Validation of the Mental Health Continuum Short Form – Revised

Examination of Model Fit, Internal Consistency, and Convergent Validity using Resilience and Self-Esteem

Bachelor Thesis

by Max Böggemann

Supervisors:

Dr. Marion Sommers-Spijkerman

Prof. Dr. Gerben J. Westerhof

June 27, 2019

Abstract

The most widely used instrument to measure well-being is the Mental Health Continuum – Short Form (MHC-SF). The original version was revised, since Köhle (2010) found that participants had problems with the comprehension of items. This cross-sectional, online survey study aimed to validate the revised version - the MHC-SF-R - by testing the factorial structure, internal consistency, and convergent validity of the MHC-SF-R and MHC-SF in a student sample (N=108). Confirmatory factor analysis (CFA) indicated insufficient fit of the MHC-SF-R items to a four-factor model (emotional well-being (WB), psychological WB, societal WB, relational WB). Results showed good internal consistency for the total scale. However, the societal WB subscale performed poorly in terms of factor loadings and internal consistency. Evidence for convergent validity was found by comparing the scores of the MHC-SF-R to related measures of well-being. Firstly, self-esteem with the Rosenberg's Self-Esteem Scale and secondly, resilience with the Brief Resilience Scale. Results of the MHC-SF-R and the MHC-SF were compared, and both scales showed poor model fit. The emotional WB subscale of the MHC-SF-R performed better in terms of internal consistency and convergent validity. Moreover, the relational WB subscale performed good as well. However, the social WB subscale and the psychological WB subscale of the MHC-SF remain the better choice for practical usage. By implementing further refinements, the MHC-SF-R could become a reliable and valid instrument to measure well-being, and possibly obtain higher validity for measuring social well-being compared to the original version, the MHC-SF.

Table of Contents

Introduction

Mental health is defined by the World Health Organization (2014) as:

"*...a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community."*

Therefore, to be mentally healthy means to be in a state of well-being which is more than just the absence of illness. Vella-Brodrick and Klein (2010) explain that well-being is a multifaceted construct for which they state the definition of Ryan and Deci (2001), who characterize well-being as a construct concerned with optimal experience and functioning. The approach of positive psychology concentrates on assessing the resources for *well-being* (e.g. positive emotions and strengths) compared to the traditional pathological approach which concentrates on the *illness*. Positive psychology implies that the absence of mental illness does not equal the presence of full mental health and that those are not two ends on the same continuum, but that they are related in some ways. Evidence for this claim was for instance found by Karaś, Cieciuch, and Keyes (2014) and Petrillo, Capone, Caso, and Keyes (2015).

In the paper of Vella-Brodrick and Klein (2010), the authors describe the dual category approach of measuring mental well-being. Firstly, the hedonic approach which has at its core the subjective experience of happiness or pleasure, also referred to as *subjective or emotional well-being*, which dominated the research on well-being measures for a long time. Secondly, the eudaimonic approach which focuses on the degree to which a person is fully functioning and has at its core personal growth and meaning also referred to as *psychological* and *social well-being*. In contrast to the hedonic approach, it focuses on fulfilling one's potential instead of evaluating one's feelings and experiences. Moreover, Keyes (2009) explains that the social dimension is concerned with optimal functioning in a social context.

The Mental Health Continuum – Short Form

The Mental Health Continuum – Short Form is the most widely used scale to measure the tripartite model of well-being, consisting of emotional well-being (WB), psychological WB, and social WB (Mental Treatment Centers, 2018). The Mental Health Continuum Short Form (MHC-SF) assesses the mental well-being of individuals in either a continuous scoring or in a categorical way as

flourishing (high level of well-being), languishing (low level of well-being), or as moderate (Keyes, 2009). The MHC-SF was derived from the MHC – Long Form (MHC-LF). It measures emotional WB with six items derived from the Bradburn's (1969) affect balance scale, and a single item for the quality of life overall based on Cantril's (1965) self-anchoring items, the six dimensions of Ryff's (1989) model of psychological WB with 18 items, and the five dimensions of Keyes' (1998) model of social well-being with 15 items.

The MHC-SF consists of 14 items which were found to be representative for each facet of WB and hence, forms a more efficient way of measuring well-being. The *emotional WB* subscale consists of three items concerning happiness, interest in life, and satisfaction. The *social WB* subscale consists of five items concerning social contribution, social integration, social actualization, social acceptance, and social coherence. The *psychological WB* subscale measures self-acceptance, environmental mastery, positive relations with others, personal growth, autonomy, and purpose in life with six items (Keyes, 2009).

MHC-SF: Psychometric Properties

The psychometric properties of the MHC-SF were examined in several countries and in many studies. Evidence for the three factor structure of the MHC-SF (emotional, psychological, and social WB) have been found in national representative samples in the Netherlands (Lamers, Westerhof, Bohlmeijer, ten Klooster, & Keyes, 2011), in South Africa (Keyes et al., 2008), in Poland (Karaś, et al., 2014), in Italy (Joshanloo, Capone, Petrillo, & Caso, 2017), and furthermore, in a large sample (N = 2907) of adolescents (Keyes, 2005).

Furthermore, good internal consistency (> .80) and discriminant validity in adolescents aged 12-18 were found (Keyes, 2005; Keyes, 2006). Moreover, good internal consistency was found for adults in the Netherlands (Lamers et al., 2011; Westerhof & Keyes, 2010), and in South Africa (Keyes et al., 2008). The Polish and Italian adaptation found good internal consistency as well, with a Cronbach's $\alpha > .80$ (Karaś, et al., 2014; Petrillo, et al., 2015). Additionally, the Italian version showed convergent validity with several related well-being measures, good divergent and discriminant validity, and moderate test-retest reliability (Petrillo, et al., 2015). Lamers et al. (2011) found a test-retest reliability over three successive 3-month periods of .68 and .65 for 9 months.

Although the MHC-SF showed good values in terms of reliability and validity, several items of the social WB subscale were found to be rather abstract and confusing for participants (G. J. Westerhof, personal communication, 2019 March 7). Köhle (2010) also reports in her interview study about such feedback from participants. She examined weaknesses of the MHC-SF through a three-step interview procedure (including the 'think-aloud' method) with 30 people living in Enschede (the Netherlands). Participants reported to have difficulties answering the items of the social WB scale and they reported that some item's formulation was confusing, especially when asked about the frequency of feelings.

