Revision of the Mental Health Continuum – Short Form: Testing Psychometric Properties

Lara Wenglorz

July 2, 2019

Supervisors
Dr. Marion Sommers-Spijkerman
Prof. Dr. Gerben Westerhof

University of Twente
BMS Faculty
Department of Psychology

Abstract

Measurement instruments for screening people's well-being are intensively used in practice. For several years, the Mental Health Continuum -Short Form (MHC-SF), which includes 14 items, is commonly used. However, interpretation problems with the items, especially for the subscale of social well-being, are present. Therefore, a revised version (MHC-SF-R), including 19 items, was developed. To investigate psychometric properties of the MHC-SF-R, a validation study in a sample of University students was conducted (N=108). For assessing convergent validity, two additional constructs, theoretically related to well-being, were added, that were self-compassion and self-esteem. These were measured with the Self-compassion Scale – Short Form (SCS-SF) and the Rosenberg Self-esteem Scale (RSES). This cross-sectional study was conducted online with a questionnaire survey design. Confirmatory factor analysis (CFA) demonstrated poor fit to a four-factor model of the MHC-SF-R with emotional well-being, societal well-being, psychological well-being and relational well-being, especially present for societal well-being. CFA did not support a three-factor structure of the MHC-SF. Internal consistency of both versions was good for the total scale (α=.89). Except for the societal well-being dimension of the MHC-SF-R (α=.65*)*, reliability was desirable for all subscales, ranging from .74 to .88. Convergent validity of both versions was supported by moderate to high correlations with the two additional measures. All in all, the MHC-SF-R was not recommended for practical use since the dimension of societal well-being could not reach acceptable results. However, as the other three dimensions proved to be reliable and valid, it was advised to promote research on a revised version of the MHC-SF, supplementing it by qualitative designs.

Keywords: well-being, self-esteem, self-compassion, psychometric properties, validation

Table of Contents

Introduction

Treating mental health solely as the absence of a mental illness is no longer the common view. Most prominent, it is referred to the definition of the World Health Organization (WHO), stating "health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO, 2001, p.1). The domain of positive psychology is supporting this approach by looking at Psychology not only from a negative angle, but also focusing on positive ideas for enhancing people's well-being. It is no more about treating illnesses to heal, but also paying attention to positive facets. Key activities are the focus on strengths and growth to enhance people's possibilities for striving towards full potential (Seligman, 2002). Thus, psychologists' attention today is not solely focused on psychopathology, but also on well-being, often referred to as an "optimal psychological experience and functioning" (Deci & Ryan, 2008, p.1).

For screening purposes, several psychological tests measuring people's well-being are commonly used. Research gives rise to a frequent use of the Mental Health Continuum – Short Form (MHC-SF), which is proved to be a reliable and valid measurement instrument for well-being, over time (Keyes, 2009; Lamers, Westerhof, Bohlmeijer, ten Klooster, & Keyes, 2011). While its extended version, namely the Mental Health Continuum in its Long Form (MHC-LF) consists of 40 items, the short form only includes the most representative items for each category. This results in a test consisting of 14 items and measuring well-being along three subcomponents, namely the emotional well-being, the psychological well-being and the social well-being (Keyes, 2009). Emotional well-being refers to experiencing positive emotions, happiness and the overall satisfaction with one's life. Psychological well-being addresses a person's sense of self-acceptance, the relation with other people, personal growth, autonomy, the purpose of life and environmental mastery. Lastly, the dimension of social well-being measures well-being across aspects like social acceptance, social contribution, social integration, social growth and social interest (Lamers et al., 2011).

Although the MHC-SF is available in two forms, this study solely focuses on the short form. That is because, in general, short forms show some benefits compared to long forms of a test, when it comes to brief screening. If a short form has good psychometric properties, it has the advantages of briefness and efficiency. Both, the researchers and the participants are spending less time on conducting and scoring the test outcome. This is especially practical if more than one test has to be completed or if one specific test is not addressing the core interest, but only administered for complementary insights (Bowling, 2005). Due to its manageable length of 14 items, the MHC-SF is often used for briefly assessing a person's

well-being, especially in the setting of a first screening, for example, at a workplace or in clinical practice (Rogoza, Truong Thi, Rozycka-Tran, Piotrowski, & Zemojtel-Piotrowska, 2018).

Research and evaluation show that the MHC-SF is in fact a useful instrument for measuring a person's mental well-being, as it shows good psychometric properties. Several studies identified a three-factor structure of the MHC-SF, including emotional well-being, social well-being and psychological well-being (Keyes, 2009). Research shows a number of findings on reliability. Keyes (2009) reported good internal consistency (α>.80), confirmed by Perugini and colleagues (2017) (α=.89), up to Karaś, Cieciuch and Keyes (2014), who even reported excellent reliability (α=.91). Thus, reliability of the MHC-SF is proven to be good. Same goes for all three subscales, ranging from .83 for emotional- and psychological well-being to .74 for social well-being (Keyes, 2009). Test-retest reliability assessed over a period of nine months is good (α=.65) and convergent validity for the total scale of well-being is supported by significant correlations to several related constructs, reported by Lamers et al. (2011) and supported by Perugini et al. (2017). All in all, the MHC-SF was found a reliable and valid measure for a person's well-being.

