frontiers | Research Topics

Measurement in health psychology, volume II

Edited by
Paola Gremigni, Giulia Casu, Antonio de Padua Serafim and Victor Zaia

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-6799-9
DOI 10.3389/978-2-8325-6799-9

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Measurement in health psychology, volume II

Topic editors
Paola Gremigni — University of Bologna, Italy
Giulia Casu — University of Bologna, Italy
Antonio de Padua Serafim — University of São Paulo, Brazil
Victor Zaia — Faculdade de Medicina do ABC, Brazil

Citation
Gremigni, P., Casu, G., Serafim, A. d. P., Zaia, V., eds. (2025). *Measurement in health psychology, volume II*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6799-9

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 13 August 2025
DOI 10.3389/fpsyg.2025.1658599

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Pietro Cipresso,
University of Turin, Italy

*CORRESPONDENCE
Victor Zaia
victorzaia@gmail.com

RECEIVED 02 July 2025
ACCEPTED 30 July 2025
PUBLISHED 13 August 2025

CITATION
Zaia V, Casu G, Serafim AdP and Gremigni P (2025) Editorial: Measurement in health psychology, volume II.
Front. Psychol. 16:1658599.
doi: 10.3389/fpsyg.2025.1658599

Editorial: Measurement in health psychology, volume II

Victor Zaia[1]*, Giulia Casu[2], Antonio de Padua Serafim[3] and Paola Gremigni[2]

[1]Postgraduate Program in Health Sciences, Centro Universitário FMABC, Santo André, Brazil, [2]Department of Psychology, University of Bologna, Bologna, Italy, [3]Institute of Psychology, University of São Paulo, São Paulo, Brazil

KEYWORDS

health psychology, psychometrics, measurement, validation, scale development

Editorial on the Research Topic
Measurement in health psychology, volume II

Introduction

Disruptive events such as the COVID-19 pandemic, climate-related disasters, and escalating violence across diverse sociocultural contexts have had profound repercussions on the mental health of populations. In this context, Health Psychology is a multidisciplinary field that integrates perspectives from biology, behavioral science, and the social sciences to examine how these domains interact to influence health and disease. This field employs theoretical models, evidence-based methodologies, and clinical expertise to assess, manage, and prevent a broad spectrum of conditions impacting both physical and mental health. The core objectives of Health Psychology include the promotion of health, the prevention of illness, the facilitation of treatment and recovery processes, and the identification of key biological and psychosocial determinants associated with health outcomes, disease onset, and functional impairments (Abraham et al., 2024; APA, 2025).

Measurement plays a pivotal role in Health Psychology, underpinning the rigorous assessment of psychological constructs, health-related behaviors, and individual responses to illness and healthcare. Since its emergence as a distinct discipline, Health Psychology has prioritized the design and evaluation of interventions aimed at enhancing health outcomes—by mitigating risk behaviors, promoting healthy lifestyles, and supporting the management of chronic conditions. The effectiveness and scientific validity of these interventions depend on the use of psychometrically sound instruments capable of capturing the complex psychological and social dimensions of health and disease (Freedland, 2021; Scholz, 2019). Grounded in the biopsychosocial model (Bolton, 2022)—a foundational theoretical framework in Health Psychology—measurement practices play a critical role in elucidating the determinants and consequences that shape both physical and mental health.

It is in this context that, and 3 years after the publication of the inaugural issue of *Measurement in health psychology, volume II* (Casu et al., 2022), that we are pleased to introduce the second volume of this Research Topic, reflecting its significant scientific impact and the sustained demand for methodological innovation. This initiative underscores the persistent need for methodological advancement in a field of psychology that is fundamentally grounded in empirical evidence. Health Psychology equires assessment tools that are not only psychometrically robust—exhibiting strong validity, reliability, and sensitivity—but also conceptually rigorous and contextually relevant. Within this framework, the Research Topic *Measurement in health psychology, volume II* was conceived to provide an updated and integrative overview of contemporary approaches to psychological measurement in health-related domains. This second volume brings together contributions that advance theoretical frameworks, refine measurement techniques, and address emerging challenges in the development and validation of psychological instruments. These contributions are relevant across both research and applied settings. Our aim is to foster the production of reliable, valid, and actionable data in Health Psychology, with meaningful implications for scientific research, clinical practice, and health policy formulation.

This Research Topic comprises 15 peer-reviewed articles authored by 66 scholars from diverse countries and regions. The included studies focus on the development, validation, and refinement of self-report instruments designed to assess psychological, behavioral, and socio-environmental factors that influence health trajectories and the course of illness.

Measurement in clinical and healthcare contexts

Eight articles within this Research Topic focus on the development and validation of measurement tools tailored for clinical and healthcare settings. These contributions address screening instruments, outcome assessments, and context-specific health challenges.

Zaia et al. introduced the *Infertility Concern Questionnaire* (ICQ), validated among Brazilian couples undergoing fertility treatment. The ICQ captures two core dimensions—parental desire and communication about infertility—within a dyadic framework. Confirmatory factor analysis (CFA) supported a two-factor structure, and measurement invariance across partners confirmed its applicability to both sexes, offering preliminary evidence of its validity in fertility care.

In the field of adolescent mental health, Jiang et al. validated the *Depression Anxiety Stress Scale for Youth* (DASS-Y) in a sample of Chinese adolescents. The three-factor structure (depression, anxiety, stress) was confirmed via CFA, and the scale demonstrated strong internal consistency, test-retest reliability, and both convergent and criterion-related validity, supporting its use in routine school-based mental health screening.

Tai et al. compared the EQ-5D-3L and EQ-5D-5L in patients with knee osteoarthritis following unicompartmental knee arthroplasty. Their findings demonstrated the superior validity, reliability, and discriminant ability of the EQ-5D-5L, supporting its use for a more precise and reliable assessment of quality of life 1 year post-surgery in this population.

Ráczová et al. evaluated the *Developmental monitoring and screening method for the 11th check-up in primary care* (S-PMV11). This is a screening tool designed for monitoring the developmental difficulties based on parental assessment in primary pediatric care in Slovakia. Using Rasch modeling and Guttman scaling, the study confirmed unidimensionality and strong scalability, highlighting the tool's utility for early detection of developmental problems in pediatric care.

Tan et al. validated the Chinese version of the *Somatosensory Amplification Scale* (SSAS-C), which measures the tendency to perceive somatic sensations as unusually intense and alarming. The one-factor model showed good fit, and the SSAS-C was found to mediate the relationship between alexithymia and somatization, underscoring its relevance in psychosomatic assessment.

Santibáñez-Palma et al. developed and validated the *Environmental Confinement Stressors Scale* (ECSS-20) designed to assess the psychological impact of COVID-19-related confinement. The scale comprises four domains—economic stressors, social activity restrictions, habitability, and exposure to virtual media—and demonstrated strong factorial validity, internal consistency, and gender invariance, supporting its use in post-pandemic health surveillance.

Ay et al., validated the Turkish version of the *MisoQuest*, a questionnaire assessing sound sensitivity and its emotional consequences. Results indicated that anxiety significantly mediated the relationship between misophonia and quality of life, underscoring the clinical relevance of recognizing this often-overlooked condition and its broader implications for psychological wellbeing in the context of clinical assessment.

Finally, Liu et al. validated the *Oncology Nurses Health Behaviors Determinants Scale* (HBDS-ON) in China. The scale assesses six determinants of health behavior among nurses—perceived threat, benefits, barriers, self-efficacy, cues to action, and access to personal protective equipment. It demonstrated strong internal consistency for the six-factor model, offering a promising tool for enhancing occupational health interventions in clinical settings.

Psychological wellbeing and quality of life

Four studies included in this Research Topic focused on the assessment of constructs related to psychological wellbeing, quality of life, and mental health promotion. These contributions highlight the importance of using culturally sensitive and developmentally appropriate tools to capture individuals' subjective experiences and value systems.

Huang et al. validated the *Short Quality of Life Scale* (QOLS-6) among bank employees in China. Despite its brevity, confirmatory factor analysis revealed a multidimensional structure with high internal consistency. QOLS-6 scores were significantly associated with indicators of psychological distress and job satisfaction, confirming its utility as a rapid and effective tool for assessing quality of life in occupational health and broader applied contexts.

Ma et al. explored the relationship between lifestyle behaviors and subclinical depression in undergraduate students, identifying behavioral threshold through ROC analysis. Their findings support the development of concise screening tools incorporating sleep quality, sedentary behavior, and light-intensity physical activity as key factors in regulating depressive symptoms. The study highlights the potential of behavioral monitoring in mental health promotion.

Lei et al. developed and validated the *Chinese Mental Health Values Scale* (CMHVS), comprising seven culturally grounded dimensions: Expected Self, Relating to Others, Life Principles, Family, Purpose and Meaning, Achievement, and Communication.

The scale demonstrated excellent reliability and convergent validity, offering a novel tool for assessing mental health attitudes and informing culturally responsive interventions.

Finally, Fernandes et al. validated the *PERMA-Profiler* for use with Brazilian adolescents, based on Seligman's wellbeing model. The five-factor structure—Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment—was supported, with strong internal consistency and measurement invariance across age and gender groups. Significant associations with emotional and psychological outcomes support the *PERMA-Profiler* as a promising tool for assessing wellbeing in developmental contexts.

Personal characteristics and coping strategies

Three studies in this Research Topic explored individual traits and psychological mechanisms that shape responses to adversity, meaning-making processes, and the management of relational and emotional demands. These contributions underscore the relevance of assessing dispositional and coping-related dimensions within the framework of Health Psychology.

Chen et al. assessed the psychometric properties of the *Existential Fulfillment Scale* (EFS) among Chinese university students. The Chinese adaptation revealed a two-factor structure—Self-acceptance and Self-breakthrough—reflecting culturally nuanced interpretations of personal growth. The scale demonstrated strong internal consistency, test-retest reliability, and meaningful associations with indicators of wellbeing and depression, supporting its use in evaluating existential aspects of psychological functioning in young adults.

Gazzellini et al. validated the *Schema Coping Inventory* (SCI) in an Italian adult sample. Grounded in schema therapy theory, the SCI measures three maladapting coping styles: surrender, avoidance, and overcompensation. CFA confirmed the original three-factor structure, and the scale showed good internal consistency. Convergent and divergent validity were supported through correlations with psychopathological symptoms and other coping constructs, positioning the SCI as a valid and reliable tool for identifying maladaptive coping patterns in both clinical and non-clinical populations.

Finally, Glavač et al. examined the *Capacity to Love Inventory* (CTL-I) in a Slovenian non-clinical sample. The original six-factor structure was replicated with satisfactory model fit and reliability. CTL-I subscales showed expected inverse correlations with dysfunctional personality traits and strong positive associations with dispositional mindfulness. These findings support the CTL-I as a valuable tool for assessing relational capacity and emotional maturity, reinforcing the connection between interpersonal functioning and psychological wellbeing.

Conclusions

This second volume of *Measurement in health psychology, volume II* presents a diverse array of contributions that reflect the field's increasing methodological sophistication and conceptual maturity. Collectively, the studies emphasize the pivotal role of measurement in advancing both theoretical insight and clinical application within Health Psychology. Across domains including clinical assessment, psychological wellbeing, and coping mechanisms, researchers introduced and validated psychometric tools that are not only psychometrically robust but also culturally attuned and contextually appropriate.

A hallmark of this Research Topic is its consistent methodological rigor. The majority of studies employed Confirmatory Factor Analysis (CFA), with several also utilizing Exploratory Factor Analysis (EFA) and Rasch modeling. Tests for measurement invariance across gender, age, and dyadic relationships were frequently reported, reinforcing the generalizability and robustness of the proposed instruments. Notably, many contributions also focused on the development of brief yet psychometrically sound tools, addressing the growing demand for time-efficient assessments in both clinical and research settings.

In total, the 15 instruments presented in this volume expand the repertoire of tools available for assessing psychological constructs central to understanding health, illness, and adaptation. These tools offer practical solutions for measuring mental health outcomes, behavioral patterns, emotional resilience, and socio-environmental stressors, thereby contributing to a more precise and evidence-based approach in Health Psychology research and practice, in addition to relevant social impact.

This Research Topic does not intend to exhaust the field of *Measurement in health psychology, volume II*. Rather, it represents a meaningful step forward—one that we hope will inspire continued collaborative efforts to refine, validate, and disseminate assessment tools that can support the development of more effective, evidence-informed interventions and policies.

Author contributions

VZ: Writing – review & editing, Writing – original draft. GC: Writing – review & editing, Writing – original draft. AS: Writing – review & editing, Writing – original draft. PG: Writing – review & editing, Writing – original draft.

Acknowledgments

The Guest Editors express their heartfelt appreciation to all contributing authors for their valuable and insightful work, which has significantly enriched this Research Topic. We are equally grateful to the reviewers, whose thoughtful, rigorous, and constructive feedback played a crucial role in enhancing the quality and clarity of the published articles.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission.

This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Abraham, C., Conner, M., Jones, F., and O'Connor, D. (2024). *Health Psychology (Third edition*). Abingdon: Routledge.

