Psychometric evaluation of the mental health continuum – short form in Swedish adolescents

Fredrik Söderqvist^{1,2} · Peter Larm³

Accepted: 10 March 2021 / Published online: 17 March 2021
© The Author(s) 2021

Abstract

The Mental Health Continuum – Short form (MHC-SF) is a self-report measure that has been increasingly used to monitor mental well-being at the population level. The aim of this study was to evaluate, for the first time, the psychometric properties of the MHC-SF in a population of Swedish adolescents. First, the evaluation was performed by examining face validity and test–retest reliability obtained in a pre-study. Then using data from the Survey of Adolescent Life in Vestmanland 2020 ($n = 3880$) we performed confirmatory factor analysis on different factor structures based on theory and previous research. Model-based estimates were calculated for assessing the internal reliability of the factor structure with the best fit. Convergent validity was assessed by bivariate as well as model-based correlations, and test–retest reliability was evaluated by intra-class correlation coefficients. The results show that the MHC-SF is best described with a bifactor model consisting of a dominant general well-being factor and three specific group factors of emotional, social and psychological well-being. Its overall reliability was high to very high, while the reliability of its subscales was low. A practical implication of the latter is that the subscales should not be used on their own because they are more likely to reliably measure the general well-being factor than the specific group factors. Test–retest reliability of the total scale was acceptable, and convergent validity was supported. In conclusion, we consider the Swedish MHC-SF to be a psychometrically sound instrument for monitoring overall mental well-being in Swedish adolescents.

Keywords Positive mental health · Positive epidemiology · Well-being · Adolescence

Introduction

Conventionally, epidemiological surveillance has focused mainly on diseases and their risk factors (VanderWeele et al., 2020). This is perhaps particularly true for mental health (World Health Organization, 2013) where it has been often tacitly assumed that mental well-being would prevail in the absence of pathology (Huppert & So, 2013). However, a growing body of evidence shows that high levels of well-

being—reflecting not merely the absence of poor mental health—independently predict less subsequent mental illness and have a range of positive effects on individuals and society (Chida & Steptoe, 2008; Diener et al., 2010; Huppert, 2009; Wood & Joseph, 2010). Thus, there is a strong incentive to also include measures of well-being in the monitoring of public mental health.

The Mental Health Continuum – Short form (MHC-SF) is a self-report measure that has been increasingly used to measure well-being at the population level (Keyes, 2009a). It was developed by Keyes from his original Mental Health Continuum – Long form to create a version more efficiently administered in epidemiological surveillance (C. L. Keyes, 2002). The MHC-SF includes both the dominating traditions of well-being in the literature – hedonism and eudaimonism – with items and components that rather well reflect the well-known and cited WHO definition of mental health, i.e., as a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stress of life, can work productively and fruitfully, and is able to contribute to his or her community (World Health Organization, 2013). This positive definition of mental health includes both hedonic and

✉ Fredrik Söderqvist
fredrik.soderqvist@regionvastmanland.se

Peter Larm
peter.larm@su.se

¹ Centre for Clinical Research, Uppsala University, County Hospital, Västerås, Sweden

² University Health Care Research Center, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

³ Department of Public Health Sciences, Stockholm University, Stockholm, Sweden

eudaimonic well-being and acknowledges the importance of psychological as well as social aspects of human functioning.

The psychometric characteristics of the MHC-SF have previously been evaluated and used in many studies around the world. A number of research works tested the structural validity of MHC-SF using confirmatory factor analysis (CFA) for the evaluation of model fit (Donnelly et al., 2019; Dore et al., 2017; Guo et al., 2015; Karas et al., 2014; C. L. Keyes et al., 2008; Lamers et al., 2011; Machado & Bandeira, 2015; Orpana et al., 2017; Petrillo et al., 2015; Singh et al., 2015). Using this approach alone for evaluation, however, has some limitations when measures show multidimensional item responses (Reise, 2012). First, multidimensionality may jeopardize the interpretation of the total scale score as an indicator of a single construct, which is seldom addressed. Second, it is often unclear by former studies to what extent the use of subscale scores is justified, and, if so, third, how informative the subscales are in addition to the total scale score. Consequently, some researchers suggest that the proposed correlated three-factor structure of the MHC-SF, first proposed by Keyes, is problematic, both from a theoretical point of view (de Bruin & du Plessis, 2015) and from an empirical perspective, because it often produces only marginally acceptable fit indices (Jovanovic, 2015).

