translated the translated version into English. A research committee consisting of Afrikaans and Setswana-speaking academics that were also competent in English compared the original and back-translated versions of the scale and addressed any discrepancies by revising the translated version (Brislin, 1973; Van de Vijver & Leung, 1997). The translated scales were administered to small pilot samples from the target populations to investigate if the items were clear and culturally acceptable. Participants were recruited by their lecturers during class lectures and participation was voluntary. After giving written informed consent, volunteers completed the questionnaires in class or at home.

Data Analysis

IBM SPSS Statistics 22 was used to obtain descriptive statistics, while Mplus 7.31 (Muthén & Muthén, 1998-2014) was applied to conduct confirmatory factor analysis. Both the robust maximum likelihood (MLR) and the robust weighted least squares (WLSMV) estimators were used. For model parameterisation, the latent factor variances were fixed to one (Van de Schoot, Lugtig, & Hox, 2012) and missing data were handled by full information maximum likelihood estimation.

Guidelines suggested by Byrne (2012) were used to assess single group model fit. In terms of the global fit of the models, the comparative fit index (CFI) and root mean square error of approximation (RMSEA) with its 90% confidence interval will be reported. For the CFI, earlier recommendations were that values of .90 or larger are indicative of good fit, but more recently a revised cut-off value of .95 was suggested. For the RMSEA, values of 0.05 or less suggest good fit and values smaller than 0.08 indicate reasonable fit. Results for the chi-square test will also be reported, but since the test is very sensitive to sample size it will not be used in decision making. To assess the local fit of the items, modification indices (MI's) and corresponding expected parameter change (EPC) indices were considered, where large values indicate misfit. In addition, small R^2 -values (percentage of the item variance explained by the model) and small factor loadings on the intended factor were interpreted as indications of local misfit. Omega coefficients were calculated as an indication of the factor reliabilities. We followed Johnston and Finney (2010) by calculating the reliability in the presence of a method effect as the systematic variance associated with the substantive factor divided by the sum of the systematic variance associated with the substantive factor, the sum of the items' unstandardized error variance, and the systematic variance associated with the method effect.

Once a baseline model with sufficient fit was established for the separate groups, a multi-group analysis was conducted to assess measurement invariance. Three increasingly strict levels of measurement equivalence were evaluated (Byrne & Van de Vijver, 2010). First, configural invariance was assessed, where the number of factors and the pattern of freely estimated and fixed parameters are constrained to be the same across the groups, but the parameter values are allowed to vary. Second, metric invariance was assessed, where the factor loadings of the different groups are constricted to be equal. Third, scalar invariance was evaluated, where the item intercepts are constricted to be equal across the groups in addition to the factor loadings. Once scalar invariance is established, the latent factor means

of the groups can be compared (Van de Vijver & Leung, 2011). To determine if each subsequent level of equivalence holds, the difference between the chi-square tests of the nested models (the likelihood ratio test) is often used. Although we will report this test for completeness, it is overly sensitive to large sample sizes. Therefore ΔCFI and ΔRMSEA (i.e., the difference between the CFI- and RMSEA-values of the nested tests, respectively) will be used for interpretation, where $|\Delta\text{CFI}|$ -values of .01 or less and $|\Delta\text{RMSEA}|$ -values of 0.015 or less will be used as an indication of measurement equivalence (F. F. Chen, 2007; Cheung & Rensvold, 2002). In addition, Akaike's information criterion (AIC) will be reported, where lower values indicate better fit (Byrne, 2012).

Ethical Considerations

This study formed part of the FORT 3 research-project, "The prevalence of levels of psychosocial health: dynamics and relationships with biomarkers of (ill) health in South African social context" (Wissing, 2008/2012), which received approval from the North-West University's Health Research Ethics Committee with project number NWU 00002-07-A2. Participation was voluntary and written informed consent was obtained from all participants prior to participation. Data were gathered and captured anonymously.

Results

Descriptive Statistics

After reversing the scores of reversed-phrased items, item mean scores ranged from 3.77 (item 4) to 5.83 (item 18), standard deviations from 1.29 (item 3) to 2.07 (item 7), skewness statistics from -1.17 (item 12) to 0.14 (item 3), and kurtosis statistics from -1.14 (item 4) to 0.91 (item 12) for the English group. For the Afrikaans group, mean scores ranged from 3.84 (item 4) to 6.22 (item 12), standard deviations from 1.09 (item 6) to 2.01 (item 16), skewness-values from -1.83 (item 12) to 0.11 (item 4), and kurtosis-values from -1.23 (item 16) to 4.01 (item 12). For the Setswana group, mean scores ranged from 3.78 (item 11) to

5.70 (item 1), standard deviations from 1.29 (item 13) to 2.06 (item 7), skewness-values from -1.35 (item 1) to 0.25 (item 11), and kurtosis-values from -1.26 (item 7) to 1.23 (item 1).

Single Group CFA

We started by fitting a one-factor model to the data in each group, since a total score over all items, indicative of general basic psychological need satisfaction, is often used in the literature (Model 1). This model did not fit the data well, and many of the problems lay with the reversed-phrased items. We therefore incorporated a negative-worded method effect (Model 2). Since this model still lacked good fit for all groups, we fitted a three-factor model with Autonomy, Competence, and Relatedness as correlated factors (Model 3). Once again the model did not fit the data well, and the reversed-phrased items seemed to be the source of many of the problems. We therefore added a negative-worded method effect (Model 4).

Although the global fit improved, it was still insufficient and several areas of local misfit were suggested by the MI's, EPC-values, factor loadings, and R^2 -values. We proceeded by iteratively removing the one item with the largest misfit based on both empirical and substantive evidence, where after the model was fitted again. This process was repeated one item at a time until the model fit was adequate and/or there was no clear further evidence of items with both empirical and substantive misfit. The best-fitting model for the English and Afrikaans groups was similar (see results below) and was labelled Model 5_Eng/Afr. The best-fitting model for the Setswana sample was labelled Model 5_Tswana. The global fit indices for Models 1 to 5 are presented in Table 1 and graphical representations of the models in Figure 1². The process followed for each group to establish Model 5 will now be presented.

English group. When fitting Model 4 to the data from the English sample using the MLR estimator, a large MI (34.14) and EPC-value (1.27) suggested that there was a substantial correlation between the residuals of items 7 (“I pretty much keep to myself and do not have a lot of social contacts”) and 16 (“There are not many people that I am close to”).

The content of these items appear to be very similar, and therefore we deemed one of the items to be redundant. Since item 7 had a larger R^2 -value (.27 vs .20), a larger factor loading on the Relatedness factor (.39 vs .34), and seems to be more representative of relatedness as conceptualised in basic psychological needs theory, we removed item 16, resulting in a model with improved fit (CFI = .805, RMSEA = 0.062). In this model, it was suggested that item 14 (“People I interact with on a daily basis tend to take my feelings into consideration”), which was designed to measure autonomy, also loaded on the Relatedness factor (for Relatedness by BPNS14, MI = 30.92 and EPC = 2.02). Considering the interpersonal component apparent in the item’s formulation, it appears theoretically reasonable that this item taps both autonomy and relatedness. We removed the item, which resulted in a model with CFI = .844 and RMSEA = 0.055. In this model, the residuals of item 4 (“I feel pressured in my life”) correlated substantially with the residuals of some of the other items that were not intended to measure autonomy (for BPNS4 with BPNS18, MI = 25.48 and EPC = -0.79; for BPNS4 with BPNS19, MI = 9.20 and EPC = 0.57). Upon examination of the item content, the item’s conceptual ambiguity became evident – one may feel pressured due to frustrated needs for autonomy (the target psychological need), or for relatedness or competence (not the target psychological needs). We therefore removed item 4 and obtained a scale with fit indices CFI = .869 and RMSEA = 0.052. In this model, a residual correlation between item 5 (“People I know tell me I am good at what I do”) from the Competence subscale and item 6 (“I get along with people I come in contact with”) from the Relatedness subscale was suggested (MI = 26.04, EPC = 0.63). Although item 5 was intended to measure competence, a relatedness component is implied by the “People I know tell me” phrase contained within the item. Due to the lack of conceptual clarity for item 5, we removed the item which resulted in a model with fit indices CFI = .902 and RMSEA = 0.046. In this model, item 11 (“In my daily life, I frequently have to do what I am told”) had a non-significant factor loading on the intended

factor, Autonomy. When applying the WLSMV estimator and for the other language versions of the scale, MI's suggested that the item also loaded on the Competence and Relatedness factors. This item appears to be conceptually muddled and not uniquely representative of autonomy. We removed the item, obtaining a model with fit indices CFI = .907 and RMSEA = 0.046. In this model, a loading of the relatedness item 12 ("People in my life care about me") on the Competence factor was suggested (MI=10.62, EPC=1.36). According to basic psychological needs theory, competence concerns a sense of efficacy related to one's internal and external environment. Possibly, experiencing care by others may relate to a sense of efficacy concerning the external environment, leading to the double loading of item 12. We removed the item, resulting in a model with fit indices CFI = .913 and RMSEA = 0.046. In this model, item 3 ("Often, I do not feel very competent") had a low R^2 -value (.14), which means that only 14% of the variance contained within the item is explained by the model. Removal of the item improved the model fit substantially (CFI = .932, RMSEA = 0.042). At this point no additional high MI's and EPC-values that made substantive sense were detected, all factor loadings were significant, and no R^2 -values were exceptionally small.

Using the WLSMV estimator, removal of the same seven items and, in addition, removal of item 8, were suggested. In particular, first item 14 was removed (for Relatedness by item 14, MI = 54.50 and EPC = 1.76) resulting in a model with CFI = .861 and RMSEA = 0.077; then item 16 (for Competence by item 16, MI = 16.57 and EPC = -1.95; for Autonomy by item 16, MI = 8.06 and EPC = -0.42) resulting in a model with CFI = .878 and RMSEA = 0.074; then item 8 (for Competence by item 8, MI = 16.94 and EPC = 5.27; for Relatedness by item 8, MI = 15.29 and EPC = 0.82) resulting in a model with CFI = .882 and RMSEA = 0.071; then item 11 (for Competence by 11, MI = 13.29 and EPC = 2.99; for Relatedness by 11, MI = 6.66 and EPC = -0.25; the factor loading of item 11 was also non-significant) resulting in a model with CFI = .893 and RMSEA = 0.071; then item 4 (the item's factor

loading was non-significant) resulting in a model with CFI = .908 and RMSEA = 0.884; then item 12 (for Competence by item 12, MI = 14.01 and EPC = 2.88; for Autonomy by item 12, MI = 8.12 and EPC = 0.38) resulting in a model with CFI = .910 and RMSEA = 0.072; then item 5 (for Relatedness by item 5, MI=15.29 and EPC = 0.55) resulting in a model with CFI = .938 and RMSEA = 0.061; and finally item 3 (small R^2 -value of .17) resulting in a model with CFI = .944 and RMSEA = 0.062. Our theoretical justification for the removal of items 14, 16, 11, 4, 12, 5, and 3 was similar to our reasoning when the MLR estimator was applied. For item 8 (“I generally feel free to express my ideas and opinions”), one may feel free to express your ideas and opinions because one feels competent (non-intended Competence factor), because one feels comfortable in your relationships with others (non-intended Relatedness factor), or because one experiences a sense of volition (intended Autonomy factor). The theoretical ambiguity of the item justified its removal.

Considering the results using the MLR and WLSMV estimators together, we decided to remove all items that were problematic for both estimators in order to suggest a best-fitting model for the English sample. In this model, labelled Model 5_Eng/Afr, items 3, 4, 5, 11, 12, 14, and 16 were removed. The global fit of this model is displayed in Table 1 and a graphical representation thereof in Figure 1. Using the MLR estimator, the reliabilities were .61, .59, and .58 for Autonomy, Relatedness, and Competence, respectively.