APA (2025). *What is Health Psychology?* Available online at: https://societyforhealthpsychology.org/about/what-is-health-psychology/ (Accessed June 17, 2025).

Bolton, D. (2022). Looking forward to a decade of the biopsychosocial model. *BJPsych. Bull.* 46, 1–5. doi: 10.1192/bjb.2022.34

Casu, G., de Padua Serafim, A., Zaia, V., and Gremigni, P. (2022). Editorial: measurement in health psychology. *Front. Psychol.* 13:937700. doi: 10.3389/fpsyg.2022.937700

Freedland, K. E. (2021). Health psychology's 40th anniversary. *Health Psychol.* 40:823. doi: 10.1037/hea0001139

Scholz, U. (2019). It's time to think about time in health psychology. *Appl. Psychol. Health Well Being* 11, 173–186. doi: 10.1111/aphw.12156

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 27 March 2024
DOI 10.3389/fpubh.2024.1349514

Check for updates

OPEN ACCESS

EDITED BY
Paola Gremigni,
University of Bologna, Italy

REVIEWED BY
Elisabete Carolino,
Escola Superior de Tecnologia da Saúde de Lisboa (ESTeSL), Portugal
Abbas Mardani,
Zanjan University of Medical Sciences, Iran

*CORRESPONDENCE
Lan Zhang
✉ zhang800519@126.com

RECEIVED 04 December 2023
ACCEPTED 15 March 2024
PUBLISHED 27 March 2024

CITATION
Liu Y, Zhang L, Li S, Li H and Huang Y (2024) Psychometric properties of the Chinese version of the oncology nurses health behaviors determinants scale: a cross-sectional study.
Front. Public Health 12:1349514.
doi: 10.3389/fpubh.2024.1349514

Psychometric properties of the Chinese version of the oncology nurses health behaviors determinants scale: a cross-sectional study

Yuxiu Liu[1], Lan Zhang[2]*, Shuzhen Li[3], Hua Li[1] and Yuqi Huang[1]

[1]School of Nursing, Jinzhou Medical University, Jinzhou, China, [2]Nursing Department, First Affiliated Hospital of Jinzhou Medical University, Jinzhou, China, [3]Heze Home Economics College, Heze, China

Objective: To test the validity and reliability of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON) in oncology nurses, the Chinese version was developed.

Methods: The Brislin double translation-back translation approach was employed to forward translation, back translation, synthesis, cross-cultural adaptation, and pre-survey, resulting in the first Chinese version of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON). A convenience sample technique was used to select 350 study participants in Liaoning, Shandong, and Jiangsu, China, who satisfied the inclusion and exclusion criteria, to assess the validity and reliability of the scale.

Results: The Chinese version of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON) had six subscales (perceived threat, perceived benefits, perceived barriers, self-efficacy, cues to action, and personal protective equipment availability and accessibility), including 29 items. The average scale level was 0.931, and the content validity level of the items varied from 0.857 to 1.000. Each Cronbach's α coefficient had an acceptable internal consistency reliability range of 0.806 to 0.902. $X^{2/df}$=1.667, RMSEA=0.044, RMR=0.018, CFI=0.959, NFI=0.905, TLI=0.954, and IFI=0.960 were the model fit outcomes in the validation factor analysis. All of the model fit markers fell within reasonable bounds.

Conclusion: The Chinese version of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON) has good reliability and validity and can be used as a tool to assess the influencing factors of chemotherapy exposure for oncology nurses in China.

KEYWORDS

oncology nurses, chemotherapy exposure, health beliefs, factor analysis, psychometric evaluation

1 Introduction

In recent years, the utilization of antineoplastic drugs has become increasingly prevalent in cancer treatment, mirroring the annual rise in cancer cases (1). However, the deleterious effects of these drugs on normal bodily functions have raised significant concerns, particularly among medical professionals who are regularly exposed to them (2). Among these professionals, oncology nurses stand out as a primary cohort frequently exposed to chemotherapy drugs, placing them at the highest risk of such exposures (3). Their exposure often occurs through various routes, including skin or mucosal contact, inadvertent ingestion, accidental needle pricks, or inhalation (4). Not only does this pose serious health risks to oncology nurses, consequently diminishing their quality of life, but it also directly impacts the quality of nursing care provided.

Adherence to safe chemotherapy-handling guidelines during chemotherapy procedures is crucial to prevent chemotherapy exposure (5). Research indicates that nurses often fail to fully comply with these guidelines when handling chemotherapy drugs (6). For instance, they may neglect to utilize essential personal protective equipment such as eye masks, chemo-specific gowns, and respiratory screens (4, 7).

Various factors hinder oncology nurses' adherence to safe chemotherapy-handling guidelines, including their knowledge of these guidelines, workload, interpersonal influences, and support from workplace management (2). Interestingly, nurses' health beliefs regarding chemotherapy exposures tend to have a greater impact on their adherence to the guidelines than their knowledge of such exposures (5).

The Health Belief Model (HBM) is a widely used framework for elucidating and predicting self-determined health behaviors aimed at preventing illnesses or other health issues. Within this model, health beliefs encompass perceived susceptibility to acquiring a disease, perceived severity of the condition, perceived benefits of engaging in health behaviors, perceived barriers to adopting such behaviors, and perceived self-efficacy in carrying out preventive health actions (8). Individuals' preventive health actions are influenced by their health beliefs and cues to action (9). Moreover, the HBM illustrates how health attitudes are shaped by moderating factors such as age, gender, education, and expertise (10, 11).

Consequently, health beliefs regarding chemotherapy exposure and cues to adhere to protective measures serve as pivotal factors influencing oncology nurses' compliance with chemotherapy-related occupational protective measures. By systematically evaluating these influential factors, we can establish a foundation for guiding nurses in adhering to chemotherapy-related occupational protective measures, thus facilitating the development of accurate intervention programs in the future.

The development of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON) was grounded in the Health Belief Model (5). Accordingly, this study aims to translate the HBDS-ON into Chinese, assess the scale's validity and reliability, and investigate its clinical utility among Chinese oncology nurses. This endeavor seeks to furnish a dependable and efficient tool for comprehensively assessing oncology nurses' health beliefs concerning chemotherapy exposure, as well as the factors influencing adherence to guidelines and cues for adherence.

2 Materials and methods

2.1 Participants

In 2023, a cross-sectional study was conducted between March 1st and May 31th. Oncology nurses from Class Grade hospitals in Liaoning, Shandong, and Jiangsu provinces were selected as the study participants using convenience sampling. The inclusion criteria were as follows: (1) Nurses holding a valid nurse professional qualification certificate. (2) Willingness to participate voluntarily in the study. (3) Exposure to chemotherapy drugs for a duration of at least 1 year. Nurses were excluded if they: (1) Were absent from clinical duties due to reasons such as academic pursuits, vacation, pregnancy, or maternity leave during the survey period. (2) Did not engage in the allocation and administration of chemotherapy drugs before and after maternity leave and lactation. (3) Were involved in teaching or advanced studies within the surveyed hospitals.

To ensure the reliability of the analysis results, considering a sample size of 5 to 10 times the number of variables (12), a minimum of 200 cases were required for confirmatory factor analysis using a structural equation model (13). Accounting for a potential sample loss rate of 20%, 350 oncology nurses were targeted for inclusion in this study, and all participants provided informed consent before participation.

2.2 Translation and cultural adaptation of the HBDS-ON

This study utilized email communication to reach out to the original developers of the scale and secure authorization for translating the HBDS-ON scale into Chinese. The Brislin translation model was rigorously adhered to throughout the translation process (5).

2.2.1 Forward translation

The original English version of the scale was translated into two Chinese versions, T1 and T2, by two native Chinese speakers (Care, Dr., and Nursing Postgraduate). The research team then convened to review and discuss the two Chinese versions of the questionnaire, addressing any contentious points and integrating them into the initial draft of the Chinese version, T3.

2.2.2 Back translation

T3 was translated into English versions ET1 and ET2 by two native English teachers who had no prior exposure to the questionnaire. Subsequently, these translations were merged to create the English version ET3. The ET3 version was then submitted to the original author for feedback. Following a discussion and modification process involving members of the research team, the Chinese version of the scale, C1, was finalized.

2.2.3 Cultural adaptation

The culturally adaptation of the Chinese version, C1, was carried out using the Delphi method. Seven experts, comprising three nursing management experts and four clinical oncology nursing experts, were recruited. All experts held associate senior titles and possessed bachelor's degrees or higher (3 bachelor's, 3 master's, and 1 doctorate). The average years of professional experience among the experts were (25.57 ± 7.59) years, with all having more than 15 years of experience. Leveraging their clinical experience and professional expertise, the experts evaluated the context, cultural adaptability, and language expression of the items in the Chinese scale, C1, relative to the original scale. Subsequently, the research team incorporated expert feedback to modify the questionnaire, resulting in the formation of the revised Chinese version of the scale, C2.

2.3 Measurement and instruments

2.3.1 General information questionnaire

The researchers created the general information questionnaire. Age, gender, educational background, professional experience in oncology, and technical title were among the data gathered.

2.3.2 Chinese version of the HBDS-ON

The scale used to measure oncology nurses' health behavior determinants that influence adherence to chemotherapy-handling guidelines was developed by American scholar Dania Abu-Alhaija in 2022 (5). It consists of 29 entries with 6 dimensions, including perceived threat, perceived benefits, perceived barriers, self-efficacy, cues to action, and personal protective equipment availability and accessibility. The Likert 5 scale scoring method was adopted: 1 = "strongly disagree," 2 = "disagree," 3 = "Neither agree nor disagree," 4 = "agree," and 5 = "strongly agree." The HBDS-ON is a scale used to measure individual health beliefs and actions based on the Health Belief Model. The HBDS-ON is a scale measuring individual health beliefs and actions based on the Health Belief Model, consisting of a range of responses from "strongly disagree" to "strongly agree" and designed to assess the feasibility of implementing health behaviors (10, 14).

2.4 Data collection

2.4.1 Pre-survey

In March 2023, 30 cases were selected from the cancer unit of a Grade-Three hospital in Liaoning Province. Informed consent was obtained from all participants, and the investigator provided them with an explanation of the study's objectives and relevance. Participants were then asked to rate the acceptability of each item in terms of language, meaning, and content. All clinical oncology nurses reported that the items on the scale were easy to understand and had a clear format. Consequently, no changes were made to the questions in the Chinese version of the HBDS-ON for Oncology Nurses.

2.4.2 Formal investigation

The Questionnaire Star platform was utilized for the survey. Prior to administering the survey, informed consent was obtained from the nursing department of each hospital. The head nurse of the oncology department was also contacted to communicate the research purpose and instructions for completing the questionnaire. Subsequently, the questionnaire was distributed to the WeChat group of each department.

At the onset of the questionnaire, guidance language was provided to clarify that the survey was solely for scientific research purposes. Additionally, participants were assured of the anonymity and voluntary nature of data collection. Participants completed the questionnaire online, and submission was only possible after all questions were answered.

Two weeks later, a subset of 30 oncology nurses randomly selected from the initial participants were re-surveyed using the same questionnaire to analyze test–retest reliability.

2.5 Data analysis method

SPSS 25.0 and AMOS 23.0 statistical software were used to analyze the data.

2.5.1 Project analysis method

The critical ratio method and correlation coefficient method were used to screen the items on the scale. (1) Critical ratio method: 350 questionnaires were sorted according to the total score from high to low, and an independent sample t-test was used to compare whether the difference between the high group (the first 27%) and the low group (the last 27%) was statistically significant, and statistically significant items will be retained (15). (2) Correlation coefficient method: The Pearson correlation coefficient method was used to calculate the correlation coefficient between 29 items and the total amount table, and the items with a very weak correlation ($r < 0.3$) with the total score of the scale were deleted (16).

2.5.2 Validity test method

(1) Content validity: Each item on the Chinese HBDS-ON scale was scored by seven experts, with 1 = irrelevant, 2 = weakly correlated, 3 = strongly correlated, and 4 = very correlated. Based on the expert evaluation results, the Item Level Content Validity Index (I-CVI) and the Scale Level Content Validity Index/Average (S-CVI/Ave) were computed. Generally, I-CVI $\geq$ 0.78 and CVI/Ave $\geq$ 0.9 (17). (2) Construct validity: Confirmatory Factor Analysis (CFA) was employed to assess the scale's structural validity. The model fitting indexes mainly included Chi-square freedom ratio (χ^2/df), Root-Mean-Square Error of Approximation (RMSEA), Root Mean Square Residual (RMR), Comparative Fit Index (CFI), Norm Fit Index (NFI), Tuck-Lewis Index (TLI), and Value added Fit Index (IFI). It is generally believed that $\chi^2/df < 3$, RMSEA<0.08, RMR < 0.05, CFI, NFI, TLI, IFI > 0.9 meet the model fitting requirements (18). (3) Convergent and discriminant validity: Average Variance Extracted (AVE), Combined Reliability (CR), and correlation coefficients between observed variables were calculated based on the results of CFA, and when AVE > 0.50, CR > 0.70,convergent validity was in the acceptable range (10). The discriminant validity was standardized when the correlation coefficients of all dimensions of the scale were less than the square root of AVE (10).