More recently, some of these limitations were addressed by studies that tested additional methods and models (de Bruin & du Plessis, 2015; Echeverria et al., 2017; Ferentinos et al., 2019; Hides et al., 2016; Jovanovic, 2015; Longo et al., 2020; Luijten et al., 2019; Reinhardt et al., 2020; Santini et al., 2020; Žemojtel-Piotrowska et al., 2017). The bifactor model, for example, allows for evaluation of the relevance and reliability of subscale scores after controlling for the variance of the general factor (Reise, 2012). Thus, if the contribution of a subscale, assumed to measure a specific latent construct, is small in relation to that of the general factor for the corresponding subscale items, then using the subscale as a measure of that specific latent factor can be questioned (Gignac & Watkins, 2013). Santini et al. studied the Danish MHC-SF through the use of bifactor modelling and found that subscale scores were unreliable, explaining very low variance beyond that explained by the general factor (Santini et al., 2020). This finding is well in line with the study of Žemojtel-Piotrowska et al. who also obtained similar results when they tested four different CFA models of MHC-SF, derived from theory and previous research. The 38-nation study with n-size larger than 8000 (mean age: 21.55 years) (Žemojtel-Piotrowska et al., 2017) showed that the bifactor model comprising one general factor and three uncorrelated factors of emotional, social and psychological well-being yielded the best fit.

Psychometric studies that compared the fit of the bifactor model with that of other proposed models have, in most cases, favoured the bifactor solution (de Bruin & du Plessis, 2015; Ferentinos et al., 2019; Hides et al., 2016; Longo et al., 2020;

Luijten et al., 2019; Monteiro et al., 2020; Reinhardt et al., 2020; Rogoza et al., 2018; Santini et al., 2020), including studies using exploratory structural equation modelling techniques (Ferentinos et al., 2019; Longo et al., 2020; Reinhardt et al., 2020; Rogoza et al., 2018). However, only one study (Reinhardt et al., 2020) (to our knowledge) has tested the bifactor model on adolescents exclusively. Furthermore, despite its widespread use in both America and Europe, no psychometric evaluation of the MHC-SF has yet been carried out in a Swedish population. The aim of the present study was, therefore, to evaluate the MHC-SF and its psychometric properties in Swedish adolescents. More specifically, the evaluation focused on structural and convergent validity, measurement invariance, dimensionality, and internal and test–retest reliability.

Methods

Data Collection and Study Population

First, the evaluation was performed by examining face validity and performing test–retest stability obtained in a pre-study, then factor structures and reliability were tested using data from a whole population study. Thus, two sources of data were used for the evaluation that required separate ethical approval. The original aim of the pre-study was to test a public health survey. The test was carried out at an upper secondary school in Västmanland County, Sweden. In total, 93 students participated by first completing the test survey and later attending focus group interviews to discuss the survey. Based on comments from the interviews, and preliminary psychometric testing, some minor adjustments were made, and then the MHC-SF was administered a second time two weeks after conducting the first test survey for further testing and evaluation, e.g., of test–retest stability. Written informed consent to participate was collected from all students prior to the study, and ethical approval was acquired from the regional review board at Uppsala (Dnr 2015/321).

The other source of data was the cross-sectional whole population study using Survey of Adolescent Life in Västmanland (SALVe) County in Sweden carried out during the early spring of 2020 prior to the Corona pandemic. In total, 5480 adolescents in the ninth grade of compulsory school or second year of upper secondary school were invited to participate in the study by completing a comprehensive health survey about living conditions, life habits and various aspects of health including scales of mental well-being. The survey was web based and carried out during school hours under exam-like conditions in a classroom context. Prior to logging on to the survey, all students watched a video with information about the study and its purpose and that participation in the study was voluntary and could be cancelled at any time for

any reason. Consent to participate was given by completing the questionnaire. All students participated anonymously; i.e., no names or personal identification number was collected. Both municipal and private schools were included in the study. In total, 3880 adolescents participated corresponding to a response rate of 71%, of which 5.4% did not respond to all items of the MHC-SF. Since no data imputation for missing values was used, analyses were based on total 3669 adolescents with a mean age of 16.23 years (range = 14–18 years), 51% of which were females. Ethical approval was acquired from the ethics review authority (Dnr 2019–05620).

Measurement Scales

In the SALVe study, two well-being scales were used: MHC-SF of primary interest for this evaluation (Keyes, 2009a) and the Short Warwick–Edinburgh Mental Well-being Scale (SWEMWBS) previously evaluated in Scandinavian adolescents for assessing convergent validity (Haver et al., 2015; Ringdal et al., 2018). The MHC-SF comprises 14 items of hedonic and eudaimonic well-being. Respondents are asked to rate how often in the past month they have experienced these signs of well-being on a 6-point Likert scale (never, once or twice a month, about once a week, two or three times a week, almost every day or every day). The more often the signs are experienced, the higher the rating, and the better the well-being. The SWEMWBS comprises seven items of hedonic and eudaimonic well-being. Similar to the MHC-SF, the SWEMWBS and its longer version, the WEMWBS, have been widely used to measure public mental health (Child Outcome Resource Consortium, n.d.). Respondents are asked to rate how often in the past two weeks they have experienced the seven signs of well-being on a 5-point Likert scale (none of the time, rarely, some of the time, often or all the time). The more often the signs experienced, the higher the rating, and the better the well-being.