Therefore, the revised version of the MHC-SF (MHC-SF-R) was developed to counteract those problems. In the revised version, participants are asked about the past week (compared to the last month in the original version), which could be better since participants had difficulties when asked about the frequency of specific feelings and those are easier to recall for the last week. Furthermore, the revised version made only minor changes (for instance shortening) to the items of the emotional and psychological WB scales. One major change is the splitting of the social WB scale into two dimensions, the societal WB subscale and the relational WB subscale. The societal WB subscale is the rephrased version of the social WB subscale of the MHC-SF and the relational WB subscale was newly developed and added, which resulted in five more items for the MHC-SF-R compared to the original version.

Aim and Hypotheses

This study aimed to validate the MHC-SF-R by examining the factorial structure, the internal consistency, and convergent validity. All three properties are also examined for the MHC-SF, for comparison purposes. As Poulin, Lemoine, Poirier, and Lambert (2005) explain in their validation study, researchers usually base their results of a validation on the correlations to other measures of mental health, clinical rating, or discrimination between target groups. Therefore, the convergent validity of the MHC-SF-R is examined by computing the correlations with two related constructs of well-being, psychological resilience and self-esteem.

Resilience is described by Hjemdal, Friborg, Stiles, Rosenvinge, and Martinussen (2006) as a dynamic concept capturing positive adaptation and even growth in the face of stress and trauma. Several studies examined the relationship between resilience and well-being. Fuller-Iglesias, Sellars, and Antonucci (2008) found that resilience contributes to overall life satisfaction

and mental well-being in old age. Moreover, other papers report about the relationship between emotional WB and resilience (Cummins & Wooden, 2014), resilience and mental health (Davydov, Stewart, Ritchie, & Chaudieu, 2010; Hjemdal, et al., 2006), and resilience and positive psychological WB (Davydov, et al., 2010; Patel, & Goodman, 2007).

Self-esteem is defined as a stable sense of self-worth and an important part of the self-concept besides self-efficacy and self-identification (Rosenberg, 1965). Mann, Hosman, Schaalma, and De Vries (2004) summarize research findings and found that positive self-esteem is for instance associated with mental well-being, happiness, adjustment, academic achievements, and satisfaction. Other studies found that high self-esteem does lead to greater happiness (Baumeister, Campbell, Krueger, & Vohs, 2003), that self-esteem is related to lower symptoms of depression and anxiety (Henriksen, Ranøyen, Indredavik, & Stenseng, 2017), that there is a relationship between low self-esteem and poor social functioning and school dropout (Mann, et al., 2004), and that high self-esteem is linked to job satisfaction, good social relations, and increased levels of well-being (Orth, Robins, & Widaman, 2012).

This results in the following hypotheses:

1. The factorial structure shows evidence for the quadripartite model of emotional WB, psychological WB, societal WB, and relational WB.

2.1 The MHC-SF-R shows good internal consistency.

2.2 The MHC-SF-R shows equal or better internal consistency compared to the MHC-SF.

3.1 The MHC-SF-R shows convergent validity with a moderate to strong, positive correlation to the constructs: resilience and self-esteem.

3.2 The MHC-SF-R shows equal or stronger convergent validity compared to the MHC-SF.

Method

Design

This study was conducted as a cross-sectional, online survey design. Ethical approval was obtained by the Ethics committee of the Behavioral, Management and Social sciences (BMS) department of the University of Twente (Registration number: 190441).

Participants

The target group of this study were university students (from multiple universities in at least four countries). 125 participants in this study were recruited by convenience sampling. Recruitment was done by advertising it through friends, colleagues, and other people who were known to study at a university (from 12.04.2019 till 06.05.2019). Furthermore, this study was presented in the test subject pool system 'SONA' and undergraduate psychology students of the University of Twente were granted 'SONA credits' (required for graduation) for completing the survey. One inclusion criterion was access to a device with internet connection, since the participants needed to complete the survey online. Further, inclusion criteria were the given informed consent, that they are above 18, that they are enrolled at a University, and that their English proficiency level is above intermediate (advanced, fluent, and native speaker). Participants who did not meet these criteria were excluded and they were redirected to the end of the survey.

125 Participants opened the link to the survey and 9 were excluded due to unmet inclusion criteria (7 because they were below the required English proficiency level and 3 because they were not enrolled at a university). Moreover, 7 were excluded because they did not finish the survey (data was not saved for those participants). One participant was excluded because his scores differed substantially which resulted in the data deviating from normal distribution (see Analysis). The final sample consisted of 108 participants. The participants' age ranged from 18 to 31 years (M= 22.03, SD= 2.65). Moreover, approximately 55% reported they were female and 45% to be male. Further, approximately 14% were Dutch, 73% German, and 13% of another nationality.

In addition, demographic characteristics for both groups were assessed regarding substantial differences. Firstly, age was not substantially different for the groups, with M= 21.60 (22.47) and SD= 2.27 (2.95) for the MHC-SF (MHC-SF-R). Further, there was no substantial difference found in the distribution of nationalities for the groups, with approximately 13% (15%)

being Dutch, 75% (72%) German, and 13% (13%) being of another nationality for the MHC-SF (MHC-SF-R). Lastly, differences for gender was examined and no substantial difference was found for the groups, with approximately 44% (47%) reported to be male and 56% (53%) to be female for the MHC-SF (MHC-SF-R).

Materials

Since the survey was available online, a computer with internet access was required. Furthermore, several questionnaires were used. Firstly, the MHC-SF-R, consisting of 19 items, was used to measure well-being with four subscales, emotional WB, psychological WB, societal WB, and relational WB. It uses a Likert-scale scoring (Never: 0 points – (Almost) always: 5 points). The total scores range is 0 to 95 points, a higher score indicates a higher level of well-being.

Secondly, the MHC-SF, consisting of 14 items, was used. It measured well-being and includes three subscales, emotional WB, psychological WB, and social WB (Keyes, 2009). Its psychometric properties were examined in many studies conducted in many countries (Joshanloo, et al., 2017; Karaś, et al., 2014; Keyes, 2005; Keyes, 2006; Keyes et al., 2008; Lamers, et al., 2011; Petrillo, et al., 2015; Westerhof & Keyes, 2010). The MHC-SF uses a 6-point Likert-scale scoring (Never: 0 points - Every day: 5 points). The range is 0 to 70 points, a higher score indicates a higher level of well-being.