2.5.3 Reliability test method

Reliability is an indicator of the accuracy or consistency of a measuring instrument's response to a measurement and reflects the degree of truth of the measured characteristic (19). In this study, internal consistency and retest reliability were used for reliability testing. Cronbach's α coefficient were calculated for the Chinese version of the scale and each dimension to evaluate the internal consistency of the scale. Thirty oncology nurses were selected according to the inclusion and exclusion criteria, and the Chinese version of the scale was used for the test–retest reliability, with a 2-week interval between the 2 measurements. The Intraclass Correlation Coefficient (ICC) was used to test the retest reliability of the two measurement scores with the aim of measuring the stability and consistency of the scale over time. The split-half reliability was then obtained by dividing the scale items into two halves and calculating the correlation between the results of each half.

2.5.4 Ethical consideration

The Research Ethics Review Board of Jinzhou Medical University First Affiliated Hospital (JZMULL2023046) granted ethical approval for this study. Prior to commencing the survey, participants were briefed on the study's objectives, the content of the survey, its voluntary nature, and the confidentiality of their personal information.

Consent was obtained by having participants complete the survey voluntarily.

3 Results

3.1 Cross-cultural adaptation results

Based on the actual situation and language conventions in China, the items in the initial draft of the Chinese version were thoroughly reviewed and adjusted from the perspectives of semantics, idiomatic expressions, experience, and concepts. Following the feedback from experts, the research team implemented the following modifications: (1) In items 5, 8, 13, 15, 18, 19, 27, 28, and 29, the phrase "when performing chemotherapy operations" was revised to "when performing chemotherapy operations (including preparation, administration, handling of medical waste after medication, management of patient secretions and excretions, etc.)." This modification was made to facilitate nurses' comprehension and prevent ambiguity, given that handling medications, such as preparation, patient administration, and disposal, typically involves exposure to chemotherapy drugs (20). (2) The statement "Personal protective equipment in our work area is stored in an inaccessible location" was amended to "Personal protective equipment in our work area is stored in a difficult-to-access location."

3.2 Sample characteristics

There were 350 subjects in this survey. The basic information of the subjects, such as gender, age, working experience in oncology, education, and technical title, is displayed in Table 1 for more details on the sample's demographic features.

3.3 Item analysis

The critical ratios (CR) of the 29 items of the scale in this study ranged from 7.834 to 11.876, all of which were greater than 3.000. $p < 0.001$ indicates that the items were well differentiated. Every item on the scale had a positive correlation with the overall score (r) (0.495 ~ 0.670), all >0.4, and $p < 0.001$, indicating a moderate correlation between each item and the scale. Eliminating item by item, the Chinese scale's Cronbach's α coefficient varied between 0.930 and 0.932, which does not exceed Cronbach's α value of the scale (0.933). In conclusion, each item of the scale was retained (Table 2).

3.4 Validity

3.4.1 Content validity

Seven experts evaluated the correlation between items and measured content through a 4-level scoring method (1 to 4 corresponding to "irrelevant" to "very relevant"). The content validity index S-CVI/Ave of the Chinese version of HBDS-ON was 0.931, and the item content validity index I-CVI was 0.857–1.000.

TABLE 1 Oncology nurse participants' demographics (*N* = 350).

Demographic variables	Categories	N	%
Gender	Female	345	98.6
	Male	5	1.4
Ages (years)	20~<30	125	35.7
	30~<40	177	50.6
	40~<50	33	9.4
	≥50	15	4.3
Oncology experience (years)	1 year to <5 years	158	45.1
	5 year to <10 years	106	30.3
	10 year to<15 years	57	16.3
	15 year to<20 years	20	5.7
	≥20 years	9	2.6
Education	Secondary	1	0.3
	College	40	11.4
	Undergraduate	288	82.3
	Postgraduate	21	6
Technical title	Nurse	35	10
	Nurse practitioner	180	51.4
	Nurse-in-charge	125	35.7
	Vice professor of nursing	9	2.6
	Professor of nursing	1	0.3

3.4.2 Construct validity

The survey data underwent confirmatory factor analysis (CFA) using Amos 23.0 software to construct the model. Subsequently, a structural equation model was derived (Figure 1). The goodness-of-fit index is presented in Table 3.

According to the calculation formula, the combined reliability (CR) values of the six dimensions of the Chinese HBDS-ON scale were found to be 0.903, 0.823, 0.900, 0.894, 0.916, and 0.816, respectively, all exceeding the threshold of 0.80. Additionally, the average variance extracted (AVE) values were 0.651, 0.607, 0.644, 0.586, 0.575, and 0.692 for each dimension, respectively, all surpassing the criterion of 0.50. These results indicate that the convergent validity met the established criteria as outlined in Table 4. Furthermore, the square roots of the AVE values were found to be greater than the corresponding correlation coefficients, and the discriminant validity was within acceptable limits, as presented in Table 5. In summary, based on the above findings, it can be concluded that the Chinese version of the HBDS-ON scale demonstrates good structural validity.

3.5 Reliability

The Chinese HBDS-ON scale's Cronbach's α coefficient was 0.933, whereas each subscale's Cronbach's α coefficient varied from 0.806 to 0.902, which was higher than English version (Cronbach alpha between 0.70 and 0.88), while the stability of the translated scale was found to be at a very favorable level (5). In addition, after 2 weeks, the

TABLE 2 Item analysis for the Chinese version of the HBDS-ON.

Item	Critical ratio	Correlation coefficient between item and total score	Cronbach's alpha if the item delete
PT-1	8.434	0.576	0.931
PT-2	10.225	0.587	0.932
PT-3	8.971	0.608	0.931
PT-4	10.188	0.609	0.931
PT-5	10.226	0.593	0.931
PBe-1	8.53	0.522	0.931
PBe-2	8.132	0.524	0.931
PBe-3	9.539	0.543	0.932
PBa-1	10.297	0.606	0.932
PBa-2	10.899	0.612	0.931
PBa-3	9.276	0.613	0.931
PBa-4	10.392	0.637	0.931
PBa-5	10.765	0.605	0.931
SE-1	9.899	0.579	0.930
SE-2	10.033	0.593	0.931
SE-3	9.205	0.546	0.931
SE-4	9.932	0.565	0.931
SE-5	11.284	0.615	0.931
SE-6	9.957	0.604	0.931
CA-1	10.652	0.619	0.931
CA-2	10.021	0.577	0.931
CA-3	10.287	0.631	0.931
CA-4	11.565	0.670	0.931
CA-5	11.546	0.621	0.930
CA-6	11.876	0.634	0.930
CA-7	11.031	0.616	0.930
CA-8	11.159	0.628	0.930
PPE-1	10.214	0.586	0.931
PPE-2	7.834	0.495	0.930

PT, perceived threat; PBe, perceived benefits; PBa, perceived barriers; SE, self-efficacy; CA, cues to action; PPE, personal protective equipment availability and accessibility.

test–retest reliability (ICC) was found to be 0.860.The split-half reliability was 0.796.

4 Discussion

In recent years, the increased utilization of antineoplastic drugs has led to a heightened risk of chemotherapy exposure among oncology nurses, resulting in continuous harm to their bodies (21, 22). Therefore, establishing a safe working environment to ensure the occupational protection of oncology nurses during chemotherapy is of utmost importance.

Currently, in China, the management of occupational safety protection related to antineoplastic drugs is still in its nascent stage, with no established guidelines or standards. However, hospitals conduct training sessions for oncology nurses on chemotherapy occupational protection knowledge, attitudes, and practices, and formulate related occupational protection measures. Studies have shown that the level of knowledge, attitudes, and behaviors regarding chemotherapy occupational protection among nursing staff in China is generally low to moderate, indicating a need for improvement (23). Additionally, there persists a misconception among some nursing staff that wearing protective gear is only necessary during the process of preparing chemotherapy drugs (24), leading to an increased risk of chemotherapy exposure for oncology nurses.

Utilizing a scientifically validated instrument to assess the factors influencing oncology nurses' compliance with chemotherapy-related occupational safeguards is essential. Hence, the introduction of the HBDS-ON scale, which measures variables based on the Health Belief Model that influence oncology nurses' adherence to chemotherapy handling guidelines.

The HBDS-ON scale assesses six areas: perceived threat, perceived benefits, perceived barriers, self-efficacy, cues to action, and availability and accessibility of personal protective equipment. It provides a valid and comprehensive description of the factors associated with oncology nurses' compliance with chemotherapy-related occupational safeguards (25).

The Chinese version of the HBDS-ON scale was developed in this study by adhering to Brislin's traditional translation paradigm, including forward translation, back translation, synthesis, cross-cultural adaptation, and pre-survey. According to item analysis, the Chinese version of the HBDS-ON demonstrated high item discrimination, with each item exhibiting a strong correlation with the scale. Even after the removal of each item, the Cronbach's α coefficient remained consistent with the original value of the Chinese version of the scale (0.933). These findings suggest that all 29 items of the Chinese version of the HBDS-ON can be retained with good discrimination.

In this study, both the structural validity and content validity of the HBDS-ON scale's Chinese version were thoroughly examined and assessed. Structural validity evaluates how well the theoretical assumptions of the scale align with the actual measurements, while content validity assesses whether the items adequately meet the needs and objectives of the measurement. The Chinese version of the HBDS-ON scale demonstrated good content validity, with item level content validity index (I-CVI) scores ranging from 0.857 to 1, and the scale level content validity index/average (S-CVI/Ave) of 0.931, surpassing the content validity reference value (26). These results indicate that the content evaluated by the scale is highly esteemed by professionals, and the scale's straightforward language makes it suitable for use by oncology nurses.

Structural validity was examined using confirmatory factor analysis (CFA) for the Chinese version of the HBDS-ON scale. The results indicated a well-fitting structural equation model. Convergent validity, which assesses whether items measuring the same underlying constructs are grouped together, was confirmed as the CR values exceeded 0.8 and the AVE values exceeded 0.5 for all six subscales, indicating convergence (27). Discriminant validity, which tests whether items measuring different constructs are not grouped together, was confirmed as the square root of the AVE was greater than the correlation coefficient of each specific subscale, indicating good differentiation among the subscales (28).

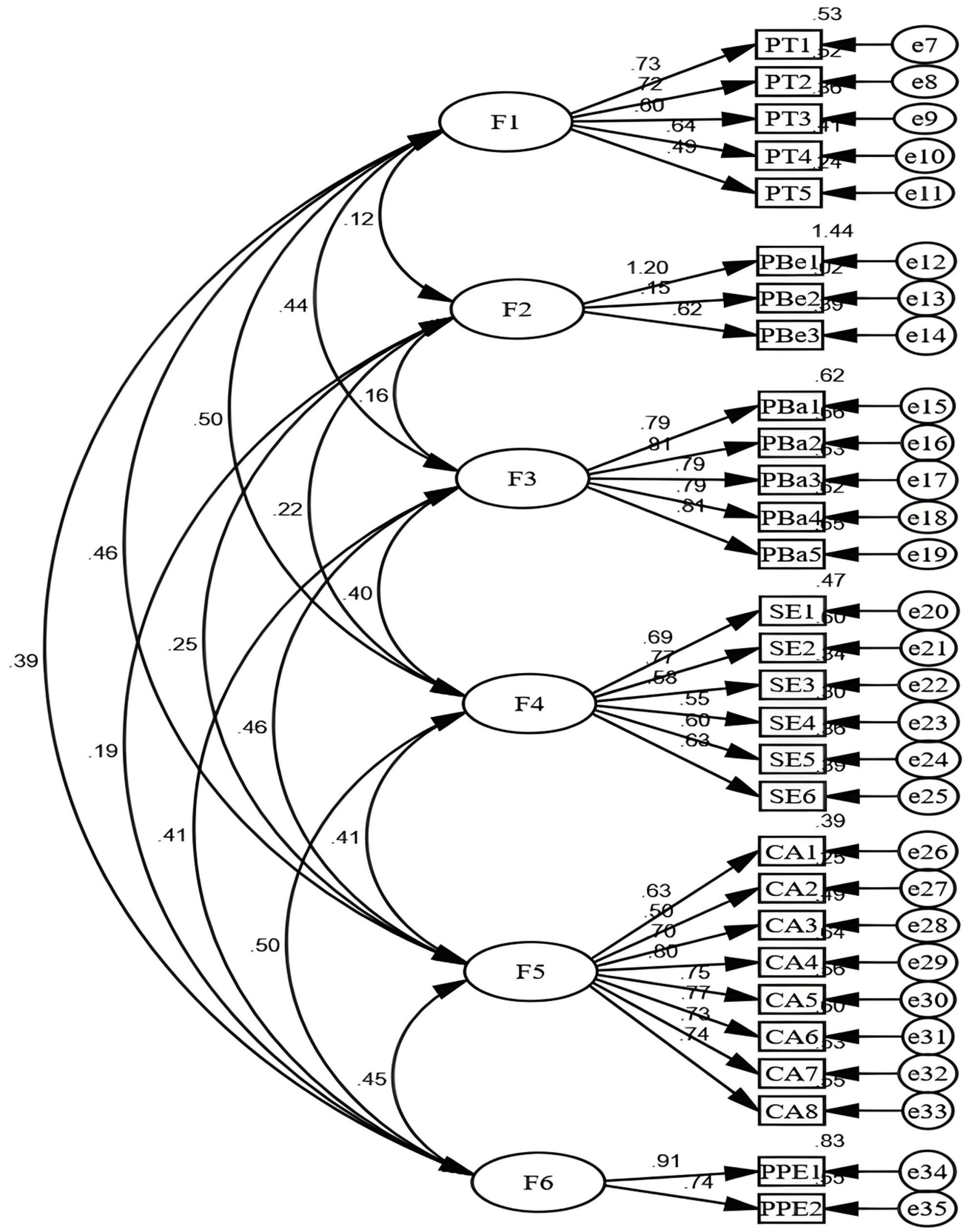

FIGURE 1
Standardized six-factor structural model ofthe Chinese HBDS-ON scale (*N* = 350). PT: Perceived Threat, PBe: Perceived Benefits, PBa: Perceived Barriers, SE: Self-Efficacy, CA: Cues to Action, PPE: Personal Protective Equipment availability and accessibility.

