

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/189389/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Piolanti, Antonio, Glasauer, Christina, Shenderovich, Yulia , Wienand, Dennis, Simon, Judit, Shimbov, Bojan and Foran, Heather M. 2026. Ordinal scaling properties of the WHO-5 among young adolescents in Moldova and North Macedonia: a Mokken Scale Analysis. Health and Quality of Life Outcomes 10.1186/s12955-026-02624-7

Publishers page: <https://doi.org/10.1186/s12955-026-02624-7>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Ordinal scaling properties of the WHO-5 among young adolescents in Moldova and North Macedonia: a Mokken Scale Analysis

Received: 19 May 2026

Accepted: 23 August 2026

Published online: 28 August 2026

Cite this article as: Piolanti A., Glasauer C., Shenderovich Y. *et al.* Ordinal scaling properties of the WHO-5 among young adolescents in Moldova and North Macedonia: a Mokken Scale Analysis. *Health Qual Life Outcomes* (2026). <https://doi.org/10.1186/s12955-026-02624-7>

Antonio Piolanti, Christina Glasauer, Yulia Shenderovich, Dennis Wienand, Judit Simon, Bojan Shimbov & Heather M. Foran

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Ordinal Scaling Properties of the WHO-5 Among Young Adolescents in Moldova and North Macedonia: A Mokken Scale Analysis

Antonio Piolanti¹, Christina Glasauer¹, Yulia Shenderovich^{2,3}, Dennis Wienand⁴, Judit Simon^{4,5}, Bojan Shimbov⁶ & Heather M. Foran¹

Affiliations:

1. Department of Health Psychology, University of Klagenfurt, Universitätsstraße 65-67, 9020 Klagenfurt, Austria.
2. Centre for Development, Evaluation, Complexity and Implementation in Public Health Improvement (DECIPHer), School of Social Sciences, Cardiff University, Maindy Road, Cathays, Cardiff, CF24 4HQ, UK.
3. Wolfson Centre for Young People's Mental Health, Cardiff University, Maindy Road, Cathays, Cardiff, CF24 4HQ, UK.
4. Department of Health Economics, Center for Public Health, Medical University of Vienna, Kinderspitalgasse 15, 1090 Vienna, Austria.
5. Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK.
6. Institute of International Economics & Department of Economics, University Jaume I, Castellón de la Plana, Spain

Corresponding author:

Antonio Piolanti

Health Psychology Unit, Institute of Psychology, University of Klagenfurt, Universitätsstraße 65-67, 9020 Klagenfurt, Austria

E-mail: antonio.piolanti@aau.at

Telephone: +43 463 2700 1676

Background: The World Health Organization Five Well-Being Index (WHO-5) is commonly applied to assess subjective well-being. However, evidence on its ordinal measurement properties among young adolescents, particularly in low- and middle-income countries, is scant. In particular, there is a lack of studies that have comprehensively evaluated key Mokken model assumptions, including scalability, monotonicity and invariant item ordering. Evaluating these properties is important as they support the ordinal interpretation of WHO-5 scores and clarify whether the items form an invariant hierarchy.

Methods: Baseline data from the FLOURISH project, a multi-country study evaluating a family-focused prevention program in Moldova and North Macedonia, were used. The sample comprised 653 adolescents aged 10–14 years. The ordinal scaling properties of the WHO-5 were examined using Mokken Scale Analysis, a non-parametric item response theory approach. Specifically, overall and item-level scalability were evaluated using Loevinger's H coefficients. Single monotonicity was examined under the Monotone Homogeneity Model, while invariant item ordering under the Double Monotonicity Model was assessed using nonintersection diagnostics and the manifest invariant item ordering method. The accuracy of the estimated ordering was evaluated using the H^T coefficient. Analyses were also conducted separately by gender and country.

Results: The WHO-5 demonstrated strong overall scalability ($h = 0.53$, 95% CI: 0.49–0.57), with item-level scalability coefficients ranging from 0.50 to 0.56. Analyses revealed only minor and infrequent deviations from single monotonicity, with all diagnostic values below established thresholds. In addition, no meaningful violations of the nonintersection assumption were identified in the overall analysis. However, the H^T coefficient was 0.12 and remained below 0.30 across the gender- and country-specific analyses, indicating that the estimated item ordering was too inaccurate to be useful.

Conclusions: The findings corroborate evidence for the scalability and monotonicity of the WHO-5 and support the use of its summed score to distinguish adolescents with relatively lower or higher well-being at a given assessment, while cautioning against interpreting the group-level item ordering as a hierarchy that is accurately reproduced in individual adolescents' response patterns.