Afrikaans group. Using MLR and working from Model 4 in one-by-one removing problematic items resulted in the removal of item 14 (for Relatedness by item 14, MI = 21.09 and EPC = 1.66; removal of item 14 resulted in a model with CFI = .909 and RMSEA = 0.049), item 11 (for Competence by item 11, MI = 17.13 and EPC = -2.66; for Relatedness by item 11, MI = 7.55 and EPC = -0.80; removal of items 14 and 11 resulted in a model with CFI = .934 and RMSEA = 0.043), item 5 (for item 5 with item 6, MI = 15.56 and EPC = 0.21; for Relatedness by item 5, MI = 11.94 and EPC = 0.78; removal of items 14, 11, and 5

resulted in model with CFI = .944 and RMSEA = 0.041), item 16 (for item 16 with item 7, MI = 14.04 and EPC = 0.53; removal of items 14, 11, 5, and 16 resulted in a model with CFI = .954 and RMSEA = 0.028), item 4 (for Competence by item 4, MI = 12.25 and EPC = 2.33; removal of items 14, 11, 5, 16, and 4 resulted in a model with CFI = .965 and RMSEA = 0.034), and item 12 (for Competence by item 12, MI = 12.74 and EPC = -0.24; removal of items 14, 11, 5, 16, and 4 resulted in a model with fit CFI = .971 and RMSEA = 0.032). Based on the global fit indices, we were satisfied with the fit of this model to our data and no further major points of local misfit were identified. Removal of each of the items was substantively justified in the same way as for the English group.

Using the WLSMV estimator and working from Model 4, we removed item 14 (for Relatedness by item 14, MI = 63.77 and EPC = 1.55; removal of item 14 resulted in a model with CFI = .932 and RMSEA = 0.073), item 11 (for Competence by item 11, MI = 38.72 and EPC = -3.20; for Relatedness by item 11, MI = 22.50 and EPC = -0.56; removal of items 14 and 11 resulted in a model with CFI = .950 and RMSEA = 0.066), item 5 (for Relatedness by item 5, MI = 40.43 and EPC = 0.81; removal of items 14, 11, and 5 resulted in a model with CFI = .960 and RMSEA = 0.061), item 4 (for Competence by item 4, MI = 15.16 and EPC = 2.10; for Relatedness by item 4, MI = 13.05 and EPC = -0.45; removal of items 14, 11, 5, and 4 resulted in a model with CFI = .963 and RMSEA = 0.061), item 12 (for Competence by item 12, MI = 20.75 and EPC = 0.87; for Autonomy by item 12, MI = 19.26 and EPC = 0.67; removal of items 14, 11, 5, 4, and 12 resulted in a model with CFI = .964 and RMSEA = 0.062), and item 3 (for Relatedness by item 3, MI = 18.02 and EPC = -0.36; for Autonomy by item 3, MI = 14.66 and EPC = -0.96; removal of items 14, 11, 5, 4, 12, and 3 resulted in a model with CFI = .966 and RMSEA = 0.062). Based on the global fit indices, we were satisfied with the fit of this model and no further major points of local misfit were identified. Removal of each item was substantively justified in the same way as for the English group.

Considering the results using the MLR and WLSMV estimators together, items 4, 5, 11, 12, and 14 were problematic for both estimators. In addition, item 16 was problematic when the MLR estimator was applied and item 3 when the WLSMV estimator was applied. Since items 3 and 16 were also problematic for the English sample, we decided to remove these items for the best-fitting model for the Afrikaans sample. This model, labelled Model 5_Eng/Afr, had items 3, 4, 5, 11, 12, 14, and 16 removed. The global fit of this model is presented in Table 1 and a graphical representation thereof in Figure 1. Using MLR, the reliabilities were .66, .69, and .57 for Autonomy, Relatedness, and Competence, respectively.

Setswana group. For the Setswana group, the three-factor models (Models 2 and 4) did not outperform the one-factor models (Models 1 and 3), regardless of whether the MLR or WLSMV estimator was used. In addition, the three-factor models produced a warning that the covariance matrix was not positive definitive because the correlations between the latent factors were larger than one. We explored whether this issue would be resolved if we removed problematic items as indicated by indices of local misfit, but the problem remained. We concluded that a three-factor model was not suitable for the Setswana sample, with a one-factor model being more appropriate. The one-factor model with a negative-worded method effect (Model 3) performed considerably better than the model without a method effect (Model 1) and became the baseline from which areas of local misfit were identified.

When applying the MLR estimator, we removed items 17, 11, 9, 10, and 3 one-by-one, as these items attained small R^2 -values in each consecutive model (.05, .13, .14, .14, and .17, respectively). From this model, items 15 and 16 were removed one at a time, since they displayed non-significant factor loadings on the Basic Psychological Need Satisfaction factor. The resulting model obtained a fit of $CFI = .976$ and $RMSEA = 0.028$.

Using the WLSMV estimator, we one-by-one deleted items 17, 9, and 11 due to low R^2 -values in the consecutive models (.05, .18, and .15, respectively). From this point, items

16, 15, and 3 were removed one at a time, since these items attained non-significant factor loadings on the Basic Psychological Need Satisfaction factor in the consecutive models. The resulting model obtained a fit of $CFI = .972$ and $RMSEA = 0.049$.

Considering the results from the MLR and WLSMV estimators together, we removed the items that were problematic for both estimators to obtain a best-fitting model for the Setswana group. In this model, labelled Model 5_Tswana, items 3, 9, 11, 15, 16, and 17 were removed. The model's fit and graphical representation are displayed in Table 1 and Figure 1, respectively. The reliability of the Basic Psychological Needs factor was .64 using MLR.

Multi-group CFA

Measurement invariance. Measurement invariance was investigated using the MLR estimator (see Table 2). The WLSMV estimator could not be used, because certain items did not have values in all categories for all groups, which is a requirement for assessing measurement invariance using this estimator. Since the three-factor model did not fit the Setswana sample, we could not establish measurement invariance for this sample and invariance was only investigated between the other two groups. The best-fitting model for these groups, Model 5_Eng/Afr, was used to assess measurement invariance. The model showed sufficient configural and metric invariance, but scalar invariance was not established. Based on the modification indices, we one-by-one allowed the intercepts of items 9 and 18 to be estimated freely. At this point, partial scalar invariance was attained.

Comparison of mean scores. When the average latent factor scores were compared between the English and Afrikaans groups for the final model for which partial scalar invariance was found, statistically and practically significant differences ($z = 6.07, p < .001, d = 0.44$) were found for the Relatedness factor between the English group (reference group, $M = 0.00, SD = 0.00$) and Afrikaans group ($M = -0.63, SD = 0.10$), with the English group scoring lower on the construct (d -values of close to 0.50 or more are deemed as medium

effect sizes and are indicative of practically significant differences, Cohen, 1988). The differences on the Autonomy and Competence factors were insignificant.

Discussion

The present study explored the psychometric properties of three language versions of the BPNS. After incorporating a negative-worded method effect and removing several problematic items, a three-factor model distinguishing between autonomy, relatedness, and competence displayed a fair fit for the English sample and a good fit for the Afrikaans sample. For the Setswana sample, a one-factor model with a negative-worded method effect fitted the data best after removal of problematic items. The resulting factors' reliabilities were poor. Configural, metric, and partial scalar invariance were established between the English and Afrikaans versions. Several aspects of these findings deserve further attention.

Construction of the BPNS

Similar to Johnston and Finney (2010) and Sheldon and Hilpert (2012) we found that the original 21-item version of the scale did not fit the data well for any of the three groups in the present study and that incorporation of a negative-worded method effect as well as removal of several items were necessary to obtain satisfactory fit. For many of the removed items, the problems with the items were clearly evident from the item content as described in the Results section: Some items tapped more than one of the needs simultaneously, others were redundant, and some were not really representative of the intended constructs. More specifically, Johnston and Finney (2010) found that a three-factor model with a negative-worded method effect where items 4, 11, 14, 16, and 20 were removed exhibited a good fit to the data. This corresponds with the findings for the English and Afrikaans samples in the present study, except that item 20 was not problematic in this study, while items 3, 5, and 12 additionally displayed problems. In the Setswana sample, where a one-factor model with a negative-worded method effect outperformed the other models, items 11 and 16 were the

only problematic items that overlapped with the findings of Johnston and Finney (2010), while items 3, 9, 15, and 17 additionally displayed problems.

The fact that there were so many items that exhibited misfit in the present study and that there was considerable overlap with the findings of Johnston and Finney (2010), point out that the issues with the scale may have been due to problems in the construction phase of the original instrument rather than simply cultural variation or translation issues. Different possible sources of error will be explored. First, it seems as there was a lack of substantive grounding when some of the items were formulated. For example, an item such as item 14 (“People I interact with on a daily basis tend to take my feelings into consideration”) that clearly taps relatedness while autonomy is the item’s target construct, makes one question whether the necessary attention was given to the substantive phase of test construction, where the construct and subconstructs are explicitly conceptualised (Simms, 2008). Second, this scale was an adaptation of a domain-specific measure of basic psychological need satisfaction at work (Gagné, 2003). The findings stress the necessity of first conceptualising the construct for the new context and then, from that point of theoretical depth, formulating or adapting the items. Third, the BPNS does not distinguish between need satisfaction and need frustration or thwarting, a distinction that has been suggested repeatedly in recent literature (Vansteenkiste & Ryan, 2013). Scales that measure both aspects may yield better results (cf., B. Chen, Vansteenkiste, et al., 2015; Longo et al., 2016; Sheldon & Hilpert, 2012).

The psychometric properties of the BPNS in its original form have now been shown to be problematic in multiple samples (all three groups in the present study, Johnston & Finney, 2010; Sheldon & Hilpert, 2012). However, the scale has been used to assess domain-general basic psychological needs for many years. This places question marks over the trustworthiness of the conclusions drawn in those studies. Findings need to be reconfirmed using sound measurement instruments before they can be applied responsibly to practice.

Universality Assumption of SDT

One of the major claims of basic psychological needs theory is that the three needs are universal, that the satisfaction of the needs are essential for growth and psychological well-being, and that the frustration thereof will lead to psychological ill-being, regardless of factors such as culture, age, and gender (Deci & Ryan, 2000; Ryan & Deci, 2000b). In this study, the three-factor model did not fit the Setswana version of the scale due to very high intercorrelations between the latent constructs and, in addition, the fit was only marginally acceptable for the English sample, even after removing several problematic items. Different explanations can be sought to interpret these findings. First, the problems with the Setswana version of the scale may have been due to translation issues. However, the fact that the model fit was still only marginal for the English version of the scale, make us query whether translation was the main problem. Another possibility is that the three basic psychological needs, as theorised by SDT, are not as clearly distinguished in an African context as in other contexts where confirmation for the three-factor model was found (e.g., B. Chen, Vansteenkiste, et al., 2015; Sheldon & Hilpert, 2012). We found it particularly noteworthy that the proposed three-factor model fitted the Afrikaans sample much better than the other two samples. In South Africa, native language Afrikaans speakers often have a strong Western heritage. The participants in the English and Setswana groups, on the other hand, mostly spoke an African language at home suggesting an African heritage. This finding may therefore suggest that the three-factor model is more appropriate for people from a Western background. Yet another possibility is that the same three basic psychological needs do exist and are distinguishable in an African context, but that they manifest or operate in ways that was not captured by the BPNS. Ryan and Deci (2000b) explained that, even though SDT claims that the three basic psychological needs are universal, their modes of expression and avenues for satisfaction may differ for different cultures or developmental phases. For

example, the need for autonomy for someone from a collectivistic culture may be satisfied when he follows the advice of important people in his life, while the same need for someone from a more individualistic culture may be satisfied when he makes his own decisions and expresses his personal opinions (B. Chen, Vansteenkiste, et al., 2015). In that sense, an item such as item 8 (“I generally feel free to express my ideas and opinions”) may be a typical expression of the need for autonomy for someone from an individualistic culture, but may be less appropriate for someone from a more collectivistic culture.

These findings call us back to the drawing board, where an emic approach should be applied to study basic psychological needs and their modes of expression within an African context. In addition, exploring the psychometric properties of two recently developed measures of domain-general basic psychological need satisfaction, namely the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS, B. Chen, Vansteenkiste, et al., 2015) and the Balanced Measure of Psychological Needs (BMPN, Sheldon & Hilpert, 2012) will be a worthwhile endeavour. As many of the problems inherent in the BPNS were addressed in the construction of these scales, such studies may shed light on whether the issues with the BPNS in the Setswana and English groups were mainly due to poor scale construction of the BPNS, or due to a cultural insensitive or inappropriate underlying theory.

These findings highlight that merely transferring a theory from one context or country to another is not good practice. Basic psychological needs theory is widely claimed to be universal, and yet the entire African continent has been neglected in research studying this claim. When a theory is not applicable in a certain context, measurement instruments and interventions based on this theory will be inappropriate and ineffective. The importance of cross-cultural sensitivity when it comes to theory, measurement, and practice is imperative.