TABLE 3 The Chinese version of the HBDS-ON scale model fit index.

Item	CMIN/DF	RMSEA	RMR	CFI	NFI	TLI	IFI
Reference value	<3.000	<0.080	<0.050	>0.900	>0.900	>0.900	>0.900
model	1.667	0.044	0.018	0.959	0.905	0.954	0.960

The reliability of the Chinese version of the HBDS-ON was assessed using the test–retest (ICC), split-half, and internal consistency reliability methods. The Cronbach's α coefficient greater than 0.8 is considered excellent, between 0.6 and 0.8 is good, and less than 0.6 is poor (29). The translated scale demonstrated highly favorable stability and the Cronbach's α coefficient of 0.933, which is higher than the results of the English version. For the subscales, the Cronbach's α coefficient ranged from 0.806 to 0.902. The split-half value was 0.796. Test–retest reliability is considered good if the ICC is >0.75, better if the ICC is ≥0.4 and ≤ 0.75, and poor if the ICC is <0.4 (29). The ICC value measured after 2 weeks was 0.860, indicating that the Chinese version of the scale has good retest reliability. Therefore, the scale can

 10.3389/fpubh.2024.1349514

TABLE 4 Convergent validity of the Chinese HBDS-ON scale.

Subscale	Item	Unstandardized factor loading	Standardized factor loading	CR	AVE
Perceived threat	PT-1	1.000	0.769	0.903	0.651
	PT-2	1.047	0.793		
	PT-3	0.980	0.791		
	PT-4	1.094	0.865		
	PT-5	0.940	0.812		
Perceived benefits	PBe-1	1.000	0.802	0.823	0.607
	PBe-2	0.910	0.732		
	PBe-3	1.096	0.802		
Perceived barriers	PBa-1	1.000	0.791	0.900	0.644
	PBa-2	0.947	0.818		
	PBa-3	0.915	0.801		
	PBa-4	0.873	0.790		
	PBa-5	0.956	0.811		
Self-efficacy	SE-1	1.000	0.725	0.894	0.586
	SE-2	1.490	0.820		
	SE-3	1.050	0.690		
	SE-4	1.058	0.766		
	SE-5	1.425	0.798		
	SE-6	1.156	0.786		
Cues to action	CA-1	1.000	0.753	0.916	0.575
	CA-2	0.960	0.723		
	CA-3	1.172	0.793		
	CA-4	1.701	0.802		
	CA-5	0.990	0.745		
	CA-6	1.027	0.765		
	CA-7	1.029	0.741		
	CA-8	1.160	0.743		
Personal protective Equipment availability and accessibility	PPE-1	1.000	0.915	0.816	0.692
	PPE-2	0.770	0.739		

TABLE 5 Discriminant validity of the Chinese HBDS-ON scale.

Subscale	PPE availability and accessibility	Cues to action	Self-efficacy	Perceived barriers	Perceived benefit	Perceived threat
PPE availability and accessibility	0.832	–	–	–	–	–
Cues to action	0.435	0.759	–	–	–	–
Self-efficacy	0.497	0.435	0.765	–	–	–
Perceived barriers	0.398	0.456	0.442	0.802	–	–
Perceived benefits	0.450	0.497	0.441	0.426	0.779	–
Perceived threat	0.402	0.452	0.496	0.425	0.394	0.807

The diagonal blue numbers are AVE square root values. PPE, personal protective equipment.

be used to assess factors affecting oncology nurses' adherence to chemotherapy practice guidelines for chemotherapy safety with good temporal stability.

In conclusion, the Chinese version of the HBDS-ON scale exhibits strong structural validity, content validity, and reliability, making it a reliable tool for assessing factors influencing oncology nurses' adherence to chemotherapy handling guidelines for safe chemotherapy treatment.

5 Limitations

The present study has several limitations that need to be acknowledged. Firstly, although the sample size met the criteria for the study, the study was conducted in only three regions and may not be fully representative of oncology nurses across the country. This limitation may hinder the broad applicability and generalizability of the scale. In addition, principal component analysis was not used in this study to determine whether the structure of the original questionnaire was retained. Future studies should conduct further validation and include multi-center large sample studies to explore the applicability of the scale in different regions.

6 Conclusion

Additionally, while this study effectively introduced the English version of the Oncology Nurses Health Behaviors Determinants Scale (HBDS-ON) to China and followed the Brislin translation process meticulously, there may still be some cultural nuances and linguistic differences that were not fully addressed. Further validation and adaptation of the scale may be necessary to ensure its relevance and appropriateness within the Chinese context. Despite these limitations, this study provides a reliable instrument for evaluating the factors influencing oncology nurses' adherence to occupational safety precautions related to chemotherapy. By doing so, it contributes to the provision of excellent care for patients by ensuring the health and safety of oncology nurses and lays the groundwork for future research endeavors aimed at developing interventions and training materials in this critical area.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Jinzhou Medical University First Affiliated Hospital's Research Ethics Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. LZ: Funding acquisition, Methodology, Project administration, Resources, Supervision, Visualization, Writing – review & editing. SL: Investigation, Methodology, Resources, Writing – review & editing. HL: Investigation, Methodology, Software, Writing – review & editing. YH: Data curation, Formal analysis, Investigation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Fazel SS, Keefe A, Shareef A, Palmer AL, Brenner DR, Nakashima L, et al. Barriers and facilitators for the safe handling of antineoplastic drugs. *J Oncol Pharm Pract*. (2022) 28:1709–21. doi: 10.1177/10781552211040176

2. Polovich M, Clark PC. Factors influencing oncology nurses' use of hazardous drug safe-handling precautions. *Oncol Nurs Forum*. (2012) 39:E299–309. doi: 10.1188/12.ONF.E299-E309

3. Callahan A, Ames NJ, Manning ML, Touchton-Leonard K, Yang L, Wallen R. Factors influencing Nurses' use of hazardous drug safe-handling precautions. *Oncol Nurs Forum*. (2016) 43:342–9. doi: 10.1188/16.ONF.43-03AP

4. Boiano JM, Steege AL, Sweeney MH. Adherence to safe handling guidelines by health care workers who administer antineoplastic drugs. *J Occup Environ Hyg*. (2014) 11:728–40. doi: 10.1080/15459624.2014.916809

5. Abu-Alhaija D, Miller E, Bakas T, Shaughnessy E. The development and the content validation of the oncology nurses health behaviors determinants scale. *Semin Oncol Nurs*. (2022) 38:151317. doi: 10.1016/j.soncn.2022.151317

6. Colvin CM, Karius D, Albert NM. Nurse adherence to safe-handling practices: observation versus self-assessment. *Clin J Oncol Nurs*. (2016) 20:617–22. doi: 10.1188/16.CJON.617-622

7. Menonna-Quinn D, Polovich M, Marshall B. Personal protective equipment: evaluating usage among inpatient and outpatient oncology nurses. *Clin J Oncol Nurs*. (2019) 23:260–5. doi: 10.1188/19.CJON.260-265

8. Kumar A, Praveena PK, Barik RR. Adherence to COVID-19 appropriate behavior among small scale workers in unorganized sector in Rajasthan by applying health belief model and generalized social beliefs. *Int J Commun Med Public Health*. (2021) 8:2805. doi: 10.18203/2394-6040.ijcmph20211936

9. Champion VL, Skinner CS. The health belief model In: K Glanz, BK Rimer and K Viswanath, editors. *Health behavior and health education: theory, research, and practice*. San Francisco, CA: John Wiley & Sons, Inc (2008). 45–65.

10. Rosenstock IM. Historical origins of the health belief model. *Health Educ Monogr*. (1974) 2:328–35. doi: 10.1177/109019817400200403

11. Becker MH, Maiman LA. Sociobehavioral determinants of compliance with health and medical care recommendations. *Med Care*. (1975) 13:10–24. doi: 10.1097/00005650-197501000-00002

12. Zheng C, Yang Z, Kong L, Gao Z, Lu T, Zhang H. Psychometric evaluation of the Chinese version of the elderly-constipation impact scale: a translation and validation study. *BMC Public Health*. (2023) 23:1345. doi: 10.1186/s12889-023-16231-4

13. Wang T, Molassiotis A, Chung BPM, Tan JY. Psychometric assessment of the Chinese version of the problems and needs in palliative care questionnaire-short version in advanced cancer patients. *BMC Palliat Care*. (2019) 18:68. doi: 10.1186/s12904-019-0450-5

14. Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. *Nurs Res*. (2001) 50:300–4. doi: 10.1097/00006199-200109000-00008

15. Yang Z, Chen F, Liu S, Dai M, Zhang H. Psychometric properties of the Chinese version of the brief-mindful self-care scale: a translation and validation study. *Front Psychol*. (2021) 12:715507. doi: 10.3389/fpsyg.2021.715507

16. Xie YD, Li XY, Liu Q, Huang R, Li T, Fang YX, et al. Cross-cultural validation and psychometric testing of the debriefing experience scale (DES): a cross-sectional study. *BMC Med Educ*. (2022) 22:272. doi: 10.1186/s12909-022-03332-8

17. Wang Y, Gao Q, Liu J, Zhang F, Xu X. Translation and validation of the Chinese version of the maternal postpartum stress scale. *BMC Pregnancy Childbirth*. (2023) 23:685. doi: 10.1186/s12884-023-05990-y

18. Wang Y, Li S, Zou X, Xu L, Ni Y. Cross-cultural adaptation and validation of the Chinese version of the loneliness scale for older adults. *Geriatr Nurs*. (2022) 48:190–6. doi: 10.1016/j.gerinurse.2022.10.004

19. Liu K, Yu X, Cai Y, Tu D. Psychometric properties and measurement invariance of the LIVES daily hassles scale in Chinese samples. *Eval Health Prof*. (2023) 46:277–86. doi: 10.1177/01632787231164782

20. Leso V, Sottani C, Santocono C, Russo F, Grignani E, Iavicoli I. Exposure to antineoplastic drugs in occupational settings: a systematic review of biological monitoring data. *Int J Environ Res Public Health*. (2022) 19:3737. doi: 10.3390/ijerph19063737

21. Friese CR, Himes-Ferris L, Frasier MN, McCullagh MC, Griggs JJ. Structures and processes of care in ambulatory oncology settings and nurse-reported exposure to chemotherapy. *BMJ Qual Saf*. (2012) 21:753–9. doi: 10.1136/bmjqs-2011-000178

22. Abu-Alhaija D, Bakas T, Shaughnessy E, Miller E. The factors that influence chemotherapy exposure among nurses: an integrative review. *Workplace Health Saf*. (2023) 71:212–27. doi: 10.1177/21650799221140583

23. Xiaomin S, Yinping Z, Huanhuan W, and Bei Z. Clinical nurses' chemotherapeutic occupational protection knowledge. *Chin. J. Nurs. Educ*. (2015) 12:623–626. doi: 10.3761/j.issn.1672-9234.2015.08.017

24. Lijing Z, Jing W, Cuiping Z. Survey on the situation of training on occupational protection from chemotherapy and countermeasures. *J Qilu Nurs*. (2015) 21:26–28. doi: 10.3969/j.issn.1006-7256.2015.05.013

25. Abu-Alhaija D, Miller E, Shaughnessy E, Bakas T. Psychometric testing of the oncology nurses health behaviors determinants scale: a cross-sectional study. *Semin Oncol Nurs*. (2023) 39:151515. doi: 10.1016/j.soncn.2023.151515

26. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. (2006) 29:489–97. doi: 10.1002/nur.20147

27. Zhou S, Li X. Parental worry about children: scale development and validation among Chinese parents of preschool children. *Fam Process*. (2022) 61:1767–80. doi: 10.1111/famp.12741

28. Xiao H, Song J, Han X, Ye Z, Serier KN, Belon KE, et al. Assessing hedonic hunger among Chinese adults using the power of food scale: psychometric properties and cross-cultural invariance between China and the US. *Eat Behav*. (2023) 48:101703. doi: 10.1016/j.eatbeh.2023.101703

29. Sobral-Monteiro-Junior R, Coutinho LA, Rodrigues VD, Machado FSM, Soares WD, Oliva HNP, et al. Validity and reliability of SEIS-3: an instrument for subjective measuring of strength in older adults. *Methods X*. (2023) 12:102512. doi: 10.1016/j.mex.2023.102512

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 16 April 2024
DOI 10.3389/fpsyg.2024.1361645

Check for updates

OPEN ACCESS

EDITED BY
Antonio de Padua Serafim,
University of São Paulo, Brazil

REVIEWED BY
Celia Sofia Moreira,
University of Porto, Portugal
Caterina Novara,
University of Padua, Italy

*CORRESPONDENCE
Mert Huviyetli
mert.huviyetli.21@ucl.ac.uk

RECEIVED 26 December 2023
ACCEPTED 01 April 2024
PUBLISHED 16 April 2024

CITATION
Ay E, Huviyetli M and Çakmak E (2024) The mediating role of anxiety in the relationship between misophonia and quality of life: findings from the validated Turkish version of MisoQuest.
Front. Psychol. 15:1361645.
doi: 10.3389/fpsyg.2024.1361645

The mediating role of anxiety in the relationship between misophonia and quality of life: findings from the validated Turkish version of MisoQuest

Ezgi Ay[1,2], Mert Huviyetli[3]* and Eda Çakmak[1]

[1]Faculty of Health Sciences, Department of Audiology, Baskent University, Ankara, Türkiye, [2]Department of Interdisciplinary Neuroscience, Ankara University Graduate School of Health Sciences, Ankara, Türkiye, [3]Ear Institute, University College London, London, United Kingdom

Introduction: Misophonia is a disorder characterized by decreased tolerance to certain sounds or their associated stimuli, and many measurement tools have been developed for its diagnosis and evaluation. The aims of the current study were to develop the Turkish version of MisoQuest, a fully validated misophonia questionnaire, to evaluate the relationships between misophonia, anxiety, and quality of life, and to examine the mediating role of anxiety in the relationship between misophonia and quality of life.