Although the MHC-SF has good psychometric properties, it still faces some limitations. Especially the use in clinical practice gave rise to some problems. According to G. Westerhof (personal communication, February 14, 2019) many people face problems with interpreting specific items of the instrument. Since the test is short in length, a clear and easy understanding of each single item is very important for a reliable test outcome. Additionally, Köhle (2010) discovered specific limitations of the MHC-SF in a qualitative study. First, the main problem occurred with the subscale of social well-being, as it includes terms like "community", "society" and "important contribution" that are hard to interpret for many respondents. Another possible reason for participant's interpretation difficulties with the dimension of social well-being is explained by the fact that people often find it easier to rate qualities based on their own person, than broaden it to the whole society. When looking at society, people have to put themselves in a much broader spectrum, which makes answering more complicated. Furthermore, the period of estimating feelings over the last month may be regarded as too long, since many people cannot remember how they specifically felt in the past. Another limitation is the Likert-scale. Although the Likert-scale is well established, especially for expressing agreement, people can have problems when it comes to rate their feelings on such a precise scale (Boone & Boone, 2012). With regard to all aforementioned

limitations of the MHC-SF, an updated version of the test has potential to contribute to a clearer understanding for people in practice.

For tackling these problems, a new version of the MHC-SF was developed. In contrast to the original version, which is composed of 14 items, the new version includes 19 items. Adding five more items is due to the fact that it splits the scale of social well-being into two further dimensions, namely a societal dimension and a relational dimension. With regard to content, the dimension of emotional well-being differs not from the original version. However, the dimension of psychological well-being shows changes, namely that statements are shortened and phrased in a more simplified way. Most content changes appeared in the dimension of social well-being. The societal dimension is addressing a person's understanding of himself/herself in the context of society. The relational dimension is focusing on a narrower scale than society, looking at personal social contacts. All in all, items were shortened. Whereas the original test is asking a person to asses all items on the basis of their experiences in the past month, the new version is limiting the time period to only the past week. Furthermore, both tests are differing in their scale of frequency. In the original version, a participant can indicate the frequency choosing from never, once or twice, about once a week, about 2 or 3 times a week, almost every day and every day. The new version offers different dimensions, namely never, rarely, sometimes, regularly, often and (almost) always. Looking at factorial structure, the revised version of the MHC-SF (MHC-SF-R) was designed along four subscales, wherefore a four-factor structure, including emotional well-being, psychological well-being, social well-being and relational well-being, is expected in this study. With regard to these changes, the primary aim of this study is to test the MHC-SF-R in order to explore factorial structure and to determine its reliability and validity. Following, it is especially important to assess whether the revised version can overcome the aforementioned limitations and therefore can show even better reliability and validity than the original MHC-SF.

One partial aim of this research is to assess validity of the construct of well-being measured in the MHC-SF-R. Thus, convergent validity is measured by including further constructs that are known to be related to well-being. Focusing on the dimension of psychological well-being, self-compassion emerged as an upcoming concept and often found to be related to well-being (Neff, 2011). Neff, one of the main representatives of the concept of self-compassion, defines self-compassion as "being touched by and open to one's own suffering, not avoiding or disconnecting from it, generating the desire to alleviate one's suffering and to heal oneself with kindness" (Neff, 2003a, p. 87). In other words, it describes

careful attention to one's own feelings and experiences and treating it with compassion. Preliminary findings in the area of self-compassion and well-being are mostly showing similar results by detecting a significant positive relation, meaning that people who have a high sense of self-compassion display higher levels of well-being. That is because people high in self-esteem tend to show a more caring attitude towards themselves and focus on positive and pleasant experiences and emotions, more easily (Zessin, Dickhäuser, & Garbade, 2015). Hall, Row, Wuensch & Godley (2013) found a relation between psychological well-being and self-compassion. Findings of Zessin et al. (2015) can confirm the existence of this relationship between self-compassion and well-being, that is especially strong for psychological well-being. Another concept that is found to be related to well-being and thus used in this research is self-esteem. Self-esteem can be described as a personal evaluation of one's own worth in regard to skills, abilities and valuableness. There it is important that the matter of evaluation is significant to a person (Neff, 2011). Research on self-esteem shows that high self-esteem is in fact related to good mental health and psychological well-being. A review performed by Baumeister and colleagues (2003) concluded a strong relationship between self-esteem and positive factors related to better well-being. In conclusion, self-compassion and self-esteem are used for assessing convergent validity of the MHC-SF-R in this study.

Finally, as the primary aim of this paper is to validate the psychometric properties of the MHC-SF-R, it is closely looked at three main issues. First, the factorial structure of the MHC-SF-R is examined. Second, reliability of the MHC-SF-R is tested. And lastly, convergent validity of the MHC-SF is tested, through examining the relationship between well-being and self-compassion, as well as between well-being and self-esteem. For self-compassion, the relationship is expected to be especially high for the dimension of psychological well-being. As a secondary aim, performance of the MHC-SF-R and the MHC-SF are compared, assessing reliability, convergent validity and factorial structure of the MHC-SF. In the following, four hypotheses are stated. As a first hypothesis, confirmation of a four-factor structure of the MHC-SF-R is expected. Second, internal consistency of the MHC-SF-R was expected to be good (α>.80). Third, good convergent validity was expected. For self-esteem, a moderate positive correlation is expected. For self-compassion, a strong positive correlation is expected, highest for the subscale of psychological well-being. Fourth, the MHC-SF-R is expected to perform at least equally to the MHC-SF on internal consistency and convergent validity. In line with previous research, it is expected to confirm a three-factor

structure in the MHC-SF. Providing information on reliability and validity of the MHC-SF-R would lay a starting point in establishing this new version for reliably evaluating well-being.