The English and Setswana Versions of the BPNS: Working within Multicultural Groups

In the present study, the samples that completed the English and Setswana versions of the scale both consisted of a majority of native African language speakers. Exploring the similarities and differences between the findings for these two groups is therefore an important endeavour.

The overall performance of the English version scale was better when compared to the Setswana version in this study. One could have argued that the differential performance was due to cultural differences between the samples, but both the English and Setswana samples had a strong African heritage. The discrepancy may therefore rather be explained by the fact that African language speakers in South Africa often receive their primary and secondary education in English. Completing scales such as the BPNS in English may therefore require less cognitive strain than completing scales in an African language (such as Setswana), which could have influenced the psychometric properties of the scale.

An interesting finding was that the negative-worded method effect had the greatest impact for the Setswana sample (see the difference in fit between Models 1 versus 2 and Models 3 versus 4, respectively, for the three groups). One possible interpretation is that response style could have influenced the way in which reversed-phrased items were answered. According to Weijters and Baumgartner (2012), ambiguity due to reversal of items can be particularly prominent in cross-cultural research, as some cultures, such as East Asian cultures, tend to view statements that may seem oppositional from a Western perspective in a more compatible and inclusive way. In addition, acquiescence bias (i.e., an inclination to (dis)agree with items irrespective of the contents) may also be culturally influenced (Johnson, Kulesa, Cho, & Shavitt, 2005) and could have impacted the way in which reversed-phrased items were answered. Of note is that the impact of the reversed-phrased items was smaller for the English sample when compared to the Setswana sample, while African participants formed the majority of both samples. This finding counters the argument that culture-

informed response style caused the differential impact of the negative-worded method effect. The question becomes whether the way in which reversed items are expressed in Setswana are different from the way in which such items are expressed in English, that is, whether the difference lies more on a linguistic than a cultural level? Another possibility may be related to the fact that many African language speaking South Africans receive their formal education in English. The cognitive strain involved in responding to reversed items in Setswana could therefore have been more than responding to such items in English, resulting in the discrepancy. Further research is needed to explore the operation of reversed-phrased items in an African context – both from a linguistic and a cultural perspective.

These findings alert us to the complexities involved in working not only between different cultures, but also within multicultural contexts. Cultures and languages become increasingly entwined, resulting in multifarious systems of varying degrees of similarity and diversity. Theory, measurement and practice should move beyond merely attempting to comprehend and address cross-cultural diversity, to a point where the variability within contexts are also better understood and attended to in applications. Adding an item that assesses language of primary and secondary education to socio-demographic questionnaires may aid interpretation.

Measurement Invariance

Another point for discussion could be the measurement invariance of the scale. However, we deem any conclusions in this regard to be premature, as invariance was assessed on a scale that was significantly cut and may lack the necessary conceptual comprehensiveness to justify conclusions.

Limitations

Although this study made important contributions, it was not without limitations. A student sample was used, which restricts the generalisability of the findings to other cohorts

of the population. Future research should explore whether the findings replicate in other demographic and multicultural groups. The test construction problems with the BPNS as evident from the present study and the studies of Johnston and Finney (2010) and Sheldon and Hilpert (2012) make it difficult to determine whether findings were mainly due to issues with the scale as such, or whether they were due to problems with the underlying theory of basic psychological needs. Studying the nature and manifestation of basic psychological needs in an African context from an emic perspective, as well as exploring the psychometric properties of recently developed instruments that address some of the issues raised for the BPNS, would yield a better understanding of what these constructs entail in diverse settings.

Conclusions

The 21-item BPNS portrayed problems for all three language versions in the present study. Although removing problematic items and incorporating a negative-worded method effect improved the scale's psychometric properties, some issues were still evident. This not only confirmed findings from previous studies that highlighted problems with the particular scale, but also raised questions regarding the universality assumption of basic psychological needs theory and/or assumptions about denotations or manifestations of main constructs in various cultural contexts. In addition, this study alerted us to the linguistic and cultural complexities when working in diverse and multicultural communities and the issues involved in reversed items. Implications for theory, measurement, and practice were contemplated.

Footnotes

¹These scales were retrieved from the Self-Determination Theory website, at

<http://www.selfdeterminationtheory.org/questionnaires/>.

²Due to space limitations, we were not able to include the standardised pattern coefficients of the final models in this paper. Please contact the first author to obtain these results.

References

- Brislin, R. W. (1973). *Cross-cultural research methods*. New York, NY: John Wiley.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Byrne, B. M., & Van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*(2), 107-132.
doi:10.1080/15305051003637306
- Chen, B., Van Assche, J., Vansteenkiste, M., Soenens, B., & Beyers, W. (2015). Does psychological need satisfaction matter when environmental or financial safety are at risk? *Journal of Happiness Studies, 16*, 745-766. doi:10.1007/s10902-014-9532-5
- Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van Der Kaap-Deeder, J., . . . Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*, 216-236.
doi:10.1007/s11031-014-9450-1
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.
doi:10.1207/S15328007SEM0902_5
- Chirkov, V. I., Ryan, R. M., & Sheldon, K. M. (Eds.). (2011). *Human autonomy in cross-cultural context: Perspectives on the psychology of agency, freedom, and well-being*. New York, NY: Springer.
- Church, A. T., Katigbak, M. S., Locke, K. D., Zhang, H., Shen, J., Vargas-Flores, J., . . . Ching, C. M. (2012). Need satisfaction and well-being: Testing self-determination

theory in eight cultures. *Journal of Cross-Cultural Psychology*, 44(4), 507-534.

doi:10.1177/0022022112466590

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum, NJ: Hillsdale.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268.

doi:10.1207/S15327965PLI1104_01

Deci, E. L., & Ryan, R. M. (2008a). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology*, 49(1), 14-23. doi:10.1037/0708-5591.49.1.14

Deci, E. L., & Ryan, R. M. (2008b). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, 49(3), 182-185.

doi:10.1037/a0012801

Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former Eastern Bloc country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin*, 27(8), 930-942.

Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation & Emotion*, 27(3), 199-223.

Gunnell, K. E., Crocker, P. R. E., Wilson, P. M., Mack, D. E., & Zumbo, B. D. (2013). Psychological need satisfaction and thwarting: A test of basic psychological needs theory in physical activity contexts. *Psychology of Sport & Exercise*, 14, 599-607.

doi:10.1016/j.psychsport.2013.03.007

- Hodge, K., Lonsdale, C., & Ng, J. Y. Y. (2008). Burnout in elite rugby: Relationships with basic psychological needs fulfilment. *Journal of Sports Sciences, 26*(8), 835-844. doi:10.1080/02640410701784525
- Iardi, B. C., Leone, D., Kasser, T., & Ryan, R. M. (1993). Employee and supervisor ratings of motivation: Main effects and discrepancies associated with job satisfaction and adjustment in a factory setting. *Journal of Applied Social Psychology, 23*(21), 1789-1805.
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36*(2), 264-277. doi:10.1177/0022022104272905
- Johnston, M. M., & Finney, S. J. (2010). Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the Basic Needs Satisfaction in General Scale. *Contemporary Educational Psychology, 35*(4), 280-296. doi:10.1016/j.cedpsych.2010.04.003
- Kashdan, T. B., Mishra, A., Breen, W. E., & Froh, J. J. (2009). Gender differences in gratitude: Examining appraisals, narratives, the willingness to express emotions, and changes in psychological needs. *Journal of Personality, 77*(3), 691-730. doi:10.1111/j.1467-6494.2009.00562.x
- Kasser, T., Davey, J., & Ryan, R. M. (1992). Motivation and employee-supervisor discrepancies in a psychiatric vocational rehabilitation setting. *Rehabilitation Psychology, 37*(3), 175-188. doi:10.1037/h0079104
- Longo, Y., Gunz, A., Curtis, G. J., & Farsides, T. (2016). Measuring need satisfaction and frustration in educational and work contexts: The Need Satisfaction and Frustration Scale (NSFS). *Journal of Happiness Studies, 17*(1), 295-317. doi:10.1007/s10902-014-9595-3

- Meyer, B., Enström, M. K., Harstveit, M., Bowles, D. P., & Beevers, C. G. (2007). Happiness and despair on the catwalk: Need satisfaction, well-being, and personality adjustment among fashion models. *Journal of Positive Psychology, 2*(1), 2-17.
doi:10.1080/17439760601076635
- Milyavskaya, M., Philippe, F. L., & Koestner, R. (2013). Psychological need satisfaction across levels of experience: Their organization and contribution to general well-being. *Journal of Research in Personality, 47*(1), 41-51. doi:10.1016/j.jrp.2012.10.013
- Muller, F. H., & Louw, J. (2004). Learning environment, motivation and interest: Perspectives on self-determination theory. *South African Journal of Psychology, 34*(2), 169-190. doi:10.1177/008124630403400201
- Muthén, L. K., & Muthén, B. O. (1998-2014). Mplus (Version 7.31). Los Angeles, CA: Muthén and Muthén.
- Ng, J. Y. Y., Ntoumanis, N., Thøgersen-Ntoumani, C., Deci, E. L., Ryan, R. M., Duda, J. L., & Williams, G. C. (2012). Self-determination theory applied to health contexts: A meta-analysis. *Perspectives on Psychological Science, 7*(4), 325-340.
doi:10.1177/1745691612447309
- Philippe, F. L., Koestner, R., Beaulieu-Pelletier, G., & Lecours, S. (2011). The role of need satisfaction as a distinct and basic psychological component of autobiographical memories: A look at well-being. *Journal of Personality, 79*(5), 905-937.
doi:10.1111/j.1467-6494.2010.00710.x
- Roman, N. V., Davids, E. L., Moyo, A., Schilder, L., Lacante, M., & Lens, W. (2015). Parenting styles and psychological needs influences on adolescent life goals and aspirations in a South African setting. *Journal of Psychology in Africa, 25*(4), 305-312. doi:10.1080/14330237.2015.1078087

- Rothmann, S., Diedericks, E., & Swart, J., P. (2013). Manager relations, psychological need satisfaction and intention to leave in the agricultural sector. *South African Journal of Industrial Psychology*, 39(2), 1-11. doi:10.4102/sajip.v39i2.1129
- Ryan, R. M., & Deci, E. L. (2000a). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Inquiry*, 11(4), 319-338. doi:10.1207/S15327965PLI1104_03
- Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. doi:10.1037/110003-066X.55.1.68
- Ryan, R. M., Huta, V., & Deci, E. L. (2008). Living well: A self-determination theory perspective on eudaimonia. *Journal of Happiness Studies*, 9(1), 139-170. doi:10.1007/s10902-006-9023-4
- Sheldon, K. M., & Hilpert, J. C. (2012). The Balanced Measure of Psychological Needs (BMPN) Scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion*, 36, 439-451. doi:10.1007/s11031-012-9279-4
- Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, 2(1), 414-433. doi:10.1111/j.1751-9004.2007.00044.x
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492. doi:10.1080/17405629.2012.686740
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver

- (Eds.), *Cross-cultural research methods in psychology* (pp. 17-45). New York, NY: Cambridge University Press.
- Van den Broeck, A., Vansteenskiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the Work-related Basic Need Satisfaction Scale. *Journal of Occupational and Organizational Psychology*, 83(4), 981-1002. doi:10.1348/096317909X481382
- Van Zyl, L. E., & Rothmann, S. (2012). Beyond smiling: The evaluation of a positive psychological intervention aimed at student happiness. *Journal of Psychology in Africa*, 22(3), 369-384. doi:10.1080/14330237.2012.10820541
- Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, 23(3), 263-280. doi:10.1037/a0032359
- Vlachopoulos, S. P., & Michailidou, S. (2006). Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The Basic Psychological Needs in Exercise Scale. *Measurement in Physical Education & Exercise Science*, 10(3), 179-201.
- Weijters, B., & Baumgartner, H. (2012). Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research*, 49(5), 737-747. doi:10.1509/jmr.11.0368
- Wissing, M. P. (2008/2012). *The prevalence of levels of psychosocial health: Dynamics and relationships with biomarkers of (ill) health in South African social contexts (FORT3)*.