Methods: The reliability of the Turkish version of MisoQuest was conducted using data from 548 participants (Mean age=28.06±9.36). Then, the relationships between misophonia, anxiety, and quality of life were evaluated in a separate sample of 117 participants (Mean age=25.50±6.31) using the State–Trait Anxiety Inventory (STAI) and the Short Form 36 (SF-36) questionnaire.

Results: The results showed that the Turkish version of MisoQuest has good psychometric properties. Close-to-moderate positive correlations were found between misophonia and anxiety, and weak negative correlations were found between misophonia and quality of life. Anxiety mediated the relationships between misophonia and quality of life.

Discussion: These results emphasize that misophonia may be an important problem affecting people's quality of life and reveal the mediating role of anxiety on this effect.

KEYWORDS

misophonia, anxiety, quality of life, MisoQuest, validity and reliability

Introduction

Some individuals experience abnormal emotional, behavioral, and physiological reactions to specific everyday sounds (Ferrer-Torres and Giménez-Llort, 2022; Swedo et al., 2022). This is initially introduced as "Selective Sound Sensitivity Syndrome" (Bernstein et al., 2013), now commonly referred to as "misophonia" (Jastreboff and Jastreboff, 2001). Unlike sounds typically perceived as disturbing, individuals with misophonia suffer from decreased tolerance toward ordinary, innocuous, and mostly human-induced sounds such as chewing, sniffing, and breathing sounds (Schröder et al., 2013). However, stimuli that cause aversive reactions, also called "triggers," are not limited to human-produced sounds. Studies have reported that all kinds of sounds, regardless of their source, as well as visual stimuli and repetitive movements, may trigger misophonic reactions (Dozier, 2015; Hansen et al., 2021). These reactions are irrespective of the physical characteristics of the stimuli (e.g., intensity or spectrum of sounds); instead, they are associated with various factors, including the individual's psychological profile, previous experiences with the sound, and the context in which the sound is encountered (Jastreboff and Jastreboff, 2001).

The most common emotional responses to triggers experienced by individuals with misophonia include varying levels of anger, irritation, stress, anxiety, and disgust (Dozier, 2015; Jager et al., 2020). Besides these emotional reactions, sympathetic overactivity (fight-or-flight response) such as muscle tension, feeling of pressure in the chest, arms, and whole body, increase in heart rate and body temperature, physical pain and breathing difficulties have been reported (Edelstein et al., 2013; Dozier, 2015). Behavioral reactions involve coping strategies that individuals with misophonia use to reduce their exposure to trigger stimuli, such as escaping or avoiding situations where the trigger may be encountered, seeking to discontinue the triggering stimuli, and mimicking or reproducing the triggers (Edelstein et al., 2013). In extreme cases, verbal, or physical violence toward the source of the triggers has also been reported (Schröder et al., 2013; Tunç and Başbuğ, 2017). These maladaptive reactions can cause serious negative effects on sufferers' social life, interpersonal relationships, performance of work or academic tasks, psychological status, and quality of life (Schröder et al., 2013; Jager et al., 2020).

There are different approaches in the literature that consider misophonia as a component of decreased sound tolerance, a symptom associated with various psychiatric disorders, or a new psychiatric disorder (Schröder et al., 2013; Jastreboff and Jastreboff, 2014). Different approaches also lead to differences in the criteria proposed for the diagnosis of misophonia. In the criteria published by Schröder et al. (2013), misophonic triggers are only sounds produced by other people, anger is the dominant reaction, and aversive reactions cannot be explained by other psychopathologies. Dozier et al. (2017) expanded these criteria and included all kinds of sounds and stimuli from different modalities as misophonic triggers and emphasized the immediate physical reflex response. Finally, Jager et al. (2020) expanded and updated the diagnostic criteria proposed by Schröder et al. (2013). Although there are various criteria proposed for the diagnosis of misophonia, there are no definitive criteria in international official diagnostic systems. Different questionnaires and measurement tools have been developed to assess the misophonia. The Amsterdam Misophonia Scale (A-MISO-S) developed by Schröder et al. (2013) and the Misophonia Questionnaire (MQ) developed by Wu et al. (2014) are commonly used tools in research and clinics. Recently developed scales for misophonia assessment include the MisoQuest (Siepsiak et al., 2020a), Duke Misophonia Questionnaire (DMQ) (Rosenthal et al., 2021), Selective Sound Sensitivity Syndrome Scale (S-Five) (Vitoratou et al., 2021), Berlin Misophonia Questionnaire-Revised (BMQ - R) (Remmert et al., 2022), and New York Misophonia Scale (NYMS) (Barahmand et al., 2023).

Differences in assessment and diagnostic criteria of misophonia are a barrier to generalizing findings, comparing results across studies, and estimating the prevalence of misophonia in the general population (Swedo et al., 2022). Considering misophonia as a significant social problem and conducting further research on its diagnosis and treatment seems essential for individuals' social functioning and quality of life (Siepsiak and Dragan, 2019). MisoQuest developed by Siepsiak et al. (2020a) is a fully validated questionnaire developed to assess misophonia. MisoQuest was created based on the diagnostic criteria proposed by Schröder et al. (2013). Due to its small number of items and single-factor structure, it is considered a rapid and effective tool for screening misophonia. Therefore, the present study aims to evaluate the psychometric properties of MisoQuest in the Turkish population, to examine the relationships between misophonia and state–trait anxiety and health-related quality of life and to assess the mediating role of anxiety on the relationship between misophonia and quality of life. The findings will provide information on the generalizability of previous findings to different populations, contributing to a better understanding of the relationship between misophonia and anxiety and the impact of misophonic symptoms on individuals' lives.

Methods

Participants

For the Turkish adaptation of MisoQuest, 548 participants aged 18–63 were included in the study. The mean age of the participants was 28.06 ± 9.36 (female = 27.26 ± 8.87, male = 31.19 ± 10.58). 20% of the participants were male (n = 111) and 80% were female (n = 437). The education level of 77% of the participants (n = 420) was college or above, and 23% (n = 128) was high school or below. The questionnaire was delivered to the participants online via social media applications. Participants with self-reported psychiatric diagnoses and hearing problems were excluded.

Following the validity and reliability study of the Turkish version of MisoQuest, data were collected from a different sample of 122 participants to evaluate the relationship between misophonia and state–trait anxiety and quality of life. The MisoQuest, State–Trait Anxiety Inventory (STAI) and 36-item Short-Form Health Survey (SF-36) scales were delivered online to the participants at this stage. Of the 122 people who participated in the study, three people were excluded because they stated that they were diagnosed with anxiety disorder and two people were diagnosed with panic disorder, and the study continued with 117 people. The mean age was 25.50 ± 6.31 (female = 24.87 ± 5.84, male = 29.18 ± 7.79). The demographic characteristics of the participants are given in Table 1.

Measurement tools

MisoQuest

The MisoQuest is the first fully validated misophonia questionnaire with good psychometric values and excellent reliability (Cronbach's alpha = 0.955). This scale, which includes a single-factor and 14 items, was developed by Siepsiak et al. (2020a) based on the misophonia diagnostic criteria published by Schröder et al., with some modifications. Unlike the criteria of Schröder et al., MisoQuest includes items related to all kinds of sounds, not only human-produced sounds. The scale only assesses aversive responses to sounds; sensitivities in other sensory modalities are not included. Each item has a 5-point Likert-type response category ranging from "1– Strongly disagree," "2 – Disagree," "3 – Undecided," "4 – Agree" and "5 – Strongly agree." There is no reverse item in the MisoQuest and the total score ranges from 14 to 70.

TABLE 1 Demographic characteristics of participants.

	n = 548	n = 117
Age	M = 28.06, SD = 9.36 Range = 18–63	M = 25.50, SD = 6.31 Range = 18–46

	n (%)	n (%)
Gender		
Female	437 (80)	100 (86)
Male	111 (20)	17 (14)
Education		
High school and below	128 (23)	29 (25)
University and above	420 (77)	88 (75)
Marital status		
Married	186 (34)	31 (27)
Single	362 (66)	86 (73)

M, mean; SD, standard deviation.

State–Trait Anxiety Inventory (STAI)

The STAI, developed by Spielberger et al. (1970), is one of the most commonly used tools to assess anxiety. It is a 40-item self-assessment questionnaire consisting of two subscales that provide separate measures of two components of anxiety: state and trait anxiety. The first twenty items measure situational or state anxiety (STAI-S), and the second twenty items measure underlying or trait anxiety (STAI-T). The State-Anxiety Scale evaluates how respondents' feel about anxiety "right now, at this moment" through four scales: "1 - Not at all," "2 – Somewhat," "3 - Moderately so," and "4 - Very much so." The Trait-anxiety Scale assesses how people "generally feel" about anxiety with four scales: "1 Almost never," "2 – sometimes," "3 - Often," and "4 - Almost always." There are 10 reversed items on the state scale (items 1, 2, 5, 8, 10, 11, 15, 16, 19, and 20), and seven on the trait scale (items 21, 26, 27, 30, 33, 36, and 39). The range of scores is from 20 to 80, the higher the score indicating greater anxiety. In this study, the Turkish version of STAI adapted by Öner and Le Compte (1983) was used. The Cronbach alpha value reported by Öner and La Compte is between 0.83–0.87 for the Trait Anxiety Scale and between 0.94–0.96 for the State Anxiety Scale.

36-item Short-Form Health Survey (SF-36)

The SF-36 is a standard measurement tool for assessing health-related quality of life and was developed by Ware and Sherbourne (1992). It is a Likert-type scale made up of 36 items, divided into eight dimensions: (i) Physical Functioning (PF), which assesses whether health conditions interfere with the ability to perform daily life activities; (ii) Physical Role Functioning (RP), which measures functional limitations due to health problems; (iii) Emotional Role Functioning (RE) which evaluates functional limitations by emotional problems; (iv) Social Functioning (SF), which impacts in quantity and quality of social activities induced by mental and physical problems; (v) Mental Health (MH), which measures aspects of depressive and anxiety; (vi) General Health (GH), which evaluates individual health status and its development tendency; (vii) Bodily Pain (BP), which measures degrees of pain to daily activities; and (viii) Vitality (VT), which is a subjective assessment of energy and tiredness. The scale has no total score; sub-dimension scores can range from 0 to 100, and higher scores mean better health status. The scores in these 8 domains can be reduced to two general components: The Physical Component Summary (PCS) score is calculated using the four physical health dimensions: PF, RP, BP and GH. The Mental Component Summary (MCS) score is calculated using the four mental health perceptions: VT, SF, RE and MH. In this study, the Turkish version of the SF-36 adapted by Koçyigit et al. (1999) was used. Cronbach's alpha values reported by Koçyiğit et al. range between 0.73 and 0.76 for the subscales.

Procedure

For the adaptation of MisoQuest to Turkish, written permission was first obtained from the developers of the questionnaire. MisoQuest, originally written in Polish, was translated from Polish to Turkish by a native Turkish translator who speaks Polish fluently. The Turkish translation was shared with the developers of the original version of MisoQuest, and a back translation into Polish was provided by a translator who speaks Turkish and Polish. An audiologist and a psychologist examined the Turkish translation of the questionnaire. Then, the questionnaire was applied to 5 individuals, and final adjustments were made in line with their feedback.