We used the Swedish version of the MHC-SF made publicly available by Keyes (2009b). Based on the input from the focus group interviews, three minor adjustments were made in the phrasing of some items to better suit the adolescent context. First, examples of what the contribution to society in item 4 might consist of were added; second, the last part of item 6, in English ‘for people like you’, was deleted because it raised negative reactions in all focus groups; and third, item 10 was adjusted to improve grammar. Beyond this, there was a strong consensus among participants that the questions in the MHC-SF were interesting and relevant to the topic. No backtranslation of the MHC-SF was made.

Data Analysis

Based on theory and previous studies, four CFA models were tested using maximum likelihood estimation: a

single-factor model consisting of all items; a two-factor correlated model testing the concept of hedonic and eudaimonic well-being (i.e., items 1–3 and 4–14, respectively); a three-factor correlated model of emotional, social and psychological well-being (i.e., items 1–3, 4–8 and 9–14, respectively) as originally proposed by Keyes; and a bifactor model with one general and three specific (i.e., not correlated) factors of emotional, social and psychological well-being.

Model fit was evaluated using three different fit indices: the root mean square error of approximation (RMSEA), i.e., the standard deviation of the prediction errors; the comparative fit index (CFI) as a measure of the discrepancy between the data and the hypothesized model while adjusting for the issues of sample size inherent in the chi-squared test; and standardized root mean square residual (SRMR) as an absolute measure of fit in the difference of the observed and predicted correlations. RMSEA and SRMR values close to 0.06 were considered a close fit, and a CFI value above 0.90 was considered acceptable and above 0.95 satisfactory (Hu & Bentler, 1999). Configural, metric and scalar invariance of the MHC-SF were evaluated for gender and age using the Δ CFI, Δ RMSEA and Δ SRMR with cut-off values of ≥ 0.010 , ≥ 0.015 and ≥ 0.030 , respectively, as indicators of pronounced difference in fit between the nested and unconstrained models (Chen, 2007; Cheung & Rensvold, 2002).

Convergent validity, as the degree to which MHC-SF and SWEMWBS—both measures of hedonic and eudaimonic well-being—are in fact related, was assessed correlating the total scale scores as well as identified latent constructs from the CFA. A strong correlation with a coefficient above 0.7 supported convergent validity (Cohen, 1988).

Reliability, as a property of the data generated by the scale of interest here applied on Swedish adolescents, was assessed by calculating the model-based estimate coefficient omega (ω) and omega hierarchical (ω_H) for the general and the specific (group) factors (ω_s). To further aid in the interpretation of each factor’s importance, and assessment of dimensionality, the explained common variance (ECV) was calculated as previously detailed by Reise (Reise, 2012). The conventional and more frequently used coefficient alpha (that assumes tau-equivalence) was calculated for comparison with coefficient omega estimates (Deng & Chan, 2017) and other studies that used alpha. Recognizing that all cut-off values are arbitrary, the same recommendations were considered for omega as have been suggested for alpha; the values above 0.7 were considered acceptable, while those above 0.8 and closer to or above 0.9 indicated high to very high reliability (Kline, 2000). Finally, dimensionality was also assessed by conducting exploratory factor analysis (EFA) using maximum likelihood estimation and the ratio of the first and second eigenvalue, as suggested by Slocum-Gori and Zumbo (Slocum-

Gori & Zumbo, 2011). A ratio of ≥ 4 was considered indicative of essential unidimensionality.

Test–retest reliability of the MHC-SF total scale was assessed by calculating intra-class correlation coefficient (ICC) estimates and their 95% confidence intervals for single and average measures based on a two-way mixed-effects model where people effects were random and measure effects were fixed.

All analyses that required the use of statistical software were carried out using SPSS and AMOS for Windows versions 26 and 27, respectively.

Results

Data from the pre-study on 86 of the 93 participating students (62% females) that completed the test survey on two different occasions with two weeks in between were used to calculate test–retest reliability. Calculation of ICC's of the total MHC-SF scale scores for single and average measures indicated high to very high test–retest reliability with correlation coefficients of 0.86 (95% confidence interval [CI] 0.80–0.91) and 0.93 (95% CI 0.89–0.95), respectively.