Method

Design

The current research involved a cross-sectional study employing a questionnaire survey design. The ethics review board of the Faculty of Behavioral, Management and Social sciences of the University of Twente approved this research (190441).

Participants and Procedure

The target group of this study were students. These were recruited via convenience sampling across students in the period between 30th April 2019 and 6th May 2019. Recruitment strategies included the use of the service "Sona Systems" of the University of Twente. Sona Systems is a service provided by the University of Twente used by undergraduate psychology students. The system includes a variety of studies in which students can participate to earn so-called Sona credits. In the course of their studies, students are expected to collect 15 Sona credits. Additional strategies were the recruitment via the personal network and the social media platform Facebook via personal profiles.

The survey was generated with the service "Qualtrics", an online software enabling to create online studies for collecting data. Following, data was collected with the help of this service. The study was conducted online. First, participants were informed about the aim of the study and contact details of the researcher were provided. Participants were informed that they can quit the research at any time, without naming any reason. If participants gave informed consent, they were screened for eligibility. Exclusion criteria included not being a student, being younger than 18 or having insufficient English language proficiency (i.e. beginner or intermediate level). Then, questions exploring further demographics, including gender and nationality, followed. Afterwards, three questionnaires followed, including, depending on the group, the MHC-SF-R or the MHC-SF, the SCS-SF and the RSES. To overcome response bias, participants were randomly allocated to either of two groups, one completing the MHC-SF-R and the other completing the MHC-SF. This was done by presenting every other respondent with the particular version of the survey, either including the MHC-SF-R, or the MHC-SF. Every participant could complete the survey on his/her own, using his/her own electrical device in his/her own environment. Finally, the participant was

thanked for the participation and informed about the possibility to get information about the outcome of the research, if wanted.

Nine participants were excluded due to not fulfilling all inclusion criteria, as six participants did not have sufficient English proficiency and three were no students. Finally, the study included 109 participants (46 % male, 54 % female). Respondents were between 18 and 31 years old (M_{age}= 22.16, *SD* = 2.99). The majority of the sample was German (*n*=80), followed by Dutch (*n*=15) and Other (*n*=14). The sample was split due to randomization of the groups, which resulted in two comparable groups with group one (*n*=54) completing the MHC-SF-R and group two (*n*=55) completing the MHC-SF. Both groups did not differ significantly in age, gender, nationality or level of English proficiency. Group one included 47% female and 53% male with a mean age of 22.47 years, whereas group two included 44% female and 56% male with a mean age of 21.6 years.

Measures

Well-being. The first measure was the Mental Health Continuum – Short Form revised (MHC-SF-R) (see Appendix). This test consisted of 19 items and asked participants to indicate the experienced feeling during the past week. Items were rated on a 6-point Likert scale, ranging between 0 (*never*) to 5 (*(almost) always*). Emotional well-being was measured in three items (1-3), being identical to the original version. The dimension of social well-being was split into two, namely a societal dimension (item 4-8, 19) and a relational dimension (item 15-18). Psychological well-being was measured using six items (item 9-14), all of which were rephrased. A mean score was calculated, possibly ranging from 0-5. A higher score reflected a higher well-being. For assessing scores of the subscales, a mean of scores on all belonging items was calculated.

The second measure that was used was the Mental Health Continuum – Short Form (MHC-SF) (Keyes, 2009). It consisted of 14 items, three measuring emotional well-being, five measuring social well-being and six measuring psychological well-being. Respondents were asked to rate the frequency of a feeling during the last month, according to a 6-point Likert frequency scale, ranging between 0 (*never*) to 5 (*every day*). Scores on all items were added and a mean score was calculated. Respondents could reach a mean score between 0 and 5, a higher score indicated higher well-being. Mean scores were also computed for each subscale, including the appropriate items. In previous research, reliability and validity were shown to be good for the total scale (α=.89), the emotional well-being subscale (α=.83) and

the psychological well-being subscale (α=.83). For the subscale of social well-being, reliability was acceptable (α=.74) (Lamers et al., 2011).

Self-compassion. The Self-compassion Scale – Short Form (SCS-SF) was used for measuring self-compassion (Raes, Pommier, Neff, & Van Gucht, 2011). The SCS-SF consisted of 12 items, formulated as statements and asked the participant to indicate how often, in general, he/she behaves in the particular manner. Participants indicated the frequency according a 5-point Likert scale ranging between 1 (*almost never*) to 5 (*almost always*). Reliability of the SCS-SF proved to be high in the current study (α=.84), as well as in prior studies (α=.86) (Strauss et al., 2016). A mean score was calculated, possibly ranging between 1 and 5. The higher the mean score, the higher the level of self-compassion.

Self-esteem. For measuring self-esteem, the Rosenberg Self-Esteem Scale (RSES) was used (Rosenberg, 1965). It included ten items in form of statements to which the respondent was asked to express his/her agreement on a 4-point Likert-scale from 1 (*strongly disagree*) to 4 *(strongly agree*). A mean score was calculated, where scores could fall in a range between 1 and 5. Higher scores were reflecting higher self-esteem. For the RSES, internal consistency proved good in previous research (α=.86), that was also verified by the current study (α=.86) (Tinakon & Nahatai, 2012).

Demographics. Demographic questions included age (in years), gender (i.e. female, male or other) and nationality (i.e. Dutch, German or other).