Statistical analysis

For the item analysis of MisoQuest, corrected item-total correlation was used. Confirmatory factor analysis (CFA) was used to evaluate the construct validity of MisoQuest. CFA was performed using maximum likelihood estimation and robust versions of the fit indices because of the non-normal distribution of data. The goodness-of-fit of the model was examined with the ratio of chi-square value to degrees of freedom (X^2/df), root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis Index (TLI). For these indices, the following cut-off values were used to indicate the goodness of model fit: $X^2/\text{df} \leq 3$, RMSEA ≤ 0.06, CFI ≥ 0.95, TLI ≥ 0.95 (Hu and Bentler, 1999). Internal consistency was tested with Spearman's rank correlation coefficient. Cronbach's alpha coefficient and Spearman-Brown coefficient were used in reliability

TABLE 2 Corrected item-total correlation and factor loadings.

Misoquest items (n = 548)	Corrected item-total correlation	F1	F2	F3
I1	0.60	0.66		
I2	0.58	0.63		
I3	0.63	0.66		
I4	0.62	0.56		
I5	0.55	0.59		
I7	0.70	0.56		
I6	0.64		0.47	
I8	0.60		0.54	
I9	0.55		0.64	
I12	0.64		0.63	
I14	0.50		0.46	
I10	0.66			0.46
I11	0.59			0.52
I13	0.49			0.64

F1, emotional reactions; F2, anger and avoidance; F3, functionality.

TABLE 3 Model fit indices for confirmatory factor analysis.

(n = 548)	χ^2 (df)	χ^2 / df	RMSEA	CFI	TLI
One-factor model	338.02 (77)	4.39	0.09	0.89	0.87
Three-factor model	200.45 (74)	2.71	0.06	0.95	0.94

χ^2 / df, reduced Chi-square; RMSEA, root mean square error of approximation; CFI, comparative fit index; TLI, Tucker Lewis index.

analyses. To evaluate normality of distribution Kolmogorov–Smirnov test was used. Pearson's and Spearman's rank correlations coefficients analysis were conducted to assess the relationship between misophonia and state anxiety, trait anxiety and quality of life. A value of $p < 0.05$ was considered statistically significant. The PROCESS macro for SPSS, a bootstrapping technique developed by Hayes (2017) was used to evaluate the mediating role of anxiety on the relationships between misophonia and quality of life. The number of bootstrap resamples was set at 5000. If the upper and lower limits of the 95% confidence interval did not include zero, the mediation effect was considered statistically significant. Statistical analyses were performed using R software (Version 4.3.3.), IBM SPSS 25.0, and IBM SPSS AMOS 25.0 programs (SPSS Statistics Version 25.0. IBM Corp., Armonk, NY).

Results

Item-total analysis

The corrected item-total correlation coefficients were between 0.49 and 0.70. These coefficients are shown in Table 2 for each item.

Confirmatory factor analysis

First, the single-factor model of the Turkish version of the MisoQuest was evaluated. As a result of confirmatory factor analysis, the following goodness-of-fit indices were obtained: X^2 = 338.02 (df = 77), X^2/df = 4.39, RMSEA = 0.09, CFI = 0.89, and TLI = 0.87. These indices indicated a poor fit between the model and the data. A parallel analysis was conducted to determine the appropriate factor structure. This analysis suggested a 3-factor structure. For the 3-factor structure, the following goodness-of-fit indices were obtained: X^2 = 200.45 (df = 74), X^2/df = 2.71, CFI = 0.95; TLI = 0.94; RMSEA = 0.06. These indices indicated that the 3-factor model good fit between the model and the data. The goodness of fit values of the testing models are given in Table 3. The confirmatory factor analysis model for MisoQuest is given in Figure 1.

Factor loadings

Factor loadings for all items ranged from 0.46 to 0.66. The first factor (1.-5. and 7. items) was named "Emotional Reactions," the second factor (6., 8., 9., 12., 14. items) was named "Anger and Avoidance," and the third factor (10., 11., 13. items) was named "Functionality." Factor loadings are shown in Table 2.

Reliability analysis

According to the internal consistency analysis results, the correlations between the total score of MisoQuest and the sub-factors were significant and varies between 0.79 and 0.89. Correlation

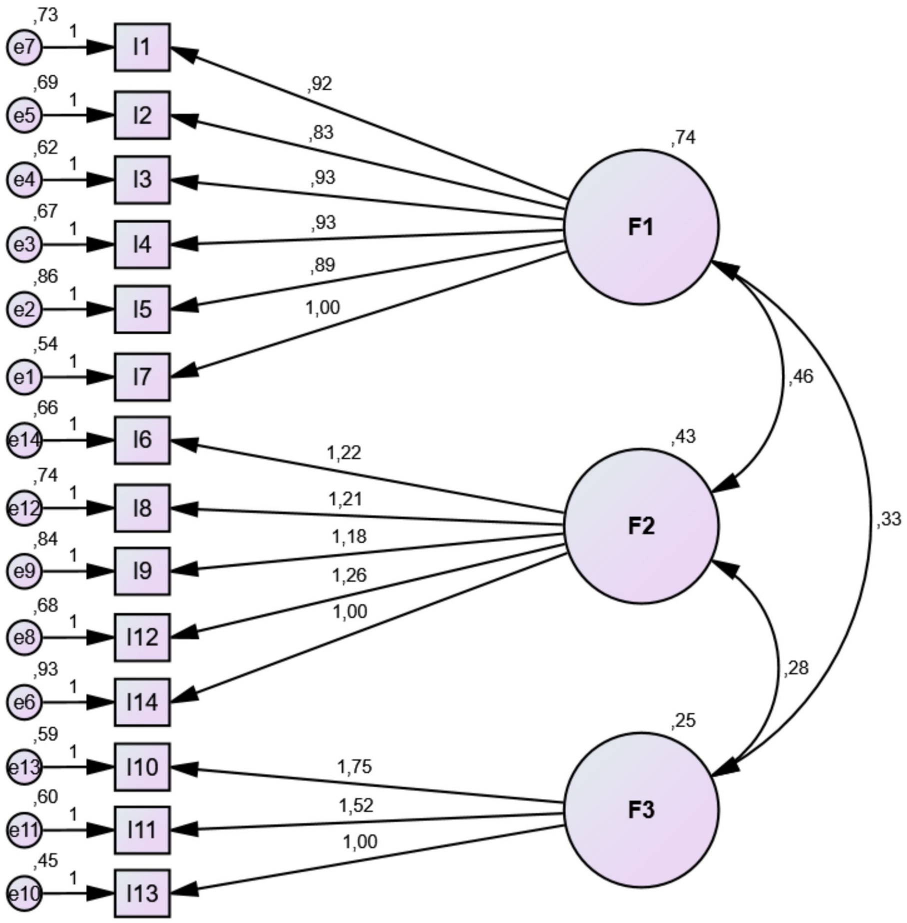

FIGURE 1
Confirmatory factor analysis model for MisoQuest.

TABLE 4 The correlation between the total score and sub-factors of the MisoQuest.

(n = 548)	Mean	SD	F1	F2	F3	MisoQuest total
MisoQuest total	36.33	10.51				
F1: Emotional reactions	15.91	5.16	–	0.639*	0.577*	0.891*
F2: Anger and avoidance	13.65	4.31		–	0.648*	0.882*
F3: Functionality	6.76	2.51			–	0.791*

*p < 0.001.

coefficients are shown in Table 4. The Cronbach's alpha value for the entire questionnaire is 0.90. It was 0.85 for the "Emotional Reactions," 0.79 for the "Anger and Avoidance," and 0.72 for the "Functionality." Spearman-Brown coefficient is found to be 0.92 for the whole MisoQuest and 0.86, 0.78 and 0.74 for the sub-factors, respectively.

Second order confirmatory factor analysis

After the first-order confirmatory model was confirmed, the second-order confirmatory factor analysis was conducted to identify whether all factors can be placed under a higher-order factor. This analysis reveals that three-factor measures a higher-order underlying construct, with very good reliability (0.84). The second-order latent variable was defined as "Misophonia." Thus, it is possible to consider that the total score of MisoQuest measures a latent misophonia variable manifested through three sub-factors of MisoQuest.

The relationship between misophonia and anxiety and quality of life

Statistically significant and close to moderate positive correlations were found between the MisoQuest total and sub-factors scores, and "STAI-State anxiety" scores (r = 0.349, r = 0.289, r = 0.384, and r = 0.286, respectively) as well as "STAI-Trait anxiety" scores (r = 0.334, r = 0.317, r = 0.335, and r = 0.247, respectively). Statistically significant weak negative correlations were found between the MisoQuest total, "Emotional Reactions" and "Anger and Avoidance" scores and "SF-36 Vitality (VT)" scores ($r = -0.194$, $r = -0.182$, $r = -0.245$, respectively), "SF-36 Social Functioning (SF)" scores ($r = -0.242$, $r = -0.244$, and $r = -0.248$, respectively), and "SF-36 Mental Health (MH)" scores ($r = -0.241$, $r = -0.215$, and $r = -0.270$, respectively). A significant weak correlation was found between the "Emotional Reactions" score and the "SF-36 Role Emotional (RE)" score ($r = -0.183$). The Pearson correlation coefficients between MisoQuest with STAI and SF-36 are given in Table 5. The correlation between the MisoQuest total score and other scales is shown as scatter plots in Figure 2.

Mediation model

The mediating role of state and trait anxiety on the relationship between misophonia and quality of life was examined separately. The

TABLE 5 Pearson correlation coefficients between MisoQuest with STAI and SF-36.

(n = 117)	MisoQuest total	F1	F2	F3
STAI – State anxiety	**0.349***	**0.289**	**0.384**	**0.286**
STAI – Trait anxiety	**0.334***	**0.317**	**0.335**	**0.247**
SF-36 Physical functioning (PF)	−0.086	−0.151	0.006	−0.048
SF-36 Role physical (RP)	−0.101	−0.121	−0.074	−0.075
SF-36 Bodily pain (BP)	−0.043	−0.029	−0.091	−0.004
SF-36 General health (GH)	−0.154	−0.150	−0.154	−0.110
SF-36 Vitality (VT)	**−0.194***	**−0.182***	**−0.245**	−0.050
SF-36 Social functioning (SF)	**−0.242**	**−0.244**	**−0.248**	−0.123
SF-36 Role emotional (RE)	−0.176	**−0.183***	−0.175	−0.119
SF-36 Mental health (MH)	**−0.241**	**−0.215***	**−0.270**	−0.183

F1, emotional reactions; F2, anger and avoidance; F3, functionality.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
Bold font indicates significant correlations.

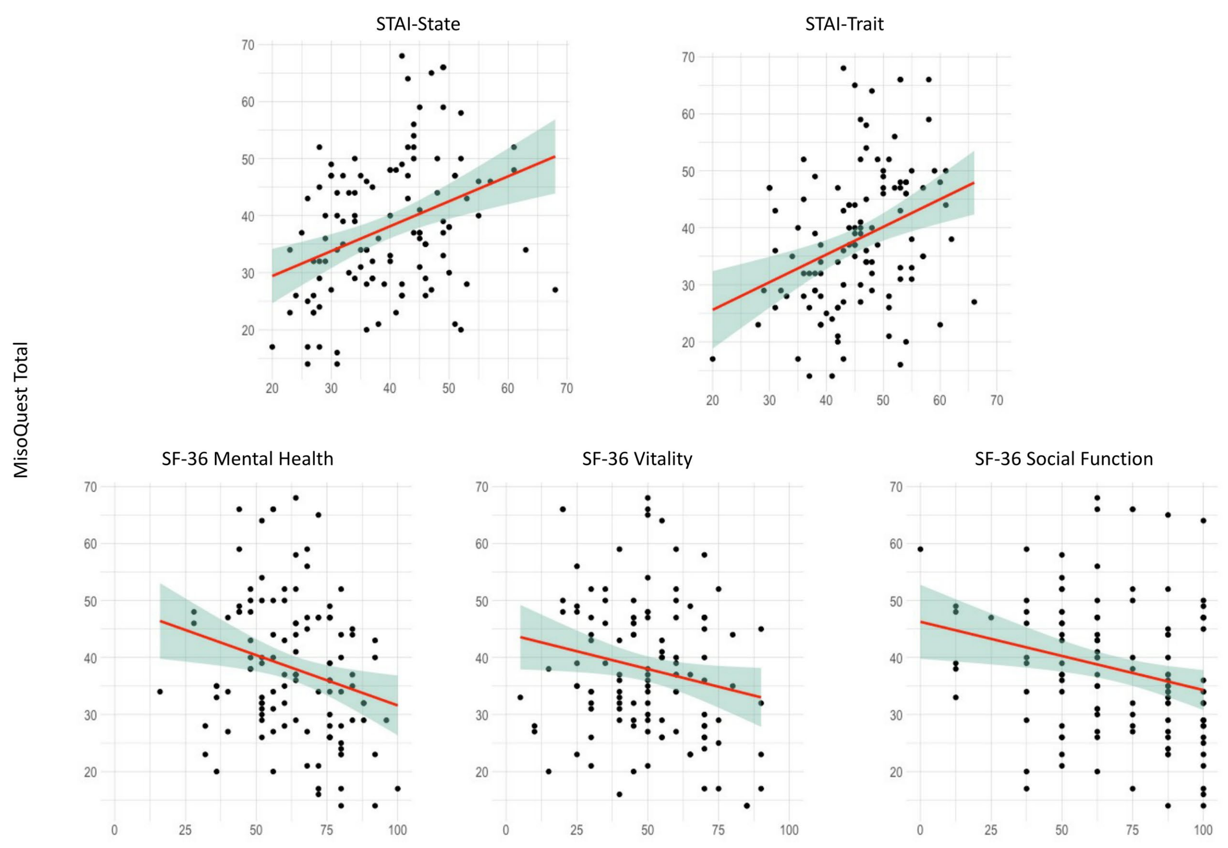

FIGURE 2
Scatter plots of correlations between MisoQuest with STAI and SF-36.