Thirdly, the Brief Resilience Scale (BRS), consisting of 6 items, was used to measure resilience. It uses a 5-point Likert-scale scoring (Strongly disagree: 1 point – Strongly agree: 5 points) and half of the items are reversed scored. The range is 6 to 30 points, a higher score means a higher level of resilience. The BRS was chosen to examine the level of resilience due to its free accessibility and its good psychometric properties, including convergent validity with other resilience measures (Smith, et al., 2008). Moreover, Windle, Noyes, and Bennett (2011) assessed many measurement scales for resilience and the BRS was one of the three best scoring tests concerning psychometric properties. Additionally, the results of this study indicated good internal consistency for the BRS (Cronbach's $\alpha = .84$).

Finally, the Rosenberg's Self-Esteem Scale (RSES), consisting of 10 items, was used to measure the level of self-esteem (Rosenberg, 1989). The RSES uses a 4-point Likert-scale scoring (Strongly disagree: 1 point – Strongly agree: 4 points) and half of the items are reversed scored. The range is 10 to 40 points, a higher score means a higher level of self-esteem. The RSES proved

to have good psychometric properties (e.g. Franck, De Raedt, Barbez, & Rosseel, 2008; Rizwan, Aftab, Shah, & Dharwarwala, 2012; Sinclair, et al., 2010). Moreover, the results of this study found good internal consistency for the RSES (Cronbach's $\alpha = .86$).

Procedure

Firstly, after clicking on the surveys link, participants were redirected to the online survey which was provided via Qualtrics. The first screen was the welcome screen which introduced them to the topic of the study and the general procedure. Afterwards, the informed consent was presented, and the survey only continued when they agreed to it (otherwise they were redirected to the end of the survey, due to unmet inclusion criteria). They were also screened for being underage, for their English proficiency level, and if they are currently enrolled at a University. If the participants met all inclusion criteria, gender and nationality were asked in the next step.

Afterwards, the survey continued and showed either the MHC-SF-R or the MHC-SF (determined by the randomization option of Qualtrics), for the purpose of comparing both questionnaires. Randomization was chosen to reduce the chance of systematic differences of the two groups. Afterwards, the RSES and then the BRS were presented. The survey took approximately 15 minutes to complete and was available to participants from 12 April 2019 until 6. May 2019.

Analysis

The data was analyzed using Lisrel 10.10 (factorial structure) and SPSS 25 (internal consistency and convergent validity). In order to analyze the results of this study, several analyses were conducted. Firstly, the data was tested for normality. The Shapiro-Wilk test showed that the scores of the RSES and BRS are normally distributed, and after removing one outlier, it also indicated normality for the psychological WB subscale of the MHC-SF-R. For all other scales, the Shapiro-Wilk test did not indicate normality. However, since the Shapiro-Wilk test is strict in terms of showing normal distribution, Q-Q-plots showed that the data for those scales is approximately normal distributed.

To examine the factorial structure of the MHC-SF-R, confirmatory factor analysis (CFA) was done to determine if there is evidence for the quadripartite model of well-being. For the CFA each item was restricted to load on one factor. Factor loadings >.5 were seen as acceptable and >.7

as ideal (Hair, Black, Babin, & Anderson, 2014). Moreover, the normal maximum likelihood estimation method was used, since robust maximum likelihood estimation method could not be used due to the limited number of participants. The used indices are the model Chi-Square (χ 2) and root mean square error of approximation (RMSEA). The Chi-Square test indicates poor fit when the χ 2 value is large and the P-value is below the significance level (Agresti, 2007). The significance level $p < .05$ was used. For the RMSEA, a cut-off proximate to .06 indicates good fit (Hu & Bentler, 1999). In addition, the factorial structure of the MHC-SF was also examined.

Moreover, a Reliability Analysis was conducted to examine the internal consistency (Cronbach's α) of the MHC-SF-R. Cronbach's α was used as a general reliability coefficient to determine if the items have different underlying properties or qualities. A value of ≥.70 was viewed as acceptable, a value of ≥ 0.80 as good, and if it was ≥.90, it was seen as excellent (Cicchetti, 1994). In addition, intercorrelations of subscales were examined through correlation analysis (Pearson) for the MHC-SF-R to further measure internal consistency. Correlations <.10 were considered weak, correlations between .10 and .30 were considered small, and correlations between .30 and .50, and .50 and 1.00 were considered moderate and strong, respectively (Cohen, 1988). The arbitrary cut-off point ≤.70 was used to determine if subscales are sufficiently distinct but related. Internal consistency and intercorrelations were also examined for the MHC-SF, for comparison purposes.

Furthermore, convergent validity was examined by calculating Pearson correlations between the scores of the MHC-SF-R and scores of the BRS and the RSES. Therefore, a correlation analysis (2-tailed) was conducted and the significance level $p < .05$ was used. For the MHC-SF, convergent validity was also examined.

Results

Factorial Structure

The findings (see Figure 1) demonstrated poor fit of the four-factor model of the MHC-SF-R to the data (χ 2 (3, N = 53) = 226.37, $p < .05$; RMSEA: .102). Factor loadings ranged from .17 to .92 and the societal WB subscale showed the lowest factor loadings with items 4,6,7, and 8 being

below the .5 cut-off point. Further, item 13 of the psychological WB subscale showed a factor loading below the .5 cut-off point. The emotional WB subscale of the MHC-SF-R showed high factor loadings. The emotional WB subscale of the MHC-SF showed slightly lower factor loadings in this regard. For the MHC-SF, the findings (see Figure 2) indicate similar poor fit for its three-factor model (χ 2 (2, N = 55) = 117.34, p <.05; RMSEA: .103) with factor loadings ranging from .56 to .84.

