Psychology Research and Behavior Management

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Measurement Properties and Optimal Cutoff Point of the WHO-5 Among Chinese Healthcare Students

Nongnong Yang[1,*], Haiyan Ma[1,2], Mengyi Huang[1,*], Esther Yuet Ying Lau[3–5], Daniel Yee Tak Fong[6], Meng Wang[7], Pengqiao Wang[1], Siyi Xu[8], Jiale Xu[1], Chen Jiang[1], Yi Luo[9], Runtang Meng[1,2,*]

[1]School of Public Health, Hangzhou Normal University, Hangzhou, Zhejiang, People's Republic of China; [2]Engineering Research Center of Mobile Health Management System, Ministry of Education, Hangzhou, Zhejiang, People's Republic of China; [3]Sleep Laboratory, Department of Psychology, The Education University of Hong Kong, Hong Kong, People's Republic of China; [4]Centre for Psychosocial Health, The Education University of Hong Kong, Hong Kong, People's Republic of China; [5]Centre for Religious and Spirituality Education, The Education University of Hong Kong, Hong Kong, People's Republic of China; [6]School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China; [7]Ophthalmology Center, Ningbo Yinzhou No.2 Hospital, Ningbo, Zhejiang, People's Republic of China; [8]School of Nursing, Hangzhou Normal University, Hangzhou, Zhejiang, People's Republic of China; [9]School of Nursing, Ningbo College of Health Sciences, Ningbo, Zhejiang, People's Republic of China

*These authors contributed equally to this work

Correspondence: Runtang Meng; Haiyan Ma, School of Public Health, Hangzhou Normal University, No. 2318, Yuhangtang Road, Yuhang District, Hangzhou, Zhejiang, 311121, People's Republic of China, Email mengruntang@hznu.edu.cn; mahaiyan@hznu.edu.cn

Purpose: The World Health Organization-Five Well-Being Index (WHO-5) is widely used to assess subjective well-being. Nevertheless, measurement invariance and optimal cutoff point of the WHO-5 have not been examined in Chinese samples. We aimed to assess measurement properties of the Chinese version of the WHO-5 (WHO-5-C) among healthcare students.

Patients and Methods: A two-wave longitudinal assessment was conducted among 343 Chinese healthcare students from September to November 2022. Measurement properties of the WHO-5-C were assessed through structural validity using confirmatory factor analysis (CFA), measurement invariance using multigroup CFA (MGCFA) and longitudinal CFA (LCFA), convergent validity using correlation analysis with the Self-Rated Health Questionnaire (SRHQ) and Patient Health Questionnaire-4 (PHQ-4), reliability using internal consistency and test–retest reliability, and optimal cutoff point using receiver operating characteristic (ROC) analysis.

Results: The WHO-5-C demonstrated satisfactory structural validity with comparative fit index (CFI) of 0.968 at baseline and 0.980 at follow-up, and adequate measurement invariance in different sociodemographic variables at baseline (gender, age, major, home location, being only child, monthly household income, part-time job, physical exercise, hobby, frequency of visiting home, and stress coping strategy) (CFI changes [ΔCFI] = −0.009–0.003) and over a week (ΔCFI = −0.006–0.000). The WHO-5-C also had good internal consistency (Cronbach's α = 0.907–0.934; McDonald's ω = 0.908–0.935) and test–retest reliability (intraclass correlation coefficient [ICC] = 0.803). Convergent validity was supported by moderate correlations of the WHO-5-C with the SRHQ and PHQ-4. The optimal cutoff point of the WHO-5-C was found to be 50, with an area under the ROC curve of 0.882 at baseline data, with sensitivity of 0.803 and specificity of 0.762 at follow-up.

Conclusion: The WHO-5-C demonstrated adequate measurement properties, especially concerning cross-sectional and longitudinal measurement invariance, with a recommended optimal cutoff point of ≥ 50 for assessing adequate level of psychological well-being in healthcare students.

Keywords: WHO-5 Well-Being Index, subjective well-being, measurement properties, diagnostic performance, observational longitudinal design, healthcare students

Introduction

Subjective well-being (SWB) refers to an individual's ability to develop their potential, to work productively and creatively, to build strong and positive relationships with others, and to contribute to their community.[1] SWB involves multiple dimensions

Received: 27 August 2023
Accepted: 26 November 2023
Published: 22 December 2023

and its contribution to all aspects of human life, eg, optimistic attitude, positive affect, psychological resilience, and happiness.[2] Positive SWB is recognized as having crucial consequences for better physical outcomes in both healthy populations and patients suffering from various diseases, such as fewer self-perceived symptoms, fewer pain sensations, and longer survival.[3–5] As such, positive SWB has been regarded as a conducive health asset given its explicit associations with salutary outcomes of mental health,[6] behaviors,[7] and disease progression and rehabilitation.[8] Therefore, identifying SWB carries clinical significance for public and psychosomatic health.

There has been increasing evidence showing that college students are exposed to excessive academic pressures that may affect their academic performance, lifestyle, physical health, and even psychosomatic disorders.[9–11] Healthcare students are expected to provide professional support in health services for individuals and societies in the future[12] and are at risk of physical and mental disorders.[13,14] Worldwide, approximately one-third of healthcare students suffer from negative emotions, which is notably higher than general population and non-healthcare students.[15,16] A meta-analysis of mental health problems among healthcare students in China showed that the prevalence of depression, anxiety, and suicidal ideation were 29%, 21%, and 11%, respectively, underlining the severity of the issue and urgency to understand their well-being more.[17] Therefore, instruments with a positive focus can suggest to respondents that such programs are for individuals with pleased SWB and in this way support, rather than detract from, these initiatives.

There are numerous approaches in explaining and assessing well-being, and several instruments have been developed to monitor well-being.[18] To date, the popular instruments for assessing SWB state are mostly symptom-based, eg, stress, anxiety, depression and bipolar disorders.[19–21] Existing instruments developed to evaluate well-being are mostly based on a single perspective.[18] For instance, one of the most widely used for evaluating SWB is the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS), a 14-item instrument developed to assess hedonic and eudaimonic elements of mental health.[22] Despite previous researches have endorsed the WEMWBS presented appropriate capacity for mental well-being in Chinese samples, it does not consider purpose in physical condition which is an essential component of overall well-being.[23,24] Likewise, other instruments introduced into China to determine SWB include the Personal Wellbeing Index (PWI),[25] Psychological Well-Being Scale (PWBS),[26] etc., and most only assess satisfaction level in terms of health-related quality of life. It is necessary to unearth an instrument that can capture a comprehensive conception of SWB, including affective-emotional aspects and physiological feelings.

The positively worded World Health Organization-Five Well-Being Index (WHO-5) is one of the measures from a positive perspective, aiming to measure overall wellness and cover healthy elements psychologically and physically.[27] As a multidimensional screening instrument, the WHO-5 initially comprised twenty-eight items and progressively shortened to five positive items to produce a brief and comprehensive scale of well-being.[27,28] To date, there are more than thirty official language versions of the WHO-5 that have been endorsed by the World Health Organization (WHO).[27] The scale has been adapted and validated across numerous cultures, including Arabic,[29] Brazilian,[30] Chinese,[31] German,[32] Icelandic,[33] Japanese,[34] Malay,[35] Polish,[36] Portuguese,[30] Sinhala,[37] Spanish,[38] Swahili,[39] Swedish,[40] and Turkish.[41] The original English version of the WHO-5,[42] as well as the Japanese[34] and Polish[36] versions, had a single-factor structure with excellent reliability and validity in adolescents with type 1 diabetes and a cutoff point below 50 was established to identify worsened well-being. Moreover, the WHO-5 has revealed correlations with adverse health conditions among different samples of diverse cultures, including Spanish, Swedish, and Turkish.[38,40,41] The WHO-5, overall, has sufficient validity to measure SWB in patients and general populations across study fields, and the scale has been extensively applied in endocrinology, psychiatry, and clinical as well as positive psychology.[28,37] Therefore, the WHO-5 demonstrated adequate effectiveness as both an outcome measure in clinical trials and a screening instrument in community settings.

The WHO-5 has been adapted and validated in Chinese university students[43] and medical educators.[44] Measurement properties of the Chinese version of the WHO-5 (WHO-5-C) were assessed in a recent multi-site study among a sample of university students and reported a stable one-factor structure and adequate internal consistency, factorial validity, construct validity and concurrent validity.[43] According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guideline,[45] it would be integral to further ascertain test–retest reliability, measurement invariance, measurement error, and recommended cutoff point of the WHO-5-C, intending to select the most appropriate outcome measurement instruments in research and clinical practice. This study aimed to achieve the goal of validating the WHO-5-C as a screening and monitoring instrument of SWB using data collected from healthcare

students. This two-wave longitudinal study specifically assessed structural validity, measurement invariance, convergent validity and reliability; identified the optimal cutoff point by conducting receiver operating characteristic (ROC) analysis; and determined cross-sectional and longitudinal measurement invariance with respect to sociodemographic variables and approximately one-week interval, respectively. We hypothesized that the WHO-5 can serve as both an assessment and a screening instrument for routine monitoring of SWB in Chinese samples.