Structural Validity

The four tested CFA models and their respective fit indices based on the SALVe data are summarized in Table 1. The single-factor model yielded the poorest fit, followed by the two- and three-factor correlated models, none of which showed an overall acceptable fit. The only model with a good fit was the bifactor model, as illustrated by the AMOS input diagram in Fig. 1. The fit indices for the bifactor model in Table 1 show an RMSEA close to 0.06, a CFI above 0.95 and an SRMR well under 0.06, indicating that the sample correlation matrix is well recovered.

The standardized factor loadings of the bifactor model are shown in Table 2 along with omega estimates for assessing

reliability and the percentage of common variance due to the general factor as a model-based index of unidimensionality. The loadings on the general factor are all moderate to strong ranging from 0.43 for Social actualization to 0.8 for Interest and Life satisfaction. The loadings on the group factors (i.e., emotional well-being [EW], social well-being [SW] and psychological well-being [PW]), however, were mostly weak with exceptions for items 6, 7 and 8 on SW. Item 5, Social integration, was statistically non-significant, and item 4, Social contribution, was weak but significant.

Dimensionality and Internal Reliability

The degree to which item response data were unidimensional versus multidimensional is shown by the ECV in Table 2 to be 0.73, and the degree to which raw scores reflect a common dimension, as opposed to mostly error, is shown by the model-based reliability index ω_H to be 0.79, thus indicating high reliability of the general factor. The model-based sibling to ω_H , the ω index—including the general as well as the specific group factors—and analogous to the popular coefficient α , showed high to very high reliability of 0.88. Thus, when all sources of common variance are included in the estimation of reliability, only a small difference is seen between ω and α . However, large differences become evident as the viability of subscale score is evaluated by controlling for the variance due to the general factor using the model-based subscale specific estimate omega hierarchical (ω_S). In Table 2, for example, the coefficient α for the emotional well-being subscale is estimated to be as high as 0.9, whereas only 0.15 according to ω_S —differences seen for all the subscales indicating poor reliability of the group factors (EW, SW and PW). The ECV of 0.73 indicating that MHC-SF is essentially unidimensional with a rather dominating general well-being (GW) factor was also supported by EFA as the ratio of the first (6.2) and second (1.5) eigenvalue was above 4.

Measurement Invariance

Measurement invariance was investigated by multigroup analyses of gender and age. The fit indices between models were compared; the fit indices of the model constrained on loading parameters (Metric invariance) were compared with those of the model constrained on structural parameters (Configural invariance); and the fit indices of the model constrained on intercepts and measurement residuals (Scalar invariance) were compared with those of the model constrained on the structural parameters (Configural). As shown in Table 3, no or only minor changes in fit are seen between the different models indicating overall measurement invariance. The only notable decrease in fit is seen for the CFI for gender between the models of scalar and configural invariance ($\Delta\text{CFI} = 0.012$).

Table 1 The four tested confirmatory factor analysis models and their fit indices

Fit indices	One-factor	Two-factor	Three-factor	Bifactor
<i>Df</i>	77	76	74	63
<i>RMSEA</i>	0.135	0.117	0.102	0.061
<i>CFI</i>	0.825	0.872	0.904	0.971
<i>SRMR</i>	0.079	0.072	0.089	0.027

Df = degrees of freedom; *RMSEA* = root mean square error of approximation; *CFI* = comparative fit index; *SRMR* = standardized root mean square residual

Fig. 1 Impute diagram of the bifactor model with one general (GW) and three specific uncorrelated group factors (EW, SW, PW)

However, no corresponding non-invariance in fit was seen in the other indices ($\Delta\text{RMSEA} = 0.005$; $\Delta\text{SRMR} = 0.003$).

Convergent Validity

Convergent validity was assessed with a bivariate correlation of the total scale scores of MHC-SF and SWEMWBS. Since both these instruments are designed to measure hedonic and eudaimonic aspects of mental well-being, the correlation between the two should be strong. Indeed, both Pearson's correlation coefficient and Spearman's rho yielded an estimate of 0.766 in support of convergent validity. Using AMOS, and the bifactor model in Fig. 1, we also correlated the GW factor of MHC-SF with a corresponding GW factor identified for SWEMWBS—first via EFA and then tested with single-factor CFA model that yielded an acceptable to good fit ($\text{RMSEA} = 0.079$; $\text{CFI} = 0.975$; $\text{SRMR} = 0.0244$). A correlation coefficient of 0.86 was found between the two modelled GW factors.

Discussion

Given the increased use of measuring well-being, it is important to psychometrically evaluate these instruments that cast light on how the concepts of well-being and positive mental health can be described and understood. The aim of this study was therefore to evaluate the MHC-SF for the first time in Swedish adolescents.

