Evaluation of a Practice Friendly Version of the Mental Health Continuum - Short Form

Melissa Vollmer

s2875896

University of Twente

BMS Faculty, Department of Psychology

First Supervisor: Dr. Erik Taal

Second Supervisor: Prof. Dr. Gerben Westerhof

Abstract

The Mental Health Continuum - Short Form (MHC-SF) is one of the most prominent measurement tools evaluating mental well-being. Despite its widespread application, participants report difficulties with estimating the frequency of their emotions. Therefore, the response format was changed by simplifying the answer options and shortening the reference period. The current research focuses on evaluating the psychometric properties of the revised questionnaire (MHC-SF-R) and comparing it to the MHC-SF. Overall, 1377 randomly selected participants contributed to this research. An evaluation of the factor structure through an exploratory factor analysis was conducted. Furthermore, the internal consistency was assessed through Cronbach's alpha and the item-total correlations. Lastly, we investigated participants responses by conducting an analysis of Skewness and Kurtosis, an independent samples t-test and an evaluation of floor and ceiling effects. Results indicated that both versions of the questionnaire show similar internal consistency. The response distribution of the MHC-SF-R is improved as normality and a reduction in floor and ceiling effects was achieved. However, the MHC-SF-R appears to show no clear relationship between the factor structure and theoretical background. In addition, a variety of cross-loading factors are present, indicating possible issues with the construct validity of the revised questionnaire. Based on the internal consistency, the MHC-SF-R appears to be a reliable measure. Lastly, the new answer format shows improved representation of the items dimensionality. In conclusion, while the psychometric properties of the MHC-SF-R show promise, we want to caution against the use of the MHC-SF-R until its validity has been ensured.

Evaluation of a Practice Friendly version of the Mental Health Continuum - Short Form

Introduction

The past years have increasingly brought mental health into the focus of public attention. Especially in the light of the current pandemic, many people are experiencing a decline in their mental health, frequently related to social isolation and reduced well-being (Vindegaard & Benros, 2020). Under a traditional paradigm of mental health, social or emotional factors would not have been considered influential (Mechanic, 1999). However, newer research suggests that emotional, social and psychological factors strongly influence mental health and well-being (Ryan & Deci, 2001). In the following, we will explore the evolving definition of mental health and illuminate the role that well-being plays in it.

Evolving Definition of Mental Health

Despite the topic of mental health being in the public eye, a unanimous definition does not exist. One of the earliest definitions is called the Medical Model of Mental Health, where mental health is characterized by the absence of mental illness (Mechanic, 1999). As the name suggests, it arose from a purely medical approach to psychology. According to this model, when mental illness is reduced or eliminated, mental health is achieved (Mechanic, 1999).

It was soon noted that the medical model is not in line with the public opinion of what mental health entails, as many people feel that it is more than the pure absence of illness. As a result, in 2004 the World Health Organization (WHO) defined mental health as “a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community (Mental Health: Strengthening Our Response, 2022).” This definition hones in on the experience of mental health as going beyond the mere presence or absence of mental illness, as it suggests that even with mental illness people could lead a life that feels worthwhile and that is filled with joy and pleasure.

In order to further conceptualize this phenomenon, the Two-Continuum Model was created (Westerhof & Keyes, 2010). The model identifies two factors, or continua, that

contribute to mental health. The first continuum is the presence or absence of mental illness. In line with the medical model, this continuum includes the burden of mental illness and its impact on daily life. The second continuum is referred to as mental well-being. It involves the person's subjective experience of their emotional state and functioning within the context of their life (Keyes, 2002). When considering these two continua, it can be seen that complete mental health can be achieved through the presence of mental well-being in the absence of mental illness (Westerhof & Keyes, 2010). This underlines the importance of discriminating between mental illness and well-being and indicates the need for assessment of both factors when determining a person's mental health.

Measuring Mental Well-Being

After illustrating the evolving definition of mental health, the practical question regarding the assessment of mental well-being remains. Research shows that it consists of three separate dimensions (Ryan & Deci, 2001). The first dimension comprises of emotional well-being, referring to the person's experiences of positive emotions and life satisfaction (Simsek, 2011). The second dimension is called psychological well-being, including factors such as self-acceptance, feeling a purpose in life, autonomy, and personal growth (Ryff, 1989). Lastly, social well-being makes up the third dimension. It encompasses feelings of social acceptance and integration. An important aspect within this dimension is feeling as though one contributes to society in a meaningful way (Keyes, 1998). All in all, it can be concluded that mental well-being requires emotional, psychological and social well-being to be present.

Starting in the 1960s, a variety of assessment tools have been developed for evaluating mental well-being. One instrument designed for the measurement of mental well-being is called the Mental Health Continuum - Long Form (MHC-LF) (Keyes, 1998; Keyes, 2002). This questionnaire measures the three factors of well-being through the administration of 40 items. It has been evaluated as a valid and reliable tool for assessing mental well-being

(Keyes, 2005). Despite its outstanding psychometric properties, the usability in psychological research was limited as the administration of the questionnaire requires a considerable investment of time and energy from the participants.

As such, the Mental Health Continuum - Short Form (MHC-SF) was developed to create a more accessible, shorter and less time intensive measurement instrument (Keyes et al., 2008). It includes 14 items, of which three are measuring emotional well-being, six psychological well-being and five social well-being. The psychometric qualities of the MHC-SF have been evaluated as satisfactory across all scales (Keyes et al., 2008; Lamers et al., 2011; Petrillo et al., 2015). As the MHC-SF is more practical due to its shorter duration and improved usability, an evaluation of psychometric properties is crucial. In the following section, reliability and validity of the MHC-SF will be discussed.

Psychometric Properties of the MHC-SF

First, the internal consistency of the questionnaire was assessed, determining whether the items measure the same construct. An internal consistency of $\alpha > .80$ was found for the total scale of the MHC-SF in samples of Dutch adolescents and Italian adults (Lamers et al., 2011; Petrillo et al., 2015). Additionally, on a subscale level, high internal consistency was demonstrated for the emotional well-being subscale ($\alpha=.83$) and the psychological well-being subscale ($\alpha=.83$) (Lamers et al., 2011). The internal consistency of the social well-being subscale was evaluated as adequate ($\alpha=.74$) (Lamers et al., 2011). Overall, the internal consistency of the MHC-SF has been found to be satisfactory on both total scale and subscale levels. This means that per scale, the included items appear to measure the same construct.

Second, test-retest reliability was assessed over a 9-month period. It was found to be moderate, indicating sensitivity to change and stability of the measurement tool (Lamers et al., 2011; Petrillo et al., 2015). This indicates that results obtained by the MHC-SF may vary when the questionnaire is administered again after 9 months, as compared to the first assessment.

Third, the construct validity of the MHC-SF was evaluated through the examination of discriminant, convergent and structural validity. This assessment determined whether the questionnaire truly measures the concept of mental well-being. It was achieved through evaluation of the factor structure, discriminant and convergent validity and will be discussed in the following paragraphs.

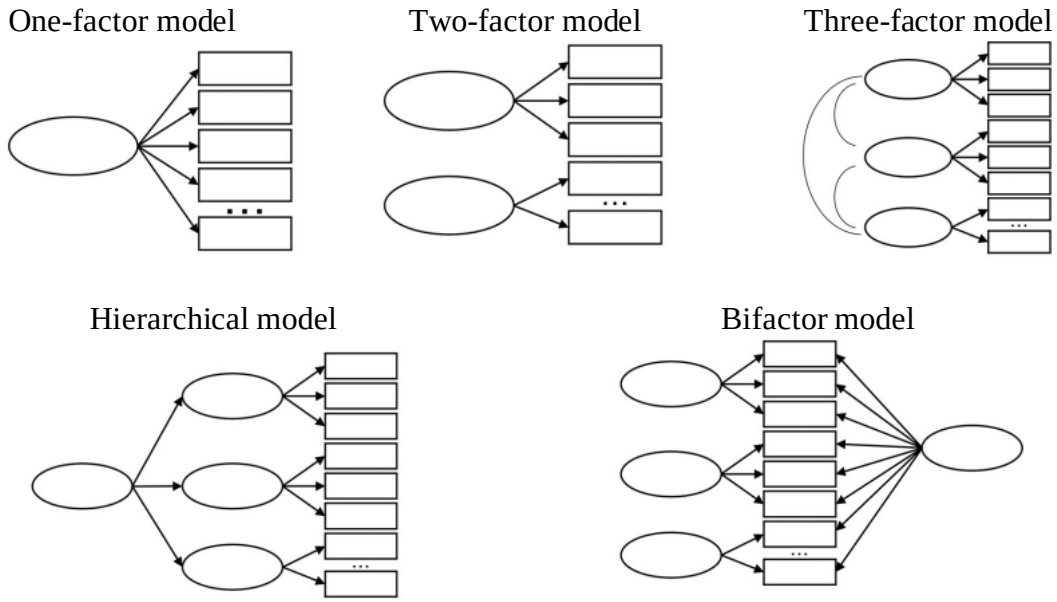
To begin, we examined the factor structure in order to evaluate structural validity. It determines whether the scores on the scale can actually represent the constructs dimensionality, thus assessing whether the chosen scale is an appropriate way of measuring the construct. Therefore, the question arises as to what factor solution should be chosen for the data. A meta-analysis by Iasiello et al. (2022) has identified large differences in the models used to explain the obtained data. The most prominent solutions found were a one-factor model, two-factor model, three-factor model, hierarchical model and bifactor model (Iasiello et al., 2022). The theoretical background behind these five factor solutions will now be explored further. Please view Figure 1 for a schematic representation of the discussed factor solutions.

When a one-factor solution was identified as being most appropriate for their data (Iasiello et al., 2022), this indicates that rather than having the three subscales of emotional, social and psychological well-being represented in the model, mental well-being would be the predominant factor explaining all the variance. Alternatively, a two-factor model could be used to explain the data obtained with the MHC-SF (Iasiello et al., 2022). Frequently when a two-factor solution is chosen, one factor explains the majority of the variance, followed by a second factor representing an underlying construct (Iasiello et al., 2022). It could be argued that the first factor may represent mental well-being, while the second factor represents an additional, underlying concept (Iasiello et al., 2022). Alternatively, a three-factor solution is frequently chosen to explain the data. This model is most in line with the theoretical background of the questionnaire, with each factor representing one of the subscales (Iasiello

et al., 2022). Further, the hierarchical model presents an alternative factor solution (Iasiello et al., 2022). It is made up of the three-factor model described above, with an additional, second order mental well-being factor. In this case it is assumed that items load on three factors representing the three types of well-being, which in turn load on a single factor, made up of mental well-being (Iasiello et al., 2022). Lastly, a bifactor model was proposed, integrating the one-factor model and the three-factor model (Iasiello et al., 2022). With this factor solution it is assumed that both types of models are present and influence the items independently.

Figure 1

Possible factor solutions of the MHC-SF



Note. This figure was adapted based on Iasiello, M., van Agteren, J., Schotanus-Dijkstra, M., Lo, L., Fassnacht, D. B., & Westerhof, G. J. (2022). Assessing mental wellbeing using the mental health continuum—short form: a systematic review and meta-analytic structural equation modelling. *Clinical Psychology: Science and Practice, (2022).*

In order to determine which factor solution is the most appropriate Iasiello et al. (2022) used structural equation modeling on the data obtained from 26 studies evaluating the MHC-SF. They concluded that a two factor model would be the most appropriate from a purely statistical standpoint, with the majority of the variance being made up by one factor. Simultaneously, concerns about the accuracy of such a solution were raised (Iassiello et al.,

2022; van Zyl & Olckers, 2019). It was questioned why a single factor solution achieved rather poor fit, while a two factor solution appeared most appropriate for the data. Based on the assumption that mental well-being would be the primary factor explaining most of the variance in a one-factor and two-factor model, similar fit indexes should be expected. However, this was not found in the review. Instead, psychometric research suggests that a questionnaire designed to assess a complex phenomenon is rarely explained by a single factor model (Reise et al., 2014). As the Mental Health Continuum was designed based on the three dimensions of mental well-being, it assesses a complex phenomenon and thus is not well explained by a one-factor solution (Iasiello et al., 2022).

Furthermore, psychometric issues were raised when it comes to a two-factor solution. It was noted that this factor model is more inaccurate in terms of its ability to accommodate for erroneous factor solutions (Reise et al., 2014). Thus, many authors have warned against the use of the two-factor model, regardless of it's frequently superior fit (Iassiello et al., 2022; Luciano et al., 2020; van Zyl & Olckers, 2019). Additionally, it was suggested that the use of a two-factor model may not contribute to a better understanding of the construct measured by the questionnaire (Iassiello et al., 2022; Sellbom & Tellegen, 2019). As an alternative, Iassiello et al. (2022) proposed the use of a three-factor model or hierarchical factor structure, as they showed appropriate fit for the data and were in line with the literature.