Materials and Methods

Participants and Procedures

This study was a two-wave longitudinal assessment and a convenience sample from a university in Hangzhou, China, participated in a paper-based survey in simplified Chinese from September to November 2022 upon informed consent. We recruited healthcare students who were enabled to read and write simplified Chinese and had the ability to conduct the response process freely. Respondents were excluded if they: 1) had difficulty understanding or writing Chinese; 2) were on leave or out of school; or 3) disagreed to participate in the study. The self-report paper-and-pencil survey was administered during breaks between classes or evening self-study and required about 10 minutes to complete the questionnaires. The trained investigators were responsible for implementing assessment and ensuring its onsite quality control. All participants were required to complete two-wave longitudinal measurements with an average interval of 7 days + 3 hours, given that 1) the appropriate interval varies from an hour to a year depending on the task, but generally speaking, a retest interval of 2 to 14 days is usual;[46] and 2) reproducibility of health status measures intended for longitudinal use may best be measured at intervals of 1 to 2 weeks.[47] Respondents were compensated with 2 CNY (1 CNY ≈ 0.150 US dollars) each. The valid questionnaires in this study (N = 343) reached recommended sample size for factor analyses: 1) the sample size of 300 is considered as good;[48] 2) the appropriate minimum for sample size include from 3 to 20 times the number of variables, and sample size above 200 is suggested when the variables-to-factors is six.[49]

The current study was conducted in accordance with the Declaration of Helsinki[50] and approved by the Institutional Review Board of the School of Public Health, Hangzhou Normal University, China (Reference No. 20210014). All healthcare students freely consented to respond to the questionnaires and provided their informed consents before inclusion into the assessments. The authors confirmed full respect and protection of individual privacy rights before, during, and after the data collection and processing.

Measures

Sociodemographic Variables

We collected sociodemographic variables, including gender (male, female), age, major (clinical medicine, preventive medicine), home location (urban, rural, suburban), being only child (yes, no), monthly household income (< 10,000 CNY, ≥ 10,000 CNY), part-time job (yes, no), physical exercise [exercise goals to improve health (yes, no)], hobby (yes, no), frequency of visiting home [frequently (once per week, twice per week, and once per month); occasionally (once per quarter, once per semester, and once per academic year)], and self-reported preferred stress coping strategy (emotion-focused, solution-focused, avoidance coping).

World Health Organization-Five Well-Being Index (WHO-5)

The WHO-5 developed by the WHO is a brief self-report instrument used to assess SWB. The WHO-5 was translated into simplified Chinese in 2007 and is available on the WHO-5 official website.[27,51] Respondents were asked how well each of the five statements applies to them when considering the last two weeks. The five items of the scale cover positive emotions (in good spirits, feeling relaxed), vitality (feeling active, waking up refreshed and rested), and being interested in things.[27] Each item was scored on a six-point Likert scale, ranging from 0 (at no time) to 5 (all the time). Multiplying the original score by four is usually recommended because quality-of-life-related scales are often converted to a percentage scale.[27,28] The final score from 0 representing the worst imaginable well-being to 100 representing the best imaginable well-being. All items of the WHO-5-C are loaded on a potential factor with adequate reliability (Cronbach's α = 0.810–0.850, McDonald's ω = 0.820–0.860) and validity (comparative fit index [CFI] = 0.974, Tucker-Lewis index [TLI] = 0.947, and root mean square error of approximation [RMSEA] = 0.080).[43]

Patient Health Questionnaire-4 (PHQ-4)

The PHQ was designed to facilitate recognition and diagnosis of the most common mental disorders in primary care patients.[52] The PHQ-4 is a validated instrument for detecting anxiety and depression two weeks prior to assessment administration. The PHQ-4 consists of two ultra-short scales, one is the Generalized Anxiety Disorder-2 (GAD-2) for anxiety detection[53] and another is the PHQ-2 that reflects depression.[54] Each item of the PHQ-4 is scored on a four-point Likert scale, ranging from 0 (not at all) to 3 (nearly every day), with higher scores indicating more severe symptom levels.[55,56] The Chinese version of the PHQ-4 (PHQ-4-C) is publicly available and has shown good internal consistency and test–retest reliability (Cronbach's α = 0.870–0.904, McDonald's ω = 0.894–0.904, intraclass correlation coefficient [ICC] = 0.697).[57,58]

Self-Rated Health Questionnaire (SRHQ)

A simple two-item questionnaire was applied to assess the self-perceived health of all participants, with one item estimating their physical health and one assessing their mental health. Each item was rated on a five-point Likert scale, from 1 (excellent) to 5 (extremely poor), with higher scores indicating worse overall health.[59,60] The SRHQ was demonstrated to have satisfactory internal consistency and test–retest reliability in our previous studies (Cronbach's α = 0.706–0.857, ICC = 0.565–0.710).[59–61]

Statistical Analysis

Data analysis was performed using R (version 4.1.3) and JASP (version 0.16.1). The R packages used in this analysis were "*lavvan* (0.6–11)",[62] "*MBESS* (4.9.2)",[63] "*irr* (0.84.1)",[64] "*semTools* (0.5–6)",[65] and "*pROC* (1.18.2)".[66] We matched data according to student ID, and missingness in the present study ranged from 0.29% to 1.46% (< 5%). We used mean imputation for continuous variables and median imputation for categorical variables.[67] Means, standard deviations (SDs), skewness, and kurtosis were used to assess multivariate normality. We selected maximum likelihood estimation (MLE) method to evaluate all confirmatory factor analysis (CFA) results given 1) data of the WHO-5-C is nonnormality (skew < 2, kurtosis < 7), and 2) the number of ordered categories is more than five.[68] All measurement properties of the WHO-5-C were assessed in accordance with the requirements of the COSMIN guideline.[69]

Structural Validity

We assessed structural validity of the WHO-5-C by CFA to evaluate the extent to which the scores reflected underlying dimensions. Since the WHO-5 has been widely recognized as a single-factor scale, we evaluated structural validity of the scale based on a single-factor model. A satisfactory model fit was indicated for CFI and TLI ≥ 0.900, RMSEA ≤ 0.080, and standardized root mean square residual (SRMR) ≤ 0.080.[70,71]

Measurement Invariance

To assess measurement invariance of the WHO-5-C with regard to demographic variables and time, we conducted a series of multi-group CFA (MGCFA) and longitudinal CFA (LCFA). This method involved progressively constraining parameters to be equal between subgroups (sociodemographic variables) and across time intervals, and then comparing changes in fit indices to determine whether the relationship between observed variables and underlying traits was equivalent.[72] Herein, a series of nested models with increasing constraints were established, including the configural (same pattern of factors), metric (same pattern of factors and loadings), scalar (same pattern of factors, loadings, and item thresholds), and strict (same pattern of factors, loadings, item thresholds, and residual variances) models. Measurement invariance was defined as the fit statistic not significantly changing in an iterative procedure of progressively strict constraints. We considered the change in CFI (ΔCFI) to be an applicable metric for measurement invariance, with a change ≤ 0.010 indicating an appropriate measurement invariance.[73,74]

Convergent Validity

To evaluate convergent validity of the WHO-5-C, we calculated Pearson correlation coefficients of the WHO-5-C with the SRHQ and PHQ-4-C, as these instruments are developed to assess an individual's subjective feelings, considering mental and physical health often influence and accompany each other.[75] We additionally analyzed inter-item and item-total correlations. We hypothesized that the WHO-5-C would be moderately correlated with the PHQ-4-C and SRHQ in expected directions, given that these instruments measure interrelated constructs. The absolute magnitude of correlation coefficient was categorized as very

strong correlation ($r > 0.900$), strong correlation ($r = 0.700–0.900$), moderate correlation ($r = 0.400–0.700$), and weak correlation ($r < 0.400$).[76]

Reliability

We assessed internal consistency of the WHO-5-C using Cronbach's α and McDonald's ω coefficients, with coefficient ≥ 0.700 considered satisfactory.[77] To measure test–retest reliability, we used ICC between two separate times and coefficient below 0.400, between 0.400 and 0.590, between 0.600 and 0.740, and greater than 0.750 indicating poor, fair, good, and excellent reliability, respectively.[78] Moreover, standard error of measurement (SEM) was calculated as additional evidence to determine measurement accuracy when evaluating test–retest reliability.[79]

Sensitivity and Specificity

To evaluate the WHO-5-C as a screening scale for SWB, area under the ROC curve (AUC), optimal cutoff point, sensitivity, and specificity were calculated using the ROC analysis. As the PHQ-2 was an ultra-brief and useful screening instrument in depression with excellent operating characteristics, the score ≥ 3 is used as a criterion to determine the optimal cutoff point of the WHO-5-C.[52] The optimal cutoff point was obtained from the point closest to the top left-hand corner of ROC curve in baseline data and examined in follow-up data. The AUC value ranges from 0.500 to 1.000, with higher value indicating better prediction. A value greater than or equal to 0.800 was regarded as a good discrimination.[80,81]

Results

Sociodemographic Variables

The sample consisted of 216 females (62.974%) with an average age of 19.650 (SD = 1.414) years, ranging from 17 to 23 years. Tables 1 and S1 summarized the participants' sociodemographic variables and scores of the measured scales at baseline and follow-up.