Keywords: *WHO-5, well-being, adolescents, psychometrics, Mokken Scale Analysis, Eastern Europe.*

Background

The World Health Organization Five Well-Being Index (WHO-5) is one of the most widely used instruments for the assessment of subjective well-being in population-based and clinical research.¹ The scale is a brief self-report measure designed to capture central aspects of positive health, such as vitality and engagement with daily life. Although the WHO-5 was originally developed for adult populations, its application has increasingly extended to adolescent samples. This extension is particularly relevant because early adolescence involves significant changes in emotional regulation, self-concept, and interpersonal relationships that may adversely influence well-being

and positive development.² Furthermore, research has shown that the COVID-19 pandemic placed a substantial burden on adolescents' well-being, particularly in low- and middle-income countries (LMICs)^{3,4}, where limited resources constrain response capacity.⁵ In these contexts, the promotion of adolescent well-being through preventive and early intervention approaches has become a key public health priority. Given its brevity and cost-free availability, the WHO-5 represents a practical instrument for assessing subjective well-being levels in low-resource and implementation settings.⁶ Consequently, ensuring sound psychometric properties of the scale among young adolescents is fundamental to support valid interpretation of well-being scores.

Existing psychometric studies in adolescent populations have largely relied on classical test theory and have supported good internal consistency and a unidimensional structure of the WHO-5.⁷⁻¹⁵ However, these approaches provide limited information regarding the ordinal structure of the instrument. Non-parametric item response theory approaches, such as Mokken Scale Analysis (MSA), are well suited to address this limitation.¹⁶ MSA evaluates whether items satisfy key ordinal assumptions, including scalability (hierarchical ordering along a single dimension), monotonicity (higher trait levels correspond to higher response probabilities), and invariant item ordering (stable relative ordering across the trait continuum).^{16,17} Establishing these properties is important for assessing well-being with the WHO-5 in both research and clinical practice, as they indicate whether total scores can be used to order individuals according to their relative well-being and whether individual items form a consistent hierarchy.

To date, only a few studies have applied MSA to the WHO-5 in adolescent populations. One study involved bilingual ethnic-minority adolescents in Norway and Denmark⁷, another involved adolescents receiving psychiatric care in Sweden⁸, while a third included a broad school-based adolescent sample in Luxembourg⁹. All three studies reported strong scalability, while the Luxembourg study also found no meaningful violations of single monotonicity. The current evidence therefore supports the scalability of the WHO-5 in adolescents. However, evidence concerning the broader set of MSA assumptions, particularly invariant item ordering, remains limited among young adolescents and in lower-resource settings. Accordingly, the present study evaluates scalability, monotonicity and invariant item ordering among young adolescents from two LMICs in Southeastern Europe, Moldova and North Macedonia.

Methods

Study setting and participants

The present analysis uses baseline data from Phase 2 of the FLOURISH project, a multi-country trial implemented in the Republic of North Macedonia and the Republic of Moldova.¹⁸ The FLOURISH project evaluates a family-based intervention aimed at improving adolescent mental health and well-being in diverse community settings.

The sample included adolescents aged 10–14 years who participated in the study alongside a primary caregiver. In North Macedonia, participants were recruited through collaborating schools. In Moldova, recruitment was conducted via youth-friendly primary healthcare services and schools across the country, covering broad catchment areas. Recruitment was open to all adolescents; however, a proportionate universalism approach was adopted, whereby all families were eligible but additional efforts were made to engage more vulnerable groups. Eligibility criteria required adolescents to reside with the participating caregiver for a minimum of four nights per week and to have sufficient proficiency in the local language (Macedonian or Romanian) to complete study questionnaires. Adolescents were excluded through a self-reported exclusionary statement if they had severe psychiatric, cognitive, or physical conditions that could substantially interfere with study participation.

Baseline data were collected prior to intervention delivery using tablet-based, computer-assisted self-report questionnaires. Trained assessors administered the assessments in private settings to ensure confidentiality, while remaining available to provide assistance when necessary.

Measure of adolescent well-being

Adolescent well-being was assessed using the World Health Organization Five Well-Being Index (WHO-5), a self-report scale designed to capture subjective well-being over the preceding two weeks.¹⁹ The WHO-5 includes five positively phrased items assessing key components of well-being. Responses are recorded on a six-point scale ranging from 0 (“At no time”) to 5 (“All of the time”), with higher scores reflecting higher levels of well-being.