One study opting to use the three-factor model to explain the obtained data was conducted by Lamers et al. (2011). Their confirmatory factor analysis showed that this was indeed the model with the best fit for the obtained data. It was found that the three factors together explained 58% of the variance (Lamers et al., 2011). Additionally, it was confirmed that these components represented the three subscales emotional, psychological and social well-being (Lamers et al., 2011).

When evaluating these results in the context of the structural validity of the MHC-SF, more than 50% of the variance is explained when using a statistically and theoretically

appropriate factor model (Taylor, 1990). We concluded that the structural validity can be considered good. Thus, it can be assumed that the scores of the MHC-SF represent an adequate way to measure the dimensionality of mental well-being.

Furthermore, the discriminant validity was assessed. This subtype of construct validity evaluates whether the measured construct is weakly related to other constructs to which there should theoretically be no relation. It was assessed by comparing the MHC-SF, as a measure of mental well-being, with measures of mental illness. A moderate negative correlation with measures of mental illness was established. Therefore, it can be concluded that the discriminant validity is strong (Lamers et al., 2011; Petrillo et al., 2015). This indicates that the results obtained by the MHC-SF differ from other, theoretically unrelated constructs.

Lastly, the convergent validity was assessed by comparison of the different subscales with related measurements. This type of validity evaluates whether the MHC-SF obtains similar results than other measurement tools designed to assess similar concepts. As such, to test convergent validity, related concepts were identified for each of the three subscales. Specifically, emotional well-being was correlated with a measurement of life satisfaction, psychological well-being was compared with self-esteem and social well-being was assessed against social engagement (Lamers et al., 2011). Overall, the convergent validity of the MHC-SF was evaluated as moderately acceptable (Lamers et al., 2011). Thus, it can be concluded that the mental well-being measurements obtained by the MHC-SF correlate with other, related concepts.

In summary, the MHC-SF shows satisfactory psychometric properties across the assessed types of reliability and validity. When it comes to the questionnaires reliability, the internal consistency was evaluated as satisfactory on an item and scale level. Additionally, the test-retest reliability was assessed as moderate over a 9 month period. It can be concluded that the MHC-SF is a consistent measurement, achieving reproducible results. Moreover, the questionnaires construct validity was reviewed. The discriminant and convergent validity

were evaluated as moderate. Further, a good structural validity was found as the factor structure aligned well with the theoretical background of the MHC-SF. It can be concluded that the MHC-SF has good construct validity and appears to measure mental well-being.

Revision of the MHC-SF

Despite the short form of the Mental Health Continuum being developed to improve the usability of the MHC-LF in practice, the response format of the MHC-SF has been criticized. The MHC-SF asked the participants to answer 14 questions about their mental well-being on a 6 point scale utilizing approximate ranges (Geisen, 2020). It included the answer options (0) not at all, (1) once or twice a month, (2) about once a week, (3) 2 or 3 times a week, (4) almost every day and (5) every day. The instructions asked the participants to indicate the frequency of particular emotions over the course of a one month reference period. Many participants experienced filling in the questionnaire as difficult, as they were unable to recall the exact frequency with which the emotion had arisen.

As such, a research team from the University of Twente, the Netherlands, revised the answer format with the goal of improving the usability of the questionnaire. Since the participants gave the feedback that it was difficult to recall exact frequencies, this component was changed. The result was a new, 6 option response format, featured in the MHC-SF-R. Specifically, the participants could choose between (0) not at all, (1) rarely, (2) sometimes, (3) regularly, (4) often and (5) (almost) all the time. By using vague quantifiers (Geisen, 2020), it allowed participants to choose an answer option that felt right to them, rather than having to remember the exact frequency of their emotions. Additionally, the reference period was shortened from one month to one week, further improving recall (Martin, 2006).

By changing the response format of the questionnaire, it has become necessary to assess whether the two questionnaire versions can be considered equal. Already in 1979, it was established that changing the response format of a questionnaire changes the way that participants respond to the items, even if the items themselves remained unchanged (Frisbie &

Brandenburg, 1979). With this study we aim to shine a light on the effect that the revision had on the psychometric properties of the questionnaire.

One factor that may have been influenced by the new response format is the presence of floor and ceiling effects, referring to the extent to which participants select the highest or lowest score of the scale. A floor effect is indicated when over 15% of participants choose the lowest response option, while a ceiling effect is present when over 15% of participants choose the highest response option (McHorney & Tarlov, 1995; Terwee et al., 2007). Floor and ceiling effects hold significance when developing a new scale as they limit the ability to distinguish between participants at the upper (ceiling effect) and lower (floor effect) end of the scale (Murugappan et al., 2022). Their presence indicate that additional items outside the bounds of the provided scale are missing to adequately measure the construct at hand. As such the reliability of the scale is reduced (Terwee et al., 2007). Previously, the potential presence of floor and ceiling effects in the data collected with the MHC-SF has not been evaluated.

The Current Study

In the current study, we will evaluate the psychometric properties of the MHC-SF-R and compare those with the MHC-SF. In this context, we aim to answer the following research questions:

(1) What factor structure arises from the MHC-SF-R and how does it compare to that of the MHC-SF?

(2) What is the internal consistency of the MHC-SF-R and how does it compare to the internal consistency of the MHC-SF?

(3) Do participants answer questions differently depending on the response format they are presented with?

Method

Design and Procedure

This non-experimental, exploratory research uses the data of the Longitudinal Internet studies for Social Sciences (LISS) panel. The research institute *Centerdata*, a part of the Dutch Tilburg University, was responsible for the data collection. The LISS panel collects data from over 5000 households in the Netherlands, creating a simple random sample of the Dutch population as part of a longitudinal study on living in the Netherlands. Participants are questioned on a variety of topics such as health, social integration, political opinions, and personality. Additionally, the LISS panel was used to collect data with the intention of validating the psychometric properties of the MHC-SF-R (Westerhof & ten Klooster, 2020).

For the purpose of validating the MHC-SF-R, data was collected from a total of 2719 participants. Four conditions, each with a different revision of the questionnaire, were created and randomly assigned to the participants. The first group was asked to complete a questionnaire with the revised answering format and questions that were phrased in a simpler way. The second group was given the original MHC-SF. The third group was given the original items, but the revised response format. The fourth and final group was assigned to the revised phrasing of the question, with the original response format. For the purpose of this study, we will be using the data from group two and three. Group two will be referred to as MHC-SF, while group three will be referred to as MHC-SF-R. To specify, both groups have been given the original item formulation, while the group MHC-SF answered the questions using the original response format and the group MHC-SF-R was using the new response format.

Participants

From the 2719 participants, 1377 belonged to groups two and three. More precisely, 650 participant responses were collected using the MHC-SF, while 727 participant responses

were collected using the MHC-SF-R. For a representation of the demographic data, please view table 1.

Participants indicated their sex (male, female), age, marital status, highest educational degree, and occupation. About half of the participants were female (52%), and underwent intermediate (MHC-SF: 24%, MHC-SF-R: 22%) or higher (MHC-SF: 25%, MHC-SF-R: 24%) vocational education. The majority of participants were either married (MHC-SF: 45%, MHC-SF-R: 43%) or had never been married (41%) and were currently employed (MHC-SF: 50%, MHC-SF-R: 44%). One thing to note is that there is an overrepresentation of older adults within the group selected for this study, with the largest group made up of participants above 65 (MHC-SF: 23%, MHC-SF-R: 26%). This was done in order to allow for within-group comparisons in later life.

When conducting a chi-square analysis, no significant difference could be found between the participants that filled out the MHC-SF compared to the participants answering the MHC-SF-R as it pertains to demographic variables.

Table 1

Participant Demographics

	MHC-SF		MHC-SF-R		$\chi 2$ (df), p
	n	%	n	%	
Total	650		727		
Gender					$\chi 2$ (1) = 0.01, p = 0.92
Male	312	48	347	47.7	
Female	338	52	380	52.3	
Age (years)					$\chi 2$ (5) = 8.24, p = 0.14
15-24	86	13.2	108	14.9	
25-34	120	18.5	145	19.9	
35-44	85	13.1	65	8.9	
45-54	87	13.4	99	13.6	
55-64	120	18.5	120	16.5	
65 and older	152	23.4	190	26.1	
Education*					$\chi 2$ (8) = 3.79, p = 0.88
primary school	16	2.5	23	3.2	
vmbo	105	16.2	134	18.4	
havo/vwo	75	11.5	94	12.9	
mbo	156	24	163	22.4	
hbo	162	24.9	176	24.2	
wo	107	16.5	112	15.4	
other	29	4.3	11	1.5	
Marital Status					$\chi 2$ (4) = 4.27, p = 0.37
Married	298	45	312	42.9	
Separated	2	0.3	3	0.4	
Divorced	57	8.8	78	10.7	
Widowed	24	3.7	39	5.4	
Never married	269	41.4	295	40.6	
Work Status					$\chi 2$ (12) = 17.02, p = 0.15
Employed	323	49.7	320	44	
Self-employed	33	5.1	28	3.9	
Unemployed	10	1.5	14	1.9	
Student	66	10.2	86	11.8	
Homemaker	38	5.8	33	4.5	
Retired	120	18.5	165	22.7	
Unfit for work	33	5.1	34	4.7	
Other	27	4.1	47	6.5	

Note. Dutch levels of Education: vmbo (intermediate secondary education), havo/vwo (higher secondary education), mbo (intermediate vocational education), hbo (higher vocational education), wo (university)

Instruments

Mental Health Continuum - Short Form

The MHC-SF is a 14 item questionnaire measuring mental well-being. It includes subscales for all three facets of mental well-being, namely emotional well-being (EWB), psychological well-being (PWB) and social well-being (SWB). Participants were instructed to select the frequency in which they experienced certain emotions within the past month. Responses could be indicated on a 6 category response scale, including the options (0) not at all, (1) one or two times per month, (2) about once a week, (3) two or three times a week, (4) almost every day, (5) every day. The participants were asked to answer 3 questions on EWB, one item from this scale is: "How often in the past month did you have the feeling that you are happy?". Then, 5 questions on SWB were asked, for example: "How often in the past month did you have the feeling that you contributed something important to society?" Lastly, 6 questions on PWB were asked. One example of this subscale is the question: "How often in the past month did you have the feeling that your life has a direction or meaning to it?" For a full list of items in the Dutch original and English translation, please view appendix 1.

Mental Health Continuum - Short Form, Revised

The MHC-SF-R utilized the same 14 items of the MHC-SF, described above. However, the revision involved different instructions for the participants. They were asked to indicate how frequently they felt a certain way, within the past week. This was indicated through the use of 6 answer categories, namely (0) never, (1) rarely, (2) sometimes, (3) regularly, (4) often, (5) (almost) always.

Once the participants answered the questionnaire with either the response format used by the MHC-SF or that used by the MHC-SF-R, the means were calculated from their answers. This was done on a total scale level by including all 14 items and on a subscale level.

Data Analysis

The data was analyzed using IBM SPSS Statistics for Windows (Version 28). First, the data was prepared for analysis by screening for missing cases and items. Participants that did not answer all the items they were assigned to, were excluded from the analysis. Then, the questionnaire was scored in accordance with the test manual. For each participant, a mean result for the questionnaire was calculated, as well as means for each of the three subscales. Additionally, demographic data was analyzed using descriptive statistics. We determined the amount, frequency and percentages for the variables gender, age, education, marital status and work status. These variables were analyzed using a chi-square test in order to determine any differences between the groups.

Then, we tested for normality of the data set by determining skewness and kurtosis of the MHC-SF and MHC-SF-R on a total scale and on a subscale level. As a normal distribution was found, we opted for the use of parametric tests.

To answer the first research question regarding the factor structure, an exploratory factor analysis with direct oblimin rotation was performed for the MHC-SF and the MHC-SF-R. In order to determine the number of factors included in the factor structure, we considered the eigenvalues-greater-than-one criterion (Kaiser, 1960) and performed a scree-test (Cattell, 1966). Based on the underlying factor structure of the MHC-SF and MHC-SF-R, the two versions were compared.

The second research question, pertaining to the internal consistency of the questionnaires, was analyzed by determining Cronbach's alpha of the total scale and of the subscales for both the MHC-SF and the MHC-SF-R. We then compared the obtained alpha values of both questionnaire versions. Additionally, we analyzed the item-total correlations on an item and subscale level, in order to determine to what extent the items contributed to the internal consistency of the scales. Lastly, we compared the determined item-total correlations between the two questionnaire versions.

In regards to the third research question focusing on the participants' response pattern, we calculated the skewness and kurtosis for each item of the two questionnaire versions, as well as the total scale and subscales. In this way, the response distribution and normality was analyzed. We then compared the results of the two versions. Next, an independent sample t-test was conducted to compare the means of the two questionnaires. Further, we examined the participants' response pattern. The presence of floor and ceiling effects was assessed by analyzing the number of responses indicating the lowest (0) and the highest (5) response category on the items. We determined the items that had a response rate of 15% or higher on the lowest or highest response option by using frequency analysis. We then compared these frequencies by creating a cross-table. More concretely, the response category belonging to the floor or ceiling effect was coded as 1, while all other answer options were coded as 0. This allowed us to use a chi-square test to determine the differences between the questionnaire versions. Additionally, we analyzed the data obtained from the subscales and the total scale for floor and ceiling effects and compared the two versions of the questionnaire.