Table 1 Sociodemographic Variables (N = 343)

Variables	N (%)	Mean (SD)	
		Baseline	Follow-Up
Gender			
Male	127 (37.026)	57.732 (17.168)	57.827 (18.425)
Female	216 (62.974)	56.741 (17.873)	58.167 (18.578)
Age			
< 20 years	150 (43.732)	58.800 (17.264)	60.773 (17.468)
≥ 20 years	193 (56.268)	55.793 (17.785)	55.917 (19.030)
Major			
Clinical medicine	154 (44.898)	53.039 (17.232)	54.883 (19.518)
Preventive medicine	189 (55.102)	60.423 (17.234)	60.614 (17.246)
Home location			
Urban	156 (45.481)	58.179 (17.885)	59.256 (19.050)
Rural	101 (29.446)	56.356 (18.401)	57.188 (17.786)
Suburban	86 (25.073)	56.047 (16.135)	56.837 (18.367)
Being only child			
Yes	128 (37.318)	57.094 (16.864)	57.531 (19.246)
No	215 (62.682)	57.116 (18.058)	58.344 (18.072)

(*Continued*)

Table 1 (Continued).

Variables	N (%)	Mean (SD)	
		Baseline	Follow-Up
Monthly household income			
< 10,000 CNY	141 (41.108)	55.489 (18.663)	57.135 (19.197)
≥ 10,000 CNY	202 (58.892)	58.238 (16.767)	58.673 (18.010)
Part-time job			
Yes	33 (9.621)	56.364 (17.360)	57.333 (18.806)
No	310 (90.379)	57.187 (17.648)	58.116 (18.491)
Physical exercise			
Yes	257 (74.927)	58.646 (17.513)	59.564 (18.532)
No	86 (25.073)	52.512 (17.134)	53.488 (17.720)
Hobby			
Yes	244 (71.137)	58.639 (17.849)	59.770 (18.492)
No	99 (28.863)	53.333 (16.443)	53.778 (17.891)
Frequency of visiting home			
Frequently	51 (14.869)	63.294 (17.283)	62.824 (17.386)
Occasionally	292 (85.131)	56.027 (17.456)	57.205 (18.584)
Stress coping strategy			
Emotion-focused	189 (55.102)	58.455 (16.103)	59.111 (17.952)
Solution-focused	121 (35.277)	57.983 (18.656)	59.240 (18.757)
Avoidance coping	33 (9.621)	45.182 (18.550)	47.515 (17.798)

Abbreviations: SD, standard deviation; Mean (SD), mean scores and SDs of the Chinese World Health Organization-Five Well-Being Index.

Structural Validity

Structural validity of the WHO-5-C was explored using CFA based on a one-factor model. As shown in Table 2, the fit indices of both baseline (CFI = 0.968, TLI = 0.937, SRMR = 0.028) and follow-up (CFI = 0.980, TLI = 0.961, SRMR = 0.020) analyses indicated that the WHO-5-C had a satisfactory fit with a single-factor structure.

Measurement Invariance

Cross-sectional measurement invariance of the WHO-5-C was analyzed across healthcare students' sociodemographic variables. Tables 3 and S2 showed that the WHO-5-C was well-fixed in four nested models among all subgroups, with all

Table 2 Model Fit Indices of the Single-Factor Model for the WHO-5-C (N = 343)

	χ^2 (*df*)	*P*	CFI	TLI	RMSEA (90% CI)	SRMR
Baseline	42.372 (5)	< 0.001	0.968	0.937	0.148 (0.109, 0.190)	0.028
Follow-up	35.051 (5)	< 0.001	0.980	0.961	0.132 (0.093, 0.175)	0.020
Threshold	N/A	> 0.050	≥ 0.900	≥ 0.900	≤ 0.080	≤ 0.080

Abbreviations: WHO-5-C, Chinese World Health Organization-Five Well-Being Index; χ^2, chi-square; *df*, degrees of freedom; CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation; CI, confidence interval; SRMR, standardized root mean square residual; N/A, not applicable.

Table 3 Cross-Sectional Measurement Invariances of the WHO-5-C (N = 343)

Hypothesis	Baseline							Follow-up						
	χ^2 (*df*)	*P*	$\Delta\chi^2$ (Δ*df*)	CFI	ΔCFI	TLI	SRMR	χ^2 (*df*)	*P*	$\Delta\chi^2$ (Δ*df*)	CFI	ΔCFI	TLI	SRMR
Gender (male vs female)														
Configural	52.934 (10)	< 0.001		0.964		0.929	0.027	41.874 (10)	< 0.001		0.979		0.959	0.018
Metric	55.992 (14)	< 0.001	3.058 (4)	0.965	0.001	0.950	0.037	51.370 (14)	< 0.001	9.496 (4)	0.976	−0.003	0.965	0.044
Scalar	63.280 (18)	< 0.001	7.288 (4)	0.962	−0.003	0.958	0.042	55.458 (18)	< 0.001	4.088 (4)	0.976	0.000	0.973	0.046
Strict	79.479 (23)	< 0.001	16.200 (5)*	0.953	−0.009	0.959	0.041	65.966 (23)	< 0.001	10.508 (5)	0.972	−0.004	0.976	0.046
Age (< 20 years vs ≥ 20 years)														
Configural	49.244 (10)	< 0.001		0.967		0.934	0.025	41.665 (10)	< 0.001		0.979		0.958	0.018
Metric	57.824 (14)	< 0.001	8.580 (4)	0.963	−0.004	0.947	0.051	49.650 (14)	< 0.001	7.985 (4)	0.976	−0.003	0.966	0.043
Scalar	70.454 (18)	< 0.001	12.630 (4)*	0.956	−0.007	0.951	0.058	54.244 (18)	< 0.001	4.594 (4)	0.976	0.000	0.973	0.046
Strict	74.182 (23)	< 0.001	3.728 (5)	0.957	0.001	0.963	0.054	59.088 (23)	< 0.001	4.844 (5)	0.976	0.000	0.979	0.046
Major (clinical medicine vs preventive medicine)														
Configural	46.837 (10)	< 0.001		0.968		0.936	0.026	42.164 (10)	< 0.001		0.979		0.958	0.019
Metric	57.528 (14)	< 0.001	10.691 (4)*	0.962	−0.006	0.946	0.060	58.085 (14)	< 0.001	15.921 (4)	0.971	−0.008	0.958	0.056
Scalar	63.445 (18)	< 0.001	5.917 (4)	0.961	−0.001	0.956	0.064	70.981 (18)	< 0.001	12.895 (4)	0.965	−0.006	0.961	0.063
Strict	69.586 (23)	< 0.001	6.140 (5)	0.960	−0.001	0.965	0.062	78.824 (23)	< 0.001	7.844 (5)	0.963	−0.002	0.968	0.059
Home location (urban vs rural vs suburban)														
Configural	52.121 (15)	< 0.001		0.969		0.938	0.028	47.354 (15)	< 0.001		0.979		0.958	0.020
Metric	56.444 (23)	< 0.001	4.323 (8)	0.972	0.003	0.963	0.042	50.249 (23)	0.001	2.895 (8)	0.982	0.003	0.977	0.031
Scalar	71.353 (31)	< 0.001	14.909 (8)	0.966	−0.006	0.967	0.051	57.660 (31)	0.003	7.411 (8)	0.983	0.001	0.983	0.037
Strict	85.508 (41)	< 0.001	14.155 (10)	0.963	−0.003	0.973	0.056	69.440 (41)	0.004	11.780 (10)	0.981	−0.002	0.986	0.036
Threshold	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080

(*Continued*)

https://doi.org/10.2147/PRBM.S437219 DovePress

Table 3 (Continued).