Regarding factor structure, the results indicate that positive mental health as measured by the MHC-SF is best described by a bifactor model with a dominant general well-being factor and three specific group factors of emotional, social and psychological well-being. The relatively high proportion of ECV by the general well-being factor—73%—indicates that the MHC-SF, essentially, is unidimensional (O'Connor, 2014). The latter finding is also supported by EFA with a ratio of the first and second eigenvalues above 4 (Slocum-Gori & Zumbo, 2011). Thus, the interpretation of the total scale score as an indicator of a single construct should be valid. As a result, the specific group

Table 2 CFA of the MHC-SF with standardized factor loadings and reliability estimates

Items and Dimensions	Bifactor model				Total
	EW	SW	PW	GW	
01. Happiness	0.42			0.76	
02. Interest	0.32			0.80	
03. Life satisfaction	0.36			0.80	
04. Social contribution		0.13		0.63	
05. Social integration		−0.02*		0.70	
06. Social actualization		0.65		0.43	
07. Social acceptance		0.57		0.47	
08. Social coherence		0.69		0.44	
09. Self-acceptance			0.30	0.73	
10. Mastery			0.26	0.61	
11. Positive relations			0.18	0.70	
12. Personal growth			0.40	0.59	
13. Autonomy			0.40	0.67	
14. Purpose in life			0.26	0.78	
Model-based reliability estimates	$\omega_S=0.15$	$\omega_S=0.24$	$\omega_S=0.13$	$\omega_H=0.79$	$\omega=0.88$
	ECV _S =0.05	ECV _S =0.15	ECV _S =0.07	ECV=0.73	
Variance–covariance-based reliability estimates	$\alpha=0.90$	$\alpha=0.79$	$\alpha=0.88$		$\alpha=0.92$

EW = emotional well-being; SW = social well-being; PW = psychological well-being; GW = general well-being; ω = coefficient omega; ω_H = coefficient omega hierarchical; ω_S = specific group factor omega hierarchical; ω_R = relative omega; ECV = explained common variance; ECV_S = explained common variance for specific factor; α = coefficient alpha. * = not statistically significant

factors contribute 27% of the ECV in total. Taken together, our results on factor structure are well in line with the only other study, to our knowledge, that has been published thus far and included testing of the bifactor model in adolescents (Reinhardt et al., 2020). The results are also in line with recent psychometric evaluations in adults that compared other factor solutions to the bifactor model (de Bruin & du Plessis, 2015; Ferentinos et al., 2019; Hides et al., 2016; Longo et al., 2020; Luijten et al., 2019; Monteiro et al., 2020; Reinhardt et al., 2020; Rogoza et al., 2018; Santini et al., 2020). These studies and our study show that the bifactor has superior fit compared with previously tested models, including the correlated three-factor

model first proposed by Keyes (C. L. Keyes, 2002, 2005; Lamers et al., 2011).

To verify that the bifactor structure is valid in different groups, measurement invariance of the MHC-SF was evaluated by multigroup analysis. Overall, the results support the metric, scalar and configural invariance of the model across gender and age groups. Similar multigroup invariance of the bifactor model in adolescents was recently obtained by Reinhardt et al. (Reinhardt et al., 2020). In addition, in a large study on young adults, the metric invariance and cross-cultural replicability of the bifactor model were supported by a multigroup confirmatory

Table 3 Fit indices for testing of measurement invariance

Group	Model	Model fit indices		
		RMSEA [90% CI]	CFI	SRMR
<i>Gender</i>	Metric invariance	0.040 [0.038–0.043]	0.969	0.028
	Scalar invariance	0.044 [0.042–0.046]	0.960	0.028
	Configural invariance	0.045 [0.043–0.047]	0.957	0.031
<i>Age (14–16 vs 17–18 years)</i>	Metric invariance	0.039 [0.037–0.042]	0.971	0.035
	Scalar invariance	0.039 [0.037–0.041]	0.971	0.036
	Configural invariance	0.037 [0.035–0.040]	0.971	0.035

RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual

analysis of the bifactor structure between samples in 38 countries (Žemojtel-Piotrowska et al., 2017).