In regards to the third research question, we examined the frequencies, independent samples t-test, floor and ceiling effects and chi-square test. Based on our findings, we assessed whether the answer options influenced participant responses to the same items between the two questionnaire versions.

Results

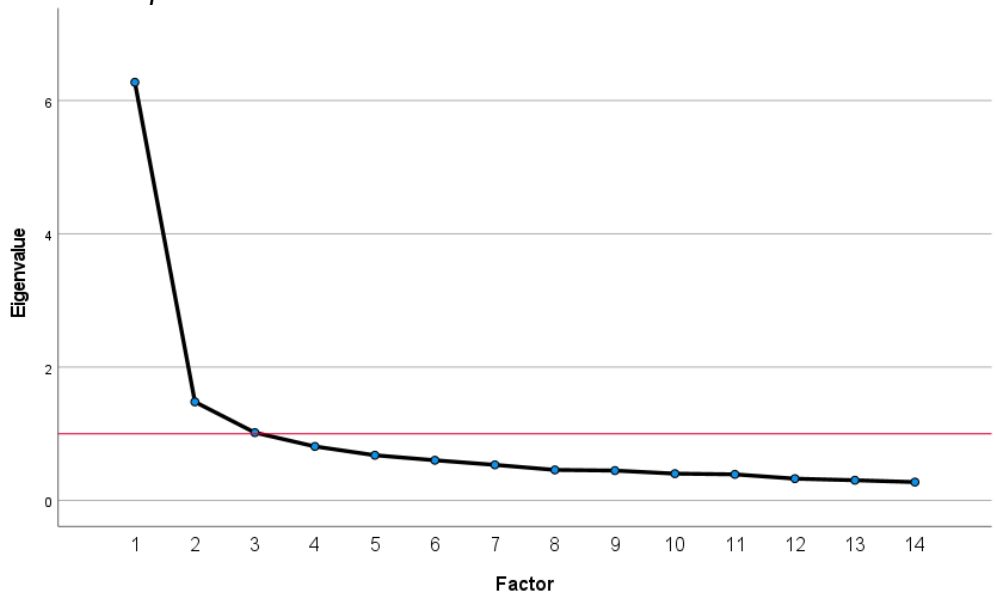
Research Question 1

An exploratory factor analysis with a direct oblimin rotation and a scree-test were performed in order to answer the first research question: What factor structure arises from the MHC-SF-R and how does it compare to that of the MHC-SF? When considering the eigenvalues-greater-than-one criterion, we determined that a three-component factor structure seems most appropriate for both the MHC-SF and MHC-SF-R. However, when analyzing the scree-plots of the MHC-SF and MHC-SF-R, using the elbow method, a sharp fall of the curve

can be observed between components one and two for both plots respectively. As such, a one-factor solution may be the most appropriate for both versions of the questionnaire. The scree-plot of the factor analysis performed on the two questionnaire versions can be seen in figures 2 and 3.

Figure 2

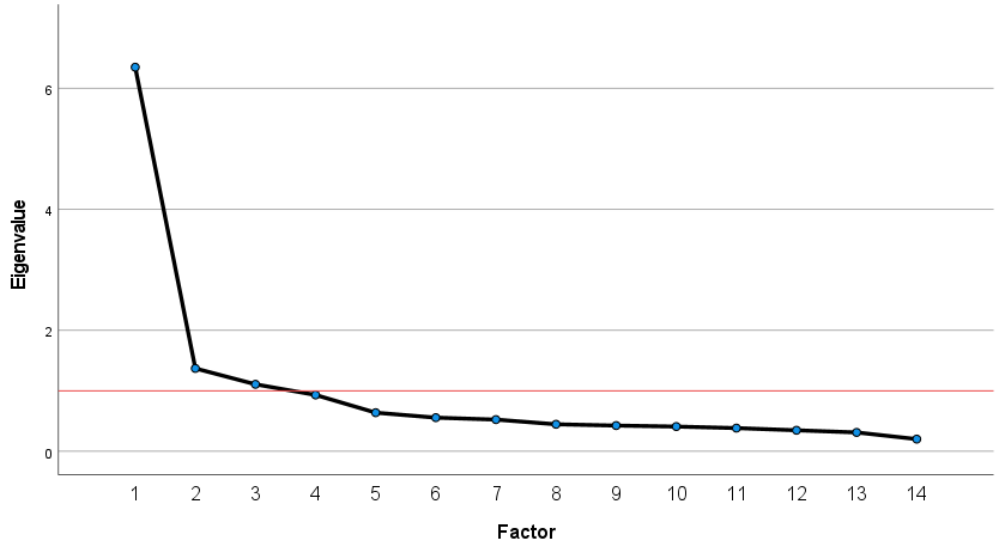
Scree-Plot of MHC-SF



Note. The red line indicates an eigenvalue of one.

Figure 3

Scree-Plot of MHC-SF-R



Note. The red line indicates an eigenvalue of one.

Despite this, we opted to retain the three-component factor structure indicated by the eigenvalues as this is most in line with the theoretical background of the questionnaire. By choosing this factor solution, we retain the representation of the three subscales emotional well-being, social well-being and psychological well-being as factors.

To continue, we examined the explained variance of the three factors. When examining the MHC-SF, it seems that almost half of the variance (44.83%) is explained by the first factor. The second factor explains 10.57% and the third and final factor explains 7.27% of the total variance. In total, a cumulative variance of 62.66% was achieved by including the three components. A similar pattern of variance can be observed from the analysis of the MHC-SF-R. The first factor explains 45.36% of the total variance, followed by the second factor explaining 9.79% and the final factor explaining 7.91% of the variance. The three components together make up 63.05% of the total variance.

MHC-SF

Further, we evaluated the factor loadings after rotation, which are represented in table 2. When considering the factor loadings, it can be observed that the three subscales seem to be represented in the structure matrix of the MHC-SF. The first factor appears to be made up of the subscale psychological well-being as the items from the scale load predominantly on this factor. Additionally, the items of the emotional well-being subscale load mostly on component 2. Therefore, it can be concluded that the second identified component represents the emotional well-being subscale. Lastly, the items of the social well-being scale strongly load on component 3, thus making up this factor. The only exception is item 8, which is loading on both factor 1 and 3. Overall, the factor analysis of the data obtained from the MHC-SF supports the selected three-factor solution.

Table 2

Factor Loadings on Identified Components after oblimin rotation

	MHC-SF			MHC-SF-R		
Items	Component 1	Component 2	Component 3	Component 1	Component 2	Component 3
Item 1 (EWB)	.07	**-.83**	.04	**.88**	.11	-.22
Item 2 (EWB)	.16	**-.75**	.06	**.78**	.07	-.01
Item 3 (EWB)	.09	**-.83**	.02	**.88**	.07	-.15
Item 4 (SWB)	-.10	-.39	**.60**	.25	**.65**	-.15
Item 5 (SWB)	.01	-.13	**.69**	.07	**.80**	-.07
Item 6 (SWB)	-.02	.04	**.84**	-.13	**.83**	.14
Item 7 (SWB)	.20	-.01	**.62**	-.03	**.51**	**.46**
Item 8 (SWB)	**.47**	.14	**.43**	-.07	.20	**.76**
Item 9 (PWB)	**.68**	-.11	.14	**.43**	.02	**.54**
Item 10 (PWB)	**.86**	-.11	-.22	**.59**	-.16	**.48**
Item 11 (PWB)	**.71**	-.21	-.06	**.61**	.03	.30
Item 12 (PWB)	**.48**	.15	.39	.23	**.42**	.10
Item 13 (PWB)	**.71**	-.01	.13	**.46**	.06	**.40**
Item 14 (PWB)	**.52**	-.27	.18	**.63**	.16	.19

MHC-SF-R

When looking at the factor loadings of the MHC-SF-R items, the representation of sub scales within the factor matrix is less evident. When assessing component 1, it can be noted that both the emotional and psychological well-being items predominantly load on this component. The exception here poses item 12, from the psychological well-being subscale, which loads mostly on component 2. In addition, items 9, 10 and 13 not only load on this factor, but also load on component 3. Thus, the emotional well-being subscale combined with the majority of the psychological well-being subscale make up component 1.

Regarding component 2, it appears that it best represents the social well-being subscale, as all items load highly on this factor. One exception is item 8, which loads predominantly on component 3. Additionally, it is of note that item 7 loads significantly on both components 2 and 3. Moreover, item 12 from the psychological well-being subscale loads on component 2.

Lastly, component 3 is made up predominantly of items that also load on the other two factors. Item 7 from the social well-being subscale, loads highest on component two, but also strongly loads on component 3. Additionally, items 9, 10 and 13 from the psychological well-being subscale also load significantly on component 1. The notable exception poses item 8 from the social well-being subscale, which exclusively loads on component 3. Based on the factor analysis, the factor structure of the MHC-SF-R remains unclear.

Differences Between MHC-SF and MHC-SF-R

When comparing the factor structure obtained from both questionnaire versions, it is evident that the data collected through the use of the MHC-SF shows a factor structure that is straightforward and in line with the theoretical background of the Mental Health Continuum. With the exception of item 8, each item loads on a single component. Moreover, these factor loadings show a clear representation of the three underlying subscales.

This clear picture is not present in the data obtained from the MHC-SF-R, as the components do not align with the underlying subscales. Additionally, a variety of items strongly load on multiple components. This underlines that the three subscales are less distinctly represented in the data of the MHC-SF-R, than they are in the MHC-SF.

Research Question 2

In order to answer the second research question (What is the internal consistency of the MHC-SF-R and how does it compare to the internal consistency of the MHC-SF?), we analyzed the reliability of the MHC-SF and MHC-SF-R on a total level and on a subscale level, considering Cronbach's alpha and item-total correlation.

Cronbach's alpha

First, we analyzed the reliability of the MHC-SF on a total and subscale level. For a full representation of the results used for the assessment of reliability, please refer to table 3. When analyzing the total scale, a high reliability ($\alpha = .90$) was found. On a subscale level, both the emotional well-being ($\alpha = .86$) and psychological well-being ($\alpha = .85$) scale can be considered as having good reliability, while the social well-being scale can be considered adequate with a Cronbach's alpha of .78.

Next, we analyzed the reliability of the MHC-SF-R. On a total scale level, an excellent reliability with a Cronbach's alpha of 0.90 was found. On a subscale level, both the reliability of emotional well-being ($\alpha = 0.88$) and psychological well-being ($\alpha = 0.85$) can be considered good, while social well-being ($\alpha = .77$) obtained an acceptable reliability.

When comparing the two versions of the questionnaire, we can observe that Cronbach's alpha did not differ much between the two versions. Additionally, we observed that the confidence intervals of the two questionnaire versions overlapped on a total scale and subscale level. As such, we can conclude that no significant differences could be found between the two versions of the measurement instrument. Both questionnaire versions can be considered reliable as high alpha values were achieved.

Table 3

Cronbach's alpha

	MHC-SF (N = 727)			MHC-SF-R (N = 650)		
	Cronbach's alpha	CI (95% lower)	CI (95% upper)	Cronbach's alpha	CI (95% lower)	CI (95% upper)
Total scale	.90	.89	.91	.90	.89	.91
EWB	.86	.84	.87	.88	.86	.89
SWB	.78	.76	.81	.77	.74	.80
PWB	.85	.83	.87	.85	.83	.87

Item-total correlations

Additionally, we examined the item-total correlations on an item and subscale level in order to determine which aspects contribute most to the questionnaire's internal consistency. The item-total correlations for the two versions of the questionnaire are shown in table 4.

To begin with, we analyzed the item-total correlations for each item with the total scale. All items have a relatively high item-total correlation, with the lowest item being item 6 of the social well-being subscale, which achieved a .53 item-total correlation in both versions of the questionnaire.

Then, we analyzed the item-total correlation on a subscale level, looking at the correlations of the individual items with the subscale they are assigned to. Just like on an item level, the three subscales all show a high item-total correlation. When looking at the item-total correlations of items 1, 2 and 3 with the emotional well-being subscale, we observed similar scores for both versions of the questionnaire. The lowest scores are achieved by item 2, with a .71 item total correlation for the EWB subscale of the MHC-SF and a score of .7 in the EWB subscale of the MHC-SF-R. All three items show excellent item-total correlation for both versions of the questionnaire. Additionally, we examined the item-total correlations of items 4 to 8 with the social well-being subscale. Here, item-total correlations appear lower than those

of the EWB subscale. Ranging from .44, achieved by item 8 of the MHC-SF-R to .64, achieved by item 6 of the MHC-SF-R. Overall, these item-total correlations can still be considered good and should allow for sufficient discrimination of high and low scores within this subscale. Lastly, we assessed the item-total correlation of items 9 to 14 with the PWB subscale. In line with the results obtained from the other subscales, the PWB subscale too shows good discrimination, with item-total correlations ranging from .49 achieved by item 12 of the MHC-SF-R to .71, achieved by item 14 of the MHC-SF-R. Overall, it can be observed that the EWB subscale seems to have the highest item-total correlation across it's items, while the SWB subscale seems to have the lowest item-total correlation across it's items and the PWB subscale falling in the middle.