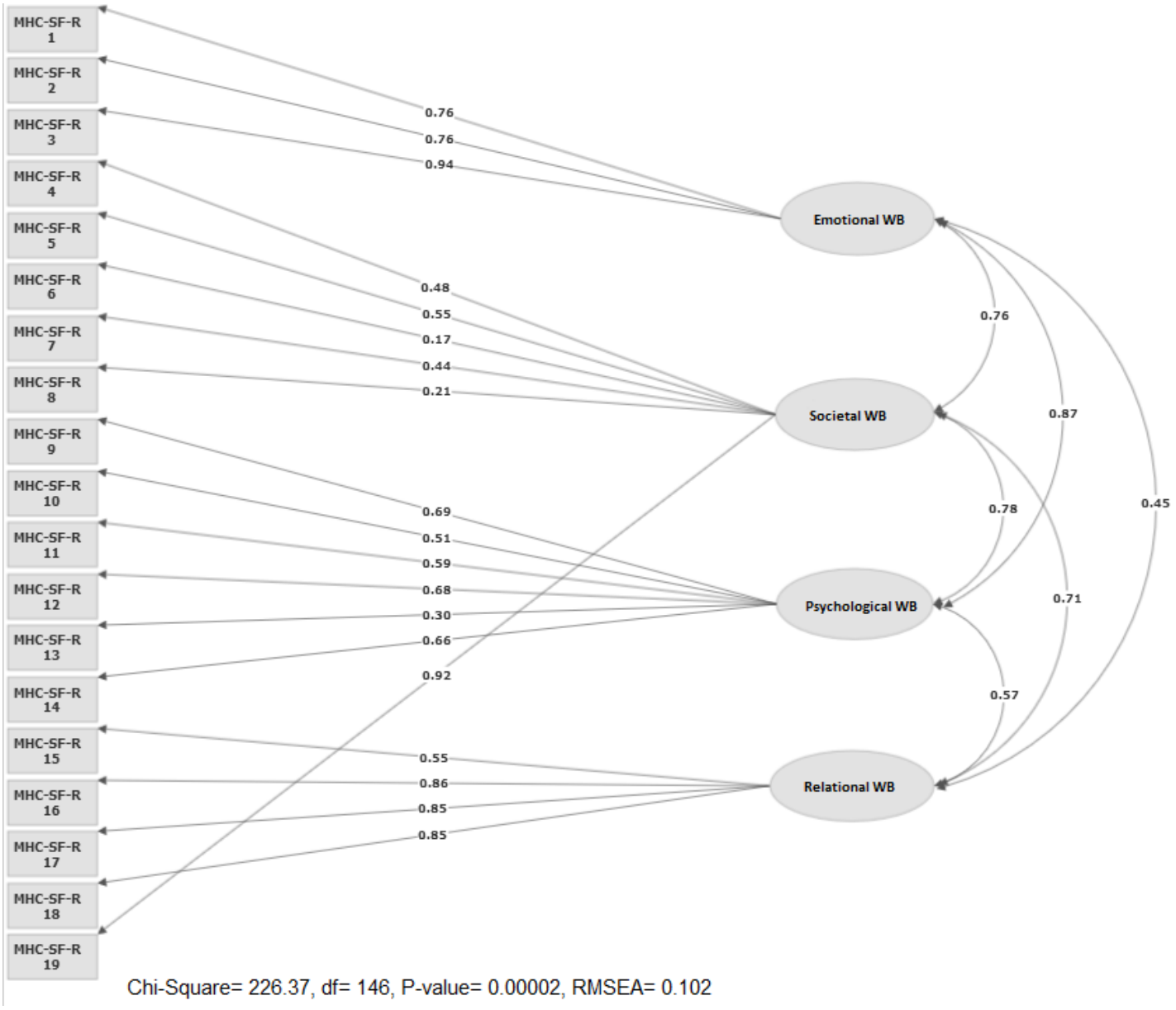

Figure 1

Confirmatory Factor Analysis of the MHC-SF-R

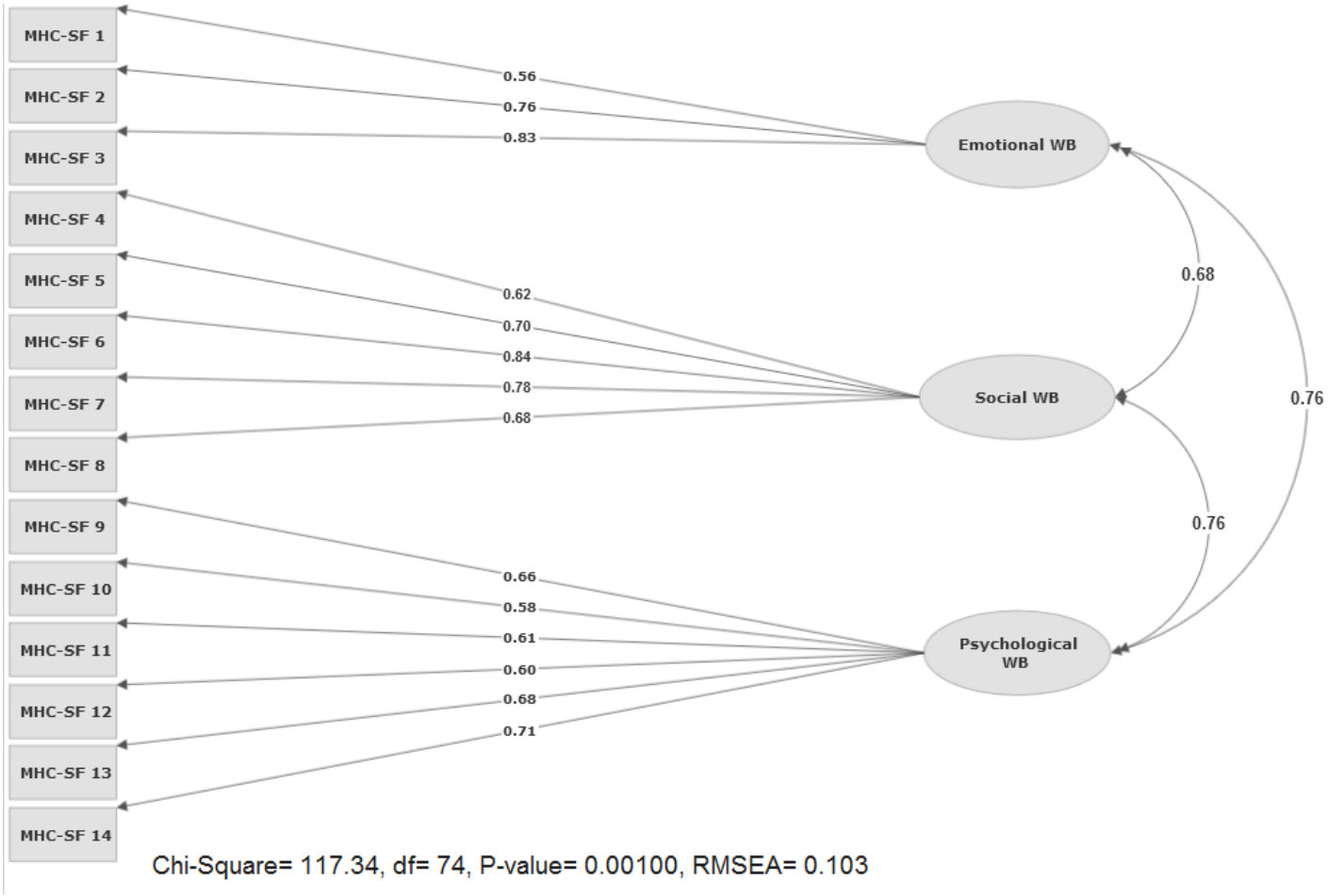

Figure 2

Confirmatory Factor Analysis of the MHC-SF

Internal Consistency

Cronbach's α of the MHC-SF-R, the MHC-SF, and their subscales can be found in Table 1. The findings indicated good internal consistency for the total scale of the MHC-SF-R and the total scale of the MHC-SF, with α = .89 for both measures. The societal WB subscale of the MHC-SF-R showed inacceptable internal consistency of <.70 and the relational WB subscale of the MHC-SF-R showed good internal consistency. Furthermore, the results indicated acceptable internal consistency for the psychological WB subscale of the MHC-SF-R. The subscales of the MHC-SF showed adequate to good internal consistency (alphas >.75).

Table 1

Internal Consistency of the MHC-SF-R and MHC-SF

(Sub)scales	Cronbach's α	(Sub)scales	Cronbach's α
MHC-SF-R	.89	MHC-SF	.89
Emotional WB	.88	Emotional WB	.77
Societal WB	.65	Social WB	.83
Psychological WB	.74	Psychological WB	.81
Relational WB	.86		

Note. 'WB' = well-being

Moreover, means, standard deviations, and intercorrelations of the subscale scores of the MHC-SF-R can be found in Table 2. Findings indicated that all subscales are related but sufficiently distinct, since the scores were positively - and moderately to strongly - correlated. The scores of the societal WB subscale showed the weakest but still moderate, positive correlation to the scores of the emotional WB subscale. The strongest, positive correlation was found between the scores of psychological WB subscale and the scores of the emotional WB subscale.

Table 2

Intercorrelations, Means, and Standard Deviations of Subscales (MHC-SF-R)

Parameter	M (SD)	Emotional WB	Societal WB	Psychological WB	Relational WB
MHC-SF-R	67.08 (11.61)				
Emotional WB	10.55 (2.93)	-			
Societal WB	19.04 (3.40)	.36**	-		
Psychological WB	22.60 (4.05)	.69**	.49**	-	
Relational WB	14.89 (3.56)	.43**	.60**	.51**	-

Note. 'WB' = well-being;
** indicates $p < .01$

Intercorrelations, means, and standard deviations of the subscale scores of the MHC-SF can be found in Table 3. Findings showed that the subscales are sufficiently distinct but related since all correlations were positive and in the moderate to strong range, but below the .70 cut-off. The associations between the psychological WB subscale scores and the scores of the other two subscales showed to be positive, strong and the results indicate a positive, moderate correlation between the scores of the social and emotional WB subscale.