For use in the present study, the WHO-5 was administered in Macedonian and Romanian. The language versions were developed through a comprehensive cultural adaptation process described elsewhere.²⁰ Briefly, this process involved forward and backward translation by bilingual mental health professionals, evaluation of semantic and conceptual equivalence, panel review and pilot testing with adolescents in both countries. In addition, qualitative think-aloud interviews were conducted to assess clarity and interpretability of the items, leading to final versions that were linguistically accessible and culturally appropriate for young adolescents.²⁰

Statistical Analysis

The ordinal scaling properties of the WHO-5 were examined using MSA.²¹ MSA assesses key assumptions related to item scalability and monotonicity, thereby evaluating the extent to which observed response patterns support an ordinal interpretation of the total score.

Scalability was evaluated using Loevinger's H coefficients, which quantify the degree to which item responses conform to a cumulative ordering structure.²² Specifically, scalability was assessed at two levels: the item level (h_i) and the overall scale level (h). Loevinger's coefficients were derived by comparing the number of observed Guttman errors, defined as instances in which a respondent endorses a more difficult item but not a less difficult item within an item pair, with the number of errors expected under item independence.^{23,24} At the item level, h_i summarizes the scalability of a single item by comparing the number of observed and expected errors across all item pairs involving that item. At the scale level, the overall scalability coefficient h summarizes scalability across all items and is based on the comparison between the total number of observed and expected Guttman errors across all item pairs. Larger Loevinger's coefficients indicate fewer observed errors relative to chance and, consequently, stronger ordinal scalability. Scalability was interpreted using established convention thresholds, with values below 0.30 indicating poor scalability, between 0.30 and 0.39 reflecting weak scalability, values between 0.40 and 0.49 indicating moderate scalability and values of 0.50 or higher representing strong scalability.^{21,25} Sampling uncertainty around the overall and item-level scalability coefficients was quantified using 95% confidence intervals, calculated from the corresponding standard errors.

Monotonicity was examined within the framework of two hierarchical models of MSA, namely the Monotone Homogeneity Model (MHM) and the more restrictive Double Monotonicity Model (DMM). The MHM is defined by three core assumptions: unidimensionality, indicating that all items reflect a single underlying construct; local independence, whereby item responses are independent conditional on the latent trait; and single monotonicity, such that the probability of endorsing an item at or above a given response category does not decrease with increasing levels of the latent trait.^{16,21} The DMM builds on the MHM by incorporating invariant item ordering (IIO), which requires that item response functions do not intersect, resulting in a consistent ordering of items by difficulty across the latent continuum.^{16,26}

In the present study, we examined only single monotonicity and IIO because a prior study²⁷ using the same analytic sample had already provided support for the unidimensional structure of the

WHO-5 and the local independence of its items. Specifically, the one-factor model showed adequate fit (CFI = .995, TLI = .990, RMSEA = .069, SRMR = .019), with standardized factor loadings ranging from 0.70 to 0.81 and residual variances from 0.34 to 0.51. Inspection of modification indices and expected parameter changes did not identify any substantial residual correlations, providing no indication of local dependence. Multi-group CFA also supported configural, threshold, metric and scalar invariance across Moldova and North Macedonia. The configural (CFI = .996, RMSEA = .063), threshold (CFI = .995, RMSEA = .045), metric (CFI = .996, RMSEA = .036) and scalar models (CFI = .995, RMSEA = .039) showed adequate fit. Relative to the preceding model, changes in CFI were $-.001$, $.001$ and $-.001$, while changes in RMSEA were $-.018$, $-.009$ and $.003$ for the threshold, metric and scalar models, respectively. These changes met prespecified criteria, and the corresponding scaled chi-square difference tests were also nonsignificant. Overall, these findings supported unidimensionality, provided no indication of local dependence and supported cross-country measurement invariance, justifying the present focus on single monotonicity and IIO.

Monotonicity-related assumptions were evaluated using item-level diagnostics. Potential violations of single monotonicity were identified through matrices summarizing local decreases in item response functions, whereas potential violations of IIO were assessed using the manifest invariant item ordering (MIIO) method, which examines pairwise intersections between manifest item response functions.²⁵ The extent and relevance of detected violations were assessed by computing *Crit* coefficients for each item. These coefficients combine information on item scalability, the number and relative frequency of violations, their maximum and cumulative magnitudes and their statistical significance.²⁸ Following established guidelines, *Crit* values below 40 were interpreted as indicating no meaningful violations, values between 40 and 80 as warranting further scrutiny, and values above 80 as reflecting substantial violations of monotonicity assumptions.