An advantage of the bifactor model compared with, for example, single- or multi-factor correlated models is that it allows for evaluation of the relevance and reliability of subscale scores after controlling for the variance of the general factor (Reise, 2012). In the present study, this was done by calculating ECV for each latent factor along with the model-based reliability estimate omega hierarchical (ω_H) for the general factor and the equivalent estimate for each subscale (ω_S). While the latter supported high reliability of the general factor alone ($\omega_H = 0.79$) and very high overall reliability in the general factor together with the group factors ($\omega = 0.88$), it also revealed very low reliability for the subscales of emotional, social and psychological well-being ω_H (0.15, 0.24 and 0.13, respectively). The latter finding is well in line with previous studies that evaluated the bifactor model on the MHC-SF and assessed model-based reliability with omega hierarchical (Reinhardt et al., 2020; Žemojtel-Piotrowska et al., 2017). Hence, the use of subscale scores alone cannot be recommended. In addition, assessment of internal reliability of these subscales by estimating coefficient α cannot be recommended since it tends to conflate multiple sources of systematic variance when data are associated with a multidimensional model (Gignac & Watkins, 2013). Nevertheless, total scale scores of the MHC-SF yielded high to very internal consistency as well as test–retest reliability as indicated by strong ICCs using data from our pre-study.

Regarding convergent validity, a strong correlation (0.76) was found in the present study between total scale scores of MHC-SF and SWEMWBS, stronger than previously reported by Clarke et al. (0.65), also in adolescents (Clarke et al., 2011). In addition, using the modelled GW factors of the two scales, we filtered out the error variances, which then amplified the correlation further (0.86), providing ample support for convergent validity.

Strengths and Limitations

One of the strengths of our study was using the pre-study conducted to test the MHC-SF with respect to face validity and test–retest validity. It gave us valuable qualitative input from adolescents in the specific age groups that later participated in the population study on which the main analyses of this evaluation were based. With input from the focus groups, we could make the necessary, yet small, linguistic adjustments, and there was a strong consensus among the informants that the content of the 14 items was relevant to them and the topic. Another strength of the present study is that its main analyses are based on a whole population study with a relatively decent participation rate. The data should thus be highly representative for adolescents at least in Region of Västmanland County, but probably also to Swedish

adolescents in general because of the size of the study. A limitation of the study is that the sample consisted of ninth graders of compulsory school and second graders of upper secondary school only. Thus, the analysis of measurement invariance across age was limited, and the psychometric characteristics of MHC-SF for younger adolescents remain unknown. Furthermore, this study was conducted on a general population. The factor structure might be different in clinical populations, for example. Future research on adolescents would benefit from including clinical populations too. In Sweden, there is also a need for a corresponding psychometric evaluation in adults and the elderly.

Conclusions

To conclude, this large population study on Swedish adolescents found that the MHC-SF is essentially unidimensional and best described by a bifactor model with a dominant general well-being factor and three specific group factors of emotional, social and psychological well-being. Its general well-being factor has high internal reliability, but the reliability of its subscales of the specific group factor is low. A practical implication of the latter is that the subscales should not be used on their own because they are more likely to reliably measure the general well-being factor than the specific group factors. Test–retest reliability was good, and convergent validity was supported. In conclusion, we consider the MHC-SF to be a psychometrically sound instrument for overall mental well-being in Swedish adolescents.

Acknowledgments We would like to thank the students who participated in the study and the school teachers for their cooperation in administering the codes for the questionnaire. We would also like to thank Madelen Lindquist for her valuable help with contacting and recruiting the participating 52 schools and Lydia Ersberg and Jim Lindström for assisting in the testing and development of the public health survey.

Code Availability All analyses that required the use of statistical software were carried out using SPSS and AMOS for Windows versions 26 and 27, respectively.

Authors' Contributions Fredrik Söderqvist was the lead researcher, who designed the study, performed statistical analysis and was responsible for manuscript writing. Peter Larm gave valuable comments on analyses and contributed to critical revision of the manuscript.

Funding Open access funding provided by Uppsala University. This research received no specific grant from any funding agency, commercial or not-for-profit sectors but was supported by the departments of Welfare and Public health and Center for Clinical Research, Region Västmanland, Sweden.

Data Availability Data are available from the corresponding author upon reasonable request.

Declarations

Ethics Approval The study was conducted in accordance with the principles of the Declaration of Helsinki, and the study protocols were approved by the regional review board at Uppsala (Dnr 2015/321) and the ethics review authority (Dnr 2019–05620).

Consent to Participate All participants watched a video with information about the study and its purpose and that participation in the study was voluntary and could be cancelled at any time for any reason. Consent to participate was given by completing the questionnaire. All students participated anonymously; i.e., no names or personal identification number was collected.

Consent for Publication Not applicable.