MisoQuest Total score was determined as the independent variable and the three sub-factor scores of SF-36 [Vitality (VT), Social Functioning (SF) and Mental Health (MH)] that had a significant correlation with MisoQuest Total were determined as the dependent variable. For both state and trait anxiety, total effect of misophonia on SF-36 (VT), SF-36 (SF), SF-36 (MH) was statistically significant. The indirect effect of misophonia on SF-36 (VT), SF-36 (SF), SF-36 (MH) was also statistically significant. However, the direct effect of misophonia on SF-36 (VT), SF-36 (SF), SF-36 (MH) was not significant. These results suggest that state and trait anxiety fully mediated the relationship between misophonia and quality of life. The results of the mediation analysis are presented in Figure 3 and Table 6.

Discussion

A growing number of studies have attempted to understand symptoms of misophonia, its pathophysiology, prevalence, and relationship with different audiological, psychological, and psychiatric factors (Siepsiak and Dragan, 2019; Ferrer-Torres and Giménez-Llort, 2022). However, standard measurement tools used in clinics and research for the diagnosis and evaluation of misophonia are still lacking (Swedo et al., 2022). The purposes of the present study are to examine the Turkish psychometric properties of MisoQuest, to evaluate the relationships between misophonia, state and trait anxiety and health-related quality of

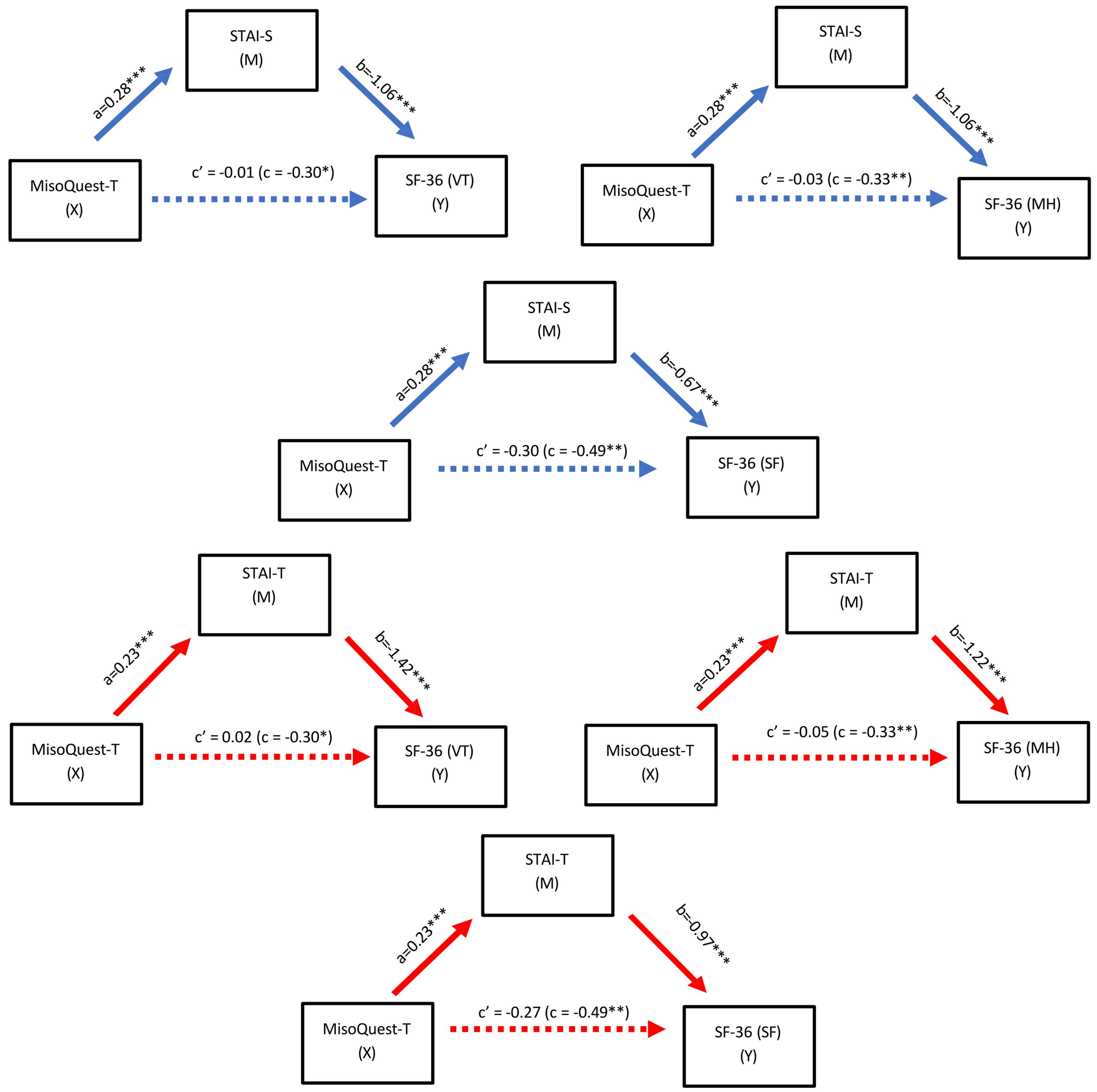

FIGURE 3
Path model of the mediating effect of state **(A)** and trait **(B)** anxiety on the relationship between misophonia and quality of life. Note. a,b,c, and c' are unstandardised regression coefficients. a path (direct effect of MisoQuest-Total scores on STAI), b path (direct effect of STAI on SF-36), c (total effect of MisoQuest-Total on SF-36), c' (direct effect MisoQuest-Total on SF-36). ****p < 0.0001, **p < 0.01, *p < 0.05.

TABLE 6 Mediation of the relationship between misophonia and quality of life by state and trait anxiety.

	95% Confidence interval						
		Effect	SE	t	*p*-value	Lower	Upper
MisoQuest-Total - STAI-State - SF- 36(VT-SF-MH)	*Direct effect* MisoQuest-total ➔ SF-36 (VT)	−0.0053	0.1427	−0.0409	0.9675	−0.2640	0.2534
	Indirect effect MisoQuest -total ➔ STAI-S ➔SF-36 (VT)	−0.2971	0.0735			−0.4480	−0.1559
	Total effect MisoQuest -total ➔ SF-36 (VT)	−0.3025	0.1427	−2.1192	0.0362	−0.5852	−0.0198
	Direct effect MisoQuest -total ➔ SF-36 (SF)	−0.3026	0.1898	−1.5945	0.1136	−0.6786	0.0733
	Indirect effect MisoQuest -total ➔ STAI-S ➔SF-36 (SF)	−0.1882	0.0764			−0.3593	−0.0607
	Total effect MisoQuest -total ➔ SF-36 (SF)	−0.4908	0.1832	−2.6791	0.0085	−0.8537	−0.1279
	Direct effect MisoQuest -total ➔ SF-36 (MH)	−0.0309	0.1057	−0.2929	0.7701	−0.2403	0.1784
	Indirect effect MisoQuest -total ➔ STAI-S ➔SF-36 (MH)	−0.2973	0.0718			−0.4414	−0.1621
	Total effect MisoQuest -total ➔ SF-36 (MH)	−0.3282	0.1235	−2.6581	0.009	−0.5729	−0.0836
MisoQuest-Total - STAI-Trait - SF- 36(VT-SF-MH)	*Direct effect* MisoQuest -total ➔ SF-36 (VT)	0.0208	0.1222	0.1704	0.8650	−0.2213	0.2630
	Indirect Effect MisoQuest -total ➔ STAI-T ➔SF-36 (VT)	−0.3233	0.0869			−0.4960	−0.1565
	Total effect MisoQuest -total ➔ SF-36 (VT)	−0.3025	0.1222	−2.1192	0.0362	−0.5852	−0.0198
	Direct effect MisoQuest -total ➔ SF-36 (SF)	−0.2664	0.1849	−1.4409	0.1524	−0.6326	0.0998
	Indirect effect MisoQuest -total ➔ STAI-T ➔SF-36 (SF)	−0.2245	0.0913			−0.4257	−0.0697
	Total effect MisoQuest -total ➔ SF-36 (SF)	−0.4908	0.1832	−2.6791	0.0085	−0.8537	−0.1279
	Direct effect MisoQuest -total ➔ SF-36 (MH)	−0.0459	0.1052	−0.4366	0.6632	−0.2544	0.1625
	Indirect effect MisoQuest -total ➔ STAI-T ➔SF-36 (MH)	−0.2823	0.0758			−0.4359	−0.1369
	Total effect MisoQuest -total ➔ SF-36 (MH)	−0.3282	0.1235	−2.6581	0.009	−0.5729	−0.0836

life, and to assess the mediating role of anxiety in the effect of misophonia on quality of life.

Confirmatory Factor Analysis (CFA) revealed a three-factor structure for the Turkish version of MisoQuest. Since the factor loadings were greater than 0.3 and goodness-of-fit indices were good and quite acceptable in the CFA, the three-factor structure was considered appropriate. Reliability analyses indicate that the Turkish version of MisoQuest has high internal consistency. Cronbach's alpha coefficient is very close to the value obtained in the original version of the questionnaire ($\alpha = 0.955$) (Siepsiak et al., 2020a). Moreover, corrected item-total correlations indicate that each item contributes sufficiently to the total score. The results demonstrate that the Turkish version of MisoQuest is valid and reliable.

In the Turkish version of MisoQuest, items related to individuals' internal emotional experiences, items related to anger response and avoidance of social environments, and items related to functionality in daily life are collected into three separate factors, unlike the original single-factor version of the questionnaire. A similar difference in factor structure between the original questionnaire and its Turkish version was also reported in the MQ (Wu et al., 2014; Sakarya and Çakmak, 2022). It is possible that the experience and expression of misophonia may vary as a result of social and cultural norms in different societies. Zhou et al. (2017), reported weaker correlations between misophonia symptoms and impairment in functioning in Chinese students than in American students. Turkey is considered a collectivist country in terms of cultural values such as commitment to the group and family (Kağitçibasi, 1996) and is similar to China in this context (Suh et al., 1998). Obtaining a different factor structure from the original questionnaire in this study may be due to these characteristics of Turkish culture. Studies that evaluate different cultures with the same standard measurement tools will increase our knowledge of the cross-cultural characteristics of misophonia.

Due to differences in assessment methods, measurement tools and diagnostic criteria, the frequency of misophonia reported in studies varies within a wide range between 3 and 55% (Jastreboff and Jastreboff, 2014; Wu et al., 2014; Zhou et al., 2017; Naylor et al., 2021). In a study conducted with depression patients using MisoQuest, it was reported that 8.5% of the patients met the misophonia criteria both in face-to-face interviews and in MisoQuest (when the cut-off value was 61 and above) (Siepsiak et al., 2020b). In another study conducted with 253 individuals using the same cut-off value, 69% of participants had self-reported misophonia, while 45% of those reporting self-reported misophonia met the MisoQuest cut-off criterion (Enzler et al., 2021). In our study, misophonia was detected in 2% of the total sample (N = 665) according to this cut-off value. Although this percentage is lower than those reported in other studies, it is close to Jastreboff and Jastreboff's (2014) estimates of 3.2%. It is thought that the higher frequencies of misophonia reported for different measurement tools may be related to the fact that these tools include more general sound sensitivities and lower severity misophonia symptoms in measurement (Siepsiak et al., 2020b). Individuals with misophonia find similar stimuli to be aversive to those without misophonia, but they experience extreme levels of aversion (Edelstein et al., 2013). Additionally, when only severe symptoms are taken into account, the misophonia frequency reported in the studies decreases considerably (Naylor et al., 2021). MisoQuest appears to detect more severe and quality of life impairing levels of misophonia, consistent with the purpose for which it was developed (Siepsiak et al., 2020b). Evaluations using different misophonia measurement tools on the same population may provide comparable findings regarding these tools.