Hypothesis	Baseline							Follow-up						
	χ^2 (*df*)	*P*	$\Delta\chi^2$ (Δ*df*)	CFI	ΔCFI	TLI	SRMR	χ^2 (*df*)	*P*	$\Delta\chi^2$ (Δ*df*)	CFI	ΔCFI	TLI	SRMR
Being only child (yes vs no)														
Configural	55.642 (10)	< 0.001		0.962		0.923	0.027	45.525 (10)	< 0.001		0.977		0.954	0.019
Metric	57.227 (14)	< 0.001	1.585 (4)	0.964	0.002	0.948	0.031	55.883 (14)	< 0.001	10.358 (4)	0.973	−0.004	0.961	0.051
Scalar	58.106 (18)	< 0.001	0.879 (4)	0.966	0.002	0.962	0.032	58.072 (18)	< 0.001	2.189 (4)	0.974	0.001	0.971	0.052
Strict	59.867 (23)	< 0.001	1.760 (5)	0.969	0.003	0.973	0.032	63.610 (23)	< 0.001	5.539 (5)	0.974	0.000	0.977	0.048
Monthly household income (<10,000 CNY vs ≥10,000 CNY)														
Configural	52.339 (10)	< 0.001		0.964		0.928	0.028	42.364 (10)	< 0.001		0.979		0.958	0.019
Metric	54.249 (14)	< 0.001	1.911 (4)	0.966	0.002	0.951	0.035	46.149 (14)	< 0.001	3.785 (4)	0.979	0.000	0.970	0.036
Scalar	59.468 (18)	< 0.001	5.219 (4)	0.965	−0.001	0.961	0.038	49.460 (18)	< 0.001	3.311 (4)	0.979	0.000	0.977	0.038
Strict	60.943 (23)	< 0.001	1.475 (5)	0.968	0.003	0.972	0.037	52.133 (23)	< 0.001	2.673 (5)	0.981	0.002	0.983	0.038
Part-time job (yes vs no)														
Configural	47.502 (10)	< 0.001		0.968		0.937	0.026	37.583 (10)	< 0.001		0.982		0.964	0.018
Metric	53.590 (14)	< 0.001	6.088 (4)	0.966	−0.002	0.952	0.038	39.405 (14)	< 0.001	1.822 (4)	0.983	0.001	0.976	0.022
Scalar	54.664 (18)	< 0.001	1.075 (4)	0.969	0.003	0.966	0.038	39.965 (18)	0.002	0.559 (4)	0.986	0.003	0.984	0.023
Strict	56.380 (23)	< 0.001	1.716 (5)	0.972	0.003	0.975	0.037	43.980 (23)	0.005	4.015 (5)	0.986	0.000	0.988	0.023
Physical exercise (yes vs no)														
Configural	48.218 (10)	< 0.001		0.967		0.934	0.027	45.160 (10)	< 0.001		0.977		0.953	0.018
Metric	49.586 (14)	< 0.001	1.368 (4)	0.969	0.002	0.956	0.031	46.897 (14)	< 0.001	1.737 (4)	0.978	0.001	0.969	0.023
Scalar	55.063 (18)	< 0.001	5.477 (4)	0.968	−0.001	0.965	0.034	50.420 (18)	< 0.001	3.523 (4)	0.979	0.001	0.976	0.025
Strict	58.785 (23)	< 0.001	3.722 (5)	0.969	0.001	0.973	0.036	55.071 (23)	< 0.001	4.651 (5)	0.979	0.000	0.982	0.028
Threshold	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080

Hobby (yes vs no)														
Configural	45.443 (10)	< 0.001		0.970		0.939	0.025	41.634 (10)	< 0.001		0.979		0.958	0.018
Metric	54.322 (14)	< 0.001	8.879 (4)	0.966	−0.004	0.951	0.047	45.428 (14)	< 0.001	3.794 (4)	0.979	0.000	0.970	0.030
Scalar	60.352 (18)	< 0.001	6.030 (4)	0.964	−0.002	0.960	0.051	50.697 (18)	< 0.001	5.269 (4)	0.978	−0.001	0.976	0.033
Strict	64.862 (23)	< 0.001	4.511 (5)	0.964	0.000	0.969	0.049	58.747 (23)	< 0.001	8.050 (5)	0.976	−0.002	0.979	0.033
Frequency of visiting home (frequently vs occasionally)														
Configural	59.381 (10)	< 0.001		0.958		0.916	0.030	44.238 (10)	< 0.001		0.978		0.955	0.020
Metric	60.465 (14)	< 0.001	1.084 (4)	0.961	0.003	0.944	0.033	55.843 (14)	< 0.001	11.604 (4)	0.973	−0.005	0.961	0.034
Scalar	67.474 (18)	< 0.001	7.010 (4)	0.958	−0.003	0.953	0.037	63.186 (18)	< 0.001	7.343 (4)	0.970	−0.003	0.967	0.036
Strict	74.911 (23)	< 0.001	7.437 (5)	0.956	−0.002	0.962	0.037	64.914 (23)	< 0.001	1.728 (5)	0.972	0.002	0.976	0.038
Stress coping strategy (emotion-focused vs solution-focused vs avoidance coping)														
Configural	64.714 (15)	< 0.001		0.957		0.914	0.030	56.192 (15)	< 0.001		0.973		0.945	0.021
Metric	72.264 (23)	< 0.001	7.550 (8)	0.958	0.001	0.945	0.047	61.663 (23)	< 0.001	5.471 (8)	0.974	0.001	0.967	0.040
Scalar	84.429 (31)	< 0.001	12.165 (8)	0.954	−0.004	0.955	0.052	72.893 (31)	< 0.001	11.230 (8)	0.972	−0.002	0.973	0.044
Strict	97.831 (41)	< 0.001	13.402 (10)	0.951	−0.003	0.964	0.060	86.189 (41)	< 0.001	13.296 (10)	0.970	−0.002	0.978	0.044
Threshold	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080	N/A	> 0.050	N/A	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.080

Note: *$P < 0.05$.

Abbreviations: WHO-5-C, Chinese World Health Organization-Five Well-Being Index; χ^2, chi-square; *df*, degrees of freedom; CFI, comparative fit index; TLI, Tucker-Lewis index; SRMR, standardized root mean square residual; Δ, change in χ^2, *df*, and CFI; N/A, not applicable.

https://doi.org/10.2147/PRBM.S437219

CFI and TLI values greater than 0.900 and SRMR values below 0.080. Meanwhile, MGCFA tests demonstrated that all ΔCFI values (ΔCFI = −0.009–0.003) were within the recommended range, indicating that the WHO-5-C had acceptable measurement invariance across different sociodemographic variables.

In LCFA tests of the WHO-5-C, all fit indices were in line with the proposed thresholds (CFI = 0.974–0.980, TLI = 0.968–0.976, REMSA = 0.070–0.081, SRMR = 0.022–0.027), and there were no substantial CFI changes in each nested model (ΔCFI = −0.006–0.000), supporting that the WHO-5-C had good longitudinal measurement invariance (Table 4).

Convergent Validity

As shown in Figure 1, inter-item coefficients of the WHO-5-C ranged from 0.535 to 0.844 and item-total coefficients ranged from 0.806 to 0.935, indicating moderate to super strong correlations. Moderate correlations were observed between the WHO-5-C and SRHQ (Time 1: $r = -0.561$, Time 2: $r = -0.573$), as well as the PHQ-4-C (Time 1: $r = -0.691$, Time 2: $r = -0.675$), demonstrating adequate convergent validity.

Reliability

Cronbach's α and McDonald's ω coefficients of the WHO-5-C at baseline were 0.907 (range of α-if-item-deleted = 0.865–0.905) and 0.908 (range of α-if-item-deleted = 0.865–0.905), respectively; and at follow-up were 0.934 (range of α-if-item-deleted = 0.906–0.930) and 0.935 (range of α-if-item-deleted = 0.904–0.931), respectively. The WHO-5-C demonstrated excellent test–retest reliability, with an ICC of 0.803 (range of ICC-if-item-deleted = 0.785–0.799), and stability of the scale was also demonstrated by SEM indices. No significant increase in indices was observed when any item was deleted. Similarly, the SRHQ and PHQ-4-C showed good internal consistencies and test–retest reliabilities (Table 5).

Sensitivity and Specificity

Based on baseline data, ROC curve analysis indicated that the WHO-5-C had an AUC value of 0.882, and the optimal cutoff point of ≥ 50 had sensitivity of 0.782 and specificity of 0.857 for screening satisfactory SWB in healthcare students (Figure 2). Sensitivity and specificity of the WHO-5-C were 0.803 and 0.762, respectively, in follow-up data when the cutoff point of 50 was applied as the threshold for predicting SWB (Table 6).