Data Analysis

All data were analyzed using SPSS 25 and LISREL 10.10. Characteristics of the sample were determined by computing descriptive statistics or frequencies for age, gender and nationality. Descriptive statistics, including mean and standard deviation, of the MHC-SF-R with all four subscales, the MHC-SF with all four subscales, the SCS-SF and the RSES were computed. Due to small sample sizes, scatterplots of the sample were investigated for testing a normal distribution. One outlier was removed to approach an approximately normal distribution, resulting in a final sample of 108 participants.

Factorial structure of the MHC-SF-R was assessed with a confirmatory factor analysis (CFA) conducted with LISREL. As data were only approximately normally distributed, maximum likelihood estimation method was used. All variables were handled as continuous. The aim was to validate the four-factor model with four latent variables being emotional well-being (EW), psychological well-being (PW), societal well-being (SW) and relational well-

being (RW). Items were restricted to load on one factor only. For evaluating model fit, four fit indices were used, namely the comparative fit index (CFI), the chi-squared test (χ^2), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). For the chi-squared test statistic, smaller values reflect a better model fit, for the other indices, acceptable model fit is obtained with RMSEA <.08, SRMR <.08, a chi-square with p>.05 and CFI ≥.90 (Hu & Bentler, 1999; Kline, 2005). For the items, low factor loadings are considered >.30, medium factor loadings for >.50 and high factor loadings are >.70 (Shevlin & Miles, 1998). Loadings between the subscales should not exceed .90 to be desirable. Loadings >.90 can be considered too high, leading to the assumption of measuring identical constructs (Kline, 2005).

For computing reliability of the MHC-SF-R and the MHC-SF, SPSS was used. Cronbach's alphas were calculated for the MHC-SF-R and all four subscales, the MHC-SF with all three subscales, the SCS-SF and the RSES. Desirable values for α are ranging from .70 to .95. Values >.70 are acceptable, >.80 good and >.90 is showing excellent internal consistency (Tavakol & Dennick, 2011).

Convergent validity of the MHC-SF-R and the MHC-SF were examined. To do so, Pearson correlations between all four scales, the MHC-SF-R with subscales, the MHC-SF with subscales, the SCS-SF and the RSES were computed. Values between .40<*r*>0.70 are desirable for showing convergent validity. Values below .40 are considered low and values above .70 are considered too high, showing too much overlap in the measured constructs (Schober, Boer & Schwarte, 2018).

All analyses were conducted two-tailed and a p-value of .05 was applied.

Results

Factor Structure

For the MHC-SF-R, CFA demonstrated poor fit of the four-factor model (χ^2 (146)=226.37; $p < .05$; RMSEA=.10; CFI= .82; SRMR=.11). Figure 1 displays the four-factor model of the MHC-SF-R. Correlations between the factors were acceptable, ranging between .45 and .87. For emotional well-being, psychological well-being and relational well-being, all factor loadings were at least medium, except for a low loading on item 13. Factor loadings ranged between .51 and .94. Items of societal well-being included low factor loadings (Figure 1).

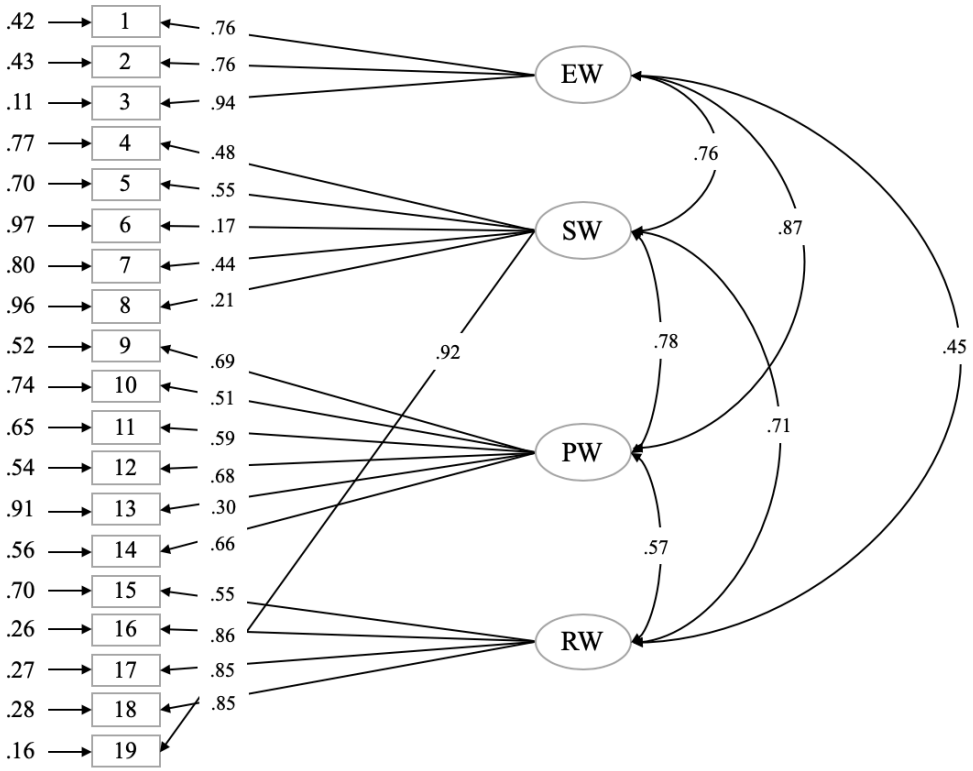

Figure 1. Confirmatory factor analysis of the MHC-SF-R with a four-factor model. EW= Emotional well-being; SW= Societal well-being; PW= Psychological well-being; RW= Relational well-being.