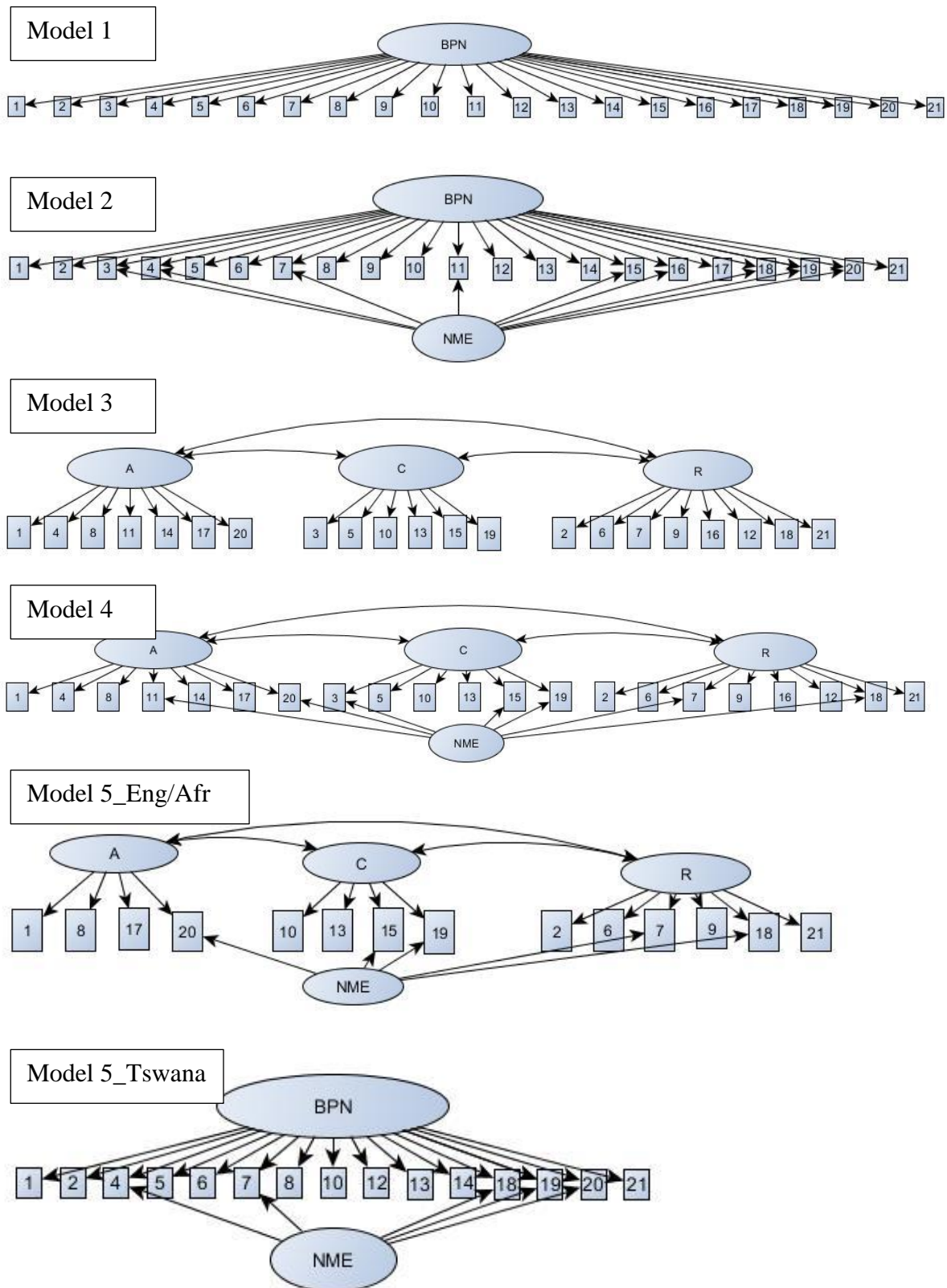


Figure 1. Competing models of the structure of the BPNS. BPN = Basic psychological needs, A = Autonomy, C = Competence, R = Relatedness, NME = Negative-worded method effect.

Table 1

Fit statistics from single-group confirmatory factor analysis

Model	MLR				WLSMV			
	χ^2	df	CFI	RMSEA (90% CI)	χ^2	df	CFI	RMSEA (90% CI)
Eng (<i>n</i> = 322)								
Model 1	593.91	189	.628	0.082 (0.074; 0.089)	845.26	189	.729	0.104 (0.097; 0.111)
Model 2	454.09	180	.748	0.069 (0.061; 0.077)	606.51	180	.824	0.086 (0.078; 0.093)
Model 3	560.88	186	.655	0.079 (0.072; 0.087)	798.89	186	.747	0.101 (0.094; 0.108)
Model 4	411.98	177	.784	0.064 (0.056; 0.072)	559.28	177	.842	0.082 (0.074; 0.090)
Model 5_Eng/Afr	108.48	69	.932	0.042 (0.026; 0.057)	177.73	69	.932	0.070 (0.057; 0.083)
Afrikaans (<i>n</i> = 478)								
Model 1	780.96	189	.735	0.081 (0.075; 0.087)	1447.77	189	.805	0.118 (0.113; 0.124)
Model 2	474.66	180	.868	0.059 (0.052; 0.065)	788.00	180	.906	0.084 (0.078; 0.090)
Model 3	678.11	186	.780	0.075 (0.069; 0.081)	1237.75	186	.837	0.109 (0.103; 0.115)
Model 4	423.53	177	.890	0.054 (0.048; 0.061)	676.98	177	.922	0.077 (0.071; 0.083)
Model 5_Eng/Afr	101.97	69	.972	0.032 (0.017; 0.044)	182.58	69	.971	0.059 (0.049; 0.069)

Setswana (<i>n</i> = 256)								
Model 1	714.97	189	.459	0.104 (0.096; 0.112)	1018.82	189	.620	0.131 (0.123; 0.139)
Model 2	354.34	180	.821	0.062 (0.052; 0.071)	390.12	180	.904	0.068 (0.058; 0.077)
Model 3	714.80	186	.456	0.105 (0.097; 0.114)	1006.50	186	.625	0.131 (0.123; 0.139)
Model 4	346.28	177	.826	0.061 (0.052; 0.071)	380.88	177	.907	0.067 (0.058; 0.076)
Model 5_Tswana	115.27	85	.953	0.037 (0.017; 0.054)	136.29	85	.972	0.049 (0.033; 0.063)

Note. MLR = robust maximum likelihood estimator; WLSMV = robust weighted least squares estimator; χ^2 = chi-square test statistic; *df* = degrees of freedom; *p* = probability value; CFI = comparative fit index; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval of the RMSEA. Model 1 = single-factor model; Model 2 = single-factor model with negative-worded method effect; Model 3 = three-factor model; Model 4 = three-factor model with negative-worded method effect; Model 5_Eng/Afr = best fitting model for the English and Afrikaans samples; Model 5_Tswana = best fitting model for the Setswana sample.

Table 2

Measurement invariance of Model 5_Eng/Afr for English and Afrikaans groups

Model	χ^2	<i>df</i>	<i>p</i>	CFI	RMSEA	AIC	Model comparison	$\Delta\chi^2$	<i>df</i>	<i>p</i>	Δ CFI	Δ RMSEA
Configural	210.32	138	<.001	.959	0.036	38635.36						
Metric	236.06	153	<.001	.953	0.037	38639.49	Configural vs Metric	25.55	15	.043	-.006	0.001
Scalar	293.31	163	<.001	.926	0.045	38678.97	Metric vs Scalar	63.28	10	<.001	-.027	0.008
Scalar_P	257.09	161	<.001	.946	0.039	38647.32	Metric vs Scalar_P	22.51	8	.004	-.007	0.002

Note. Model 5_Eng/Afr = best fitting model for the English and Afrikaans samples; χ^2 = chi-square test statistic; *df* = degrees of freedom; *p* = probability value; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike's information criterion; $\Delta\chi^2$ = likelihood ratio test (difference in chi-square between nested models); Δ CFI = difference in CFI between nested models; Δ RMSEA = difference in RMSEA between nested models. Configural = configural invariance model; Metric = metric invariance model; Scalar = scalar invariance model; Scalar_P = partial scalar invariance model where the intercepts of items 9 and 18 were allowed to be freely estimated in both groups.

Section 5: Conclusions, Implications, and Recommendations

Section 5: Conclusions, Implications, and Recommendations

The overall aim of the present study was to further the process of understanding and measuring eudaimonic well-being (EWB) across different demographic groups by implementing modern psychometric techniques. This was done by exploring in three manuscripts the cross-cultural psychometric properties of measurement instruments that operationalise three prominent theories associated with EWB.

Specific Conclusions, Implications and Recommendations from the Three Manuscripts

The first manuscript applied Rasch modelling to explore the psychometric properties of the Meaning in Life Questionnaire (MLQ, Steger, Frazier, Oishi, & Kaler, 2006) among adults from South Africa, Australia, and New Zealand. Although the MLQ in general displayed good psychometric properties according to the Rasch model, several areas for revision of the scale were indicated. In addition, significant implications of the findings for the understanding of meaning in life were highlighted. The Presence subscale of the MLQ was found to be insensitive for high levels of presence of meaning in life, while the majority of the respondents attained high scores. The significant practical implications of this finding were contemplated, for example, that correlations in correlational studies will be largely influenced by the minority of people exhibiting lower levels of presence of meaning, and that the scale would probably not be sensitive to changes in the presence of meaning in life of people on the higher end of the continuum in experimental studies or intervention evaluations. The major importance of further exploration of the nature and measurement of the presence of meaning in life, particularly at high levels, was highlighted. The results furthermore suggested that some of the response categories were redundant and that using less nuanced categories indicative of low and medium levels of meaning in life (five or six rating scale categories instead of seven) should be explored. The only reversed-phrased item (“My life has no clear purpose”) in the scale displayed problems, which pointed towards

removal of this item and highlighted the potential problems involved in negated, reversed-phrased items. No differential item functioning (DIF) was found for gender, age group, or education level, and for the Search subscale there was also no DIF for country, but the Presence subscale displayed DIF for country on four of the five items in the subscale. This highlighted that caution should be applied when scores on the Presence subscale are compared across the countries.

The second manuscript explored the suitability of applying a bifactor model to Afrikaans, English, and Setswana versions of the Mental Health Continuum Short Form (MHC-SF, Keyes, 2006, 2009; Keyes et al., 2008), whereafter the measurement invariance across the three versions of the scale was assessed. In comparison with other models that are commonly applied to the MHC-SF, a bifactor model demonstrated superior fit. The general mental health factor attained good reliability and explained the largest portion of the common variance, suggesting that interpretation and use of the scale's total scale score in subsequent analyses would be appropriate. In contrast, the subscales were not reliable and explained only a small portion of the common variance, suggesting that subscale scores should not be interpreted or used in further analyses. More comprehensive measurement of the subdimensions of overall positive mental health may be needed to obtain reliable scores that are suitable for interpretation and use in further analyses. It is of note that the Afrikaans version consistently obtained better fit indices than the other two versions (the participants who completed the Afrikaans version typically had a Western cultural heritage while many of the participants who completed the other two versions probably had an African heritage), while the Setswana version performed better than the English version (the majority of the participants who completed the English version did not indicate English as their native language, while the majority of the participants who completed the Setswana version spoke Setswana at home). These findings indicate that the model operationalised by this scale may

be more applicable to Western samples than African samples, and that the language of test administration should deserve careful consideration in the research process. A first follow-up step should be to study overall positive mental health within an African context from an emic perspective. In terms of measurement invariance across the three language versions, the scale displayed configural, partial metric, and partial scalar invariance. The average scores on the general mental health factor did not differ practically significantly across the groups, which is noteworthy given that the groups differ in terms of cultural heritage and political background. When looking at the performance of the items of the scale, the 'social contribution' item of the Social Well-being subscale displayed poor fit. This item appeared to target a different aspect (contribution towards a greater good) than the other items in the scale (all self-orientated), and removal of the item lead to substantial improvement in the model fit. This finding highlighted the lack of focus on contribution towards a greater good beyond the self in the present scale and indicated a need for more focus on this central aspect of functioning well in future eudaimonic well-being research.