Table 3

Intercorrelations, Means, and Standard Deviations of Subscales (MHC-SF)

(Sub)scales	M (SD)	Emotional WB	Social WB	Psychological WB
MHC-SF	43.82 (11.34)			
Emotional WB	10.51 (2.41)	-		
Social WB	12.91 (5.32)	.48**	-	
Psychological WB	20.40 (5.49)	.57**	.63**	-

Note. 'WB' = well-being;
** indicates $p < .01$

Convergent Validity

The findings concerning convergent validity for the MHC-SF-R are shown in Table 4. The results showed a strong positive correlation between the total scores of the MHC-SF-R and the RSES. A positive, strong association was found for the scores of the emotional WB subscale of the MHC-SF-R and the scores of the RSES. Moreover, the results showed to be non-significant between the scores of the BRS and the societal WB subscale of the MHC-SF-R. The BRS showed the strongest associations with the emotional WB subscales of both measures and the RSES the strongest with the psychological WB subscales of both measures. The scores of the MHC-SF-R showed higher correlations than the scores of MHC-SF, except the scores of the societal WB subscale of the MHC-SF-R.

Table 4

Pearson Correlations: MHC-SF-R and RSES, BRS

Subscales	RSES	BRS
MHC-SF-R	.62**	.47**
Emotional WB	.57**	.51**
Societal WB	.38**	.23
Psychological WB	.68**	.45**
Relational WB	.34*	.36**

Note. 'WB' = well-being;
* indicates p <.05; ** indicates p <.01

Results regarding convergent validity for the MHC-SF are shown in Table 5. A strong, positive association was found between the total scores of the MHC-SF and the RSES. The results indicated a positive, moderate correlation between the scores of the emotional WB subscale of the MHC-SF and the scores of the RSES and BRS. Moreover, a non-significant result was found between the scores of the BRS and the social WB subscale and the scores of the BRS and the psychological WB subscale of the MHC-SF.

Table 5

Pearson Correlations: MHC-SF and RSES, BRS

Subscales	RSES	BRS
MHC-SF	.62**	.35**
Emotional WB	.49**	.47**
Social WB	.47**	.27
Psychological WB	.60**	.26

Note. 'WB' = well-being;
** indicates p <.01

Discussion

Validation of the MHC-SF-R

This study aimed to validate the MHC-SF-R which was developed based on reported problems with the comprehensibility of items of the original version, the MHC-SF (Köhle, 2010; G. J. Westerhof, personal communication, 2019 March 7). In order to validate the MHC-SF-R, this study examined the factorial structure, internal consistency, and convergent validity. To compare the MHC-SF-R and the original version, the same properties were also examined for the MHC-SF.