Model fit of the three-factor structure for the MHC-SF was poor (χ^2 (74)=117.34; $p <$ .05; RMSEA=.10; CFI=.86; SRMR=.09). Figure 2, which shows a three-factor model of the MHC-SF, displays that correlations between all three factors were desirable, ranging between .63 and .76. Factor loadings of all fourteen items were desirable, being between .56 and .84 (Figure 2).

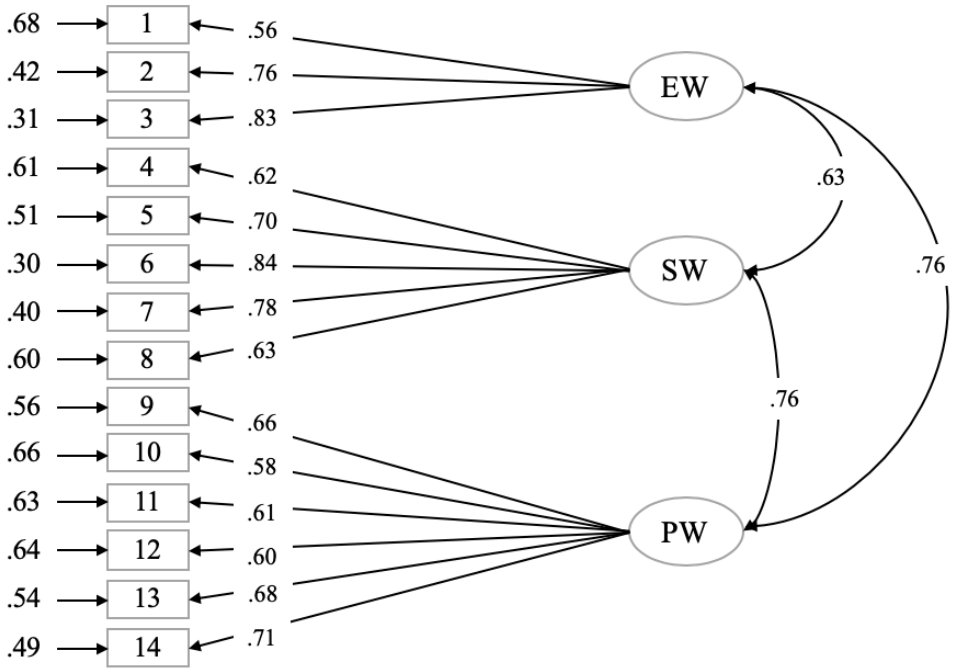

Figure 2. Confirmatory factor analysis of the MHC-SF with a three-factor model. EW= Emotional well-being; SW= Social well-being; PW= Psychological well-being.

Internal Consistency

For the MHC-SF-R, reliability was calculated to be good for the total scale (α=.89). Reliability was good for two subscales, ranging between .86 and .88 for relational- and emotional well-being and acceptable for psychological well-being (α=.74). Only the subscale of social well-being reported a questionable result (α=.65) (Table 1). Item-total correlation revealed a low correlation of item 6 (*r*=.13), which is part of the social well-being subscale. However, removing this item did not influence internal consistency of the total scale, wherefore it was not removed. For the MHC-SF, internal consistency was calculated to be good for the total scale (α=.89) and good for all three subscales, ranging between .77 and .83.

Table 1

Reliabilities, Means and Standard Deviations of the MHC-SF-R with subscales, the MHC-SF, the SCS-SF and the RSES

(Sub)scale	*N* items	Cronbach's α	*M (SD)*	*N* participants
MHC-SF-R	19	.89	3.53 (.61)	53
EW	3	.88	3.52 (.98)	53
SW	6	.65	3.17 (.67)	53
PW	6	.74	3.77 (.68)	53
RW	4	.86	3.72 (.89)	53
MHC-SF	14	.89	3.13 (.81)	55
EW	3	.77	3.50 (.80)	55
SW	5	.83	2.59 (1.10)	55
PW	6	.81	3.40 (.92)	55
SCS-SF	12	.84	3.20 (.68)	108
RSES	10	.86	3.10 (.49)	108

Note. MHC-SF-R= Mental Health Continuum – Short Form – Revised; EW= Emotional well-being; SW= Social well-being; RW= Relational well-being; PW= Psychological well-being; MHC-SF= Mental Health Continuum – Short Form; SCS-SF= Self-compassion Scale – Short Form; RSES= Rosenberg Self-esteem Scale.

Intercorrelations

Correlations between the subscales and the MHC-SF-R were significantly strong and positive for all four subscales. Between the four subscales, almost all correlations were at least moderately positive. Only the correlation between the social well-being and emotional well-being was low (r=.36) (Table 2). For the MHC-SF, strong positive correlations with all three subscales were found. Intercorrelations between all three subscales were positive and at least moderate (Table 3).

Table 2

Pearson correlations between the MHC-SF-R with subscales, the SCS-SF and the RSES

(Sub)scale	MHC-SF-R	EW	SW	PW	RW	SCS-SF	RSES
MHC-SF-R							
EW	.75**						
SW	.80**	.36**					
PW	.85**	.69**	.49**				
RW	.80**	.43**	.60**	.51**			
SCS-SF	.49**	.44**	.29*	.51**	.32*		
RSES	.62**	.57**	.38**	.68**	.34*	.59**	

Note. *N*=53. MHC-SF-R= Mental Health Continuum – Short Form Revised; EW= Emotional well-being; SW= Social well-being; PW= Psychological well-being; RW= Relational well-being; SCS-SF= Self-compassion Scale – Short Form; RSES= Rosenberg Self-esteem Scale.