After the MIIO analysis, the accuracy of the estimated item ordering was summarized using the H^T coefficient. H^T is an overall coefficient expressing the extent to which the ordering of the item totals is reflected in individual respondents' item-score patterns.²⁹ Thus, whereas MIIO evaluates whether item response functions intersect, H^T indicates how accurately the resulting invariant item ordering can be estimated. Low values occur when the item response functions lie close together, producing a less accurate ordering, and higher values occur when the functions are farther apart,

producing a more stable and accurate ordering. Using proposed thresholds, values below 0.30 indicate an item ordering that is too inaccurate to be useful, values from 0.30 to below 0.40 indicate low accuracy, values from 0.40 to below 0.50 indicate medium accuracy, and values of 0.50 or higher indicate high accuracy.²⁹

Analyses were conducted for the overall sample and separately by country and gender. All analyses were conducted in R version 4.5.2 using the mokken package, version 3.1.2. WHO-5 items were treated as ordinal variables with six ordered response categories (0–5).

Results

The study sample included 653 adolescents with a mean age of 11.9 years ($SD = 1.29$; range = 10–14 years), including 322 participants from Moldova and 331 from North Macedonia. The gender distribution was balanced, with 330 participants identifying as male (50.5%), 321 as female (49.2%) and two as diverse (0.3%). The vast majority of adolescents were currently attending school (98.8%). With respect to ethnicity, most participants identified as Macedonian (48.5%) or Moldovan (45.5%), followed by Romanian (3.2%); all other ethnic groups each represented less than 2% of the sample. Religious affiliation was predominantly Eastern Orthodox (87.7%).

Table 1 presents the response-category distributions for the five WHO-5 items. There were no missing responses for any item. Across all items, responses were concentrated in the higher categories, with “Most of the time” or “All of the time” selected by 61.4% of adolescents for feeling cheerful, 51.3% for feeling calm, 54.7% for feeling active, 41.8% for feeling rested, and 52.1% for having an interesting daily life. The item concerning waking up feeling fresh and rested showed the greatest dispersion across response categories and the highest endorsement of the lower response categories.

Table 2 reports the scalability coefficients and monotonicity diagnostics for the WHO-5. The overall scalability of the scale was $h = 0.53$. Item-level scalability coefficients ranged from $h_i = 0.50$ to 0.56. Specifically, item scalability coefficients were 0.52 for (1) “*I have felt cheerful and in good spirits*”, 0.53 for (2) “*I have felt calm and relaxed*”, 0.56 for (3) “*I have felt active and vigorous*”, 0.52 for (4) “*I woke up feeling fresh and rested*”, and 0.50 for (5) “*My daily life has been filled with things that interest me*”.

With respect to single monotonicity, two items showed no violations, two items showed one violation each and one item showed three violations. Corresponding *Crit* values ranged from 0 to 14. For double monotonicity, no violations were observed for any item, with *Crit* values of 0 for

all items. The H^T coefficient was 0.12. The estimated item ordering, from the highest to the lowest mean item score, was Cheerful, Active, Interested, Calm and Rested. Diagnostic plots for IIO are presented in Supplementary Figure S1.

Gender-specific analyses are reported in Tables 3 and 4. Among male adolescents, the WHO-5 demonstrated moderate overall scalability ($h = 0.44$, 95% CI: 0.38–0.51), with item-level coefficients ranging from 0.39 to 0.49. Among female adolescents, overall scalability was strong ($h = 0.60$, 95% CI: 0.55–0.66), with item-level coefficients ranging from 0.57 to 0.64. Single-monotonicity violations were infrequent in the male subgroup and absent in the female subgroup and corresponding *Crit* values were below 40. Similarly, although some violations of double monotonicity were identified in both subgroups, all corresponding *Crit* values were below 40. The H^T coefficients were 0.10 for male and 0.15 for female adolescents. The estimated ordering also differed between the two groups: Active, Cheerful, Calm, Interested, and Rested among male adolescents; and Cheerful, Interested, Active, Calm, and Rested among female adolescents.

Country-specific analyses are reported in the Supplementary Material (Tables S1 and S2) and are fully consistent with these findings.

Discussion

The present investigation examined the ordinal scaling performance of the WHO-5 Well-Being Index among young adolescents in a prevention-oriented sample from Southeastern Europe using Mokken Scale Analysis. This study extends previous adolescent Mokken research by jointly evaluating scalability, single monotonicity and invariant item ordering in a younger sample and a geographical and cultural context in which these properties have not previously been examined. Overall, the findings suggest that the instrument exhibits a coherent ordinal structure in this context, with consistent item scalability and no meaningful violations of single monotonicity or IIO, although the estimated invariant item ordering was insufficiently accurate to be useful.