When comparing the item-total correlations of the two versions of the questionnaire on a total and subscale level, no major differences can be observed. All items and subscales have achieved an item-total correlation of above .4 and therefore have very good discrimination.

Table 4:

Item-total correlations on an item and subscale level

	MHC-SF				MHC-SF-R			
	Total scale	EWB subscale	SWB subscale	PWB subscale	Total scale	EWB subscale	SWB subscale	PWB subscale
Item 1 (EWB)	.57	.74			.65	.79		
Item 2 (EWB)	.62	.71			.67	.70		
Item 3 (EWB)	.58	.74			.67	.80		
Item 4 (SWB)	.57		.52		.55		.50	
Item 5 (SWB)	.55		.56		.54		.57	
Item 6 (SWB)	.53		.60		.53		.64	
Item 7 (SWB)	.58		.58		.56		.58	
Item 8 (SWB)	.57		.54		.47		.44	
Item 9 (PWB)	.70			.68	.66			.66
Item 10 (PWB)	.58			.63	.61			.60
Item 11 (PWB)	.65			.66	.68			.68
Item 12 (PWB)	.54			.53	.51			.49
Item 13 (PWB)	.66			.68	.63			.69
Item 14 (PWB)	.71			.67	.73			.71

Research Question 3

In order to answer the third research question (Do participants answer questions differently depending on the response format they are presented with?) we explored the differences between the mean scores obtained on a total scale and subscale level. Moreover, we investigated the presence of floor and ceiling effects in the data and analyzed the distribution of responses.

Distribution of Responses

In order to gain insight into the way that participants answered the questionnaire, we looked at the distribution of responses on a total scale, subscale and item level. Overall, we have observed an almost exclusively normal distribution on a scale and subscale level. The Skewness and Kurtosis of the total scale and for each subscale is listed in table 5. Lastly, histograms for each item and scale are shown in appendix B.

When looking at the distribution of the total questionnaire, both versions are normally distributed. This is also the case, for the social and psychological well-being subscales, which show a normal distribution. In contrast to this, the data obtained from the emotional well-being subscale of the MHC-SF is both skewed and too peaked to be normally distributed. Please view figures 4 and 5 for the histograms of the emotional well-being subscale for each version of the questionnaire

Moreover, we explored the distribution of responses for each item, comparing the response format of the MHC-SF with that of the MHC-SF-R. First, it can be noted that for all items of the MHC-SF-R a normal distribution was achieved, as the skewness and kurtosis were within a normal range of -1 to 1. This was not the case for the MHC-SF. Here, items 2, 3 and 10 can be considered both skewed in the distribution and too peaked. Additionally, the answer distribution of item 5 and 12 should be considered too flat, while not skewed.

T-test with independent samples

When looking at the results obtained from the total questionnaire, participants scored differently between the MHC-SF and the MHC-SF-R. Please view table 6 for the outcome of the t-test. The mean score obtained from the MHC-SF-R is 0.1 points lower compared to the MHC-SF. Similar differences can be found on a subscale level.

First, a significant difference was observed between the emotional well-being subscales of the MHC-SF and MHC-SF-R. On average, participants scored 0.29 points lower on the MHC-SF-R. Moreover, the mean score of the psychological well-being scale also differed significantly between the MHC-SF and the MHC-SF-R. The scores on the MHC-SF-R were on average 0.12 points lower. One exception is the social well-being scale, which does not differ significantly between the MHC-SF and the MHC-SF-R. Additionally, this is the only subscale where the mean is higher for the MHC-SF-R as compared to the MHC-SF.

These findings show that the revision of the MHC-SF causes a significant difference between the results obtained from the two versions of the questionnaire. With the exception of the social well-being subscale, all other scales show a difference between the two questionnaire versions. However, when considering the effect sizes, we can see that these differences are rather small. The largest effect size can be found in the emotional well-being scale, with a Cohen's d of 0.31, which should still be considered small. As such, we can conclude that the present differences between the two versions of the questionnaire should be considered minimal. This indicates that despite the changed response format of the items, people answer rather similarly when comparing the MHC-SF with the MHC-SF-R.

Figure 4

Histogram of the Emotional Well-Being Subscale of the MHC-SF

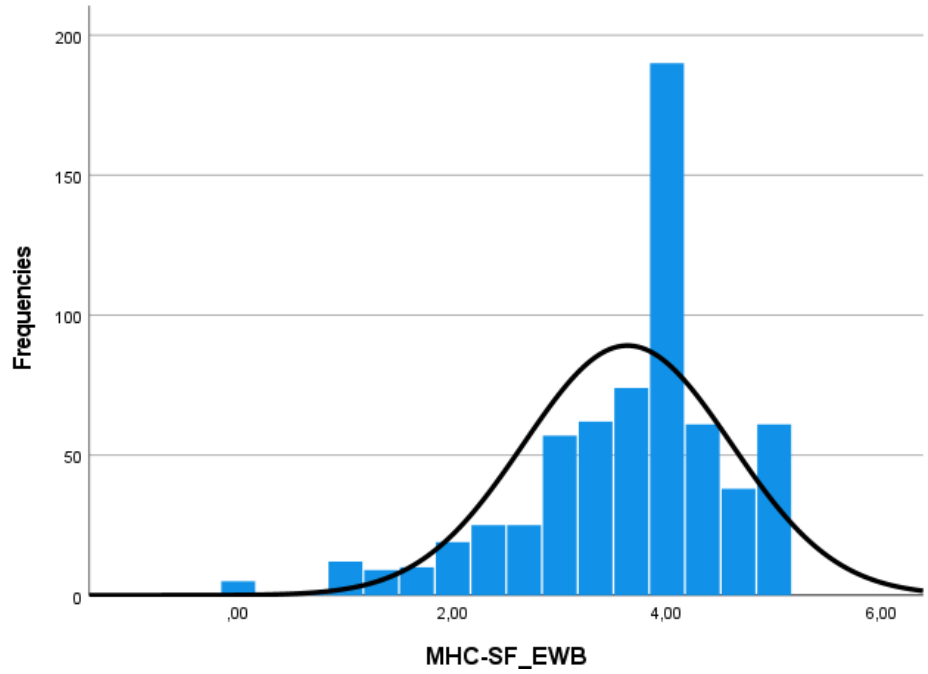


Figure 5

Histogram of the Emotional Well-Being Subscale of the MHC-SF-R

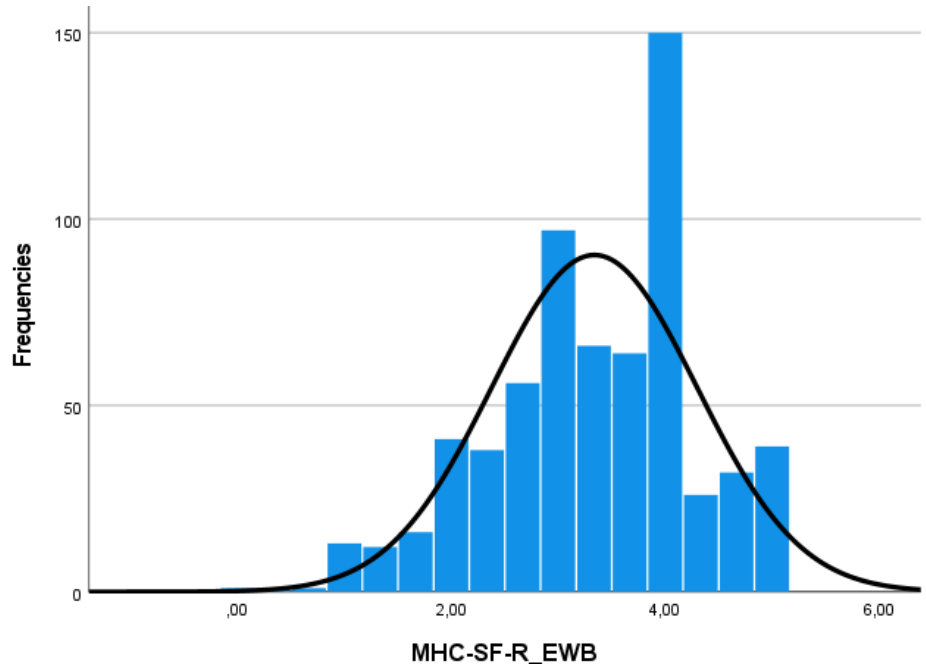


Table 5

Skewness and Kurtosis

Items/Scales	MHC-SF		MHC-SF-R	
	Skewness	Kurtosis	Skewness	Kurtosis
Item 1 (EWB)	-.92	.60	-.43	-.17
Item 2 (EWB)	**-1.18**	**1.32**	-.54	.24
Item 3 (EWB)	**-1.18**	**1.71**	-.5	.01
Item 4 (SWB)	-.84	-.96	.23	-.45
Item 5 (SWB)	.16	**-1.15**	.15	-.82
Item 6 (SWB)	.42	-.79	.33	-.18
Item 7 (SWB)	-.25	-.85	-.19	-.26
Item 8 (SWB)	-.34	-.75	-.2	-.11
Item 9 (PWB)	-.58	-.42	-.39	.11
Item 10 (PWB)	**-1.18**	**1.23**	-.68	.73
Item 11 (PWB)	-.89	.33	-.57	.73
Item 12 (PWB)	-.30	**-1.17**	0	-.65
Item 13 (PWB)	-.46	-.56	-.33	-.20
Item 14 (PWB)	-.56	-.65	-.24	-.45
Total scale	-.31	-.206	-.21	.29
EWB subscale	**-1.11**	**1.48**	-.44	-.11
SWB subscale	.09	-.51	.16	.20
PWB subscale	-.49	-.22	-.39	.45

Table 6:

Independent samples t-test

	MHC-SF		MHC-SF-R			
	M	SD	M	SD	t-test	Cohen's d
Total	2.95	0.88	2.85	0.75	t(1280.52) = 2.30 (p<.02)	0.12
EWB	3.63	0.97	3.34	0.95	t(1375) = 5.65 (p <.001)	0.31
SWB	2.33	1.06	2.37	0.84	t(1236.37) = -0.61 (p = .540)	-0.03
PWB	3.12	1.02	3.00	0.85	t(1270.73) = 2.28 (p = .023)	0.13

Floor and ceiling effects

Additionally, we examined the response distribution of the individual items, by analyzing the response frequencies for each item and scale of the MHC-SF and MHC-SF-R. To start with, we investigated the participants means achieved through the total scale and subscales. Of special interest was determining whether any floor or ceiling effects are present on a scale level. This would be the case if more than 15% of participants scored a mean of 0 (floor effect) or 5 (ceiling effect) for the total questionnaire or the three subscales. Hence, we conducted a frequency analysis and found that all percentages for the means 0 and 5 were below 15%. Please view table 7 for a representation of this data. We concluded that no floor or ceiling effects were present on a scale level.

Table 7

Percentage of highest and lowest answers achieved by participants on a scale level

Scales	MHC-SF		MHC-SF-R	
	mean = 0	mean = 5	mean = 0	mean = 5
Total scale	0	0.9	0.2	0.5
EWB subscale	0.8	9.4	0.2	6.0
SWB subscale	0.9	0.9	0.3	0.6
PWB subscale	0.2	2.3	0.6	0.8

Moreover, by determining the percentage of participants that indicated the lowest (0) or highest (5) answer option on the response scale, we identified the items where the percentage of participants indicating the highest or lowest answer option lies above 15%. In table 8, the percentages of participants choosing answer option 0 or 5 are shown for both versions of the questionnaire. We have determined that 3 floor effects (items 5, 6 and 12) and 5 ceiling effects (items 2, 3, 10, 11 and 14) are present in the data collected from the MHC-SF. In contrast, only one ceiling effect is present in the MHC-SF-R (item 2) and no floor effects could be detected.

In order to determine whether the observed differences between the two questionnaire versions are significant, we created cross-tables comparing the answer option indicative of a floor or ceiling effect with every other possible answer option. This allowed us to determine the differences between the two questionnaire versions through the use of a chi-square test for each item. Please view table 8 for an overview of the results.