Overall, findings provide no support for a correlated four-factor model due to poor model fit which means that the first hypothesis can by rejected. However, results showed good internal consistency (in form of Cronbach's α and intercorrelations of subscales) and convergent validity for the MHC-SF-R, which means that the second and third hypotheses are confirmed. Compared to the MHC-SF, the MHC-SF-R showed equal results in terms of factorial structure and internal consistency. Although, the societal WB subscale performed poorly in comparison to the social WB subscale of the MHC-SF. Slightly better results of convergent validity were found for the MHC-SF-R. The results showed that convergent validity is satisfactory for the MHC-SF-R, but reliability and factor analysis showed that the societal WB subscale should be improved. Further, there was no evidence found for a correlated three-factor model for the MHC-SF, which is notable since it is in direct contrast to prior research (Karaś, et al., 2014; Keyes, 2005; Keyes et al., 2008; Joshanloo, et al., 2017; Lamers, et al., 2011). However, no items of the MHC-SF showed insufficient factor loadings, while items of the societal WB and the psychological WB subscale of the MHC-SF-R did. Prior research (Karaś, et al., 2014; Keyes, 2005; Keyes, 2006; Keyes et al., 2008; Lamers et al., 2011; Petrillo, et al., 2015; Westerhof & Keyes, 2010) showed good internal consistency for the MHC-SF, which is in line with the findings of this study.

The results show that the societal WB subscale performed poorly in terms of factorial structure and internal consistency compared to the other subscales. Most notable is that low factor loadings for the societal WB subscale were found, with four items having insufficient loadings. Since the societal WB subscale consists mainly of reformulated items of the original social WB subscale of the MHC-SF, and each item of the social WB subscale represents one dimension of social well-being (Keyes, 2009), it could be that the reformulated items do not represent these

dimensions well. As prior research (Köhle, 2010) reported, participants had problems with the comprehensibility of items which resulted in the social WB subscale being transformed into the societal WB subscale with shorter and reformulated items for the MHC-SF-R.

For instance, item 4 ('I make a valuable contribution to our society') aimed to measure the dimension social contribution (Keyes, 2009), and showed to have insufficient factor loading. It reflects item 4 of the MHC-SF with basically one difference, the word *important* was changed to *valuable*. It could be that participants can identify more with contributing something important than something valuable to the society, which resulted in this item not loading well on the dimension social contribution. Further, item 6 ('I accept others as they are'), showed very low factor loading. It resembles best item 7 ('that people are basically good') of the MHC-SF, which according to Keyes (2009), aimed to measure the social acceptance dimension. It could be that the statement of people being basically good, which is more about the society as a whole, reflects the social acceptance dimension more than the statement 'I accept others as they are', which is more about accepting individuals.

Item 7 ('I belong to a group of people') also showed insufficient factor loading and is the reformulation of item 5 ('that you belonged to a community (like a social group, or your neighborhood)') of the MHC-SF, which aimed to measure the dimension social integration (Keyes, 2009). The reformulation of item 5 (MHC-SF) to item 7 (MHC-SF-R) resulted in the omission of the examples 'social group' and 'your neighborhood'. This made the item more general and shorter and was probably done to reduce the reported confusion of participants. However, since item 7 (MHC-SF-R) showed low factor loading and item 5 (MHC-SF) ideal factor loading, the omission of examples probably led to participants not understanding the item as much as in the original version, since it might have become too general.

Moreover, item 8 of the MHC-SF ('That the way our society works makes sense to you') aimed to measure the dimension social coherence (Keyes, 2009). The reformulated version, item 8 of the MHC-SF-R, was phrased as 'I understand how our society works'. To say that something makes (kind of) sense might be easier than saying that something is really understood, which could have resulted in more people reporting that society makes sense to them and fewer that they really understand how society works. This could be the reason for the very low factor loading of item 8 (MHC-SF-R) compared to the sufficient factor loading of item 8 (MHC-SF). The reformulations of the items could have resulted in items which measure the aimed construct - social well-being -

less well compared to the items of the original version, which means that currently the societal WB subscale of the MHC-SF-R does not contribute as much as the other subscales to the explanation of the model and performs poorly compared to the original social well-being subscale of the MHC-SF.

Additionally, item 13 ('I stand up for myself') of the psychological WB subscale of the MHC-SF-R showed low factor loading. It reflects the reformulated item 13 ('confident to think or express your own ideas and opinions') of the MHC-SF, which aimed to measure the autonomy dimension of psychological well-being (Keyes, 2009). Compared to the low factor loading of item 13 (MHC-SF-R), item 13 of the MHC-SF showed to have sufficient factor loading which could be because 'I stand up for myself' is very general and sounds like being able to defend oneself, while being confident to think or express one's own ideas and opinions is more specific and more about the character of the person. It could be that the formulation of item 13 (MHC-SF) reflects the dimension autonomy better than item 13 of the MHC-SF-R, which was indicated by the findings of this study.

Furthermore, this study showed that the newly added subscale, the relational WB subscale, performed well in terms of factor loadings and internal consistency. Since this subscale was newly developed and added, it is notable that it performed well compared to the societal WB subscale, which is the corrected version of the social WB subscale of the MHC-SF. While the societal WB subscale is more concerned with one's place and contribution to society, the relational WB subscale is more concerned with one's feelings and satisfaction in regard to the connection to social contacts and the society. As this subscale performed well, it can be assumed that this part of social WB was missing in the original version and the development of this subscale was a good contribution to the MHC-SF-R.

Moreover, the MHC-SF-R showed moderate to strong, positive correlations to the related constructs self-esteem and resilience which supports convergent validity. Though, findings indicate a non-significant association between the scores of the societal WB subscale and the BRS. The examined association between the scores of the BRS and the relational WB subscale was found to be moderate but substantially lower compared to emotional WB and psychological WB subscales. Additionally, the found association between the scores of the social WB subscale (MHC-SF) and the BRS, was also non-significant. Since the relational WB subscale and societal WB subscale of the MHC-SF-R, and also the social WB subscale of the MHC-SF, aim to measure

the construct social well-being and the findings indicate insufficient or lower correlations compared to the other subscales, the findings indicate a weaker relationship between the constructs social well-being and resilience. These findings are in line with prior research which mention the relationship between emotional well-being and resilience (Cummins & Wooden, 2014) and psychological well-being and resilience (Davydov, et al., 2010; Patel, & Goodman, 2007) specifically, but not for social well-being and resilience.