Specifically, the WHO-5 showed a strong overall scalability coefficient, indicating that the five items form a well-defined ordinal continuum of adolescent well-being. Item-level coefficients were also uniformly acceptable, suggesting that each item contributes consistently to the ordinal ordering of individuals along the latent well-being continuum. Gender- and country-specific analyses broadly supported this pattern, with moderate scalability among male adolescents and youth in North Macedonia and strong scalability among female adolescents and youth in Moldova. These results support the interpretation of the WHO-5 as a unidimensional ordinal measure in

which higher summed scores correspond to higher levels of subjective well-being. The findings are also consistent with previous Mokken analyses conducted among bilingual ethnic-minority adolescents in Norway and Denmark,⁷ adolescents receiving psychiatric care in Sweden,⁸ and school-based youth in Luxembourg,⁹ all of which reported strong scalability.

Assessment of single monotonicity revealed only a small number of minor violations across items, with all associated diagnostic values remaining well below established thresholds. This pattern was also observed in the gender- and country-specific analyses. This is consistent with the Luxembourg study, which similarly reported no meaningful violations of monotonicity.⁹ The findings therefore suggest that the WHO-5 items conform to a monotonic structure, such that higher levels of adolescent well-being are associated with an equal or greater probability of endorsing higher response categories within each item. This pattern supports the ordinal interpretation of item responses and strengthens confidence in the use of summed scores to represent ordered differences in well-being. This property is relevant for population monitoring and applied research, in which the WHO-5 may be used to distinguish adolescents with relatively lower or higher well-being at a given assessment.

Lastly, regarding double monotonicity, the MIIO diagnostics identified no meaningful intersections between the item response functions. However, the H^T coefficients were consistently below 0.30 overall and across gender and country subgroups, indicating insufficient separation between the functions and, consequently, an item ordering that was too inaccurate to be useful.

One possible explanation for the limited accuracy with which the item ordering could be estimated is the conceptual breadth of the WHO-5. Its items assess related but distinct manifestations of well-being, including positive mood, calmness, vitality, restorative sleep and engagement with daily life. These experiences may occur in different combinations while remaining indicators of a common underlying construct. For example, an adolescent may feel active and interested in daily activities while not feeling fresh and rested, whereas another may feel rested but report less positive mood or engagement. From an assessment perspective, the item ordering should therefore not be treated as a hierarchy that is accurately reproduced in individual adolescents' response patterns.

The main limitations of this study concern the generalizability of the findings. This research focused on early adolescents participating in a prevention-oriented intervention trial, which may limit applicability to older adolescents or to youth not engaged with school or health services. In addition, the study was conducted in two LMICs in Southeastern Europe. Further research is

needed to determine whether the observed ordinal scaling performance of the WHO-5 generalizes to adolescents in other cultural, socioeconomic and regional settings. Adolescent well-being is a multidimensional and culturally embedded construct, and cross-cultural differences in how well-being is interpreted may influence item functioning and scale performance.

Despite these limitations, the present study makes important contributions to the literature and has critical implications for the assessment of adolescent well-being with the WHO-5. To our knowledge, it is one of the first studies to jointly examine scalability, single monotonicity and IIO among younger adolescents from LMIC settings. The findings corroborate evidence for the scalability and monotonicity of the WHO-5 while showing that, although no meaningful violations of IIO were detected, the item ordering could not be estimated with sufficient accuracy to be useful. Accordingly, the findings support the use of the WHO-5 summed score to distinguish adolescents with relatively lower or higher well-being at a given assessment, while cautioning against treating the group-level item ordering as a hierarchy that is accurately reproduced in individual adolescents' response patterns. Further studies in other adolescent populations are needed to determine whether these findings can be replicated in different cultural and geographical contexts.

List of Abbreviations

DMM: Double Monotonicity Model

IIO: Invariant Item Ordering

LMICs: Low- and middle-income countries

MHM: Monotone Homogeneity Model

MIIO: Manifest Invariant Item Ordering

MSA: Mokken Scale Analysis

WHO: World Health Organization

WHO-5: World Health Organization Five Well-Being Index

Declarations

Ethics approval and consent to participate

The study protocol received ethical approval from three institutional review boards: the Institutional Review Board for Research Ethics at the University of Klagenfurt (approval number 2023-013), the Ethical Commission for Human Research of the Medical Faculty at Ss. Cyril and Methodius University in Skopje (approval number 03-2144/4), and the National Committee of Ethical Expertise for Clinical Trials of the Ministry of Health of the Republic of Moldova

(approval number 1476). All procedures were conducted in accordance with the ethical standards outlined in the Declaration of Helsinki as well as applicable national regulations.