Competing Interests None of the authors report any competing interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504.
- Cheung, G. W., & Rensvold, R. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255.
- Chida, Y., & Steptoe, A. (2008). Positive psychological well-being and mortality: A quantitative review of prospective observational studies. *Psychosomatic Medicine*, 70(7), 741–756. <https://doi.org/10.1097/PSY.0b013e31818105ba>.
- Child Outcome Resource Consortium (n.d.). Short Warwick-Edinburgh Mental Well-being Scale - Psychometric properties. <https://www.core.uk.net/outcome-experience-measures/short-warwick-edinburgh-mental-wellbeing-scale/>. Accessed 1 November 2020.
- Clarke, A., Friede, T., Putz, R., Ashdown, J., Martin, S., Blake, A., et al. (2011). Warwick-Edinburgh mental well-being scale (WEMWBS): Validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health*, 11, 487. <https://doi.org/10.1186/1471-2458-11-487>.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum Associates.
- de Bruin, G. P., & du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum-short form (MHC-SF). *Psychological Reports*, 116(2), 438–446. <https://doi.org/10.2466/03.02.PR0.116k20w6>.
- Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational and Psychological Measurement*, 77(2), 185–203. <https://doi.org/10.1177/0013164416658325>.
- Diener, E., Helliwell, J. F., & Kahneman, D. (2010). *International differences in well-being*. Oxford University Press.
- Donnelly, A., O'Reilly, A., Dolphin, L., O'Keeffe, L., & Moore, J. (2019). Measuring the performance of the mental health continuum-short form (MHC-SF) in a primary care youth mental health service. *Irish Journal of Psychological Medicine*, 36(3), 201–205. <https://doi.org/10.1017/ipm.2018.55>.
- Dore, I., O'Loughlin, J. L., Sabiston, C. M., & Fournier, L. (2017). Psychometric evaluation of the mental health continuum-short form in French Canadian young adults. *Canadian Journal of Psychiatry*, 62(4), 286–294. <https://doi.org/10.1177/0706743716675855>.
- Echeverria, G., Torres, M., Pedrals, N., Padilla, O., Rigotti, A., & Bitran, M. (2017). Validation of a Spanish version of the mental health continuum-short form questionnaire. *Psicothema*, 29(1), 96–102. <https://doi.org/10.7334/psicothema2016.3>.
- Ferentinos, P., Yotsidi, V., Porichi, E., Douzenis, A., Papageorgiou, C., & Stalikas, A. (2019). Well-being in patients with affective disorders compared to nonclinical participants: A multi-model evaluation of the mental health continuum-short form. *Journal of Clinical Psychology*, 75(9), 1585–1612. <https://doi.org/10.1002/jclp.22780>.
- Gignac, G. E., & Watkins, M. W. (2013). Bifactor Modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, 48(5), 639–662. <https://doi.org/10.1080/00273171.2013.804398>.
- Guo, C., Tomson, G., Guo, J., Li, X., Keller, C., & Soderqvist, 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, 198. <https://doi.org/10.1186/s12955-015-0394-2>.
- Haver, A., Akerjordet, K., Caputi, P., Furunes, T., & Magee, C. (2015). Measuring mental well-being: A validation of the short Warwick-Edinburgh mental well-being scale in Norwegian and Swedish. *Scandinavian Journal of Public Health*, 43(7), 721–727. <https://doi.org/10.1177/1403494815588862>.
- Hides, L., Quinn, C., Stoyanov, S., Cockshaw, W., Mitchell, T., & Kavanagh, D. J. (2016). Is the mental wellbeing of young Australians best represented by a single, multidimensional or bifactor model? *Psychiatry Research*, 241, 1–7. <https://doi.org/10.1016/j.psychres.2016.04.077>.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied psychology: health and well-being*, 1(2), 137–164.
- Huppert, F. A., & So, T. T. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research*, 110(3), 837–861. <https://doi.org/10.1007/s11205-011-9966-7>.
- Jovanovic, V. (2015). Structural validity of the mental health continuum-short form: The bifactor model of emotional, social and psychological well-being. *Personality and Individual Differences*, 75, 154–159.
- Karas, D., Cieciuch, J., & Keyes, C. L. M. (2014). The polish adaptation of the mental health continuum-short form (MHC-SF). *Personal and Individual Differences*, 69, 104–109.
- Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43(2), 207–222.
- Keyes, C. L. (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. M. (2009a). Atlanta: *Brief Description of the Mental Health Continuum Short Form (MHC-SF)*. <https://www.aacu.org/sites/default/files/MHC-SFEnglish.pdf>. Accessed 31 October 2020.