Discussion

The adaptation of existing well-being instruments in Chinese populations mostly focuses on construct validity and reliability, but rarely responds to measurement invariance and clinical diagnostic capability. As an example, the WEMWBS is an instrument of well-being considering positive aspects of mental health, and despite the Chinese version representing stable factor structure and reliability across populations, its measurement invariance and truncation value are not available.[23,82,83] Herein, the current study comprehensively evaluated measurement properties of the WHO-5-C in healthcare students, intending to introduce a well-being measuring tool with all-round persuasiveness. To our knowledge, this is the first study to explore measurement invariance of the WHO-5-C and determine a cutoff point for SWB identification in healthcare students. We provide strong evidence of a one-factor model with adequate validity and reliability, as well as measurement invariance, explaining its stability across different sociodemographic variables and time intervals. ROC analysis supports the WHO-5-C as a sensitive and specific instrument in assessing positive mental and physical well-being, with a recommended cutoff point of ≥ 50 indicating adequate SWB for Chinese healthcare students. Overall, our findings highlight effectiveness of the WHO-5-C and its validity in assessing SWB among healthcare students.

Structural Validity

CFA results corroborated general consensus of previous studies that the WHO-5 had a stable single-factor structure, which has also been found in most language versions: Brazilian,[30] Japanese,[34] Polish,[36] Sinhala,[37] Spanish,[38] Swedish,[40] and Turkish.[41] The WHO-5-C had adequate structural validity, and these findings were also found in another study in multiple samples.[43] In summary, our study adds to the growing body of evidence supporting structural validity of the WHO-5.

Table 4 Longitudinal Measurement Invariances of the WHO-5-C (N = 343)

Hypothesis	χ^2 (*df*)	*P*	Scaled Chi-Square Difference Test Statistics		CFI	ΔCFI	TLI	ΔTLI	RMSEA (90% CI)	ΔRMSEA	SRMR	ΔSRMR
			$\Delta\chi^2$ (Δ*df*)	*P*								
Configural	94.419 (29)	< 0.001			0.980		0.968		0.081 (0.063, 0.100)		0.022	
Metric	95.342 (33)	< 0.001	0.923 (4)	0.921	0.980	0.000	0.973	0.005	0.074 (0.057, 0.092)	−0.007	0.024	0.002
Scalar	99.464 (37)	< 0.001	4.122 (4)	0.390	0.980	0.000	0.976	0.003	0.070 (0.054, 0.087)	−0.004	0.025	0.001
Strict	125.095 (42)	< 0.001	25.631 (5)	< 0.001	0.974	−0.006	0.972	−0.004	0.076 (0.061, 0.091)	0.006	0.027	0.002
Threshold	N/A	> 0.050	N/A	> 0.050	≥ 0.900	≤ 0.010	≥ 0.900	≤ 0.010	≤ 0.080	≤ 0.015	≤ 0.080	≤ 0.030

Abbreviations: WHO-5-C, Chinese World Health Organization-Five Well-Being Index; χ^2, chi-square; *df*, degrees of freedom; CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation; CI, confidence interval; SRMR, standardized root mean square residual; Δ, change in χ^2, *df*, CFI, TLI, RMSEA, and SRMR; N/A, not applicable.

https://doi.org/10.2147/PRBM.S437219

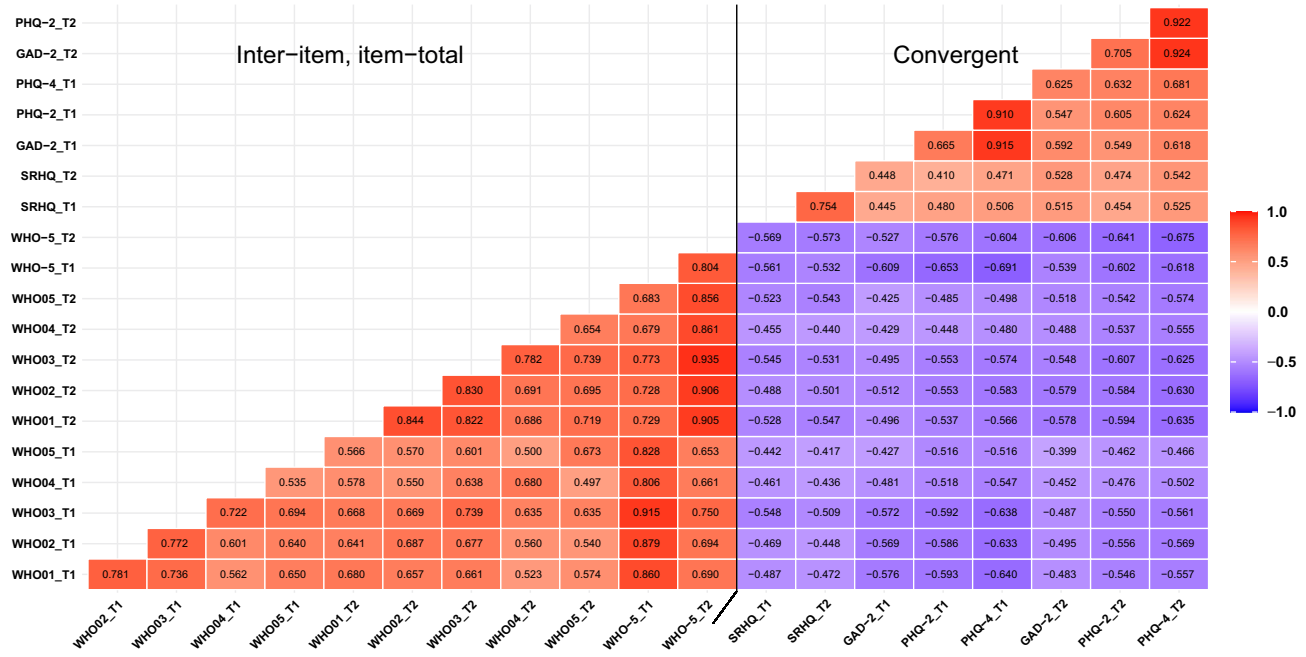

Figure 1 Inter-item and item-total correlations between the Chinese WHO-5, PHQ-4, and SRHQ (N=343).
Abbreviations: WHO-5, World Health Organization-Five Well-Being Index; WHO01-05, item 01–05 of the WHO-5; PHQ-2, Patient Health Questionnaire-2; GAD-2, Generalized Anxiety Disorder-2; PHQ-4, Patient Health Questionnaire-4; SRHQ, Self-Rated Health Questionnaire; T1, Time 1; T2, Time 2.

Measurement Invariance

Our study has validated cross-cultural validity of the WHO-5-C by demonstrating adequate measurement invariance across sociodemographic variables and time intervals. Adequate measurement invariance ensured that the scale had consistent measurement properties in a heterogeneous group and that targeted measurement was stably assessed by the scale.[84] While partial fitting indices of MGCFA tests were contradictory to expectations, likely due to a small degree of freedom,[85,86] all CFI changes among the four progressive nested models were consistently within adequate range, which is recognized as the most pivotal indicator to assess measurement invariance of a scale.[74] Only limited evidence for the WHO-5 measurement invariance was found in extant literature. The Sinhala version supported measurement invariance across gender in all four nested models.[37] Our findings extend previous ones in supporting satisfactory measurement invariance of the WHO-5-C in healthcare students both by different sociodemographic variables and over time.

Convergent Validity

Our findings demonstrated that the WHO-5-C demonstrated moderate negative correlations with both the SRHQ and PHQ-4-C, showing the WHO-5-C had sufficient convergent validity. Associations of comparable strengths with its related constructs of mental and physical health have been reported in community and patient samples,[33,87,88] and special populations like healthcare workers in multi-national studies.[89] Our evidence for construct validity of the WHO-5 in Chinese population adds support to its cross-cultural applicability as a valid instrument for measuring SWB in diverse domains and groups.

Reliability

Similar and consistent with previous studies of other language versions of the WHO-5,[29,30,33,37–39,41,90] Cronbach's α, McDonald's ω and ICC of the WHO-5-C across two separate time points were greater than 0.800 in our study, illustrating its high internal consistency and test–retest reliability. Despite the brevity of the WHO-5-C, our data further support coherence across items and stability over time in Chinese.