The study involved adolescents between 10 and 14 years of age. Prior to participation, written informed consent was obtained from caregivers for all adolescents. Additional assent and consent procedures differed according to national requirements. In Moldova, all participating adolescents provided written assent. In North Macedonia, adolescents aged 10–13 years provided assent, whereas 14-year-old participants provided written informed consent themselves, in line with local ethical regulations.

Consent for publication

Not applicable.

Availability of data and materials

The data used in the present study are available from the corresponding author upon request.

Competing interests

The authors declare no competing interests.

Funding

This study was funded by the European Union through the Horizon Europe Research and Innovation Programme under grant agreement no. 101095528, as well as by UK Research and Innovation through the UK Government's Horizon Europe funding guarantee (grant no. 10050850). The funding organizations had no role in the conception or design of the study, the collection, analysis, or interpretation of the data, nor in the writing or decision to submit the manuscript.

Authors' contributions

AP and HMF contributed to data collection. AP conceptualized the study. AP conducted statistical analysis. AP drafted the manuscript, with contributions from all authors. All authors contributed to the critical revision of the manuscript and approved the final version.

Acknowledgements

The authors would like to sincerely thank the research and implementation teams involved in the FLOURISH study, including facilitators, supervisors, assessors, and project coordinators, for their valuable contributions to the project. We are also deeply grateful to the adolescents, caregivers, and community stakeholders whose participation and collaboration made this research possible.

YS is supported by DECIPHer, which is funded by the Welsh Government through Health and Care Research Wales, and by the Wolfson Centre for Young People's Mental Health, established with support from the Wolfson Foundation.

References

1. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a Systematic Review of the Literature. *Psychother Psychosom* [Internet]. 2015;84(3):167–76. Available from: <https://karger.com/pps/article/84/3/167/282903/The-WHO-5-Well-Being-Index-A-Systematic-Review-of>
2. Blakemore SJ, Mills KL. Is adolescence a sensitive period for sociocultural processing? *Annu Rev Psychol*. 2014;65(1):187–207.
3. Bhatia A, Fabbri C, Cerna-Turoff I, Tanton C, Knight L, Turner E, et al. Covid-19 response measures and violence against children. *Bullet World Health Org*. 2020;98(9):583–3A.
4. Bhatia A, Fabbri C, Cerna-Turoff I, Turner E, Lokot M, Warria A, et al. Violence against children during the COVID-19 pandemic. *Bullet World Health Org*. 2021;99(10):730–8.
5. Dlouhy M. Mental health policy in Eastern Europe: a comparative analysis of seven mental health systems. *BMC Health Serv Res*. 2014 ;14(1).
6. Lam LT, Lam MK. Child and adolescent mental well-being intervention programme: A systematic review of randomised controlled trials. *Front Psychiatry* [Internet]. 2023 Apr 6;14. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10116571/>

7. Sirpal MK, Haugen W, Sparle K, Haavet OR. Validation study of HSCL-10, HSCL-6, WHO-5 and 3-key questions in 14-16 year ethnic minority adolescents. *BMC Fam Pract*. 2016 Jan 27;17:7. doi:10.1186/s12875-016-0405-3.
8. Blom E, bech P, Högberg G, Larsson J, Serlachius E. Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales -testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. *Health Qual Life Outcomes*. 2012 Dec 11;10.
9. Brisson R. Psychometric Evaluation and Sociodemographic Measurement Invariance of the WHO-5 Well-Being Index among Adolescents in Luxembourg. *J Pers Assess*. 2025 Oct 21;1–10.
10. Carvalho PS, Martins MV, Azevedo I, Rodrigues R, Ferreira S. World Health Organization's Well-Being Index - WHO-5: psychometric performance of the Portuguese version for adolescents. *Port J Public Health*. 2025 Jan 27;43(1):1–16.
11. de Wit M, Pouwer F, Gemke RBJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 Well-Being Index in Adolescents With Type 1 Diabetes. *Diabetes Care* [Internet]. 2007 May 2;30(8):2003–6. Available from: <http://care.diabetesjournals.org/content/30/8/2003>