The third manuscript explored the psychometric properties of English, Afrikaans, and Setswana versions of the Basic Psychological Needs Scale (BPNS, Gagné, 2003), an operationalisation of basic psychological needs theory, which is a subtheory of self-determination theory (Deci & Ryan, 2008a, 2008b). All three versions of the scale displayed poor fit. After removing several problematic items and incorporating a negative-worded method effect, a three-factor model fitted the Afrikaans version well and the English version marginally well. The three-factor model did not fit the Setswana version of the scale, and a one-factor model only displayed an adequate fit after removal of several items and the incorporation of a negative-worded method effect. The reliability indices of the factors were low. These results suggested serious problems in the initial scale construction phase. In addition, the findings placed question marks over the widely accepted claim that basic

psychological needs and their associations with psychological well-being and ill-being are universal. In fact, this theory has been studied across countries and cultures, but very few studies examined the theory in an African context. Revisiting the universality assumption of self-determination theory and an emic approach to studying basic psychological needs within the African context was indicated. The negative-worded method effect displayed the largest effect for the Setswana version. This raised questions regarding the impact of response styles, linguistic factors, and the interaction between native language, language of education, and language of test administration on how respondents responded to the reversed-phrased items.

General Conclusions, Implications and Recommendations: An Integration

From the study of the three scales using different modern statistical approaches in different contexts, cultures, and demographic groups, significant conclusions regarding the theory and measurement of EWB crystallised. These, together with the implications thereof and recommendations based upon the findings, will now be discussed.

Theory of Well-being and EWB

The findings in this study have important implications for the conceptualisation and understanding of well-being in general and EWB in particular on a theoretical level. Different models and theories suggest that positive and negative mental health, traits, or experiences are related but distinct phenomena. Two of the models used in the present study are examples thereof. The two-continua model of mental health of Keyes (2002) poses that psychological well-being and ill-being are related but distinct continua, and the MHC-SF examined in this study aims to measure the upper continuum, called positive mental health. In the study of basic psychological needs, operationalised by the BPNS in this study, it has recently been suggested that the incorporation of both need satisfaction and need frustration into the conceptualisation and measurement of basic psychological needs may improve the measurement of the needs and better explain associations with psychological well- and ill-

being (Vansteenkiste & Ryan, 2013). Research has therefore recognised the importance of distinguishing between positive versus negative mental health. What is, however, less studied, is how the dimension of positive mental health operates, how the population is distributed along the continuum, and in particular, how well different levels of this continuum are captured by the instruments available. The present study addressed the latter by conducting Rasch analyses on the MLQ. Analysis of the Presence subscale showed that the scale was insensitive for high levels of presence of meaning, while most of the respondents scored in that range. This raised the question whether presence of meaning in life should possibly be considered a quasi-trait (Reise & Waller, 2009), which contains most of its variation at the lower (absence) end of the continuum, while variation at the upper (presence) end is limited – even though most people attain such high levels of meaning in life. This finding has vast implications for correlational studies using this measure, and studies on the effectiveness of interventions in the general population where fluctuations in levels of presence of meaning are assessed using this measure. The question becomes whether a similar trend would be found for other measures of EWB. Together these models and findings suggest that more research is needed to disentangle how the presence versus absence of positive traits or experiences and the presence versus absence of negative traits or experiences operate and interlink, and, especially, to understand the variability of these dimensions at the different levels in their continua. Only once we better comprehend the nuances and functioning of EWB, especially at high levels, will we be able to establish its relationships with other indicators of biopsychosocial well-being and the impact of interventions with the necessary sensitivity. One way to achieve this end is by conducting qualitative research to understand the meaning and manifestations of constructs associated with EWB from a bottom-up perspective.

In addition to understanding the nuances of well-being and EWB at different levels, a relevant and important question in well-being research is what the impact is of overall well-being (the so-called “g-factor” of well-being) versus the impact of subsaspects thereof: Are subfacets of well-being sufficiently differentiable from the g-factor to have unique associations and predictive power when they are related to external variables? In terms of operationalisation, researchers have to decide between relying on total scores over different well-being scales or subscales versus using scores of individual scales or subscales in subsequent analyses. The drawback of using total scores is that it does not yield information on the associations between the subdimensions and the outcome variables. On the other hand, the drawback of using individual scale or subscale scores is that it holds conceptual ambiguity, as it cannot distinguish the unique contribution of each subdimension from the contribution of the overall construct that is shared by the interrelated subdimensions (F. F. Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). The bifactor model attempts to overcome the disadvantages in both approaches by allowing the researcher to model the effect of both the overall construct and the subsaspects thereof, while partialling out the common variance that the subsaspects share when testing the associations between the subfacets and external variables (F. F. Chen et al., 2012). In the present study, the bifactor model showed superior fit to the MHC-SF in three cultural groups when compared with other commonly used models. However, while the general mental health factor was reliable and explained the greatest part of the common variance, the specific factors accounted for only a small portion of response variance and exhibited low reliability, suggesting that subscale scores should not be interpreted or used in subsequent analyses. The same was found by Jovanović (2015b) and De Bruin and Du Plessis (2015) when they applied the bifactor model to the MHC-SF. In contrast, when F. F. Chen, Jing, Hayes, and Lee (2013) fitted a bifactor model to overall well-being, they found that both a general well-being factor and specific factors of psychological

(eudaimonic) and subjective (hedonic) well-being exist and that each has unique predictive power. In line with these findings, Jovanović (2015a) found support for a bifactor model of subjective well-being, with a strong general factor, but also specific factors that uniquely and reliably measure subfacets of subjective well-being. In similar vein, Wissing and Temane (2008) found that well-being has a hierarchical structure, with general psychological well-being as a higher-order factor across different cultural groups, and lower-order factors which exhibited varying nuances in relatively collectivistic versus individualistic contexts.

Gallagher, Lopez, and Preacher (2009) confirmed the hierarchical structure of well-being by showing that well-being has a general higher-order factor with hedonic, eudaimonic, and social well-being as correlated but distinct subcomponents. Joshanloo (2015) found that hedonic and eudaimonic aspects of well-being are correlated but distinct using exploratory structural equation modelling. The difference between the present study and the studies of Jovanović (2015b) and De Bruin and Du Plessis (2015), which did not find sufficient support for interpretation of subfacet scores, and other studies that yielded results in support of the unique contribution of subfacet scores, is that the former used very short subscales to operationalise the different subdimensions of well-being, while the subdimensions were measured more comprehensively in the latter group of studies. Together, these findings suggest that a strong g-factor of well-being exists, while, at the same time, unique subdimensions can be distinguished and interpreted, given that they are measured comprehensively. In addition, factor structures such as the bifactor model, which partials out the common variance between the subdimensions due to the g-factor, would probably yield more accurate results when associations with external variables are investigated, and should be considered in future well-being research.

On a more specific level, it was found in the present study that the social contribution item in the MHC-SF did not fit with the rest of the items. Considering all three scales in the

present study, this was the only item that explicitly focused on a contribution towards something larger than the self. This is somewhat surprising, given that transcending the pre-occupation with the self, contribution to others, and social consciousness has been recognised as one of the core features of meaning in life (Emmons, 2005; Reker, 1997; Schnell, 2010; Steger, Kashdan, Sullivan, & Lorentz, 2008). In addition, the work of the virtue ethicists (Fowers, 2010, 2012; Richardson, 2012) has a focus on functioning well as the realisation of virtues that is not used primarily for one's own gain, but for the benefit of society and the greater good. Recently, theories that focus on interconnectedness between all dimensions and levels of life (Delle Fave & Soosai-Nathan, 2014; Helne & Hirvilammi, 2015; Wissing, 2014) also emphasise the essence of relationality that goes beyond the self in functioning well. The key and defining aspect of self-transcendence in functioning well is underrepresented in the theories and measurement instruments of eudaimonic well-being in the present study, and many other theories and measures that fall within this domain (e.g., Ryff, 1989b; Sheldon, 2002, 2013; Waterman et al., 2010). This component deserves more attention in theory and measurement.

Test Construction

The present study pointed towards important aspects that should be taken into consideration during the conceptualisation, item formulation, and validation phases of test construction. According to guidelines on test construction (cf., DeVellis, 2012; Simms, 2008), the first and often neglected phase is the substantive phase, where the construct is conceptualised in clear and unambiguous terms. It is of utmost importance to clearly define and demarcate the overall construct and the subdimensions thereof prior to starting with item formulation. In the present study, questions were raised about whether sufficient attention was given to this phase when the BPNS was constructed. For example, the scale was initially developed for administration in a work context, but then adapted to the domain-general

context without a thorough construct clarification for the new context. In addition, some items clearly tapped more than one of the basic psychological needs, which also suggest lack of attention to delineating the subdimensions. Another important aspect to take into consideration during the substantive phase of test construction that was highlighted in the present study, involves the conceptualisation of the construct for different levels of the latent variable in order to develop well-targeted instruments where the intensity levels of the construct captured by the measure match the intensity levels in the population. This idea stems from Rasch modelling, where explicit consideration is given to a scale's targeting (Pallant & Tennant, 2007; Tennant & Conaghan, 2007). In the present study, Rasch analysis of the MLQ revealed that a very limited range of presence of meaning was targeted by the scale, and that this range did not match the range within the population that was studied. Future research should focus on not only clearly and context-sensitively defining EWB constructs during the substantive phase of scale development, but also to explicate what the constructs look like at different levels of intensity. Sufficient information regarding how EWB operates and manifests at different levels of intensity may not currently be available and qualitative studies that explore the meaning and manifestation of the constructs at different levels of intensity may be necessary.

In terms of item formulation, the study underlined some potential hazards to avoid. In both the MLQ and the BPNS the reversed-phrased items displayed problems. Most of the reversed-phrased items in these scales were merely negations of positively phrased items, often using the word "not" to obtain the reversal. Including such items is discouraged in the literature (Weijters & Baumgartner, 2012; Weijters, Cabooter, & Schillewaert, 2010), since they do not broaden the content domain tapped by the scale, and they carry the disadvantages inherent in negated items (e.g., there is increased cognitive strain required to respond to negated items) and reversed items (e.g., response styles such as acquiescence, careless

responding, and confirmation bias can influence responses to reversed items, particularly in cross-cultural studies). However, excluding reversed-phrased items in their entirety is also not advisable, as these items can, for example, control acquiescence, disrupt non-substantive responding, and broaden the content domain tapped by the instrument (Weijters & Baumgartner, 2012; Weijters et al., 2010). In addition, polar opposite items may actually represent a different dimension of the construct, as is evident in the need satisfaction / need frustration distinction recently made in basic psychological need theory (B. Chen et al., 2015; Sheldon & Hilpert, 2012; Vansteenkiste & Ryan, 2013). As recommended in the literature, scales that measure EWB should include reversed-phrased items, but then non-negated, polar opposite items that use antonymic expressions rather than negations (Weijters & Baumgartner, 2012; Weijters et al., 2010). Further investigation is needed to examine when such items represent a related but distinct dimension of the construct under study.

The Rasch analysis of the MLQ pointed out that seven response categories were too many for the respondents in this study and that five or six response categories may be more appropriate. Weijters et al. (2010) suggested that seven response categories may be appropriate for populations with high cognitive abilities, verbal skills, or questionnaire experience, such as students, but that five categories may be more apt for the general population. Future research should apply item response theory and Rasch analyses to the MLQ and other EWB measures in diverse populations to investigate whether the findings replicate and to examine the performance of different response category combinations. Only then will it be possible to make confident and accurate suggestions on the number of response categories that is appropriate in EWB measurement, possibly varying for different contexts and groups.

Theory and Measurement of EWB from a Cross-Cultural Perspective

Apart from general conclusions regarding the theory and measurement of EWB, this study has some specific implications from a cross-cultural perspective. The study highlighted that theories and their operationalisations that were found to be appropriate in one context cannot necessarily merely be transferred to another context. Both the MHC-SF and the BPNS showed the best performance in the Afrikaans sample (of whom the majority typically had a Western background) when compared with the English and Setswana samples (of whom the majority typically had an African background). Both of these scales were developed in the West and are based on theories that have their origin in a Western context, and these findings suggest that the theories appear to be more appropriate for Western than non-Western groups, and that their transferability to the African context cannot just be assumed. This is a particularly significant finding for theories that claim to be universally true, such as basic psychological needs theory, while almost no research was done on the African continent to test this theory. It is, in fact, surprising that such strong claims are made while studies from an entire continent is lacking. This reminds of Henrich, Heine, and Norenzayan (2010), who remarked that broad claims about human behaviour and psychology are often published based entirely on samples from “Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies” (p. 61), while these participants are actually unusual when compared with the majority of the human population on earth. Cross-cultural studies that include, amongst others, African collaborators, should be conducted to understand the nuances, similarities, and differences in the constructs that we work with in various contexts. This will enable well-being researchers to develop theories, measures, and applications that are cross-culturally sensitive and nuanced.