In regards to the floor effects, we determined that the two questionnaire versions do not differ much. While the MHC-SF-R does not display any floor effects, they are present for items 5, 6 and 12 from the MHC-SF. Despite the decrease in the percentage of participants choosing the lowest answer option in the MHC-SF-R, no statistically significant differences between the two versions of the questionnaire could be found for the items with a floor effect present. However, it is of note that the p-value obtained from the chi-square analysis of items 5 and 12 is close to .05, thus just barely being considered insignificant.

In comparison to this, we observed significant differences between the two versions when it came to the ceiling effects. Five ceiling effects, in items 2, 3, 10, 11 and 14, were observed in the data collected with the MHC-SF. While the percentage of participants choosing the highest answer option decreased when using the MHC-SF-R, one ceiling effect for item 2 remained present. However, through the chi-square analysis, we could determine that this reduction represents a statistically significant difference between the two questionnaire versions. In fact, this can also be observed for all other items where a ceiling effect is present in the data of the MHC-SF. For these items too, a statistically significant reduction in the number of participants choosing the highest answer option was found.

Table 8

Percentage of highest or lowest answers chosen and Chi-Square Test

	MHC-SF		MHC-SF-R		$\chi 2$ (df), p	
	0 = not at all	5 = every day	0 = not at all	5 = always	Floor Effects	Ceiling Effects
Item 1 (EWB)	2.0	13.1	1.4	8.9	$\chi 2$ (1) = 1.70, p = .19	$\chi 2$ (1) = 4.62, p = **.03**
Item 2 (EWB)	1.5	**27.4**	0.7	**16.4**	$\chi 2$ (1) = .21, p = .64	$\chi 2$ (1) = 12.87, p = **<.001**
Item 3 (EWB)	1.2	**16.9**	0.6	11.7	$\chi 2$ (1) =.19, p = .67	$\chi 2$ (1) = 4.15, p = **.04**
Item 4 (SWB)	13.1	6.3	6.7	2.3	$\chi 2$ (1) = 2.47, p = .12	$\chi 2$ (1) = 8.78, p = **.003**
Item 5 (SWB)	**18.5**	8.0	11.0	3.0	$\chi 2$ (1) = 3.65, p = .06	$\chi 2$ (1) = 12.30, p = **<.001**
Item 6 (SWB)	**23.1**	2.5	5.9	1.2	$\chi 2$ (1) = .07, p = .80	$\chi 2$ (1) = 3.35, p = .07
Item 7 (SWB)	7.4	6.3	2.6	4.8	$\chi 2$ (1) = 14.15, p = **<.001**	$\chi 2$ (1) = 1.22, p = .27
Item 8 (SWB)	8.5	8.6	2.8	5.5	$\chi 2$ (1) = 2.86, p = .09	$\chi 2$ (1) = 2.73, p = .10
Item 9 (PWB)	4.5	10.2	1.5	5.8	$\chi 2$ (1) = 4.19, p = **.04**	$\chi 2$ (1) = 6.83, p = **.009**
Item 10 (PWB)	1.7	**21.4**	1.2	13.6	$\chi 2$ (1) = 2.72, p = .10	$\chi 2$ (1) = 6.56, p = **.01**
Item 11 (PWB)	2.6	**24.9**	1.5	12.0	$\chi 2$ (1) = .28, p = .60	$\chi 2$ (1) = 29.63, p = **<.001**
Item 12 (PWB)	**16.0**	7.4	6.5	2.9	$\chi 2$ (1) = 3.77, p = .05	$\chi 2$ (1) = 6.87, p = **.009**
Item 13 (PWB)	4.5	8.8	2.8	5.6	$\chi 2$ (1) = 2.21, p = .14	$\chi 2$ (1) = 2.32, p = .13
Item 14 (PWB)	6.0	**16.0**	2.3	10.0	$\chi 2$ (1) = .63, p = .43	$\chi 2$ (1) = 7.66, p = **.006**

Differences between the MHC-SF and MHC-SF-R

Through the analysis of the response distribution and floor and ceiling effects, we determined that the new answer format leads to a difference in participant responses.

The analysis of response distribution has shown that the data collected with the revised MHC-SF-R largely features a normal distribution. In contrast, in the data collected using the MHC-SF, 5 of the 14 items were not normally distributed with three items being skewed and two items being too peaky or too flat.

Furthermore, an improvement in the distribution of responses and a reduction in floor and ceiling effects was achieved through the use of the new response format. All floor and ceiling effects identified in the data from the MHC-SF were eliminated as the number of responses for the highest or lowest response option were reduced to under 15%. The only exception poses item 2, from the emotional well-being subscale, where a ceiling effect is present in both versions of the questionnaire. Despite this, we determined that there is a significant difference between the questionnaire versions when comparing the floor and ceiling effects with a chi-square test. Overall, the analysis suggests that the revised response format causes an improvement in the distribution of responses, leading to a more normally distributed data set.

Discussion

The current study investigated the psychometric properties of two different versions of the Mental Health Continuum questionnaire. Specifically, we compared the Mental Health Continuum - Short Form (MHC-SF) with its revised version (MHC-SF-R). The differences between these two versions will be discussed below, under the header *Revision of Response Format*.

By comparing their psychometric properties with each other, we determined that both versions of the Mental Health Continuum questionnaire have excellent reliability on a total and subscale level. However, when it comes to the factor structure, the original version of the

questionnaire (MHC-SF) appears to be superior to the MHC-SF-R. When looking for a clear representation of each subscale in the factor structure, this is only the case for the MHC-SF. In contrast, when considering the way that participants responded to the questionnaire, it appears that the MHC-SF-R is superior as the items are all normally distributed, without many floor and ceiling effects.

Examination of the Factor Structure

First, we examined the factor structure that arose from the two different versions of the questionnaire. In line with its theoretical background (Iassiello et al., 2022; Keyes, 2005), a three-component solution was chosen as the most suitable for the data of both the MHC-SF and MHC-SF-R. When examining the results from the MHC-SF, the factor structure is congruent with this three component model. In contrast, this is not the case for the MHC-SF-R where the three components do not correspond with the theoretical background.

Based on our findings, we have answered the first research question: What factor structure arises from the MHC-SF-R and how does it compare to that of the MHC-SF? The original MHC-SF appears to better represent the three subscales, social well-being, emotional well-being and psychological well-being as each factor loads on a different component. In contrast, the MHC-SF-R does not show the same desirable factor structure. As such, the question arises whether questionnaires items still sufficiently measure the three underlying types of well-being when applying the new response format to the questions.

As the most appropriate fit model of the MHC-SF varies a lot between different studies assessing the psychometric properties of the MHC-SF, review studies have been conducted (Iasiello et al., 2022; Schutte & Wissing, 2017). As already mentioned in the introduction, a single-factor model, two-factor model and three-factor model have been proposed, among others (Iasiello et al., 2022). While Iasiello et al. (2022) concluded that a two-factor solution would achieve the best fit from a purely statistical standpoint, they suggest that a three-factor solution would be a more appropriate model. This was based on

psychometric limitations of the two-factor solution, allowing for an explanation of implausible and unsuitable model due to its flexibility (Reise et al., 2014; van Zyl & Olckers, 2019). This makes it inappropriate for most data sets, so too the current one. In order to best explain the data on the basis of the theoretical background, a three-factor solution was deemed to be the most appropriate (Iassiello et al., 2022).

Many previous studies already concluded that a three-factor solution represents the best fit for their data (Guo et al., 2015; Karaś et al., 2014; Keyes et al., 2008; Lamers et al., 2011; Lim, 2013; Petrillo et al., 2015; Schutte & Wissing, 2017). However, the obtained fit indexes suggest only a marginally acceptable fit, further underlining the difficulties in finding an appropriate factor solution for the data of the MHC-SF.

In line with this earlier research, the data examined in the current study did not unequivocally lend itself to a three-factor solution. While the factor analysis of the MHC-SF shows a good alignment of the three components with the subscales, this is not the case for the data obtained with the MHC-SF-R. Rather than having each subscale represented by one component, it appears that the emotional and psychological well-being subscales are more closely related to each other than in previous versions of the questionnaire, as one factor is made up predominantly by the questions from these subscales. Additionally, a variety of cross-loadings are present in the psychological well-being subscale. Frequently, items load on component one and three. Based on the theoretical background of the Mental Health Continuum, it is unclear why these factor loadings were found.

Since the two versions of the questionnaire only differ on the response format, it is likely that this is related to the changes in the factor structure. One possibility may be the use of a shorter reference period when collecting the data of the MHC-SF-R. Rather than using a one month time frame, the revised measurement uses a one week reference period. This may not have been a sufficient period of time for the experiences described in the questionnaire to occur regularly ("reference period in questions about past events," 2017). In this case, it is

likely that the factor structure of the MHC-SF-R does not align well with the theoretical background of the Mental Health Continuum, as the revised questionnaire may have misrepresented mental well-being and its three sub-types.

Overall, it may be possible that the revision of the answer format further exacerbated the existing difficulties with determining an appropriate factor solution. As the factor structure is an important component in a questionnaires validity, the high number of cross-loading items may reflect issues with the discriminant validity of the MHC-SF-R. As the items load on two different component, they may not be able to discriminate adequately between the different components of mental well-being. Thus, one may question whether the revised questionnaire is sufficiently able to represent the different attributes and facets of mental well-being after the revision. Answering this question has a direct impact on the construct validity of the MHC-SF-R and whether it can be considered an appropriate measure of mental well-being. As an examination of validity was excluded from this research, it remains unclear whether the MHC-SF-R is able to accurately measure mental well-being.

Examination of Internal Consistency

To continue, we analyzed the internal consistency of the two questionnaire versions in order to answer the second research question: What is the internal consistency of the MHC-SF-R and how does it compare to the internal consistency of the MHC-SF? We determined that the revised questionnaire has excellent reliability on a total scale and subscale level. Since the MHC-SF has good reliability as well, we can conclude that the reliability did not differ between questionnaires. This indicates that both the MHC-SF and MHC-SF-R are precise measurement instruments, enabling similar participants to score similar results. Therefore, the results of both questionnaire versions should be stable and can be considered replicable.

One reason as to why the internal consistency reliability does not change between the questionnaire versions may be that the obtained alpha value for the MHC-SF is already excellent. Thus, there is not much room for improvement of Cronbach's alpha without losing

the predictive value of the questions. If alpha were to further increase to 1, this would indicate that the items would be answered identically (Green et al., 1977). Based on the way that the items were created, it should be the case that each question represents a unique aspect of mental well-being (Keyes et al., 2008). Therefore, it follows that each item adds a unique aspect to the questionnaire, causing the alpha values to remain slightly below 1.

Additionally, our findings are in line with previous research on the Mental Health Continuum questionnaire (Keyes, 1998; Keyes, 2002; Lamers et al., 2010). When considering the stages of revision that the Mental Health Continuum has undergone in the past, it appears that the revision process did not negatively impact the questionnaire's internal consistency. Throughout all stages of the revision process, from the original Mental Health Continuum (Keyes, 1998; Keyes, 2002) to the MHC-SF (Lamers et al., 2010) to the current revision creating the MHC-SF-R, an internal consistency of around .8 has been retained for the entire scale. On a subscale level, the different versions achieved a moderate-to-excellent internal consistency (Keyes, 2002; Lamers et al., 2010). Based on the first revision, in which the items and answer format were changed, as well as previous studies reviewing the psychometric properties of other revised questionnaires (Conner et al., 2018; Dworkin et al., 2009; Kriston et al., 2010; Wagner et al., 2020), it was concluded that the internal consistency reliability is frequently retained in the revision process. Thus, the results from our analysis of internal consistency align well with the body of research on this topic.

Examination of Response Format

Lastly, we answered the third research question: Do participants answer questions differently depending on the response format they are presented with? To start with, differences can be found in the distribution of responses. While in the revised questionnaire all items were normally distributed, the original questionnaire includes five items that are not normally distributed, with differences in either skewness or kurtosis. However, on a total and subscale level, only minute differences could be found. In case of statistically significant

differences between the questionnaire versions, the effect sizes were rather small, indicating a benign interaction. As such, we can conclude that the two versions of the questionnaire are similar to each other, only differing slightly. We therefore consider both questionnaire versions to be equal in regards to their response distribution of the items.

Additionally, we determined the presence of floor and ceiling effects. The data obtained from the MHC-SF shows three floor effects and five ceiling effects. In comparison, the data of the MHC-SF-R only showed one ceiling effect for item 2 and no floor effects.

It appears that by changing the response format, participants opted to choose the lowest or highest response option significantly fewer times compared to the original response format. As the new response format features vague quantifiers, rather than approximate ranges, the need to estimate the frequency in which the emotion occurred is eliminated. This allows participants to choose an answer option that feels appropriate to them, without it needing to occur a certain amount of time. As a result, participants may have chosen the answer options located in the middle of the range more frequently, improving the distribution of responses and reducing the occurrence of floor and ceiling effects.