*p<.0.5. **p<.01.

Convergent Validity

Between the MHC-SF-R and the SCS-SF, the RSES, and the psychological well-being dimension of the SCS-SF, only significant correlations were found. For the SCS-SF, a moderately positive correlation was found, which was strongly positive for the psychological well-being dimension of the MHC-SF-R. Correlation between the MHC-SF-R and the RSES was found to be positive and strong (Table 2). Correlation between the SCS-SF and the MHC-SF was significant and a moderate positive one. The RSES and MHC-SF were significantly strongly positive correlated (Table 3).

Table 3

Pearson correlations between the MHC-SF with subscales, the SCS-SF and the RSES

(Sub)scale	MHC-SF	EW	SW	PW	SCS-SF	RSES
MHC-SF						
EW	.71**					
SW	.88**	.48**				
PW	.90**	.57**	.63**			
SCS-SF	.43**	.41**	.43**	.29*		
RSES	.62**	.49**	.47**	.61**	.59**	

Note. N=55. MHC-SF= Mental Health Continuum – Short Form; EW= Emotional well-being; SW= Social well-being; PW= Psychological well-being; SCS-SF= Self-compassion Scale – Short Form; RSES= Rosenberg Self-esteem Scale.

*p<.0.5. **p<.01.

Discussion

Key Findings

The primary aim of this study was to validate the revised version of the MHC-SF. In course of this validation, four hypotheses were tested. First, a four-factor structure of the MHC-SF-R was expected but could not be confirmed. There, the societal well-being subscale performed less well than the others. Second, the MHC-SF-R performed at least good on internal consistency and third, convergent validity of the MHC-SF-R was also evaluated to be good. However, findings differed from the proposed hypothesis. The correlation between self-compassion and well-being measured with the MHC-SF-R was lower than expected, but still strongest for the subscales of psychological well-being, which was in line with the expectation. Correlation between self-esteem and well-being measured with the MHC-SF-R was higher than expected.

The secondary aim of the study was to compare the MHC-SF-R to the original version of the MHC-SF, dealing with the fourth hypothesis, namely that the MHC-SF-R would perform at least equally to the MHC-SF on internal consistency and convergent validity. At this point, results had to be reflected very carefully. Since the sample was split, no statistical comparison could be applied to compare both versions. However, both instruments reached satisfactory results for internal consistency and convergent validity. Therefore, the fourth hypothesis was accepted. CFA could not demonstrate acceptable model fit for a four-factor

structure of the MHC-SF-R, nor for a three-factor structure of the MHC-SF. In the following, it will be further elaborated on the comparison of both versions.

All in all, findings are two-part. One the one hand, the MHC-SF-R was found to have good reliability and convergent validity. On the other hand, a four-factor structure with emotional well-being, societal well-being, relational well-being and psychological well-being could not be verified, as the subscale of societal well-being performed poor, in particular.

Reflection on Results

Since the revision of the MHC-SF was currently developed and not tested yet, there are no existing findings about its psychometric properties. Findings on the psychometric properties of the MHC-SF are in line with previous research, supporting preliminary research by Lamers et al. (2011) and others. Thus, this study could confirm good internal consistency and good convergent validity, measured with related constructs like self-compassion and self-esteem. In contrast to existing evaluations of the MHC-SF, where the social well-being subscale mostly scored lowest on internal consistency, this study found more satisfactory results for the subscale. Beside the aforementioned shortcomings of the MHC-SF-R, the total scale of well-being, measured with the MHC-SF-R and the MHC-SF were concluded to be reliable and valid for the sample of University students in this study.

Looking at the overall results, it gets visible that societal well-being was the only subscale which performed insufficiently on internal consistency. Moreover, this dimension demonstrated the most unsatisfactory results concerning factor loadings, item-total correlations and correlations to constructs measured with the SCS-SF and the RSES, making it selectively important for further investigation. This finding is especially striking, as the social subscale already had to undergo the most significant changes. To get insight into possible reasons for these problems, all items of the social dimension were explicitly examined by the researcher. Examination showed one characteristic of the items of social well-being, namely that most items are strongly dependent on how people perceive society. Partly, the statements do not even address a person directly, but ask about general assumptions about society. Therefore, it may be questionable if the content of the societal dimension even adds information about a person’s well-being, as it assumes that people’s well-being is dependent of their understanding of society. By using the term society, one remark, mentioned by Köhle (2011) could still not be overcome, namely that understanding of society is very abstract and strongly depends on personal understanding. It is left open,

whether society includes people all over the world, or whether it is restricted to a narrower environment.

Comparing both versions, emotional well-being performed better in the revised version as factor loadings of items were higher, as well as internal consistency. Since items of this scale are the same in both tests, a possible explanation for differing results may be due to the time period on which participants were asked to rate the items. In the revised version, participants rated their feelings based on the past week, instead of the past month. Thus, a shorter time period may be more appropriate, as longer time periods run the risk of recall bias in the respondents. People may find difficulties in remembering all contributing experiences from the last month, so that their answer gets inaccurate (Stull, Leidy, Parasuraman, & Chassany, 2009). Psychological well-being performed better in the MHC-SF on factor loadings and internal consistency, which may indicate that shortening of the items did not work successfully. As the items in the revised version all share the same sentence structure, it could be possible that people perceive items too similar. Thus, if people cannot clearly differentiate between two aspects, it may lead to imprecise answers (Bangor, Kortum, & Miller, 2009).