Table 5 Internal Consistency and Test–Retest Reliability of the WHO-5-C, SRHQ, and PHQ-4-C (N = 343)

Scales	Cronbach's α		McDonald's ω		ICC (95% CI)	SEM	
	Baseline	Follow-Up	Baseline	Follow-Up		Baseline	Follow-Up
WHO-5-C	0.907 (0.891, 0.922)	0.934 (0.923, 0.945)	0.908 (0.893, 0.924)	0.935 (0.924, 0.946)	0.803 (0.762, 0.838)	7.810	8.209
WHO-5-C_{01}	0.883 (0.863, 0.903)	0.915 (0.900, 0.929)	0.885 (0.865, 0.905)	0.915 (0.900, 0.930)	0.796 (0.754, 0.832)	6.578	6.875
WHO-5-C_{02}	0.877 (0.857, 0.898)	0.914 (0.899, 0.929)	0.881 (0.860, 0.901)	0.915 (0.901, 0.930)	0.799 (0.757, 0.834)	6.394	6.682
WHO-5-C_{03}	0.865 (0.843, 0.888)	0.906 (0.890, 0.921)	0.865 (0.841, 0.888)	0.904 (0.887, 0.921)	0.785 (0.741, 0.822)	6.484	6.820
WHO-5-C_{04}	0.905 (0.889, 0.921)	0.930 (0.918, 0.942)	0.905 (0.889, 0.922)	0.931 (0.918, 0.943)	0.791 (0.748, 0.828)	6.521	6.772
WHO-5-C_{05}	0.898 (0.881, 0.916)	0.930 (0.918, 0.942)	0.900 (0.883, 0.918)	0.931 (0.919, 0.943)	0.798 (0.756, 0.833)	6.347	6.712
SRHQ	0.727	0.805	N/A	N/A	0.749 (0.697, 0.793)	0.637	0.675
PHQ-4-C	0.818 (0.786, 0.849)	0.846 (0.820, 0.872)	0.820 (0.789, 0.851)	0.849 (0.823, 0.875)	0.681 (0.620, 0.734)	1.234	1.263

Note: The McDonald's ω and the 95% CI of Cronbach's α cannot be calculated due to the SRHQ containing only 2 items.

Abbreviations: WHO-5-C, Chinese World Health Organization-Five Well-Being Index; WHO-5-$C_{01\text{-}05}$, item 01–05 of the WHO-5-C was deleted; SRHQ, Self-Rated Health Questionnaire; PHQ-4-C, Chinese Patient Health Questionnaire-4; ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement, calculated as "SD × sqrt (1-ICC)".

https://doi.org/10.2147/PRBM.S437219

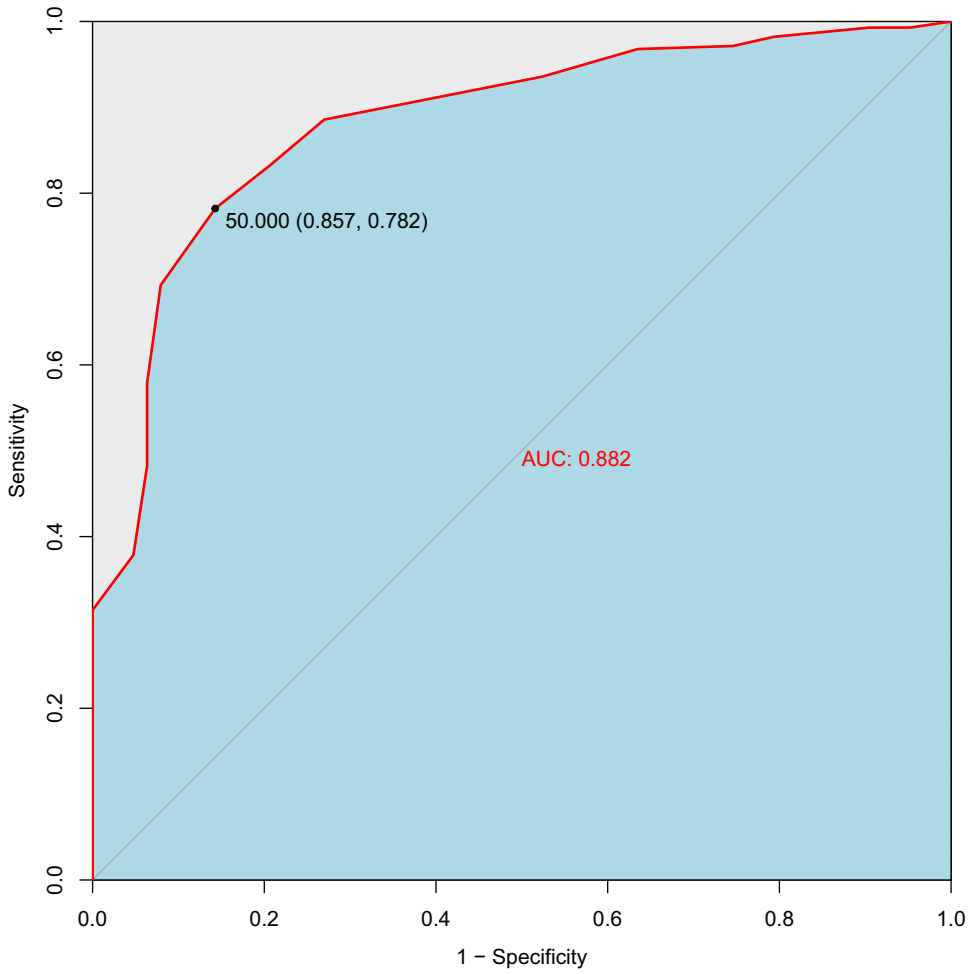

Figure 2 ROC curve of the Chinese WHO-5 for well-being (N=343).
Abbreviations: WHO-5, World Health Organization-Five Well-Being Index; ROC, receiver operating characteristic; AUC, area under the curve.

Sensitivity and Specificity

Our study showed that the WHO-5-C had a reasonable ability to distinguish those who had psychological positive components when the cutoff point was ≥ 50 (AUC = 0.882), suggesting that the scale is a suitable screening instrument for measuring adequate SWB in Chinese healthcare students. Consistent with other studies in Asia, a standard cutoff point of ≥ 13 had an excellent sensitivity/specificity trade-off in the Arabic and Japanese versions, which meant that respondents with positive mindset can be detected by this cutoff criterion.[29,34] Along the same vein, a reduction of 50% indicates reduced wellness in the Brazilian version and recommends that further clinical diagnosis for depression should

Table 6 Sensitivity and Specificity of the WHO-5-C for Identifying Well-Being in Healthcare Students (N = 343)

WHO-5-C	PHQ-2-C		Total	Sensitivity	Specificity
	< 3	≥ 3			
≥ 50	211	13	224	0.803	0.762
< 50	66	53	119		
Total	277	66	343		

Abbreviations: WHO-5-C, Chinese World Health Organization-Five Well-Being Index; PHQ-2-C, Chinese Patient Health Questionnaire-2.

be undertaken.[30] Other cutoff points for different purposes and samples have been identified in other language versions of the WHO-5. The German version of the WHO-5 was applied to screen depression in an elderly population (AUC = 0.886), and a standard cutoff point of ≥ 16 was sufficiently sensitive and specific in predicting a status of optimal well-being.[32] The fact that the definition and feeling of wellness are notably discrepant in different stages of life due to the gains and losses in health assets, social networks, economic resources and family structures.[91,92] A systematic review of the WHO-5, thoroughly, drew a conclusion that the scale had been applied successfully as a generic scale for well-being across numerous fields, and a point below 50 was defined as an obvious reduction in well-being.[28] The WHO-5, briefly, has appropriate validity both as a screening instrument for subjective psychosomatic well-being and as an outcome measurement in clinical trials.

Strengths and Limitations

The present study contributes to literature on utilization of the WHO-5 in Chinese healthcare students by providing empirical findings on its measurement properties and diagnostic performance. The study supports the application of the WHO-5-C in young adults with adequate validity and reliability, laying a solid foundation for its utilization in China. To our knowledge, our study is the first to verify cutoff point and to comprehensively examine measurement invariance of the WHO-5-C among healthcare students in terms of sociological characteristics and measurement times. As such, we provide a precise threshold for evaluating SWB state for clinical and research purposes in young adults, a vulnerable group for onset of psychosomatic health problems and developmental challenges.

This study is admittedly limited in that participants involved were all young healthcare students, rendering it difficult to generalize the findings to other age groups or disciplines. Given the promising findings this study offers, it would be highly valuable to further investigate psychometric properties of the WHO-5-C in various samples or a nationally representative and to validate the optimal cutoff point in different purposes and external samples. Despite widespread utilization of the PHQ-4-C, an instrument of positive SWB should be considered as a complementary measure to capture the comprehensive and multidimensional construct of wellness.[1] Moreover, despite the WHO-5-C is freely available, it is essential to examine its content validity with a view to providing credible evidence for the sustainable implementation of the scale in China.

Conclusion

In conclusion, this study supports the WHO-5-C as a reliable and valid instrument for capturing SWB in healthcare students, with satisfactory measurement invariance in different sociodemographic variables and over time, and a cutoff point of ≥ 50 for identifying significantly adequate SWB state. Together with the family of studies on the WHO-5 worldwide, our evidence enables research and clinical communities to apply the scale for screening well-being in public health and primary care settings.

Data Sharing Statement

Not available.