Strengths and Limitations

As this study aimed to validate the MHC-SF-R, but also examined the MHC-SF itself for the purpose of comparing both, this comparison can be viewed as a strength of this study. To prevent imbalanced and therefore biased groups, participants were randomly assigned to groups. Such bias could for example be caused by the researcher assigning the participants to the two groups while being under the influence of unconscious stereotypes. Due to the randomization of participants to the two groups, the groups should not differ in any systematic way (Suresh, 2011). Indeed, there were no substantial differences found between the two groups in regard to age, gender, or nationality (see Participants).

A disadvantage of having two groups was that examining correlations between the MHC-SF-R and MHC-SF was not possible, since different participants filled out either. Furthermore, the splitting of participants into two groups, one group for the MHC-SF-R and one for the MHC-SF, resulted in less power for the method used for examining factorial structure (confirmatory factor analysis), since the sample size was only approximately 50 for both groups. Letting participants fill out both tests - the MHC-SF-R and MHC-SF - would have been ideal, but not realistic, since both tests are similar in content and if participants fill out highly similar questionnaires in a small timespan, biased data could be the result. This is, because answering similar (or the same) items multiple times in the same survey could lead to exhaustion or boredom of participants which would result in respondent fatigue (Lavrakas, 2008a), which causes participants to answer more ‘don’t know’ options or choosing always the same answer option (straight-line answering).

Moreover, data of online surveys is always at risk to be biased by superficial answers of participants. One possibility of such bias is that participants could have given socially desirable answers (Lavrakas, 2008b). Another reason for superficial answers could be the SONA system, since this study could be conducted via SONA and participants (students) got credits for

completing the survey, this could have resulted in some answering (superficially) the survey just to get credits. These superficial answers and the corresponding bias could have led to an underestimation of reliability and validity, because if students answered based on the desire to get credits, they could have done that in a random manner. Such random answers would have resulted in data which is not based on any content of the survey, and this could have led to an underestimation of reliability and validity of the data, since data is more coherent when all participants generate it based on the same content instead of randomness.

Furthermore, as this study used convenience sampling and the sample includes only students, the sample could be biased (Mackey & Gass, 2005). Such systematic bias could be caused by a difference between the theoretical results of the entire population and the results of this study, since the sample includes only a part of the population.

Practical Implications

This study found that the MHC-SF-R still lacks support for its four-factor model. Especially the societal WB subscale, showed to be poor in terms of factor loadings and internal consistency and usage is not advised at this stage. The psychological WB subscale of the MHC-SF-R performed less well compared the psychological WB subscale of the original version, which means that it is still advised to use the psychological WB subscale of the MHC-SF. The relational WB subscale performed well, which means that it could be used together with the social WB subscale of the MHC-SF since both measure the construct social well-being partly. Moreover, the emotional WB subscale of the MHC-SF-R performed better compared to the original version and can be advised to use. Overall, the MHC-SF is still more supported by literature and this study found no substantial evidence to advice the usage of the MHC-SF-R at this stage, except for its emotional WB subscale and the added relational WB subscale, which could be a valuable contribution to measure the construct social well-being.

Directions for Future Research

Future studies could further improve the MHC-SF-R in terms of factorial structure. One step would be to repeat this study with a bigger sample and therefore, control for the limitation of having less participants per group than normally required for confirmatory factor analysis. Furthermore, a different sampling method than convenience sampling (and a sample including not only students)

could be used to prevent bias. Hence, other results could be possible with a bigger sample and different sample. Additionally, future studies could examine more properties of the MHC-SF-R. For instance, test-retest reliability could be examined to further validate the MHC-SF-R, which was examined for the MHC-SF in several studies (e.g. Lamers et al., 2011; Petrillo, et al., 2015).

Moreover, results showed that the societal WB subscale needs to be refined, since it lacks reliability and validity. By looking at the transformation of the social WB subscale of the MHC-SF to the societal WB subscale of the MHC-SF-R, corrections could be made to the formulation of individual items based on those differences, since the social WB subscale of the MHC-SF performed better. In addition, the same approach could be used to improve the psychological WB subscale, since one item of the psychological WB subscale (item 13) of the MHC-SF-R performed poorly regarding its factor loading.

Such corrections could for example be guided by the cognitive interviewing method. Latcheva (2011) used the cognitive interviewing method to improve items which performed badly in factor analysis and she used Foddy's classification of principal problems which respondents face in the answering process, to guide the cognitive interviewing. Such problems could for instance be that participants encounter problems associated with different perspectives to answer the same question or with comprehension difficulties of key concepts or overall meanings of items (Foddy, 1996). Latcheva (2011) let participants answer items based on these specific problems, for example key concepts of questions or answering the items from different perspectives to test the items' validity. Another method to examine problems participants face when answering items was used by Köhle (2010), she used a three-step interview design including the think-aloud method, which means that participants say out loud what they think while answering the questionnaire items at the same time, to examine problems with comprehensibility (e.g. what exactly is not understood of a question).

To conclude, if the societal WB subscale (and item 13 of the psychological WB subscale) is improved, the MHC-SF-R would have a more extensive measurement of social WB, since the relational WB subscale adds more dimensionality to the measurement of social well-being. This could make the MHC-SF-R superior to the original version. Hence, the MHC-SF-R could become a promising instrument to measure well-being.

Conclusion

This study found no substantial evidence for the four-factor model of the MHC-SF-R. Though, good internal consistency and evidence for convergent validity was found for the total scale as hypothesized. Further, the added relational WB subscale seems to be a promising contribution to measuring social well-being. However, the societal WB subscale lacks internal consistency and performed poorly in terms of factor loadings and therefore, requires further refinements. The results deliver not enough support to declare the MHC-SF-R a reliable and valid instrument to measure well-being at this stage. However, further refinements could lead to the development of a reliable and valid instrument to measure well-being.

References

Agresti, A. (2007). *An Introduction to Categorical Data Analysis.* Hoboken, NJ: John Wiley & Sons.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4*, 1-44.

Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.

Cantril, H. (1965). *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-284.

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

Cummins, R. A., & Wooden, M. (2014). Personal resilience in times of crisis: The implications of SWB homeostasis and set-points. *Journal of happiness studies*, *15*(1), 223-235

Davydov, D., Stewart, R., Ritchie, K., & Chaudieu, I. (2010). Resilience and mental health. *Clinical Psychology Review, 30*(5), 479-495. doi:10.1016/j.cpr.2010.03.003

Foddy, W. (1996). The in-depth testing of survey questions: a critical appraisal of methods. *Quality and Quantity, 30*, 361–370.

Franck, E., De Raedt, R., Barbez, C., & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg self-esteem scale. *Psychologica Belgica*, *48*(1), 25-35.

Fuller-Iglesias, H., Sellars, B., & Antonucci, T. (2008). Resilience in old age: Social relations as a protective factor. *Research in Human Development, 5*(3), 181-193.

Hair, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis* (Seventh ed.). Harlow, Essex: Pearson.

Henriksen, I., Ranøyen, I., Indredavik, M., & Stenseng, F. (2017). The role of self-esteem in the development of psychiatric problems: A three-year prospective study in a clinical sample of adolescents. *Child and Adolescent Psychiatry and Mental Health, 11*(1), 1-9. doi:10.1186/s13034-017-0207-y

Hjemdal, O., Friborg, O., Stiles, T., Rosenvinge, J., & Martinussen, M. (2006). Resilience predicting psychiatric symptoms: A prospective study of protective factors and their role in adjustment to stressful life events. *Clinical Psychology & Psychotherapy, 13*(3), 194-201. doi:10.1002/cpp.488

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Joshanloo, M., Capone, V., Petrillo, G., & Caso, D. (2017). Discriminant validity of hedonic, social, and psychological well-being in two italian samples. *Personality and Individual Differences, 109*, 23-27. doi:10.1016/j.paid.2016.12.036

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*, 121–140.

Keyes, C. L. M. (2005). The subjective well-being of America's youth: Toward a comprehensive assessment. *Adolescent and Family Health, 4*, 3−11.

Keyes, C. L. M. (2006). Mental health in adolescence: Is America's youth flourishing? *American Journal of Orthopsychiatry, 76*, 395–402.

Keyes, C. L. M. (2009). Atlanta: Brief description of the mental health continuum short form (MHC-SF). Retrieved from http://www.sociology.emory.edu/ckeyes/

Keyes, C. L. M., Wissing, M., Potgieter, J. P., Temane, M., Kruger, A., & van Rooy, S. (2008). Evaluation of the Mental Health Continuum Short Form (MHC-SF) in Setswana speaking South Africans. *Clinical Psychology and Psychotherapy, 15*, 181–192.

Köhle, N. (2010). *'Mag ik vragen, wat ik u moet vragen?' Een onderzoek naar de zwakheden van het Mental Health Continuum Short Form met behulp van het Three-Step Test-Interview.* (Bachelor's thesis). Retrieved from https://essay.utwente.nl/59879/

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. https://doi.org/10.1002/jclp.20741

Latcheva, R. (2011). Cognitive interviewing and factor-analytic techniques: A mixed method approach to validity of survey items measuring national identity. *Quality and Quantity, 45*(6), 1175-1199.

Lavrakas, P. J. (2008a). Respondent Fatigue. In *Encyclopedia of survey research methods.* Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

Lavrakas, P. J. (2008b). Social Desirability. In *Encyclopedia of survey research methods.* Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design*. New Jersey: Lawrence Erlbaum Associates, Inc.

Mann, M. M., Hosman, C. M., Schaalma, H. P., & De Vries, N. K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health education research*, *19*(4), 357-372.

Mental Treatment Centers. (2018, December 9). Mental well-being. Retrieved from http://www.wtrm.org/mental-treatment-centers/mental-well-being/

Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology, 102*(6), 1271–1288. doi: 10.1037/a0025558.

Patel, V., & Goodman, A. (2007). Researching protective and promotive factors in mental health. *International Journal of Epidemiology, 36*(4), 703-7.

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: An International and Interdisciplinary Journal for Quality-Of-Life Measurement, 121*(1), 291-312. doi:10.1007/s11205-014-0629-3

Poulin, C., Lemoine, O., Poirier, L., & Lambert, J. (2005). Validation study of a nonspecific psychological distress scale. *Social Psychiatry and Psychiatric Epidemiology, 40*(12), 1019-1024. doi:10.1007/s00127-005-0961-7

Rizwan, M., Aftab, S., Shah, I., & Dharwarwala, R. (2012). Psychometric Properties of the Rosenberg Self-esteem Scale in Pakistan Late Adolescent. *The International Journal of Educational and Psychological Assessment, 10*, 125-138.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton. NJ: Princeton University Press.

Rosenberg, M. (1989). *Society and the adolescent self- image* (Revised edition). Middletown, CT: Wesleyan University Press.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141-166.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well–being. *Journal of Personality and Social Psychology, 57*, 1069–1081.

Sinclair, S., Blais, M., Gansler, D., Sandberg, E., Bistis, K., & LoCicero, A. (2010). Psychometric properties of the rosenberg self-esteem scale: Overall and across demographic groups living within the united states. *Evaluation & the Health Professions, 33*(1), 56-80. doi:10.1177/0163278709356187

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. International journal of behavioral medicine, 15(3), 194-200.

Suresh, K. P. (2011). An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of human reproductive sciences*, *4*(1), 8.

Vella-Brodrick, D., & Klein, B. (2010). Positive psychology and the internet: A mental health opportunity. *Electronic Journal of Applied Psychology*, *6*(2), 30-41.

Westerhof, G. J., & Keyes, C. L. M. (2010). Mental illness and mental health: The two continua model across the lifespan. *Journal of Adult Development, 17*, 110–119.

Windle, G., Noyes, J., & Bennett, K. (2011). A methodological review of resilience measurement scales. *Health and Quality of Life Outcomes, 9*(1), 1-18. doi:10.1186/1477-7525-9-8

World Health Organization. (2014, August). Mental health: a state of well-being**.** Retrieved from https://www.who.int/features/factfiles/mental_health/en/