Performance of the social well-being dimension cannot be directly compared, as the MHC-SF-R split it into two further subscales. For the MHC-SF, social well-being performed good on internal consistency and factor loadings. On the contrary, societal well-being showed low factor loadings and internal consistency. Only looking at the societal subscale, the social well-being dimension performed better in the original version. However, for the MHC-SF-R relational well-being could reach satisfactory results for internal consistency and factor loadings. Since this dimension was completely new developed, results for the subscale of relational well-being are especially pleasant. Correlations between each version of the MHC-SF and the SCS-SF and RSES were good. In line with the expectation, psychological well-being of the MHC-SF-R correlated highest with the SCS-SF. On the contrary it is striking, that psychological well-being measured with the MHC-SF showed a low correlation with the SCS-SF. Conclusively, the results are ambiguous, and no version outperforms the other. As, for now, the societal dimension is clearly showing poor performance, the MHC-SF is considered to be more appropriate for further use.

Lastly, contrary to the expectation, self-esteem was stronger related to well-being than self-compassion. Especially for the subscales of emotional well-being and psychological well-being, scores correlated higher with self-esteem than with self-compassion. Still, in line with existing research, both constructs are in fact related to well-being and also strongly related to

each other, supporting findings by Neff (2003b) and Souza & Hutz (2016). A possible connection could lay in the hypothesis that self-esteem is a construct that is more easily understood by people, which facilitates their own assessment of self-esteem, rather than their own assessment of self-compassion. Neff & Vonk (2011) state that people tend to rate themselves high in their level of self-esteem, to defend themselves and present themselves favorable, especially if they are feeling well. Only if people experience failures, their sense of self-compassion is about to rise and gets more available to them. Following, people identify with self-esteem in positive ways, whereas self-compassion appears to be more important in negative situations (Zessin et al., 2015). Possibly, this issue may be related to a stronger relationship between well-being and self-esteem in this study, as the sample comprised students expected to be healthy on average.

Strengths and Limitations

All findings have to be regarded with sight on the strengths and limitations of the study. A strength of this study is the randomization of groups. Administering both versions of the test within the whole group could have led to response bias, like repetition bias. However, at this point it is hard to assess whether splitting the group brings more advantages than disadvantages. Ideally, both versions of the test should have been administered in the whole group, without implying response bias. As this study design cannot possibly be conducted, a randomization of two groups was preferred.

Another strength is that the current study presents first insights into a newly developed version of a commonly used measurement instrument. Although further testing is needed, the study could reach significant results and gave rise to striking aspects, like the insufficiency of the societal well-being subscale. Following, it allows to already set a focus for further research.

One limitation is an insufficient sample size for each group (n=55; n=53), which makes investigation of factorial structure with common methods, like CFA, problematic. A minimum sample size of 100 is recommended. Good sample size is reached with 500, and at best, even more than 1000 participants (MacCallum, Widaman, Zhang, & Hong, 1999).

The second limitation is the use of convenience sampling as a sampling method. Convenience sampling falls under the category of non-probability sampling which has two main disadvantages, namely selection bias and generalizability of the results. Participants were easily accessible or even related to the researcher, which resulted in a total sample that

was not random. Therefore, it is questionable whether results can be generalized to a wider population, or even to other populations (Etikan, Musa & Alkassim, 2016).

Another form of bias that this study is vulnerable to is the possibility of social desirability. This implies that participants tend to give answers that they regard as more socially favorable (Fisher & Katz, 2000). As the study was investigating well-being and a great part of the participants was personally related to the researcher, it cannot be guaranteed that participants gave honest answers. Although the study was conducted online, participants that were not familiar with the online tool of Qualtrics reported that they were afraid that their responses can be led back to them. Thus, it is possible that they did not want to disclose their true thoughts to the researcher.

The last limitation is that this study does not allow for statistically comparing performance of the MHC-SF with the MHC-SF-R, as the sample was split into two groups. Additionally, it was not possible to compare both groups on their general level of well-being to assess whether groups are approximately equal to be suited for comparison. If one group performs significantly better on well-being in general, results had to be discussed under this condition. A general measure for well-being, that is the same for both groups, should have been included in the study to assess whether both groups are comparable on their level of well-being.

Future Recommendations

For the future, several implications can be proposed. First, the MHC-SF-R should further be tested, including other populations and larger sample sizes. Possible populations include adults in various age groups, from different countries. Within future studies, the MHC-SF-R can be evaluated on its psychometric properties to gain support for good reliability and validity. Once all subscales can prove reliability and validity, it is useful to work on broader studies including the MHC-SF-R, evaluating test-retest reliability and its sensitivity to change. This is necessary to establish the new measure and to make it suitable for replacing the original MHC-SF in the long run.

Special attention should be directed to its factorial structure. If further research shows that a four-factor structure of the MHC-SF-R cannot be confirmed in several populations, two options are possible. One possibility would be to examine another factor structure, for example treating the social dimension and the relational dimension as one subscale of social well-being and trying to confirm a three-factor structure of the MHC-SF-R. Another possibility is to focus on exploratory factor analysis to get insights into other factorial

structured of the MHC-SF-R. For looking deeper into factorial structure of the MHC-SF-R, increasing sample size is especially important. Future studies should include sufficient sample size for EFA and CFA to reach accurate results and draw reliable conclusions.