- Keyes, C. L. M. (2009b). Swedish MHC-SF. https://www.researchgate.net/publication/322552541_Swedish_MHC-SF. Accessed 30 October 2020.
- 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>.
- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). Routledge.
- 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>.
- Longo, Y., Jovanovic, V., Sampaio de Carvalho, J., & Karas, D. (2020). The general factor of well-being: Multinational evidence using Bifactor ESEM on the mental health continuum-short form. *Assessment*, 27(3), 596–606. <https://doi.org/10.1177/1073191117748394>.
- Luijten, C. C., Kuppens, S., van de Bongardt, D., & Nieboer, A. P. (2019). Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF) in Dutch adolescents. *Health and Quality of Life Outcomes*, 17(1), 157. <https://doi.org/10.1186/s12955-019-1221-y>.
- Machado, W. L., & Bandeira, D. R. (2015). Positive mental health scale: Validation of the mental health continuum – Short form. *Psico-USF, Braganca Paulista*, 20(2), 259–274.
- Monteiro, F., Fonseca, A., Pereira, M., & Canavarro, M. C. (2020). Measuring positive mental health in the postpartum period: The Bifactor structure of the mental health continuum-short form in Portuguese women. *Assessment*, <https://doi.org/10.1177/1073191120910247>.
- O'Connor, Q. H. (2014). *Bifactor models, explained common variance (ECV) and the usefulness of scores from unidimensional item response theory analyses (unpublished master's thesis)*. University of North Carolina.
- Orpana, H., Vachon, J., Dykxhoorn, J., & Jayaraman, G. (2017). Measuring positive mental health in Canada: Construct validation of the mental health continuum-short form. *Health Promotion and Chronic Disease Prevention in Canada: research, policy and practice*, 37(4), 123–130. <https://doi.org/10.24095/hpcdp.37.4.03>.
- Petrillo, G., Capone, V., Caso, D., & Keyes, C. L. M. (2015). The mental health continuum-short form (MHC-SF) as a measure of well-being in the Italian context. *Social Indicators Research*, 2015(121), 1–22.
- Reinhardt, M., Horvath, Z., Morgan, A., & Kokonyei, G. (2020). Well-being profiles in adolescence: Psychometric properties and latent profile analysis of the mental health continuum model - a methodological study. *Health and Quality of Life Outcomes*, 18(1), 95. <https://doi.org/10.1186/s12955-020-01332-0>.
- Reise, S. P. (2012). Invited paper: The rediscovery of Bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
- Ringdal, R., Bradley Eilertsen, M. E., Bjørnsen, H. N., Espnes, G. A., & Moksnes, U. K. (2018). Validation of two versions of the Warwick-Edinburgh mental well-being scale among Norwegian adolescents. *Scandinavian Journal of Public Health*, 46(7), 718–725. <https://doi.org/10.1177/1403494817735391>.
- Rogoza, R., Truong Thi, K. H., Rozycka-Tran, J., Piotrowski, J., & Zemojtel-Piotrowska, M. (2018). Psychometric properties of the MHC-SF: An integration of the existing measurement approaches. *Journal of Clinical Psychology*, 74(10), 1742–1758. <https://doi.org/10.1002/jclp.22626>.
- Santini, Z. I., Torres-Sahli, M., Hinrichsen, C., Meilstrup, C., Madsen, K. R., Rayce, S. B., Baker, M. M., ten Have, M., Schotanus-Dijkstra, M., & Koushede, V. (2020). Measuring positive mental health and flourishing in Denmark: Validation of the mental health continuum-short form (MHC-SF) and cross-cultural comparison across three countries. *Health and Quality of Life Outcomes*, 18(1), 297. <https://doi.org/10.1186/s12955-020-01546-2>.
- Singh, K., Bassi, M., Junnarkar, M., & Negri, L. (2015). Mental health and psychosocial functioning in adolescence: An investigation among Indian students from Delhi. *Journal of Adolescence*, 39, 59–69. <https://doi.org/10.1016/j.adolescence.2014.12.008>.
- Slocum-Gori, S., & Zumbo, B. D. (2011). Assessing the Unidimensionality of psychological scales: Using multiple criteria from factor analysis. *Social Indicators Research*, 102, 443–461.
- VanderWeele, T. J., Chen, Y., Long, K., Kim, E. S., Trudel-Fitzgerald, C., & Kubzansky, L. D. (2020). Positive epidemiology? *Epidemiology*, 31(2), 189–193. <https://doi.org/10.1097/EDE.0000000000001147>.
- Wood, A. M., & Joseph, S. (2010). The absence of positive psychological (eudemonic) well-being as a risk factor for depression: A ten year cohort study. *Journal of Affective Disorders*, 122(3), 213–217. <https://doi.org/10.1016/j.jad.2009.06.032>.
- World Health Organization (2013). Mental health action plan 2013–2020. http://apps.who.int/iris/bitstream/10665/89966/1/9789241506021_eng.pdf. Accessed 31 October 2020.
- Żemojtel-Piotrowska, M., Piotrowski, J., Osin, E., Cieciuch, J., Adams, B., Ardi, R., et al. (2017). The mental health continuum-short form: The structure and application for cross-cultural studies—a 38 nation study. *Journal of Clinical Psychology*, 74, 1034–1052.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.