The findings of this study have implications for measurement practices in Africa and South Africa, where the majority of tests that are used have been developed in Western

contexts (Foxcroft, 2011). Careful consideration should be given to make sure that best practices are followed when tests are adapted. For example, Foxcroft (2011) provided the following guidelines for ethically acceptable test adaptation: Test developers should assess measures' psychometric qualities in the new context and check for construct equivalence, make sure that the test content is culture-appropriate, adapt tests in collaboration with cultural experts, anthropologists, and linguists, do translations in a scientific manner, and keep in mind that self-reflection items may be biased in non-Western collectivist societies where individual desires are placed below community needs. Although responsibly adapting tests developed in Western contexts can contribute towards ethical assessment in Africa, there are limitations to this approach. A serious call for qualitative, bottom-up studies that investigate well-being constructs (Delle Fave & Bassi, 2009; Delle Fave, Brdar, Wissing, & Vella-Brodick, 2013), also from an emic perspective, within African contexts is made. Only when we understand these constructs better for the African and South African population will we be able to tailor interventions appropriately.

Findings regarding the model fit of the English and Setswana versions of the MHC-SF and BPNS alerted us to additional factors that may play a role in the psychometric properties of scales, especially in the African and South African context. Firstly, we realised that it is not only home language and language of test administration that may impact the way in which people respond to scales, but also the language in which they received their formal education (in South Africa, many people that speak an African language at home received their schooling in English). The impact may be elevated when participants need to respond to complex or reversed-phrased items, which requires considerable cognitive strain. Assessing participants' language of schooling in socio-demographic questionnaires may add valuable information for the interpretation of findings. The study also drew our attention to the possibility that response styles could impact the way in which participants responded to

items, and future research needs to investigate response styles in the diverse African and South African context in more depth. In addition, it was suggested that there could be linguistic differences in the way in which items are formulated in English and Setswana, possibly because certain words or concepts do not exist in both languages. Once again, further research in this regard is indicated.

In cross-cultural studies, the aim is often to compare the mean scores of a construct for the different groups. In all three manuscripts of the present study, some form of differential item functioning or inequivalence was found between at least some of the groups involved. This finding highlights the importance of first assessing for invariance and making model adaptations before comparing the mean scores of different cultural or demographic groups.

Limitations

Even though this study made important contributions towards the conceptualisation and measurement of EWB, it is not without limitations. Firstly, this study focused on only three theories of EWB with their corresponding measurements instruments (with one of these measures, namely the MHC-SF, also having a hedonic well-being component). Future studies should expand on this by including other models and scales in rigorous cross-cultural studies. Secondly, this study applied Rasch analysis to one of the scales and confirmatory factor analysis (using the bifactor model and a negative-worded method effect respectively) to the other two scales. More studies that apply these approaches to additional scales will help to investigate whether the findings replicate across different EWB constructs. Thirdly, other modern psychometric approaches, such as other item response theory models (Embretson & Reise, 2000) or exploratory structural equation modelling (Marsh, Morin, Parker, & Kaur, 2014) should be applied in future studies to further understand the nature and measurement of EWB. Lastly, the samples in this study were restricted and generalisations cannot be made to

other groups and contexts. Conducting similar studies in other contexts will give insight into whether the findings replicate and will foster a deeper, culturally-informed understanding of the nature and measurement of EWB which can, in turn, inform applications and interventions.

Contribution

Although there was a traditional focus on hedonic well-being in the well-being literature, eudaimonic well-being enjoyed increasing attention in recent years, where the conceptualisation, measurement, associations, and enhancement of eudaimonic well-being were explored. This study contributed to this body of knowledge by applying modern psychometric techniques to investigate the psychometric properties of three instruments of EWB in different cultural and demographic contexts. The findings not only shed light on the psychometric properties and avenues for revision of the specific scales that were examined in this study in the present contexts, but also provided general pointers that will enhance the construction of instruments that intend to evaluate EWB and well-being in general in diverse contexts. The study also provided insight into EWB on a theoretical level, where new perspectives on the conceptual understanding and nuances of EWB were gained, particularly from a cross-cultural point of view. Altogether, this study furthered the science of EWB on theoretical and measurement levels, which can, in turn, be applied to develop culture-sensitive interventions that truly enhance the lives of people.

Complete Reference List

- Aina, O. F., & Morakinyo, O. (2011). Culture-bound syndromes and the neglect of cultural factors in psychopathologies among Africans. *African Journal of Psychiatry, 14*, 278-285. doi:10.4314/ajpsy.v14i4.4
- Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement, 2*(4), 581-594. doi:0.1080/14330237.2014.904088
- Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care, 42*(1), 1-16.
- Andrich, D., Sheridan, B., & Luo, G. (2014). Rumm 2030 for Windows (Version 5.52). Perth, Australia: Rumm Laboratory.
- Baumeister, R. F., & Vohs, K. D. (2002). The pursuit of meaningfulness in life. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 608-618). New York, NY: Oxford University Press.
- Baumeister, R. F., Vohs, K. D., Aaker, J. L., & Garbinsky, E. N. (2013). Some key differences between a happy life and a meaningful life. *The Journal of Positive Psychology, 8*(6), 505-516. doi:10.1080/17439760.2013.830764
- Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika, 74*(1), 137-143. doi:10.1007/S11336-008-9100-1
- Biswas-Diener, R., Kashdan, T. B., & King, L. A. (2009). Two traditions of happiness research, not two distinct types of happiness. *The Journal of Positive Psychology, 4*(3), 208-211. doi:10.1080/17439760902844400
- Bobowik, M., Basabe, N., & Páez, D. (2015). The bright side of migration: Hedonic, psychological, and social well-being in immigrants in Spain. *Social Science Research, 51*, 189-204. doi:10.1016/j.ssresearch.2014.09.011

- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New York, NY: Routledge.
- Bowen, N. K., & Guo, S. (2012). *Structural equation modeling*. New York, NY: Oxford University Press.
- Boyraz, G., Lightsey, O. R., & Can, A. (2013). The Turkish version of the Meaning in Life Questionnaire: Assessing the measurement invariance across Turkish and American adult samples. *Journal of Personality Assessment, 95*(4), 423-431.
doi:10.1080/00223891.2013.765882
- Brandstätter, M., Baumann, U., Borasio, G. D., & Fegg, M. J. (2012). Systematic review of meaning in life assessment instruments. *Psycho-Oncology, 21*(10), 1034-1052.
doi:10.1002/pon.2113
- Brislin, R. W. (1973). *Cross-cultural research methods*. New York, NY: John Wiley.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Byrne, B. M., & Van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*(2), 107-132.
doi:10.1080/15305051003637306
- Chao, R. C.-L., & Green, K. E. (2013). Rasch analysis of the Outcome Questionnaire with African Americans. *Psychological Assessment, 25*(2), 568-582.
doi:10.1037/a0032083

- Chen, B., Van Assche, J., Vansteenkiste, M., Soenens, B., & Beyers, W. (2015). Does psychological need satisfaction matter when environmental or financial safety are at risk? *Journal of Happiness Studies*, *16*, 745-766. doi:10.1007/s10902-014-9532-5
- Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van Der Kaap-Deeder, J., . . . Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, *39*, 216-236. doi:10.1007/s11031-014-9450-1
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*, *29*(4), 468-508. doi:10.1177/0049124101029004003
- Chen, F. F. (2007a). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504.
- Chen, F. F. (2007b). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504.
- Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219-251. doi:10.1111/j.1467-6494.2011.00739.x
- Chen, F. F., Jing, Y., Hayes, A., & Lee, J. M. (2013). Two concepts or two approaches? A bifactor analysis of psychological and subjective well-being. *Journal of Happiness Studies*, *14*(3), 1033-1068. doi:10.1007/s10902-012-9367-x
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255. doi:10.1207/S15328007SEM0902_5

- Chirkov, V. I., Ryan, R. M., & Sheldon, K. M. (Eds.). (2011). *Human autonomy in cross-cultural context: Perspectives on the psychology of agency, freedom, and well-being*. New York, NY: Springer.
- Church, A. T., Katigbak, M. S., Locke, K. D., Zhang, H., Shen, J., Vargas-Flores, J., . . . Ching, C. M. (2012). Need satisfaction and well-being: Testing self-determination theory in eight cultures. *Journal of Cross-Cultural Psychology, 44*(4), 507-534. doi:10.1177/0022022112466590
- Cohen, A. B. (2009). Many forms of culture. *The American Psychologist, 64*(3), 194-204. doi:10.1037/a0015308
- Cohen, A. B. (2010). Just how many different forms of culture are there? *The American Psychologist, 65*(1), 59-61. doi:10.1037/a0017793
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum, NJ: Hillsdale.
- Cohen, K., & Cairns, D. (2012). Is searching for meaning in life associated with reduced subjective well-being? Confirmation and possible moderators. *Journal of Happiness Studies, 13*(2), 313-331. doi:10.1007/s10902-011-9265-7
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. doi:10.1037/h0040957
- Crumbaugh, J. C., & Maholick, L. T. (1964). An experimental study in existentialism: The psychometric approach to Frankl's concept of noogenic neurosis. *Journal of Clinical Psychology, 20*(2), 200-207. doi:10.1002/1097-4679(196404)20:2<200::AID-JCLP2270200203>3.0.CO;2-U
- Da Rocha, N. S., Chachamovich, E., De Almeida Fleck, M. P., & Tennant, A. (2013). An introduction to Rasch analysis for psychiatric practice and research. *Journal of Psychiatric Research, 47*(2), 141-148. doi:10.1016/j.jpsychires.2012.09.014

- De Bruin, G. P., & Du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum-short form (MHC-SF). *Psychological Reports, 116*(2), 438-446. doi:10.2466/03.02.PR0.116k20w6
- Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. doi:10.1207/S15327965PLI1104_01
- Deci, E. L., & Ryan, R. M. (2008a). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology, 49*(1), 14-23. doi:10.1037/0708-5591.49.1.14
- Deci, E. L., & Ryan, R. M. (2008b). Hedonia, eudaimonia, and well-being: An introduction. *Journal of Happiness Studies, 9*(1), 1-11. doi:10.1007/s10902-006-9018-1
- Deci, E. L., & Ryan, R. M. (2008c). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*(3), 182-185. doi:10.1037/a0012801
- Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former Eastern Bloc country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin, 27*(8), 930-942.
- Delle Fave, A., & Bassi, M. (2009). The contribution of diversity to happiness research. *The Journal of Positive Psychology, 4*(3), 205-207. doi:10.1080/17439760902844319
- Delle Fave, A., Brdar, I., Freire, T., Vella-Brodrick, D., & Wissing, M. P. (2011). The eudaimonic and hedonic components of happiness: Qualitative and quantitative findings. *Social Indicators Research, 100*(2), 185-207. doi:10.1007/s11205-010-9632-

- Delle Fave, A., Brdar, I., Wissing, M. P., & Vella-Brodrick, D. A. (2013). Sources and motives for personal meaning in adulthood. *The Journal of Positive Psychology, 8*(6), 517-529. doi:10.1080/17439760.2013.830761
- Delle Fave, A., & Soosai-Nathan, L. (2014). Meaning as inter-connectedness: Theoretical perspectives and empirical evidence. *Journal of Psychology in Africa, 24*(1), 53-68. doi:10.1080/14330237.2014.904090
- DeVellis, R. F. (2006). Classical test theory. *Medical Care, 44*(11), S50-S59. doi:10.1097/01.mlr.0000245426.10853.30
- DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Thousand Oaks, CA: Sage.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*(3), 542-575.
- Diener, E. (2009). *Assessing well-being: The collected works of Ed Diener*. New York, NY: Springer.
- Efklides, A., & Moraitou, D. (2013a). Introduction: Looking at quality of life and well-being from a positive psychology perspective. In A. Efklides & D. Moraitou (Eds.), *A positive psychology perspective on quality of life* (pp. 1-14). New York, NY: Springer.
- Efklides, A., & Moraitou, D. (Eds.). (2013b). *A positive psychology perspective on quality of life*. New York, NY: Springer.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Emmons, R. A. (2005). Striving for the sacred: Personal goals, life meaning, and religion. *Journal of Social Issues, 61*(4), 731-745.
- Fischer, R., & Fontaine, J. R. J. (2011). Methods for investigating structural equivalence. In D. Matsumoto & F. J. R. Van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 179-215). New York, NY: Cambridge University Press.