The presence of floor and ceiling effects can impede the researchers ability to distinguish between participants scoring at the extreme end of the scale (Murugappan et al., 2022; Terwee et al., 2006), as their presence suggests that participants would have liked to select a higher or lower score than the bounds of the provided answer scale allows for. Due to the limitations of the scale, they probably opted to choose the highest or lowest score instead (Terwee et al., 2006). Hence, we can assume that a reduced occurrence of floor and ceiling effects indicates a better representation of the population by the questionnaire. This is particularly crucial when using the Mental Health Continuum to gain insights into the mental well-being of a specific population, as it would allow for the data to be generalized to people outside the studied sample.

Overall, we can conclude that participants seem to have answered differently depending on what version of the questionnaire they were presented with. It appears that the MHC-SF-R represents an improvement upon the MHC-SF as a normal distribution of items is achieved and only one ceiling effect is present. This suggests that the revised scale is perceived as sufficient in order to answer the questions adequately (Terwee et al., 2006).

Revision of Response Format

In the current study, we evaluated a revision of the MHC-SF questionnaire. As changes in the response format were made, it has become necessary to evaluate the effect these changes have had on the psychometric properties of the questionnaire and the response behavior of the participants.

The revision consisted of an adaptation of the response format. When using the MHC-SF, participants were asked to estimate the frequency in which certain emotions occurred within the past month. Many participants experienced this as difficult since they could not easily approximate the frequency of their emotions. Thus, a new response format was created relying on vague quantifiers. Additionally, the reference period was shortened from one month to one week. Compared to the original response format, the revised answer options were more vague. This was done to improve the user experience by reducing the difficulty connected with remembering exact frequencies. However, to this day no qualitative study has been performed to evaluate the subjective success of the review from the participants side.

When looking at the two response formats from a theoretical perspective, both types of close-ended response categories have their merits. When using a response format with approximate ranges, such as in the MHC-SF, this allows the participants to indicate estimates rather than specific frequencies. The use of this questionnaire format is most beneficial when investigating an emotion or behavior that presents itself regularly for the participants (Geisen, 2020). However, when this is not the case, or the frequency of the emotion or behavior changes throughout the reference period, the use of a response format utilizing vague

quantifiers is recommended (Geisen, 2020). By giving participants vague answer options, such as *often* or *sometimes*, it allows them to answer the question, even if the exact frequency is unknown or hard to quantify (Geisen, 2020). The downside to this response format is it's subjectivity. As different participants may have a different definition for what vague terms such as *often* entail, they may answer differently despite experiencing the emotion or behavior in the same frequency (Geisen, 2020).

When applying this to the questions asked by the Mental Health Continuum questionnaire, it may be said that some of the items are hard to quantify. One example might be the question: In the past month, how often have you felt that our society becomes better for people? It could be argued that few people, hold on during their everyday life to think about society changing for the better. While this feeling may be present for many people, it could nevertheless be difficult to quantify. As such, the new response options may be considered as more appropriate for the questions at hand. Additionally, the new format may improve the usability of the questionnaire as it eliminates the need to estimate approximate frequency ranges for the emotions.

However, it could also be argued that the reference period used by the MHC-SF is more appropriate for the questions posed by the measurement instrument. As some of the items may be hard to quantify, awarding the participants a longer timeframe to consider may allow for a greater number of such experiences to occur within the reference period. As such, it would enable the collection of more representative data. Therefore, it is difficult to draw a definitive conclusion on which answer format is superior on a theoretical basis alone.

Limitations

One of the biggest limitations of the current research study is that only the internal consistency was assessed to draw conclusions about the measurements' reliability. Other types of reliability, such as the test-retest reliability, were not examined. Inclusion of a follow-up measure would have provided insight into the stability of the data obtained from the

different versions of the questionnaire over time. However, due to the limitations of the collected data, this was not possible.

Additionally, apart from the factor structure, we largely excluded validity from our examination. This would have been insightful in order to determine whether the questionnaire actually measures well-being. Nonetheless, we opted not to include this in the research as it did not appear necessary to evaluate the validity of the questionnaire. Previous research had already established that the MHC-SF has good validity (Lamers et al., 2011). As the MHC-SF-R uses the same exact questions, we assumed that the validity of the questionnaire would not change due to the adoption of a new answer format. However, the difficulties we experienced with interpreting the factor structure of the MHC-SF-R put this into question. With a factor structure that does not reflect the theoretical background of the questionnaire and a variety of cross-loading items, the questionnaires discriminant validity may not be acceptable. This may hint at larger issues with the construct validity of the MHC-SF-R. Yet, as of right now, we are not aware of a research study that has assessed the construct validity of the MHC-SF-R.

Another shortcoming of the current study is that the revision of the answer scale took place based on the judgement of the researchers. After participant feedback was heard, indicating that answering the questions was difficult, the researchers devised a new response format in order to make the questionnaire more accessible. Nevertheless, it is unclear whether participants perceive these changes as effective for increasing comprehension of the questionnaire.

Lastly, the data analyzed in this study represents a cross-section of the Dutch population. However, the Mental Health Continuum is also used on other populations within and outside the Netherlands, such as patients with mental health problems. It is imperative that researchers employing the questionnaire outside of the studied population proceed with caution as it is unclear whether results can be generalized to other cultures and populations.

Recommendations for future research

As previously stated, we did not include all the collected data in this examination. Specifically, responses were collected from (1) the MHC-SF, made up of the original questions with the original response format, (2) the MHC-SF-R, made up of the original questions with the revised response format, (3) another questionnaire made up of revised questions with the original response format and (4) another questionnaire made up of revised questions with the revised response format. In the current study, we explored the psychometric properties of conditions one and two. Simultaneously, the psychometric properties of the other conditions have previously been and are currently assessed in parallel studies to this one. However, as of right now, the results from the different analyses have not yet been integrated in a single study. As such, we recommend this integration to be conducted in the future, creating a holistic overview of the different versions of the Mental Health Continuum questionnaire. Specifically, it would allow a conclusion to be drawn on which version of the questionnaire holds the most advantage for use in scientific practice.

Furthermore, we strongly recommend an assessment of the MHC-SF-Rs construct validity. As described above, the factor analysis illuminated an array of cross-loadings and an overall factor structure not in line with the theoretical background of the questionnaire. This raises the question whether the construct validity has been reduced through the review process. It would be prudent to evaluate this before using the MHC-SF-R in scientific or clinical practice.

Additionally, future studies should focus on collecting and analyzing samples from different populations. As of right now, the Mental Health Continuum is widely used not only within psychological research, but also clinical practice (Jankowski et al., 2022). It would be insightful to collect data from samples of patients with different psychological conditions such as depression, anxiety or personality disorders, since these clinical samples may interpret or answer the questions differently to the average population assessed within this study.

In this vein, the current study was done on a sample of households within the Netherlands. However, the Mental Health Continuum is also employed in other countries worldwide. While the MHC-SF has already been validated for a number of cultures (Fonte et al., 2020; Góngora & Solano, 2017; Keyes et al. 2008), this should also be undertaken for the MHC-SF-R. Especially, when it comes to countries outside of Europe that might have a different socio-economic structure or value system, it would be insightful to study the reliability and validity of the MHC-SF-R within this context.

Lastly, as already mentioned in the limitations, it remains unclear whether the changes made in the revision process actually accomplished the goal of improving the usability of the questionnaire. To assess this a qualitative study should be conducted in order to gain inside into the participants experiences when filling in the questionnaire. In this way it could be ensured that the MHC-SF-R indeed has better usability compared with the MHC-SF.

Conclusion and Practical Implications

While it is hard to draw a definitive conclusion on which version of the questionnaire is superior. At face value, it appears that the MHC-SF-R has some distinct advantages that qualify it for a preferred use in well-being research and clinical practice. Particularly, it stands out due to its improved usability as compared to the MHC-SF. As the answer format was changed in order to improve the usability, it has become a more convenient tool to employ in practice. Additionally, this new answer format appears to have benefited the collected data in terms of its representativeness of the population. However, as of right now we want to caution against the use of the MHC-SF-R as it remains unclear whether the questionnaire is an appropriate measure of mental well-being. Based on the obtained factor structure, we have raised concerns about the discriminant validity of the questionnaire. Until a thorough examination of construct validity has been conducted, we believe it prudent to only opt to use the MHC-SF-R in specific circumstances where the improved usability outweighs the potential risks associated with employing a measure that has not been sufficiently scrutinized.

References

A reference period in questions about past events. (2017, November 5). researchscout. https://www.researchscout.co.uk/single-post/2017/11/05/a-reference-period-in-questions-about-past-events

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Conner, B. T., Rahm-Knigge, R. L., & Jenkins, A. L. (2018). Revision and clarification of the sensitivity to punishment sensitivity to reward questionnaire. *Personality and Individual Differences*, *121*, 31–40. https://doi.org/10.1016/j.paid.2017.09.016

Dworkin, R. H., Turk, D. C., Revicki, D. A., Harding, G., Coyne, K. S., Peirce-Sandner, S., Bhagwat, D., Everton, D., Burke, L. B., Cowan, P., Farrar, J. T., Hertz, S., Max, M. B., Rappaport, B. A., & Melzack, R. (2009). Development and initial validation of an expanded and revised version of the short-form mcgill pain questionnaire (sf-mpq-2). *Pain*, *144*(1-2), 35–42. https://doi.org/10.1016/j.pain.2009.02.007

Fonte, C., Silva, I., Vilhena, E., & Keyes, C. L. (2020). The Portuguese adaptation of the mental health continuum-short form for adult population. *Community mental health journal*, *56*(2), 368-375. https://doi.org/10.1007/s10597-019-00484-8

Frisbie, D. A., & Brandenburg, D. C. (1979). Equivalence of questionnaire items with varying response formats. *Journal of Educational Measurement, 16*(1), 43–48. https://doi.org/10.1111/j.1745-3984.1979.tb00085.x

Geisen, E. (2020, June 4). *Tips for measuring behavioral frequency*. Qualtrics. https://www.qualtrics.com/blog/measuring-behavioral-frequency/

Góngora, V. C., & Solano, A. C. (2017). Assessment of the Mental Health Continuum in a sample of Argentinean Adults. *Psychology*, *8*(03), 303. https://doi.org/10.4236/psych.2017.83018

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient Alpha as an index of test Unidimensionality1. *Educational and Psychological Measurement*, *37*(4), 827-838. https://doi.org/10.1177/001316447703700403

Guo, C., Tomson, G., Guo, J., Li, X., Keller, C., & Söderqvist, F. (2015). Psychometric evaluation of the Mental Health Continuum-Short Form (MHC-SF) in Chinese adolescents – a methodological study. *Health and Quality of Life Outcomes, 13*(1). https://doi.org/10.1186/s12955-015-0394-2

IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp

Iasiello, M., Van Agteren, J., Schotanus-Dijkstra, M., Lo, L., Fassnacht, D. B., & Westerhof, G. J. (2022). Assessing mental wellbeing using the mental health continuum—Short form: A systematic review and meta-analytic structural equation modelling. *Clinical Psychology: Science and Practice*, *29*(4), 442-456. https://doi.org/10.1037/cps0000074

Jankowski, P. J., Sandage, S. J., & Crabtree, S. A. (2022). The psychometric challenges of implementing wellbeing assessment into clinical research and practice: A commentary on "assessing mental wellbeing using the mental health continuum–short form: A systematic review and meta-analytic structural equation modeling". *Clinical Psychology: Science and Practice, 29*(4), 457–460. https://doi.org/10.1037/cps0000090

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement, 20*(1), 141-151. https://doi.org/10.1177/001316446002000116

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. https://doi.org/10.1016/j.paid.2014.05.011

Keyes, C. (1998). Social well-being*. *Social Psychology Quarterly*, *61*(2), 121–140. https://doi.org/10.2307/2787065

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behaviour, 43*(2), 207–222. https://doi.org/10.2307/3090197

Keyes, C. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology, 73*(3), 539–548. https://doi.org/10.1037/0022-006x.73.3.539

Keyes, C. L., 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 & psychotherapy, 15*(3), 181-192. https://doi.org/10.1002/cpp.572

Kriston, L., Scholl, I., Hölzel, L., Simon, D., Loh, A., & Härter, M. (2010). The 9-item shared decision making questionnaire (sdm-q-9). development and psychometric properties in a primary care sample. *Patient Education and Counseling*, *80*(1), 94–99. https://doi.org/10.1016/j.pec.2009.09.034

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

Lim, Y.-J. (2013). Psychometric Characteristics of the Korean Mental Health Continuum–Short Form in an Adolescent Sample. *Journal of Psychoeducational Assessment, 32*(4), 356–364. https://doi.org/10.1177/0734282913511431

Luciano, J. V., Sanabria-Mazo, J. P., Feliu-Soler, A., & Forero, C. G. (2020). The pros and cons of bifactor models for testing dimensionality and psychopathological models: A commentary on the manuscript' A systematic review and meta-analytic factor analysis

of the depression anxiety stress scales'. Clinical Psychology: Science and Practice. Advance online publication. https://doi.org/10.1111/cpsp.12386