Acknowledgments

The authors thank the study participants and the research assistants for their time. The authors are indebted to three anonymous reviewers and a handling editor for their insightful views and constructive comments.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the Medical Research Fund of Zhejiang Province, Grant No. 2023RC073 and the Research Initiation Fund of Hangzhou Normal University, Grant No. RWSK20201003.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Beddington J, Cooper CL, Field J, et al. The mental wealth of nations. *Nature*. 2008;455(7216):1057–1060. doi:10.1038/4551057a
2. Faruk MO, Alam F, Chowdhury KUA, Soron TR. Validation of the Bangla WHO-5 Well-being Index. *Glob Ment Health*. 2021;8:e26. doi:10.1017/gmh.2021.26
3. Finan PH, Garland EL. The role of positive affect in pain and its treatment. *Clin J Pain*. 2015;31(2):177–187. doi:10.1097/ajp.0000000000000092
4. Rasmussen HN, Scheier MF, Greenhouse JB. Optimism and physical health: a meta-analytic review. *Ann Behav Med*. 2009;37(3):239–256. doi:10.1007/s12160-009-9111-x
5. Chida Y, Steptoe A. Positive psychological well-being and mortality: a quantitative review of prospective observational studies. *Psychosom Med*. 2008;70(7):741–756. doi:10.1097/PSY.0b013e31818105ba
6. Bolier L, Haverman M, Westerhof GJ, Riper H, Smit F, Bohlmeijer E. Positive psychology interventions: a meta-analysis of randomized controlled studies. *BMC Public Health*. 2013;13:119. doi:10.1186/1471-2458-13-119
7. Kubzansky LD, Huffman JC, Boehm JK, et al. Positive psychological well-being and cardiovascular disease: JACC Health Promotion Series. *J Am Coll Cardiol*. 2018;72(12):1382–1396. doi:10.1016/j.jacc.2018.07.042
8. Howell RT, Kern ML, Lyubomirsky S. Health benefits: meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychol Rev*. 2007;1(1):83–136. doi:10.1080/17437190701492486
9. Liu X, Ping S, Gao W. Changes in undergraduate students' psychological well-being as they experience university life. *Int J Environ Res Public Health*. 2019;16(16):2864. doi:10.3390/ijerph16162864
10. Holm-Hadulla RM, Klimov M, Juche T, Möltner A, Herpertz SC. Well-being and mental health of students during the COVID-19 pandemic. *Psychopathology*. 2021;54(6):291–297. doi:10.1159/000519366
11. Sheldon E, Simmonds-Buckley M, Bone C, et al. Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. *J Affect Disord*. 2021;287:282–292. doi:10.1016/j.jad.2021.03.054
12. Sattar K, Yusoff MSB, Arifin WN, Yasin MAM, Nor MZM. A scoping review on the relationship between mental wellbeing and medical professionalism. *Med Educ Online*. 2023;28(1):2165892. doi:10.1080/10872981.2023.2165892
13. Wang CY, Pan RY, Wan XY, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *Int J Environ Res Public Health*. 2020;17(5):1729. doi:10.3390/ijerph17051729
14. Cao WJ, Fang ZW, Hou GQ, et al. The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry Res*. 2020;287:112934. doi:10.1016/j.psychres.2020.112934
15. Moreira de Sousa J, Moreira CA, Telles-Correia D. Anxiety, depression and academic performance: a study amongst Portuguese medical students versus non-medical students. *Acta Med Port*. 2018;31(9):454–462. doi:10.20344/amp.9996
16. Quek TT, Tam WW, Tran BX, et al. The global prevalence of anxiety among medical students: a meta-analysis. *Int J Environ Res Public Health*. 2019;16(15):2735. doi:10.3390/ijerph16152735
17. Zeng W, Chen R, Wang X, Zhang Q, Deng W. Prevalence of mental health problems among medical students in China: a meta-analysis. *Medicine*. 2019;98(18):e15337. doi:10.1097/md.0000000000015337
18. Cooke PJ, Melchert TP, Connor K. Measuring well-being: a review of instruments. *Couns Psychol*. 2016;44(5):730–757. doi:10.1177/0011000016633507
19. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361–370. doi:10.1111/j.1600-0447.1983.tb09716.x
20. Melzack R. The short-form McGill Pain Questionnaire. *Pain*. 1987;30(2):191–197. doi:10.1016/0304-3959(87)91074-8
21. McIntyre RS, Alda M, Baldessarini RJ, et al. The clinical characterization of the adult patient with bipolar disorder aimed at personalization of management. *World Psychiatry*. 2022;21(3):364–387. doi:10.1002/wps.20997
22. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Out*. 2007;5(1):63. doi:10.1186/1477-7525-5-63
23. Taggart F, Friede T, Weich S, Clarke A, Johnson M, Stewart-Brown S. Cross cultural evaluation of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS)-a mixed methods study. *Health Qual Life Out*. 2013;11:27. doi:10.1186/1477-7525-11-27
24. Diener E. Subjective well-being: the science of happiness and a proposal for a national index. *Am Psychol*. 2000;55(1):34–43. doi:10.1037/0003-066X.55.1.34
25. Smyth R, Nielsen I, Zhai Q. Personal Well-being in urban China. *Soc Indic Res*. 2010;95(2):231–251. doi:10.1007/s11205-009-9457-2
26. Chan DW, Chan L-K, Sun X. Developing a brief version of Ryff's scale to assess the psychological well-being of adolescents in Hong Kong. *Eur J Psychol Assess*. 2019;35(3):414–422. doi:10.1027/1015-5759/a000403
27. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden, 12–13 February 1998. World Health Organization. Regional Office for Europe; 1998. Available from: https://iris.who.int/handle/10665/349766. Accessed January 30, 2021.
28. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom*. 2015;84(3):167–176. doi:10.1159/000376585
29. Sibai AM, Chaaya M, Tohme RA, Mahfoud Z, Al-Amin H. Validation of the Arabic version of the 5-item WHO Well Being Index in elderly population. *Int J Geriatr Psychiatry*. 2009;24(1):106–107. doi:10.1002/gps.2079
30. de Souza CM, Hidalgo MP. World Health Organization 5-item well-being index: validation of the Brazilian Portuguese version. *Eur Arch Psychiatry Clin Neurosci*. 2012;262(3):239–244. doi:10.1007/s00406-011-0255-x