- Fowers, B. J. (2005). *Virtue and psychology: Pursuing excellence in ordinary practices*. Washington, DC: American Psychological Association.
- Fowers, B. J. (2010). Instrumentalism and psychology: Beyond using and being used. *Theory and Psychology, 20*(1), 102-124. doi:10.1177/0959354309346080
- Fowers, B. J. (2012). Placing virtue and the human good in psychology. *Journal of Theoretical and Philosophical Psychology, 32*(1), 1-9. doi:10.1037/a0025819
- Fox, C. M., & Jones, J. A. (1998). Uses of Rasch modeling in counseling psychology research. *Journal of Counseling Psychology, 45*(1), 30. doi:0022-0167/98/53.00
- Foxcroft, C. D. (2011). Ethical issues related to psychological testing in Africa: What I have learned (so far). *Online Readings in Psychology and Culture, Unit 2*. Retrieved from <http://scholarworks.gvsu.edu/orpc/vol2/iss2/7>
- Frankl, V. E. (1963). *Man's search for meaning: An introduction to logotherapy*. New York, NY: Washington Square Press.
- Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation & Emotion, 27*(3), 199-223.
- Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality, 77*(4), 1025-1050. doi:10.1111/j.1467-6494.2009.00573.x
- Grouden, M. E., & Jose, P. E. (2014). How do sources of meaning in life vary according to demographic factors? *New Zealand Journal of Psychology, 43*(3), 29-38.
- Gunnell, K. E., Crocker, P. R. E., Wilson, P. M., Mack, D. E., & Zumbo, B. D. (2013). Psychological need satisfaction and thwarting: A test of basic psychological needs theory in physical activity contexts. *Psychology of Sport & Exercise, 14*, 599-607. doi:10.1016/j.psychsport.2013.03.007

- Hagquist, C., Bruce, M., & Gustavsson, J. P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46*(3), 380-393. doi:10.1016/j.ijnurstu.2008.10.007
- Heintzelman, S. J., & King, L. A. (2013). On knowing more than we can tell: Intuitive processes and the experience of meaning. *Journal of Positive Psychology, 8*(6), 471-482. doi:10.1080/17439760.2013.830758
- Heintzelman, S. J., & King, L. A. (2014). Life is pretty meaningful. *The American Psychologist, 69*(6), 561-574. doi:10.1037/a0035049
- Helne, T., & Hirvilammi, T. (2015). Wellbeing and sustainability: A relational approach. *Sustainable Development, 23*(3), 167-175. doi:10.1002/sd.1581
- Henderson, L. W., & Knight, T. (2012). Integrating the hedonic and eudaimonic perspectives to more comprehensively understand wellbeing and pathways to wellbeing. *International Journal of Wellbeing, 2*(3), 196-221. doi:10.5502/ijw.v2i3.3
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. doi:10.1017/S0140525X0999152X
- Hodge, K., Lonsdale, C., & Ng, J. Y. Y. (2008). Burnout in elite rugby: Relationships with basic psychological needs fulfilment. *Journal of Sports Sciences, 26*(8), 835-844. doi:10.1080/02640410701784525
- Huta, V., & Ryan, R. (2010). Pursuing pleasure or virtue: The differential and overlapping well-being benefits of hedonic and eudaimonic motives. *Journal of Happiness Studies, 11*, 735-762. doi:10.1007/s10902-009-9171-4
- Ilardi, B. C., Leone, D., Kasser, T., & Ryan, R. M. (1993). Employee and supervisor ratings of motivation: Main effects and discrepancies associated with job satisfaction and adjustment in a factory setting. *Journal of Applied Social Psychology, 23*(21), 1789-1805.

- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36(2), 264-277. doi:10.1177/0022022104272905
- Johnston, M. M., & Finney, S. J. (2010). Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the Basic Needs Satisfaction in General Scale. *Contemporary Educational Psychology*, 35(4), 280-296. doi:10.1016/j.cedpsych.2010.04.003
- Joshanloo, M. (2015). Revisiting the empirical distinction between hedonic and eudaimonic aspects of well-being using exploratory structural equation modeling. *Journal of Happiness Studies*, 1-14. doi:10.1007/s10902-015-9683-z
- Jovanović, V. (2015a). A bifactor model of subjective well-being: A re-examination of the structure of subjective well-being. *Personality and Individual Differences*, 87, 45-49. doi:10.1016/j.paid.2015.07.026
- Jovanović, V. (2015b). Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. *Personality and Individual Differences*, 75, 154-159. doi:10.1016/j.paid.2014.11.026
- Kahneman, D., Diener, E., & Schwarz, N. (1999). *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage Foundation.
- Karabatsos, G. (2001). The Rasch model, additive conjoint measurement, and new models of probabilistic measurement theory. *Journal of Applied Measurement*, 2(4), 389-423.
- Karaś, D., Cieciuch, J., & Keyes, C. L. M. (2014). The Polish adaptation of the Mental Health Continuum-Short Form (MHC-SF). *Personality and Individual Differences*, 69, 104-109. doi:10.1016/j.paid.2014.05.011

- Kashdan, T., Biswas-Diener, R., & King, L. (2008). Reconsidering happiness: The costs of distinguishing between hedonics and eudaimonia. *The Journal of Positive Psychology, 3*(4), 219-233. doi:10.1080/17439760802303044
- Kashdan, T. B., Mishra, A., Breen, W. E., & Froh, J. J. (2009). Gender differences in gratitude: Examining appraisals, narratives, the willingness to express emotions, and changes in psychological needs. *Journal of Personality, 77*(3), 691-730. doi:10.1111/j.1467-6494.2009.00562.x
- Kasser, T., Davey, J., & Ryan, R. M. (1992). Motivation and employee-supervisor discrepancies in a psychiatric vocational rehabilitation setting. *Rehabilitation Psychology, 37*(3), 175-188. doi:10.1037/h0079104
- Keyes, C. L., Fredrickson, B. L., & Park, N. (2012). Positive psychology and the quality of life. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of social indicators and quality of life research* (pp. 99-112). New York, NY: Springer.
- Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly, 61*(2), 121-140.
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health & Social Behavior, 43*(2), 207-222. Retrieved from <http://www.jstor.org/stable/3090197>
- Keyes, C. L. M. (2006). The subjective well-being of America's youth: Toward a comprehensive assessment. *Adolescent and Family Health, 4*(1), 3-11. doi:10.1037/0002-9432.76.3.395
- Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. *American Psychologist, 62*(2), 95-108. doi:10.1037/0003-066X.62.2.95

- Keyes, C. L. M. (2009). Brief description of the Mental Health Continuum Short Form (MHC-SF). Retrieved from <https://www.aacu.org/sites/default/files/MHC-SFEnglish.pdf>.
- Keyes, C. L. M., Wissing, M. P., Potgieter, J. P., Temane, Q. M., Kruger, A., & Van Rooy, S. (2008). Evaluation of the Mental Health Continuum–Short Form (MHC–SF) in Setswana-speaking South Africans. *Clinical Psychology & Psychotherapy*, *15*(3), 181-192. doi:10.1002/cpp.572
- Khumalo, I. P., Wissing, M. P., Schutte, L., & De Klerk, W. (2015). Translation and measurement invariance of the Meaning in Life Questionnaire across three languages in South Africa. Manuscript submitted for publication.
- Knoop, H. H., & Delle Fave, A. (2013). Positive psychology and cross-cultural research. In H. H. Knoop & A. Delle Fave (Eds.), *Positive psychology and cross-cultural research* (pp. 1-10). New York, NY: Springer.
- Lamers, S. M. A., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. M. (2011). Evaluating the psychometric properties of the Mental Health Continuum–Short Form (MHC-SF). *Journal of Clinical Psychology*, *67*(1), 99-110. doi:10.1002/jclp.20741
- Lim, Y.-J. (2014). Psychometric characteristics of the Korean Mental Health Continuum–Short Form in an adolescent sample. *Journal of Psychoeducational Assessment*, *32*(4), 356-364. doi:10.1177/0734282913511431
- Linacre, J. M. (2014a). *A user's guide to Winsteps Ministep Rasch-model computer programs: Program manual 3.81.0*. Beaverton, Oregon: Winsteps.com.
- Linacre, J. M. (2014b). *Winsteps (Version 3.81.0)*. Beaverton, Oregon: Winsteps.com.
- Liu, S.-S., & Gan, Y.-Q. (2010). Reliability and validity of the Chinese version of the Meaning in Life Questionnaire. *Chinese Mental Health Journal*, *24*(6), 478-482.

- Longo, Y., Gunz, A., Curtis, G. J., & Farsides, T. (2016). Measuring need satisfaction and frustration in educational and work contexts: The Need Satisfaction and Frustration Scale (NSFS). *Journal of Happiness Studies, 17*(1), 295-317. doi:10.1007/s10902-014-9595-3
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association 58*(303), 690-700.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal Of The National Cancer Institute, 22*(4), 719-748.
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85-110. doi:10.1146/annurev-clinpsy-032813-153700
- Matsumoto, D., & Yoo, S. H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Science, 1*(3), 234-250. doi:10.1111/j.1745-6916.2006.00014.x
- Meyer, B., Enström, M. K., Harstveit, M., Bowles, D. P., & Beevers, C. G. (2007). Happiness and despair on the catwalk: Need satisfaction, well-being, and personality adjustment among fashion models. *Journal of Positive Psychology, 2*(1), 2-17. doi:10.1080/17439760601076635
- Milyavskaya, M., Philippe, F. L., & Koestner, R. (2013). Psychological need satisfaction across levels of experience: Their organization and contribution to general well-being. *Journal of Research in Personality, 47*(1), 41-51. doi:10.1016/j.jrp.2012.10.013

- Morgan, J., & Farsides, T. (2009). Measuring meaning in life. *Journal of Happiness Studies*, *10*(2), 197-214. doi:10.1007/s10902-007-9075-0
- Muller, F. H., & Louw, J. (2004). Learning environment, motivation and interest: Perspectives on self-determination theory. *South African Journal of Psychology*, *34*(2), 169-190. doi:10.1177/008124630403400201
- Muthén, L. K., & Muthén, B. O. (1998-2014). Mplus (Version 7.31). Los Angeles, CA: Muthén and Muthén.
- Myers, N. D., Wolfe, E. W., Feltz, D. L., & Penfield, R. D. (2006). Identifying differential item functioning of rating scale items with the Rasch model: An introduction and an application. *Measurement in Physical Education & Exercise Science*, *10*(4), 215-240. doi:10.1207/s15327841mpee1004_1
- Ng, J. Y. Y., Ntoumanis, N., Thøgersen-Ntoumani, C., Deci, E. L., Ryan, R. M., Duda, J. L., & Williams, G. C. (2012). Self-determination theory applied to health contexts: A meta-analysis. *Perspectives on Psychological Science*, *7*(4), 325-340. doi:10.1177/1745691612447309
- Nowlis, S. M., Kahn, B. E., & Dhar, R. (2002). Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research*, *29*(3), 319-334. doi:10.1086/344431
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Oishi, S. (2010). Culture and well-being: Conceptual and methodological issues. In E. Diener, J. F. Helliwell, & D. Kahneman (Eds.), *International differences in well-being* (pp. 34-69). New York, NY: Oxford University Press.
- Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, *46*(1), 1-18. doi:10.1348/014466506X96931

- Park, C. L. (2015). Integrating positive psychology into health-related quality of life research. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, 24(7), 1645-1651. doi:10.1007/s11136-014-0889-z
- Peltzer, K. (2002). Brain fog symptoms among black South African university students. *Journal of Child and Adolescent Mental Health*, 14(2), 115-122.
doi:10.1080/16826108.2002.9632433
- Peltzer, K., Cherian, V. I., & Cherian, L. (1998). Brain fog symptoms in rural South African secondary school pupils. *Psychological Reports*, 83(3), 1187-1196.
- Petrillo, G., Capone, V., Caso, D., & Keyes, C. L. M. (2015). The Mental Health Continuum-Short Form (MHC-SF) as a measure of well-being in the Italian context. *Social Indicators Research*, 121(1), 291-312. doi:10.1007/s11205-014-0629-3
- Philippe, F. L., Koestner, R., Beaulieu-Pelletier, G., & Lecours, S. (2011). The role of need satisfaction as a distinct and basic psychological component of autobiographical memories: A look at well-being. *Journal of Personality*, 79(5), 905-937.
doi:10.1111/j.1467-6494.2010.00710.x
- Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, 24, 737-756. doi:10.1080/00207598908247842
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*.
Copenhagen, Denmark Danmarks Peadagogiske Institut.
- Reeve, B. B., & Fayers, P. (2005). Applying item response theory modelling for evaluating questionnaire item and scale properties. In P. Fayers & R. D. Hays (Eds.), *Assessing quality of life in clinical trials: Methods and practice* (pp. 55-73). New York, NY: Oxford University Press.