Martin, E. (2006). Survey questionnaire construction. *Survey methodology*, *13*, 1-13. https://doi.org/10.1016/b0-12-369398-5/00433-3

McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation - Official Journal of the International Society of Quality of Life Res*, *4*(4), 293–307. https://doi.org/10.1007/BF01593882

Mechanic, D. (1999). Mental health and mental illness: definitions and perspectives. *Mental health: Strengthening our response*. (2022, June 17). WHO | World Health Organization. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

Murugappan, M. N., King-Kallimanis, B. L., Mangir, C., Howie, L., Bhatnagar, V., Beaver, J. A., Bach, E. M., Henson, S. R., & Kluetz, P. G. (2022). Floor and ceiling effects in the EORTC QLQ-C30 Physical Functioning subscale among patients with advanced or metastatic breast cancer. *Cancer*, *128*(4), 808-818. https://doi.org/10.1002/cncr.33959

Petrillo, G., Capone, V., Caso, D., & Keyes, C. L. (2015). The Mental Health Continuum–Short Form (MHC–SF) as a measure of well-being in the Italian context. *Social indicators research, 121*(1), 291-312. https://doi.org/10.1007/s11205-014-0629-3

Reise, S. P., Cook, K. F., & Moore, T. M. (2014). Evaluating the impact of multidimensionality on unidimensional item response theory model parameters. In S. P. Reise & D. A. Revicki (Eds.), Handbook of item response theory modeling: Applications to typical performance assessment (pp. 13–40). Routledge/Taylor & Francis Group. https://doi.org/10 .4324/9781315736013-11

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*(1), 141-166. https://doi.org/10.1146/annurev.psych.52.1.141

Ryff, C. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069-1081. https://doi.org/10.1037/0022-3514.57.6.1069

Schutte, L., & Wissing, M. P. (2017). Clarifying the factor structure of the mental health continuum short form in three languages: a bifactor exploratory structural equation modeling approach. *Society and Mental Health*, *7*(3), 142–158. https://doi.org/10.1177/2156869317707793

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31*(12), 1428–1441. https://doi.org/10.1037/pas0000623

Simsek, O. (2011). An intentional model of emotional well-being: The development and initial validation of a measure of subjective well-being. *Journal of Happiness Studies*, *12*(3), 421–442. https://doi.org/10.1007/s10902-010-9203-0

Taylor, R. (1990). Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, *6*(1), 35-39. https://doi.org/10.1177/875647939000600106

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

van Zyl, L. E., & Olckers, C. (2019). The Mental Health Continuum-Short Form in organisational contexts factorial validity, invariance, and internal consistency. *European Journal of Mental Health, 14*(2), 230–259. https://doi.org/10.5708/EJMH.14.2019.2.2

Vindegaard, N., & Benros, M. E. (2020). Covid-19 pandemic and mental health consequences: systematic review of the current evidence. *Brain Behavior and Immunity*, *89*, 531–542. https://doi.org/10.1016/j.bbi.2020.05.048

Wagner, A., Nübling, M., Hammer, A., Manser, T., & Rieger, M. A. (2020). Comparing perceived psychosocial working conditions of nurses and physicians in two university hospitals in germany with other german professionals - feasibility of scale conversion between two versions of the german copenhagen psychosocial questionnaire (copsoq). *Journal of Occupational Medicine and Toxicology*, *15*(1). https://doi.org/10.1186/s12995-020-00277-w

Westerhof, G. J., & Keyes, C. L. (2010). Mental illness and mental health: The two continua model across the lifespan. *Journal of adult development, 17*(2), 110-119. https://doi.org/10.1007/s10804-009-9082-y

Appendix A

Study Materials

Questionnaire Items

Items in Dutch (used for data collection)

In de afgelopen maand, hoe vaak had u het gevoel...

1. ...dat u gelukkig was?
2. ...dat u geïnteresseerd was in het leven?
3. ...dat u tevreden was?
4. ...dat u iets belangrijks hebt bijgedragen aan de samenleving?
5. ...dat u deel uitmaakte van een gemeenschap (zoals een sociale groep, uw buurt, uw stad)?
6. ...dat onze samenleving beter wordt voor mensen?
7. ...dat mensen in principe goed zijn?
8. ...dat u begrijpt hoe onze maatschappij werkt?
9. ...dat u de meeste aspecten van uw persoonlijkheid graag mocht?
10. ...dat u goed kon omgaan met uw alledaagse verantwoordelijkheden?
11. ...dat u warme en vertrouwde relaties met anderen had?
12. ...dat u werd uitgedaagd om te groeien of een beter mens te worden?
13. ...dat u zelfverzekerd uw eigen ideeën en meningen gedacht en geuit hebt?
14. ...dat uw leven een richting of zin heeft?

Items in English

In the past month, how often have you felt…

1. ...that you were happy?
2. ...that you were interested in life?
3. ...that you were satisfied?
4. ...that you have contributed something important to society?

5. ...that you were part of a community (such as a social group, your neighborhood, your city)?
6. ...that our society becomes better for people?
7. ...that people are basically good?
8. ...that you understand how our society works?
9. ...that you liked most aspects of your personality?
10. ...that you coped well with your everyday responsibilities?
11. ...that you had warm and trusting relationships with others?
12. ...that you were challenged to grow or become a better person?
13. ...that you have confidently thought and expressed your own ideas and opinions?
14. ...that your life has a direction or meaning?

Response format MHC-SF

Response options in Dutch (used for data collection)

0. Nooit
1. Eén of twee keer per maand
2. Ongeveer 1 keer per week
3. 2 of 3 keer per week
4. Bijna elke dag
5. Elke dag

Response options in English

0. Not at all
1. Once or twice a month
2. About once a week
3. 2 or 3 times a week
4. Almost every day
5. Everyday

Response format MHC-SF-R

Response options in Dutch (used for data collection)

0. Nooit
1. Zelden
2. Soms
3. Regelmatig
4. Vaak
5. (Bijna) altijd

Response options in English

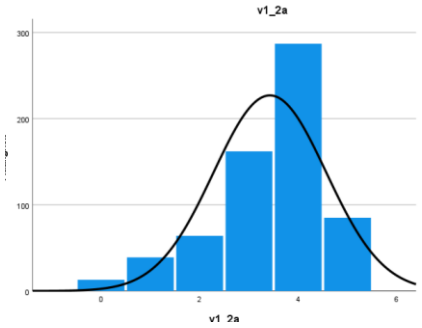
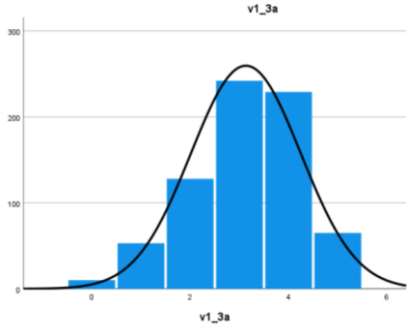
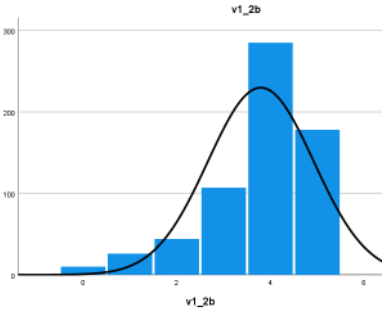
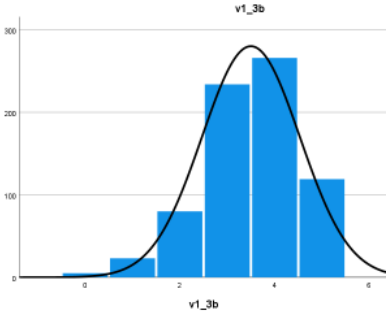
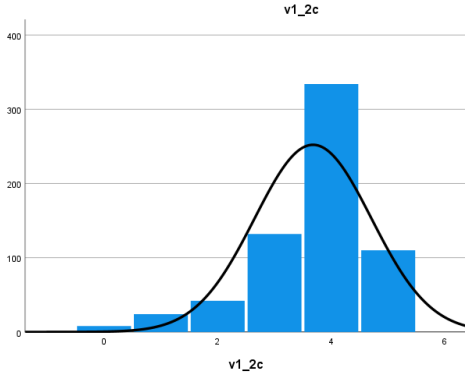
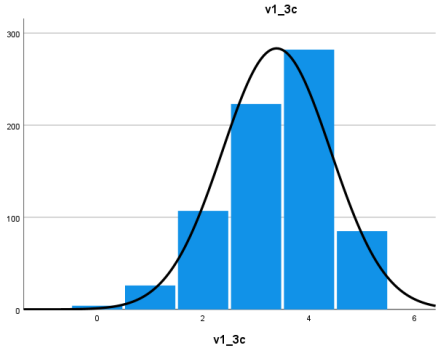
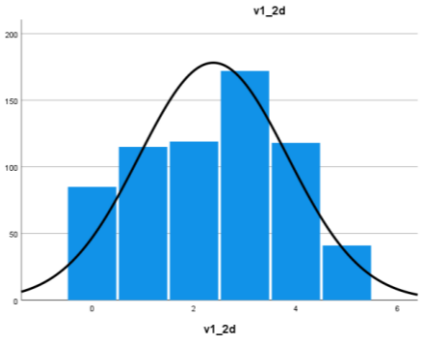
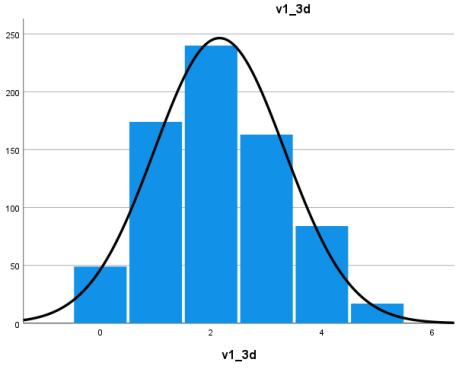
0. Not at all
1. Rarely
2. Sometimes
3. Regularly
4. Often
5. (Almost) all the time

Appendix B

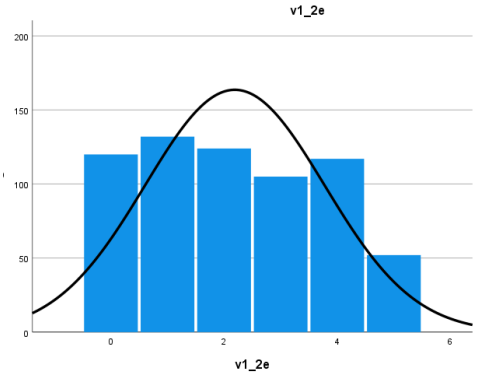
Distribution of Participant Responses

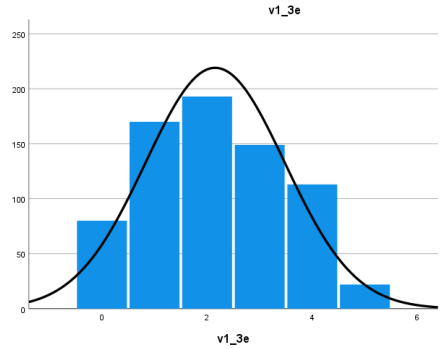
Table B1

Histogram of Participant Responses per Item

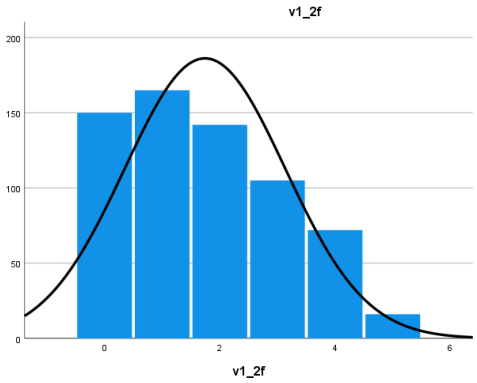
	MHC-SF	MHC-SF-R
Item 1		
Item 2		
Item 3		
Item 4		