12. Doumit R, Hallit S, Sanchez-Ruiz MJ, El Khoury-Malhame M. Psychometric validation of an Arabic version of the WHO-5 wellbeing index among Lebanese adolescents. *PLOS One*. 2025 Jul 17;20(7):e0317644.
13. Mshiu JJ, Alexander FC, Rushaigo AM, Amani E, Roche JM, Samman E, et al. Adolescents' mental well-being and associated factors: a cross-sectional survey in selected regions of Tanzania. *Discov Public Health*. 2025;22:507. doi:10.1186/s12982-025-00907-w
14. Sischa PE, Martin G, Residori C, Hammami N, Page N, Schnohr C, et al. Cross-National Validation of the WHO-5 Well-Being Index Within Adolescent Populations: Findings From 43 Countries. *Assessment* [Internet]. 2025;33(1):10731911241309452. Available from: <https://pubmed.ncbi.nlm.nih.gov/39884717/>
15. Quansah F, Hagan JE, Ankomah F, Agormedah EK, Nugba RM, Srem-Sai M, et al. Validation of the WHO-5 Well-Being Scale among Adolescents in Ghana: Evidence-Based Assessment of the Internal and External Structure of the Measure. *Children*. 2022 Jul 1;9(7):991.
16. Sijtsma K, Molenaar IW. Introduction to nonparametric item response theory. Thousand Oaks, Calif.: Sage Publications; 2002.
17. Wilson M. Constructing Measures: An item response modeling approach. Taylor & Francis; 2023.

18. Piolanti A, Mueller J, Waller F, Heinrichs N, Simon J, Shenderovich Y, et al. Family-focused intervention programme to foster adolescent mental health and well-being: protocol for a multicountry cluster randomised factorial trial (FLOURISH Phase 2). *BMJ Open*. 2025 Feb 1;15(2):e094085.
19. World Health Organization Regional Office for Europe. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting, Stockholm, Sweden, 12–13 February 1998 [Internet]. Copenhagen: WHO Regional Office for Europe; 1998. Available from: <https://iris.who.int/handle/10665/349766>
20. Piolanti A, Waller F, Raleva M, Kunovski I, Wienand D, Simon J, et al. Measuring adolescents' well-being and mental health in Eastern Europe: Cross-cultural adaptation and pilot testing of the WHO-5 and RCADS-25. *Forthcoming*.
21. Mokken R. *A Theory and Procedure of Scale Analysis: With Applications in Political Research*. Berlin: De Gruyter; 1971.
22. Loevinger J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol Bull*. 1948;45(6):507–29.
23. Franco VR, Laros JA, Bastos RVS. Theoretical and Practical Foundations of Mokken Scale Analysis in Psychology. *Paidéia (Ribeirão Preto)*. 2022;32.

24. Crichton N. Information point Mokken Scale Analysis. *J Clin Nurs*. 1999;8(4):388.
25. Hardouin JB, Bonnaud-Antignac A, Sébille V. Nonparametric Item Response Theory Using Stata. *Stata J*. 2011 Mar;11(1):30–51.

Table 1. Response-category distributions for the WHO-5 items in the overall adolescent sample

Item	0. At no time	1. Some of the time	2. Less than half of the time	3. More than half of the time	4. Most of the time	5. All of the time
(1) <i>I have felt cheerful and in good spirits</i>	8 (1.2)	70 (10.7)	54 (8.3)	120 (18.4)	250 (38.3)	151 (23.1)
(2) <i>I have felt calm and relaxed</i>	9 (1.4)	91 (13.9)	61 (9.3)	157 (24.0)	212 (32.5)	123 (18.8)
(3) <i>I have felt active and vigorous</i>	19 (2.9)	70 (10.7)	64 (9.8)	143 (21.9)	179 (27.4)	178 (27.3)
(4) <i>I woke up feeling fresh and rested</i>	62 (9.5)	120 (18.4)	90 (13.8)	108 (16.5)	151 (23.1)	122 (18.7)
(5) <i>My daily life has been filled with things that interest me</i>	13 (2.0)	94 (14.4)	71 (10.9)	135 (20.7)	186 (28.5)	154 (23.6)

Note. N = 653.