Still, interpretation problems with the items of the MHC-SF are not solved. Findings also revealed most crucial problems with the societal well-being subscale. Therefore, it is especially recommended to further work on this subscale. One possible recommendation is to go on with qualitative research for the societal well-being subscale, to get insight into reasons for unsatisfactory results. To be more precise, it is recommended to use the method of thinking aloud while future participants are completing the test. Moreover, personal interviews about the questionnaire could be conducted. Both strategies are expected to provide deeper insights into possible interpretation problems of the items.

Conclusion

All in all, it is regarded as meaningful to promote research on a revised version of the MHC-SF. However, the MHC-SF-R, as evaluated in this study, is not yet sufficiently tested to start using it in practice. Therefore, it is still recommended to use the MHC-SF for screening people's well-being, but to keep focus on research on a revised version, that is especially attentive to the dimension of societal well-being.

References

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114-123.

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.

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, *50*(2), 1-5.

Bowling, A. (2005). Just one question: If one question works, why ask several? *J Epidemiol Community Health, 59*, 342-345. doi: 10.1136/jech.2004.021204

Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: An introduction. *Journal of happiness studies*, *9*(1), 1-11.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics*, *5*(1), 1-4.

Fisher, R. J., & Katz, J. E. (2000). Social-desirability bias and the validity of self-reported values. *Psychology & marketing*, *17*(2), 105-120.

Hall, C. W., Row, K. A., Wuensch, K. L., & Godley, K. R. (2013). The role of self-compassion in physical and psychological well-being. *The Journal of psychology*, *147*(4), 311-323.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Karaś, D., Cieciuch, J., & Keyes, C. L. (2014). The polish adaptation of the mental health continuum-short form (MHC-SF). *Personality and Individual Differences*, *69*, 104-109.

Keyes, C. L. M. (2009). Atlanta: *Brief description of the mental health continuum short form (MHC-SF).*

Köhle, N. (2010). “Mag ik vragen, wat ik u moet vragen?“.

Kline, R.B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd Edition ed.). New York: The Guilford Press.

Lamers, S. M., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. (2011). Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF). *Journal of clinical psychology*, *67*(1), 99-110.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, *4*(1), 84.

Neff, K. D. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and identity*, *2*(2), 85-101.

Neff, K. D. (2003b). The development and validation of a scale to measure self-compassion. *Self and identity*, *2*(3), 223-250.

Neff, K. D. (2011). Self-compassion, self-esteem, and well-being. *Social and personality psychology compass*, *5*(1), 1-12.

Neff, K. D., & Vonk, R. (2009). Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of personality*, *77*(1), 23-50.

Perugini, M. L. L., de la Iglesia, G., Solano, A. C., & Keyes, C. L. M. (2017). The mental health continuum–short form (MHC–SF) in the Argentinean context: Confirmatory factor analysis and measurement invariance. *Europe's journal of psychology*, *13*(1), 93.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical psychology & psychotherapy*, *18*(3), 250-255.

Rogoza, R., Truong Thi, K. H., Różycka-Tran, J., Piotrowski, J., & Żemojtel-Piotrowska, M. (2018). Psychometric properties of the MHC-SF: An integration of the existing measurement approaches. *Journal of clinical psychology*, *74*(10), 1742-1758.

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

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763-1768.

Seligman, M. E. (2002). Positive psychology, positive prevention, and positive therapy. *Handbook of positive psychology*, *2*(2002), 3-12.

Shevlin, M., & Miles, J. N. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual differences*, *25*(1), 85-90.

Souza, L. K. D., & Hutz, C. S. (2016). Self-compassion in relation to self-esteem, self-efficacy and demographical aspects. *Paidéia (Ribeirão Preto)*, *26*(64), 181-188.

Strauss, C., Taylor, B. L., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical psychology review*, *47*, 15-27.

Stull, D. E., Leidy, N. K., Parasuraman, B., & Chassany, O. (2009). Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Current medical research and opinion*, *25*(4), 929-942.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, *2*, 53.

Tinakon, W., & Nahathai, W. (2012). A comparison of reliability and construct validity between the original and revised versions of the Rosenberg Self-Esteem Scale. *Psychiatry investigation*, *9*(1), 54. doi: 10.4306/pi.2012.9.1.54

WHO (2001). *Basic documents*. 43rd Edition. Geneva, World Health Organization:1.

Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta-analysis. *Applied Psychology: Health and Well-Being*, *7*(3), 340-364.

Appendix

The MHC-SF revised

The next questions concern feelings that people can have. Read each item carefully and mark the answer that best represents how often you have experienced the feeling during the PAST WEEK.

	Never	Rarely	Some-times	Regu-larly	Often	(Almost) always
I am happy	0	1	2	3	4	5
I am interested in life	0	1	2	3	4	5
I am satisfied with my life.	0	1	2	3	4	5
I make a valuable contribution to our society	0	1	2	3	4	5
I think our country is developing well.	0	1	2	3	4	5
I accept others as they are.	0	1	2	3	4	5
I belong to a group of people.	0	1	2	3	4	5
I understand how our society works.	0	1	2	3	4	5
I accept myself as I am.	0	1	2	3	4	5
I am able to master my life.	0	1	2	3	4	5
I share love and sorrow with some people.	0	1	2	3	4	5
I can develop myself.	0	1	2	3	4	5
I stand up for myself.	0	1	2	3	4	5
I feel my life has purpose.	0	1	2	3	4	5
I can mean something for others.	0	1	2	3	4	5
I am satisfied with my social contacts.	0	1	2	3	4	5
I feel connected to other people.	0	1	2	3	4	5
I can rely on other people.	0	1	2	3	4	5
I find my place in this society.	0	1	2	3	4	5