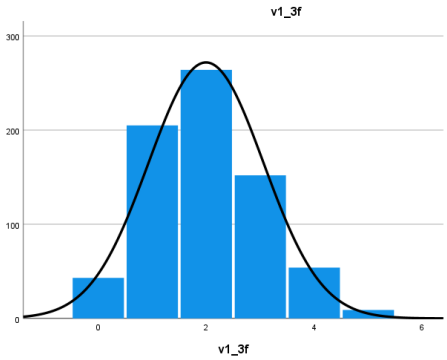
Item 5



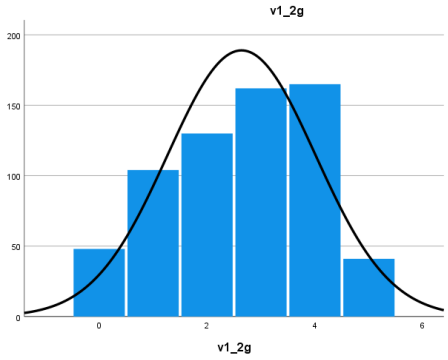


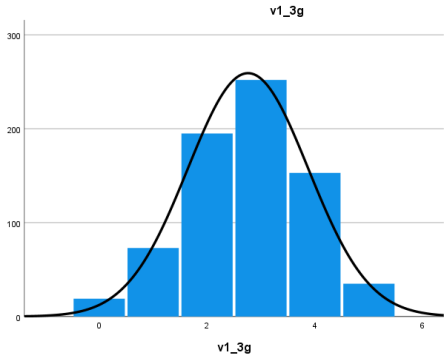
Item 6



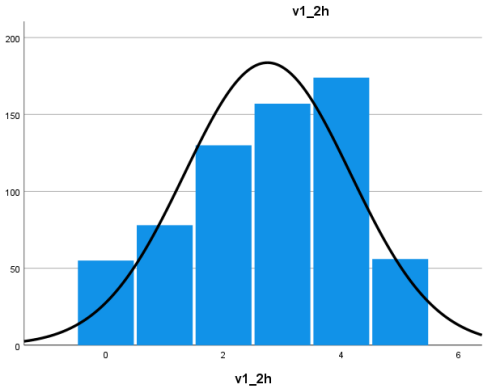


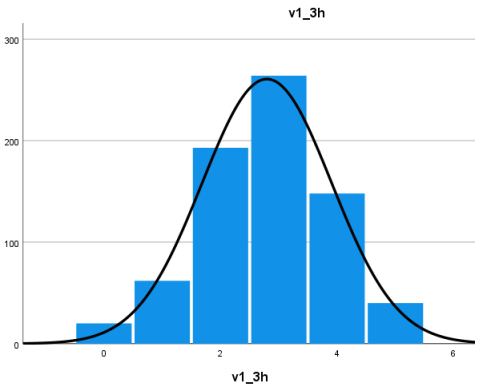
Item 7



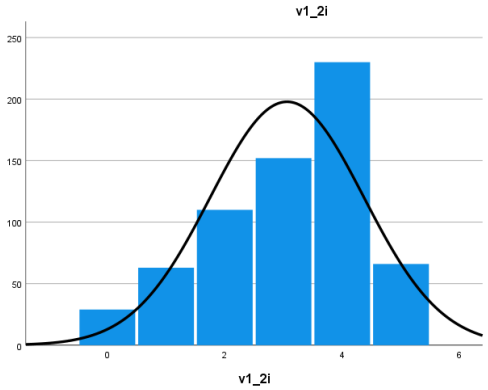


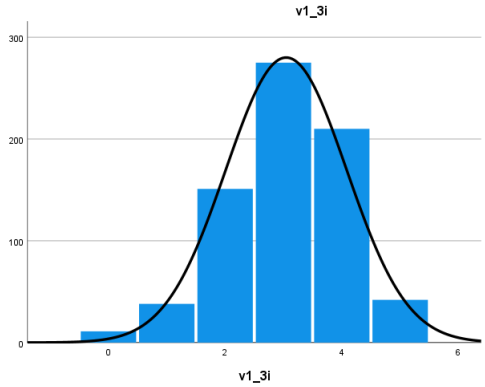
Item 8



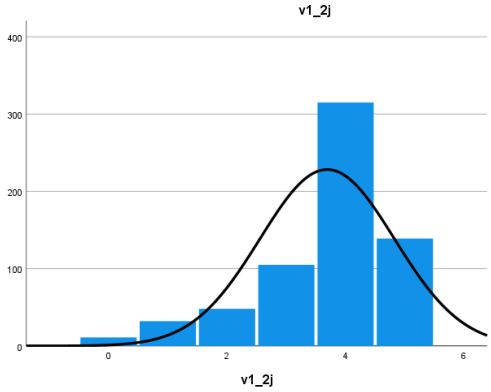


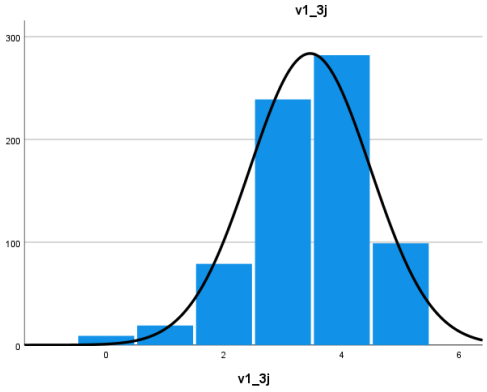
Item 9



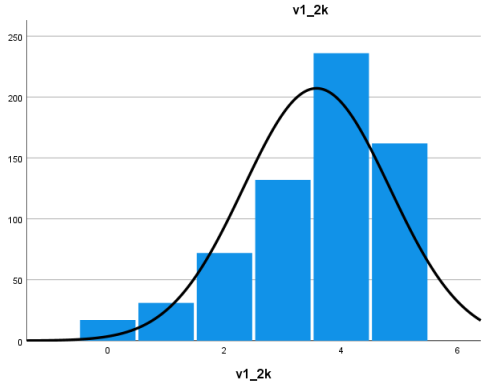


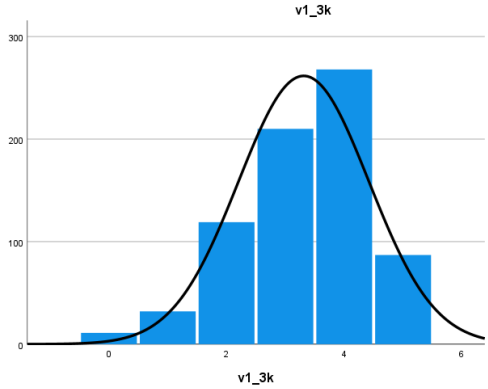
Item 10



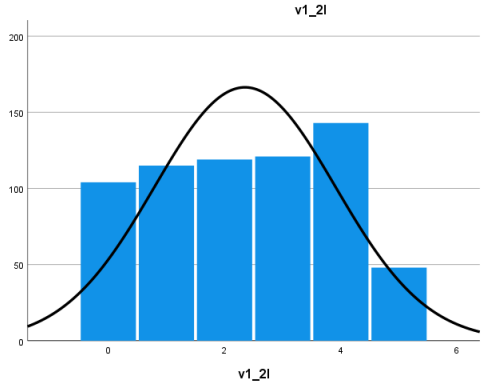


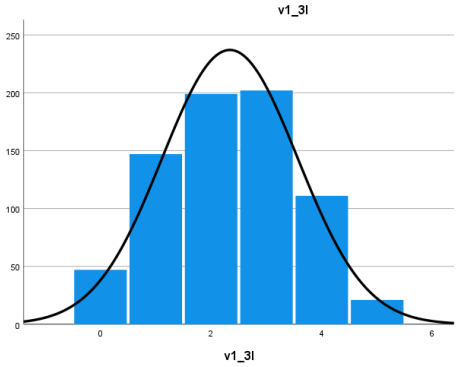
Item 11





Item 12





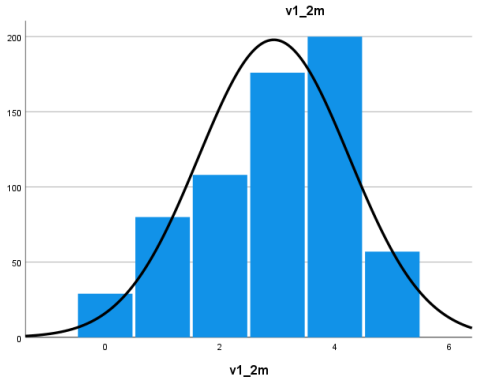
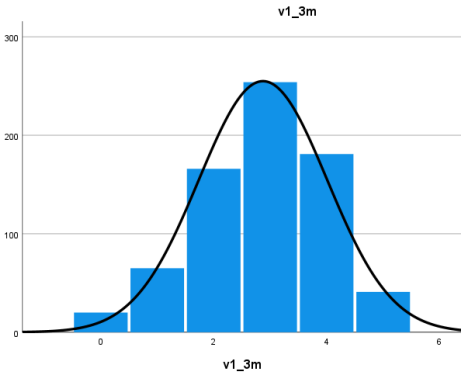
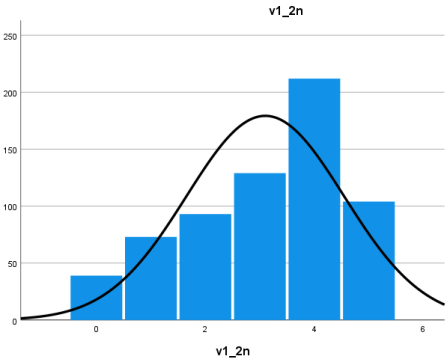
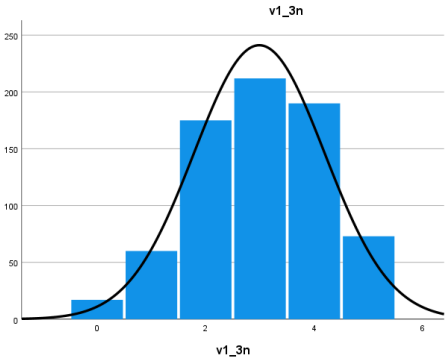
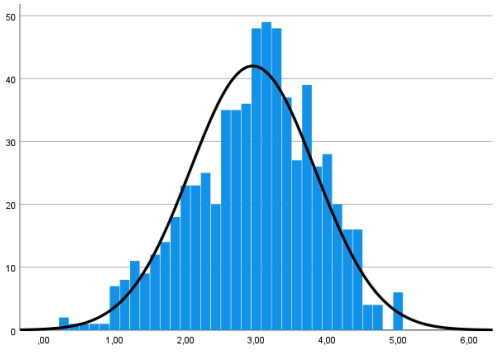
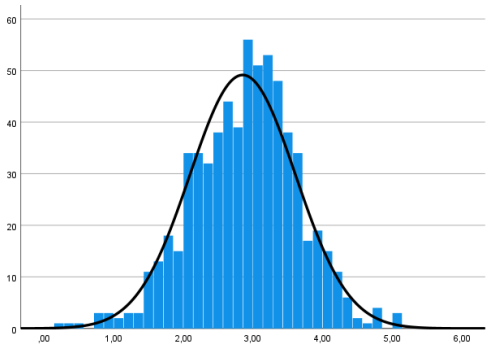
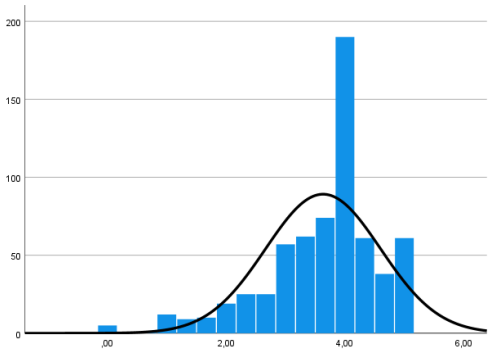
Item 13		
Item 14		

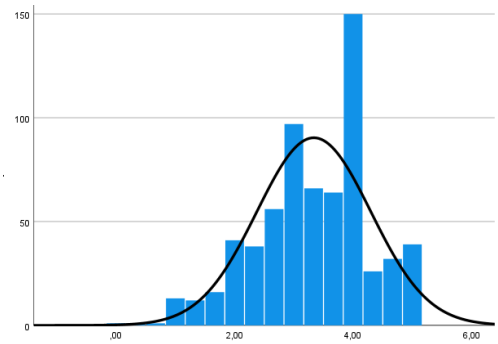
Table B2

Histogram of Participant Responses per Scale

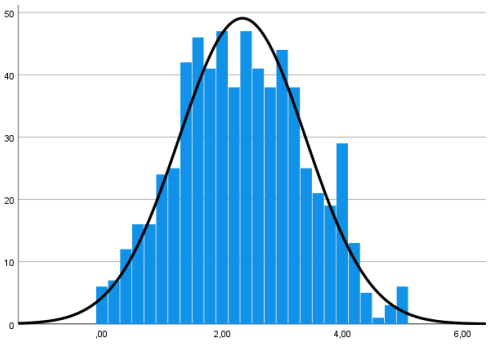
	MHC-SF	MHC-SF-R
Total Scale		

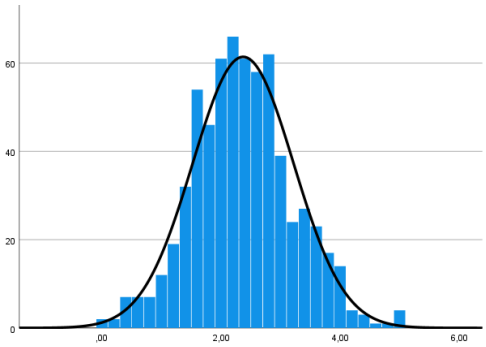
EWB subscale





SWB subscale





PWB subscale