31. Lee GL, Fan GK, Chan SW. Validation of Chinese and English versions of the Holistic Well-being Scale in patients with cancer. *Support Care Cancer*. 2015;23(12):3563–3571. doi:10.1007/s00520-015-2736-3
32. Bonsignore M, Barkow K, Jessen F, Heun R. Validity of the five-item WHO Well-Being Index (WHO-5) in an elderly population. *Eur Arch Psychiatry Clin Neurosci*. 2001;251(Suppl 2):27–31. doi:10.1007/bf03035123
33. Guðmundsdóttir HB, Olason DP, Guðmundsdóttir DG, Sigurðsson JF. A psychometric evaluation of the Icelandic version of the WHO-5. *Scand J Psychol*. 2014;55(6):567–572. doi:10.1111/sjop.12156
34. Awata S, Bech P, Yoshida S, et al. Reliability and validity of the Japanese version of the World Health Organization-Five Well-Being Index in the context of detecting depression in diabetic patients. *Psychiatry Clin Neurosci*. 2007;61(1):112–119. doi:10.1111/j.1440-1819.2007.01619.x
35. Suhaimi AF, Makki SM, Tan KA, Silim UA, Ibrahim N. Translation and validation of the Malay version of the WHO-5 Well-Being Index: reliability and validity evidence from a sample of type 2 diabetes mellitus patients. *Int J Environ Res Public Health*. 2022;19(7):4415. doi:10.3390/ijerph19074415
36. Cichoń E, Kiejna A, Kokoszka A, et al. Validation of the Polish version of WHO-5 as a screening instrument for depression in adults with diabetes. *Diabetes Res Clin Pract*. 2020;159:107970. doi:10.1016/j.diabres.2019.107970
37. Perera BPR, Jayasuriya R, Caldera A, Wickremasinghe AR. Assessing mental well-being in a Sinhala speaking Sri Lankan population: validation of the WHO-5 well-being index. *Health Qual Life Out*. 2020;18(1):305. doi:10.1186/s12955-020-01532-8
38. Lucas-Carrasco R. Reliability and validity of the Spanish version of the World Health Organization-Five Well-Being Index in elderly. *Psychiatry Clin Neurosci*. 2012;66(6):508–513. doi:10.1111/j.1440-1819.2012.02387.x
39. Chongwo E, Ssewanyana D, Nasambu C, et al. Validation of a Swahili version of the World Health Organization 5-item well-being index among adults living with HIV and epilepsy in rural coastal Kenya. *Glob Health Res Policy*. 2018;3:26. doi:10.1186/s41256-018-0081-z
40. Löve J, Andersson L, Moore CD, Hensing G. Psychometric analysis of the Swedish translation of the WHO well-being index. *Qual Life Res*. 2014;23(1):293–297. doi:10.1007/s11136-013-0447-0
41. Eser E, Çevik C, Baydur H, et al. Reliability and validity of the Turkish version of the WHO-5, in adults and older adults for its use in primary care settings. *Prim Health Care Res Dev*. 2019;20:e100. doi:10.1017/s1463423619000343
42. de Wit M, Pouwer F, Gemke RJ. Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. *Diabetes Care*. 2007;30 (8):2003–2006. doi:10.2337/dc07-0447
43. Fung SF, Kong CYW, Liu YM, et al. Validity and psychometric evaluation of the Chinese version of the 5-Item WHO Well-Being Index. *Front Public Health*. 2022;10:872436. doi:10.3389/fpubh.2022.872436
44. Chan L, Liu RKW, Lam TP, Chen JY, Tipoe GL, Ganotice FA. Validation of the World Health Organization Well-Being Index (WHO-5) among medical educators in Hong Kong: a confirmatory factor analysis. *Med Educ Online*. 2022;27(1):2044635. doi:10.1080/10872981.2022.2044635
45. Saracci R. The World Health Organisation needs to reconsider its definition of health. *BMJ*. 1997;314(7091):1409–1410. doi:10.1136/bmj.314.7091.1409
46. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford University Press; 2014.
47. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials*. 1991;12(4 Suppl):142S–158S. doi:10.1016/s0197-2456(05)80019-4
48. Kline P. *An Easy Guide to Factor Analysis*. New York: Routledge; 1994.
49. Mundfrom DJ, Shaw DG, Ke TL. Minimum sample size recommendations for conducting factor analyses. *Int J Test*. 2005;5(2):159–168. doi:10.1207/s15327574ijt0502_4
50. Association WM. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–2194. doi:10.1001/jama.2013.281053
51. The Chinese version of the WHO (Five) in characters used in PR China. Psychiatric Research Unit, WHO Collaborating Center for Mental Health. Available from: https://www.psykiatri-regionh.dk/who-5/Documents/WHO5_Chinese_PR.pdf. Accessed January 30, 2021.
52. Spitzer RL, Kroenke JBWWK. Instructions for Patient Health Questionnaire (PHQ) and GAD-7 Measures. Available from: https://www.phqscreeners.com/. Accessed January 30, 2021.
53. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a Primary Care Setting. *Cureus*. 2020;12(5):e8224. doi:10.7759/cureus.8224
54. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care*. 2003;41 (11):1284–1292. doi:10.1097/01.Mlr.0000093487.78664.3c
55. Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord*. 2010;122(1–2):86–95. doi:10.1016/j.jad.2009.06.019
56. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics*. 2009;50 (6):613–621. doi:10.1176/appi.psy.50.6.613
57. Luo Y, Fei S, Gong B, Sun T, Meng R. Understanding the mediating role of anxiety and depression on the relationship between perceived stress and sleep quality among health care workers in the COVID-19 response. *Nat Sci Sleep*. 2021;13:1747–1758. doi:10.2147/nss.S313258
58. Meng R, Dong L, Dzierzewski JM, et al. The RU_SATED as a measure of sleep health: cross-cultural adaptation and validation in Chinese healthcare students. *BMC Psychology*. 2023;11(1):200. doi:10.1186/s40359-023-01203-5
59. Zhu Y, Jiang C, Yang Y, et al. Depression and anxiety mediate the association between sleep quality and self-rated health in healthcare students. *Behav Sci*. 2023;13(2):82. doi:10.3390/bs13020082
60. Jiang C, Ma H, Luo Y, et al. Validation of the Chinese version of the Perceived Stress Scale-10 integrating exploratory graph analysis and confirmatory factor analysis. *Gen Hosp Psychiatry*. 2023;84:194–202. doi:10.1016/j.genhosppsych.2023.07.008
61. Jiang C, Zhu Y, Luo Y, et al. Validation of the Chinese version of the Rosenberg Self-Esteem Scale: evidence from a three-wave longitudinal study. *BMC Psychology*. 2023;11(1):345. doi:10.1186/s40359-023-01293-1
62. Rosseel Y. Lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48:1–36. doi:10.18637/jss.v048.i02
63. Kelley K. MBESS: the MBESS R package. 2022. Available from: https://CRAN.R-project.org/package=MBESS. Accessed January 30, 2023.
64. Gamer M, Lemon J, Singh IFP. irr: various coefficients of interrater reliability and agreement. 2019. Available from: https://CRAN.R-project.org/package=irr. Accessed January 30, 2023.
65. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. Useful tools for structural equation modeling. 2022. Available from: https://CRAN.R-project.org/package=semTools. Accessed January 30, 2023.

66. Sachs MC. plotROC: a tool for plotting ROC curves. *J Stat Softw*. 2017;79(2):1–19. doi:10.18637/jss.v079.c02
67. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biol Med*. 2013;86(3):343–358.
68. Finney SJ, DiStefano C. *Nonnormal and Categorical Data in Structural Equation Models. Structural Equation Modeling: A Second Course*. Greenwich, CT: Information Age; 2006.
69. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–745. doi:10.1016/j.jclinepi.2010.02.006
70. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis: Pearson New International Edition*. 7th ed. London, UK: Pearson Higher Education; 2014.
71. Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York, NY, USA: Guilford Publications; 2016.
72. Edwards MC, Houts CR, Wirth RJ. Measurement invariance, the lack thereof, and modeling change. *Qual Life Res*. 2018;27(7):1735–1743. doi:10.1007/s11136-017-1673-7
73. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling*. 2002;9(2):233–255. doi:10.1207/S15328007SEM0902_5
74. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev*. 2016;41:71–90. doi:10.1016/j.dr.2016.06.004
75. Doan T, Ha V, Strazdins L, Chateau D. Healthy minds live in healthy bodies - effect of physical health on mental health: evidence from Australian longitudinal data. *Curr Psychol*. 2022;42:18702–18713. doi:10.1007/s12144-022-03053-7
76. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. 2018;126(5):1763–1768. doi:10.1213/ane.0000000000002864
77. DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007;39(2):155–164. doi:10.1111/j.1547-5069.2007.00161.x
78. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
79. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231–240. doi:10.1519/15184.1
80. Walter SD. The partial area under the summary ROC curve. *Stat Med*. 2005;24(13):2025–2040. doi:10.1002/sim.2103
81. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*. 1993;39(4):561–577. doi:10.1093/clinchem/39.4.561
82. Ng SS, Lo AW, Leung TK, et al. Translation and validation of the Chinese version of the short Warwick-Edinburgh Mental Well-being Scale for patients with mental illness in Hong Kong. *East Asian Arch Psychiatry*. 2014;24(1):3–9.
83. Dong A, Chen X, Zhu L, et al. Translation and validation of a Chinese version of the Warwick–Edinburgh Mental Well-being Scale with undergraduate nursing trainees. *J Psychiatr Ment Health Nurs*. 2016;23(9–10):554–560. doi:10.1111/jpm.12344
84. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Modeling*. 2007;14(3):464–504. doi:10.1080/10705510701301834
85. Lai K, Green SB. The problem with having two watches: assessment of fit when RMSEA and CFI disagree. *Multivariate Behav Res*. 2016;51(2–3):220–239. doi:10.1080/00273171.2015.1134306
86. Shi D, DiStefano C, Maydeu-Olivares A, Lee T. Evaluating SEM model fit with small degrees of freedom. *Multivariate Behav Res*. 2022;57(2–3):179–207. doi:10.1080/00273171.2020.1868965
87. Hansen CP, Amiri M. Combined detection of depression and anxiety in epilepsy patients using the Neurological Disorders Depression Inventory for Epilepsy and the World Health Organization well-being index. *Seizure*. 2015;33:41–45. doi:10.1016/j.seizure.2015.10.008
88. Mergl R, Seidscheck I, Allgaier AK, Möller HJ, Hegerl U, Henkel V. Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. *Depress Anxiety*. 2007;24(3):185–195. doi:10.1002/da.20192
89. Lara-Cabrera ML, Betancort M, Muñoz-Rubilar A, Rodríguez-Novo N, Bjerkeset O, De Las Cuevas C. Psychometric properties of the WHO-5 Well-Being Index among nurses during the COVID-19 pandemic: a cross-sectional study in three countries. *Int J Environ Res Public Health*. 2022;19(16):10106. doi:10.3390/ijerph191610106
90. Awata S, Bech P, Koizumi Y, et al. Validity and utility of the Japanese version of the WHO-Five Well-Being Index in the context of detecting suicidal ideation in elderly community residents. *Int Psychogeriatr*. 2007;19(1):77–88. doi:10.1017/s1041610206004212
91. Mirowsky J, Ross CE. Age and depression. *J Health Soc Behav*. 1992;33(3):187–205. doi:10.2307/2137349
92. Stone AA, Schwartz JE, Broderick JE, Deaton A. A snapshot of the age distribution of psychological well-being in the United States. *PNAS*. 2010;107(22):9985–9990. doi:10.1073/pnas.1003744107