26. Sijsma K, van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *Br J Math Stat Psychol*. 2016 Dec 13;70(1):137–58.
27. Piolanti A, Waller F, Kunovski I, Shenderovich Y, Mueller J, Wienand D, et al. Structural validity and measurement invariance of the WHO-5 Well-Being Index among young adolescents in a family-focused prevention context in Southeastern Europe. Forthcoming.
28. Molenaar IW, Sijsma K, Boer P. MSP5 for Windows user's manual: A program for Mokken scale analysis for polytomous items (Version 5.0). 2000.
29. Ligtoet R, van der Ark LA, te Marvelde JM, Sijsma K. Investigating an invariant item ordering for polytomously scored items. *Educ Psychol Meas*. 2010;70(4):578-595.

Table 2. Scalability coefficients, single- and double-monotonicity diagnostics and invariant item ordering accuracy for the WHO-5 among adolescents

Scale and Items	h / h_i (95% CI)	Single Monotonicity		Double Monotonicity	
		<i>V</i>	<i>Crit</i>	<i>V</i>	<i>Crit</i>
Total Scale	0.53 (0.49-0.57)				
(1) <i>I have felt cheerful and in good spirits</i>	0.52 (0.47-0.58)	0	0	0	0
(2) <i>I have felt calm and relaxed</i>	0.53 (0.48-0.58)	0	0	0	0
(3) <i>I have felt active and vigorous</i>	0.56 (0.52-0.61)	1	0	0	0
(4) <i>I woke up feeling fresh and rested</i>	0.52 (0.47-0.57)	1	0	0	0
(5) <i>My daily life has been filled with things that interest me</i>	0.50 (0.45-0.56)	3	14	0	0
H^T				0.12	

Note. N = 653. h = overall scale scalability coefficient; h_i = item-level scalability coefficient; H^T = accuracy of invariant item ordering; CI = confidence interval; V = number of violations; $Crit$ = weighted index of the severity of violations. Estimated item ordering, from highest to lowest mean item score: Cheerful; Active; Interested; Calm; Rested. Items are presented in their original WHO-5 order.

Table 3. Scalability coefficients, single- and double-monotonicity diagnostics and invariant item ordering accuracy for the WHO-5 among male adolescents

Scale and Items	h / h_i (95% CI)	Single Monotonicity		Double Monotonicity	
		<i>V</i>	<i>Crit</i>	<i>V</i>	<i>Crit</i>
Total Scale	0.44 (0.38-0.51)				
(1) <i>I have felt cheerful and in good spirits</i>	0.44 (0.36-0.53)	1	6	3	33
(2) <i>I have felt calm and relaxed</i>	0.49 (0.42-0.56)	0	0	3	30
(3) <i>I have felt active and vigorous</i>	0.47 (0.39-0.54)	0	0	4	35
(4) <i>I woke up feeling fresh and rested</i>	0.43 (0.36-0.51)	0	0	3	33
(5) <i>My daily life has been filled with things that interest me</i>	0.39 (0.30-0.48)	1	13	1	21
H^T				0.10	

Note. N = 330. h = overall scale scalability coefficient; h_i = item-level scalability coefficient; H^T = accuracy of the invariant item ordering; CI = confidence interval; V = number of violations; $Crit$ = weighted index of the severity of violations. Estimated item ordering, from highest to lowest mean item score: Active; Cheerful; Calm; Interested; Rested. Items are presented in their original WHO-5 order.

Table 4. Scalability coefficients, single- and double-monotonicity diagnostics and invariant item ordering accuracy for the WHO-5 among female adolescents

Scale and Items	<i>h</i> / <i>h_i</i> (95% CI)	<i>Single Monotonicity</i>		<i>Double Monotonicity</i>	
		<i>V</i>	<i>Crit</i>	<i>V</i>	<i>Crit</i>
Total Scale	0.60 (0.55-0.66)				
(1) <i>I have felt cheerful and in good spirits</i>	0.59 (0.52-0.67)	0	0	0	0
(2) <i>I have felt calm and relaxed</i>	0.57 (0.51-0.64)	0	0	2	22
(3) <i>I have felt active and vigorous</i>	0.64 (0.58-0.70)	0	0	4	26
(4) <i>I woke up feeling fresh and rested</i>	0.59 (0.53-0.66)	0	0	2	16
(5) <i>My daily life has been filled with things that interest me</i>	0.60 (0.53-0.66)	0	0	0	0
<i>H^T</i>					0.15

Note. N = 321. *h* = overall scale scalability coefficient; *h_i* = item-level scalability coefficient; *H^T* = accuracy of the invariant item ordering; CI = confidence interval; *V* = number of violations; *Crit* = weighted index of the severity of violations.

Estimated item ordering, from highest to lowest mean item score: Cheerful; Interested; Active; Calm; Rested. Items are presented in their original WHO-5 order.