- Reise, S. P. (2012). Invited Paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696.
doi:10.1080/00273171.2012.715555
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544-559.
doi:10.1080/00223891.2010.496477
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology, 5*, 27-48. doi:10.1146/annurev.clinpsy.032408.153553
- Reker, G. T. (1997). Personal meaning, optimism, and choice: Existential predictors of depression in community and institutional elderly. *The Gerontologist, 37*, 709-716.
- Richardson, F. C. (2012). On psychology and virtue ethics. *Journal of Theoretical and Philosophical Psychology, 32*(1), 24-34. doi:10.1037/a0026058
- Roman, N. V., Davids, E. L., Moyo, A., Schilder, L., Lacante, M., & Lens, W. (2015). Parenting styles and psychological needs influences on adolescent life goals and aspirations in a South African setting. *Journal of Psychology in Africa, 25*(4), 305-312. doi:10.1080/14330237.2015.1078087
- Rothmann, S., Diedericks, E., & Swart, J., P. (2013). Manager relations, psychological need satisfaction and intention to leave in the agricultural sector. *South African Journal of Industrial Psychology, 39*(2), 1-11. doi:10.4102/sajip.v39i2.1129
- Ryan, R. M., Curren, R. R., & Deci, E. L. (2013). What humans need: Flourishing in Aristotelian philosophy and self-determination theory. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 57-75). Washington, DC: American Psychological Association.

- Ryan, R. M., & Deci, E. L. (2000a). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Inquiry, 11*(4), 319-338.
doi:10.1207/S15327965PLI1104_03
- Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.
doi:10.1037/110003-066X.55.1.68
- Ryan, R. M., Huta, V., & Deci, E. L. (2008). Living well: A self-determination theory perspective on eudaimonia. *Journal of Happiness Studies, 9*(1), 139-170.
doi:10.1007/s10902-006-9023-4
- Ryff, C. D. (1989a). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful ageing. *International Journal of Behavioral Development, 12*(1), 35-55.
doi:10.1177/016502548901200102
- Ryff, C. D. (1989b). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*(6), 1069-1081. doi:10.1037/0022-3514.57.6.1069
- Ryff, C. D. (2013). Eudaimonic well-being and health: Mapping consequences of self-realization. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 77-98). Washington, DC: American Psychological Association.
- Ryff, C. D., & Singer, B. H. (2008). Know thyself and become what you are: A eudaimonic approach to psychological well-being. *Journal of Happiness Studies, 9*(1), 13-39.
doi:10.1007/s10902-006-9019-0
- Schnell, T. (2009). The Sources of Meaning and Meaning in Life Questionnaire (SoMe): Relations to demographics and well-being. *The Journal of Positive Psychology, 4*(6), 483-499. doi:10.1080/17439760903271074

- Schnell, T. (2010). Existential indifference: Another quality of meaning in life. *Journal of Humanistic Psychology, 50*(3), 351-373. doi:10.1177/0022167809360259
- Schulenberg, S. E., Strack, K. M., & Buchanan, E. M. (2011). The meaning in life questionnaire: Psychometric properties with individuals with serious mental illness in an inpatient setting. *Journal of Clinical Psychology, 67*(12), 1210-1219. doi:10.1002/jclp.20841
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.
- Schutte, L., Wissing, M. P., Ellis, S. M., Jose, P. E., & Vella-Brodrick, D. A. (2016). Rasch analysis of the Meaning in Life Questionnaire among adults from South Africa, Australia, and New Zealand. *Health and Quality of Life Outcomes, 14*(1), 1-15. doi:10.1186/s12955-016-0414-x
- Sheldon, K. M. (2002). The self-concordance model of healthy goal striving: When personal goals correctly represent the person. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 65-86). Rochester, NY: University of Rochester Press.
- Sheldon, K. M. (2013). Individual daimon, universal needs, and subjective well-being: Happiness as the natural consequence of a life well lived. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 119-137). Washington, DC: American Psychological Association.
- Sheldon, K. M., & Hilpert, J. C. (2012). The Balanced Measure of Psychological Needs (BMPN) Scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion, 36*, 439-451. doi:10.1007/s11031-012-9279-4

- Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, 2(1), 414-433. doi:10.1111/j.1751-9004.2007.00044.x
- Sireci, S. G. (2011). Evaluating test and survey items for bias across languages and cultures. In D. Matsumoto & F. J. R. Van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 216-243). New York, NY: Cambridge University Press.
- Slife, B. D. (2012). Virtue ethics in practice: The Greenbrier Academy. *Journal of Theoretical and Philosophical Psychology*, 32(1), 35-42. doi:10.1037/a0026443
- Steger, M. F. (2009). Meaning in life. In S. J. Lopez, C. R. Snyder, S. J. Lopez, & C. R. Snyder (Eds.), *Oxford handbook of positive psychology* (2nd ed., pp. 679-687). New York, NY: Oxford University Press.
- Steger, M. F. (2012). Experiencing meaning in life: Optimal functioning at the nexus of well-being, psychopathology, and spirituality. In P. T. P. Wong (Ed.), *The human quest for meaning: Theories, research, and applications* (2nd ed., pp. 165-184). New York, NY: Routledge.
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53(1), 80-93. doi:10.1037/0022-0167.53.1.80
- Steger, M. F., Frazier, P. A., & Zacchanini, J. L. (2008). Terrorism in two cultures: Stress and growth following September 11 and the Madrid train bombings. *Journal of Loss & Trauma*, 13(6), 511-527. doi:10.1080/15325020802173660
- Steger, M. F., Kashdan, T. B., Sullivan, B. A., & Lorentz, D. (2008). Understanding the search for meaning in life: Personality, cognitive style, and the dynamic between seeking and experiencing meaning. *Journal of Personality*, 76(2), 199-228. doi:10.1111/j.1467-6494.2007.00484.x

- Steger, M. F., Kawabata, Y., Shimai, S., & Otake, K. (2008). The meaningful life in Japan and the United States: Levels and correlates of meaning in life. *Journal of Research in Personality, 42*(3), 660-678. doi:10.1016/j.jrp.2007.09.003
- Steger, M. F., Oishi, S., & Kashdan, T. B. (2009). Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology, 4*(1), 43-52. doi:10.1080/17439760802303127
- Steger, M. F., Shin, J. Y., Shim, Y., & Fitch-Martin, A. (2013). Is meaning in life a flagship indicator of well-being? In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 159-182). Washington, DC: American Psychological Association.
- Temane, L., Khumalo, I. P., & Wissing, M. P. (2014). Validation of the Meaning in Life Questionnaire in a South African context. *Journal of Psychology in Africa, 24*(1), 81-95. doi:0.1080/14330237.2014.904088
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism: Arthritis Care & Research, 57*(8), 1358-1362. doi:10.1002/art.23108
- Thomas, M. L. (2011). The value of item response theory in clinical assessment: A review. *Assessment, 18*(3), 291-307. doi:10.1177/1073191110374797
- Triandis, H. C. (2007). Culture and psychology: A history of the study of their relationships. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 59-76). New York, NY: Guilford Press.
- Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research:

- Introduction to the special section. *Psychological Assessment*, 21(4), 457-462.
doi:10.1037/a0017653
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492.
doi:10.1080/17405629.2012.686740
- Van de Vijver, F. J. R. (2011). Capturing bias in structural equation modeling. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 3-34). New York, NY: Routledge.
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17-45). New York, NY: Cambridge University Press.
- Van den Broeck, A., Vansteenskiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the Work-related Basic Need Satisfaction Scale. *Journal of Occupational and Organizational Psychology*, 83(4), 981-1002. doi:10.1348/096317909X481382
- Van Zyl, L. E., & Rothmann, S. (2012). Beyond smiling: The evaluation of a positive psychological intervention aimed at student happiness. *Journal of Psychology in Africa*, 22(3), 369-384. doi:10.1080/14330237.2012.10820541
- Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, 23(3), 263-280. doi:10.1037/a0032359

- Vázquez, C., Rahona, J. J., Gómez, D., Caballero, F. F., & Hervás, G. (2015). A national representative study of the relative impact of physical and psychological problems on life satisfaction. *Journal of Happiness Studies*, *16*(1), 135-148. doi:10.1007/s10902-014-9501-z
- Veenhoven, R. (2010). How universal is happiness? In E. Diener, J. F. Helliwell, & D. Kahneman (Eds.), *International differences in well-being* (pp. 328-350). New York, NY: Oxford University Press.
- Veenhoven, R. (2012). Does happiness differ across cultures? In H. Selin & G. Davey (Eds.), *Happiness across cultures: Views of happiness and quality of life in non-Western cultures* (pp. 451-472). New York, NY: Springer.
- Vittersø, J., & Søholt, Y. (2011). Life satisfaction goes with pleasure and personal growth goes with interest: Further arguments for separating hedonic and eudaimonic well-being. *The Journal of Positive Psychology*, *6*(4), 326-335.
doi:10.1080/17439760.2011.584548
- Vlachopoulos, S. P., & Michailidou, S. (2006). Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The Basic Psychological Needs in Exercise Scale. *Measurement in Physical Education & Exercise Science*, *10*(3), 179-201.
- Waterman, A. S. (1990). Personal expressiveness: Philosophical and psychological foundations. *Journal of Mind and Behavior*, *11*(1), 47-73.
- Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678-691. doi:10.1037/0022-3514.64.4.678
- Waterman, A. S. (2008). Reconsidering happiness: A eudaimonist's perspective. *The Journal of Positive Psychology*, *3*(4), 234-252. doi:10.1080/17439760802303002

- Waterman, A. S. (2013). Introduction: Considering the nature of a life well-lived — Intersections of positive psychology and eudaimonist philosophy. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 3-17). Washington, DC: American Psychological Association.
- Waterman, A. S., & Schwartz, S. J. (2013). Eudaimonic identity theory. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 99-118). Washington, DC: American Psychological Association.
- Waterman, A. S., Schwartz, S. J., & Conti, R. (2008). The implications of two conceptions of happiness (hedonic enjoyment and eudaimonia) for the understanding of intrinsic motivation. *Journal of Happiness Studies*, 9(1), 41-79. doi:10.1007/s10902-006-9020-7
- Waterman, A. S., Schwartz, S. J., Zamboanga, B. L., Ravert, R. D., Williams, M. K., Agocha, V. B., . . . Donnellan, M. B. (2010). The Questionnaire for Eudaimonic Well-Being: Psychometric properties, demographic comparisons, and evidence of validity. *The Journal of Positive Psychology*, 5(1), 41-61. doi:10.1080/17439760903435208
- Weijters, B., & Baumgartner, H. (2012). Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research*, 49(5), 737-747. doi:10.1509/jmr.11.0368
- Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(3), 236-247. doi:10.1016/j.ijresmar.2010.02.004
- Wissing, M. P. (2008/2012). *The prevalence of levels of psychosocial health: Dynamics and relationships with biomarkers of (ill) health in South African social contexts (FORT3)*.

- Wissing, M. P. (2014). Meaning and relational well-being: A reflection on the state of the art and a way forward. *Journal of Psychology in Africa, 24*(1), 183-192.
doi:10.1080/14330237.2014.904100
- Wissing, M. P., & Temane, Q. M. (2008). The structure of psychological well-being in cultural context: Towards a hierarchical model of psychological health. *Journal of Psychology in Africa, 18*(1), 45-55. doi:10.1080/14330237.2008.10820170
- Wong, P. T. P. (2010). Meaning therapy: An integrative and positive existential psychotherapy. *Journal of Contemporary Psychotherapy, 40*(2), 85-93.
doi:10.1007/s10879-009-9132-6
- Wong, P. T. P. (2011). Positive psychology 2.0: Towards a balanced interactive model of the good life. *Canadian Psychology, 52*(2), 69-81. doi:10.1037/a0022511