Although misophonia is considered a separate disorder with unique clinical features and neurophysiological mechanisms, many studies are reporting a relationship between misophonia and different psychiatric symptoms and psychopathologies (Schröder et al., 2013; Erfanian et al., 2019; Jager et al., 2020). Two studies conducted with similar methodology on samples from different cultures found significant positive correlations between misophonia symptoms assessed by MQ and anxiety, depression, and OCD symptoms (Wu et al., 2014; Zhou et al., 2017). In patients with depression, anxiety was found to be more highly correlated with the severity of misophonia measured with MisoQuest than other variables (Siepsiak et al., 2020b). Additionally, studies have shown that anxiety has a mediating role on anger outbursts associated with misophonia (Wu et al., 2014; Zhou et al., 2017). In our study, a significant, close to moderate positive correlation was obtained between MisoQuest scores and STAI-State and STAI-Trait anxiety scores. Our findings also show that anxiety has a mediating role in the relationship between misophonia and quality of life. This finding indicates that as the severity of misophonia increases, individuals' anxiety levels would increase as well, and the deterioration in their quality of life may be a function of this increase in anxiety levels. Many studies report that exposure to misophonic triggers causes anxiety symptoms (Edelstein et al., 2013). However, some studies report that exposure to these triggers elicits anticipatory anxiety associated with thinking about future misophonic situations rather than eliciting an immediate anxiety response (Jager et al., 2020). Based on the relationship between misophonia and anxiety, there are studies that suggest anxiety treatment approaches in the treatment of misophonia (Bernstein et al., 2013). Randomized controlled clinical studies are needed to better understand the causal relationship between misophonia and anxiety. These studies may also provide a better evaluation of the effectiveness of anxiety-based intervention in misophonia.

Negative reactions caused by misophonia can cause deterioration in both interpersonal relationships and work or academic task performance, and can remarkably affect the person's well-being, daily functionality, and quality of life (Edelstein et al., 2013; Wu et al., 2014; Zhou et al., 2017; Ferrer-Torres and Giménez-Llort, 2022). The impact of misophonia on an individual's quality of life can range from a mild effect to severe impairment. The greater the severity of misophonia, the greater the impact on individuals' quality of life (Jager et al., 2020). In our study, individuals' quality of life was evaluated with the SF-36 scale, and significant negative correlations were found between the misophonia total scores and the Vitality (VT), Social Functioning (SF) and Mental Health (MH) factors of the SF-36. The Role Emotional (RE) factor of SF-36 was found to be correlated with the "Emotional Reactions" factor of MisoQuest. These findings support the previous findings that have shown the negative impact of misophonia on an individual's quality of life. Moreover, no correlation was found between S-36's physical health-related factors, Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), General Health (GH), and misophonia levels. Although some people with misophonia report that they also experience unpleasant sounds as painful, the comorbidity of misophonia with any somatic problems has not been reported (Rosenthal et al., 2022). Our findings support the previously reported view that misophonia is perceived as a mental problem rather than a physical problem (Kılıç et al., 2021). However, this finding may be related to the fact that the majority of our sample consisted of normal individuals without a diagnosis of clinically significant misophonia.

Limitations

There are several limitations in this study. First, using the online self-report method in data collection may cause recruitment bias. Second, participants' psychological and audiological diagnoses were based on self-report. Evaluations made by clinical psychologists, psychiatrists, and audiologists experienced in misophonia will be important in subsequent studies, especially in detecting comorbid disorders. Third, the age, educational status, and gender distribution of the individuals in the sample was not balanced. This may affect the generalizability of the results. Fourth, the target sample of this study was individuals from the normal population, but case–control studies including individuals clinically diagnosed with misophonia will contribute to a better understanding of the differences between individuals with misophonia and the normal population.

Conclusion

An increasing number of measurement tools have been developed for misophonia in recent years, and adapting these tools to different cultures is essential in providing a standard measurement for misophonia and revealing intercultural differences. In this study, the validity and reliability of the MisoQuest in Turkish was evaluated, and it was shown that the Turkish version of the questionnaire has good psychometric properties. Developing new measurement tools with good psychometric properties for the diagnosis of misophonia and ensuring cross-cultural adaptation of these tools will make significant contributions to the creation of a standardized diagnosis and treatment protocol for misophonia. We believe that the Turkish version of MisoQuest will be a useful measurement tool that can be used both as a screening tool in clinical practice and in misophonia research. To our knowledge, our study is the first to evaluate the relationship between misophonia assessed with MisoQuest and anxiety symptoms and quality of life in the normal population. Our study found a significant relationship between the misophonia level assessed with MisoQuest and people's state and trait anxiety levels and their quality of life.

Data availability statement

The original contributions presented in the study are publicly available. This data and Turkish version of MisoQuest can be found here: https://doi.org/10.6084/m9.figshare.24903288.v2.

Ethics statement

The studies involving humans were approved by the Baskent University Clinical Research Ethics Committee (Project number: KA21/72). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

EA: Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. MH: Visualization, Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. EÇ: Visualization, Data curation, Writing – review & editing, Writing – original draft, Formal analysis.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to thank Elif Eser, Merve Topaloğlu, and Kader Hazar Delibaş for their contributions to data collection and also the reviewers for their valuable contributions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Barahmand, U., Stalias-Mantzikos, M. E., Xiang, Y., and Rotlevi, E. (2023). The New York Misophonia scale (NYMS): a new instrument to identify Misophonia in the general population. *J. Psychiatr. Pract.* 29, 269–281. doi: 10.1097/PRA.0000000000000724

Bernstein, R. E., Angell, K. L., and Dehle, C. M. (2013). A brief course of cognitive behavioural therapy for the treatment of misophonia: a case example. *Cog. Behav. Ther.* 6:e10. doi: 10.1017/S1754470X13000172

Dozier, T. H. (2015). Etiology, composition, development and maintenance of misophonia: a conditioned aversive reflex disorder. *Psychol. Thought* 8, 114–129. doi: 10.5964/psyct.v8i1.132

Dozier, T. H., Lopez, M., and Pearson, C. (2017). Proposed diagnostic criteria for misophonia: a multisensory conditioned aversive reflex disorder. *Front. Psychol.* 8:1975. doi: 10.3389/fpsyg.2017.01975

Edelstein, M., Brang, D., Rouw, R., and Ramachandran, V. S. (2013). Misophonia: physiological investigations and case descriptions. *Front. Hum. Neurosci.* 7:296. doi: 10.3389/fnhum.2013.00296

Enzler, F., Loriot, C., Fournier, P., and Noreña, A. J. (2021). A psychoacoustic test for misophonia assessment. *Sci. Rep.* 11:11044. doi: 10.1038/s41598-021-90355-8

Erfanian, M., Kartsonaki, C., and Keshavarz, A. (2019). Misophonia and comorbid psychiatric symptoms: a preliminary study of clinical findings. *Nord. J. Psychiatry* 73, 219–228. doi: 10.1080/08039488.2019.1609086

Ferrer-Torres, A., and Giménez-Llort, L. (2022). Misophonia: a systematic review of current and future trends in this emerging clinical field. *Int. J. Environ. Res. Public Health* 19:6790. doi: 10.3390/ijerph19116790

Hansen, H. A., Leber, A. B., and Saygin, Z. M. (2021). What sound sources trigger misophonia? Not just chewing and breathing. *J. Clin. Psychol.* 77, 2609–2625. doi: 10.1002/jclp.23196

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York: Guilford Publications.

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Jager, I., de Koning, P., Bost, T., Denys, D., and Vulink, N. (2020). Misophonia: phenomenology, comorbidity and demographics in a large sample. *PLoS One* 15:e0231390. doi: 10.1371/journal.pone.0231390

Jastreboff, M. M., and Jastreboff, P. J. (2001). Components of decreased sound tolerance: hyperacusis, misophonia, phonophobia. *ITHS News Lett.* 2, 1–5.

Jastreboff, P. J., and Jastreboff, M. M. (2014). Treatments for decreased sound tolerance (hyperacusis and misophonia). Seminars in hearing. *Thieme Med. Pub.* 35, 105–120. doi: 10.1055/s-0034-1372527

Kağitçibasi, Ç. (1996). The autonomous-relational self: a new synthesis. *Eur. Psychol.* 1, 180–186. doi: 10.1027/1016-9040.1.3.180

Kılıç, C., Öz, G., Avanoğlu, K. B., and Aksoy, S. (2021). The prevalence and characteristics of misophonia in Ankara, Turkey: population-based study. *BJPsych open* 7:e144. doi: 10.1192/bjo.2021.978

Koçyigit, H., Aydemir, O., Olmez, N., and Memis, A. (1999). SF-36'nın Türkçe için güvenilirligi ve geçerliligi. *Ege Fizik Tedavi ve Rehabilitasyon Dergisi* 12, 102–106.

Naylor, J., Caimino, C., Scutt, P., Hoare, D. J., and Baguley, D. M. (2021). The prevalence and severity of misophonia in a UK undergraduate medical student population and validation of the Amsterdam misophonia scale. *Psychiatry Q.* 92, 609–619. doi: 10.1007/s11126-020-09825-3

Öner, N., and Le Compte, A. (1983). *Handbook of state-trait anxiety*. Istanbul, Turkey, Bogazici University Publication.

Remmert, N., Schmidt, K. M. B., Mussel, P., Hagel, M. L., and Eid, M. (2022). The Berlin Misophonia questionnaire revised (BMQ-R): development and validation of a symptom-oriented diagnostical instrument for the measurement of misophonia. *PLoS One* 17:e0269428. doi: 10.1371/journal.pone.0269428

Rosenthal, M. Z., Anand, D., Cassiello-Robbins, C., Williams, Z. J., Guetta, R. E., Trumbull, J., et al. (2021). Development and initial validation of the duke misophonia questionnaire. *Front. Psychol.* 12:709928. doi: 10.3389/fpsyg.2021.709928

Rosenthal, M. Z., McMahon, K., Greenleaf, A. S., Cassiello-Robbins, C., Guetta, R., Trumbull, J., et al. (2022). Phenotyping misophonia: psychiatric disorders and medical health correlates. *Front. Psychol.* 13:941898. doi: 10.3389/fpsyg.2022.941898

Sakarya, M. D., and Çakmak, E. (2022). Mizofoni Ölçeği'nin Türkçe Formunun Geçerlik ve Güvenirlik Sınama Çalışması. *Psikoloji Çalışmaları* 42, 231–255. doi: 10.26650/SP2020-845239

Schröder, A., Vulink, N., and Denys, D. (2013). Misophonia: diagnostic criteria for a new psychiatric disorder. *PLoS One* 8:e54706. doi: 10.1371/journal.pone.0054706

Siepsiak, M., and Dragan, W. (2019). Misophonia-a review of research results and theoretical concepts. *Psychiatr. Pol.* 53, 447–458. doi: 10.12740/PP/92023

Siepsiak, M., Śliwerski, A., and Łukasz Dragan, W. (2020a). Development and psychometric properties of misoquest—a new self-report questionnaire for misophonia. *Int. J. Environ. Res. Public Health* 17:1797. doi: 10.3390/ijerph17051797

Siepsiak, M., Sobczak, A. M., Bohaterewicz, B., Cichocki, Ł., and Dragan, W. Ł. (2020b). Prevalence of misophonia and correlates of its symptoms among inpatients with depression. *Int. J. Environ. Res. Public Health* 17:5464. doi: 10.3390/ijerph17155464

Spielberger, C.D., Gorsuch, R.L., and Lushene, R. (1970). Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Suh, E., Diener, E., Oishi, S., and Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: emotions versus norms. *J. Pers. Soc. Psychol.* 74, 482–493. doi: 10.1037/0022-3514.74.2.482

Swedo, S. E., Baguley, D. M., Denys, D., Dixon, L. J., Erfanian, M., Fioretti, A., et al. (2022). Consensus definition of misophonia: a delphi study. *Front. Neurosci.* 16:841816. doi: 10.3389/fnins.2022.841816

Tunç, S., and Başbuğ, H. S. (2017). An extreme physical reaction in misophonia: stop smacking your mouth! *Psychiatr. Clin. Psychopharmacol.* 27, 416–418. doi: 10.1080/24750573.2017.1354656

Vitoratou, S., Uglik-Marucha, N., Hayes, C., and Gregory, J. (2021). Listening to people with misophonia: exploring the multiple dimensions of sound intolerance using a new psychometric tool, the S-five, in a large sample of individuals identifying with the condition. *Psych* 3, 639–662. doi: 10.3390/psych3040041

Ware, J. E. Jr., and Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med. Care* 30, 473–483. doi: 10.1097/00005650-199206000-00002

Wu, M. S., Lewin, A. B., Murphy, T. K., and Storch, E. A. (2014). Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. *J. Clin. Psychol.* 70, 994–1007. doi: 10.1002/jclp.22098

Zhou, X., Wu, M. S., and Storch, E. A. (2017). Misophonia symptoms among Chinese university students: incidence, associated impairment, and clinical correlates. *J. Obsessive-Comp. Rel. Disord.* 14, 7–12. doi: 10.1016/j.jocrd.2017.05.001

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 July 2024
DOI 10.3389/fpsyg.2024.1415084

OPEN ACCESS

EDITED BY
Antonio de Padua Serafim,
University of São Paulo, Brazil

REVIEWED BY
Siaw Leng Chan,
Universiti Putra Malaysia Bintulu Sarawak Campus, Malaysia
Ricardo S. S. Durães,
Methodist University of São Paulo, Brazil

*CORRESPONDENCE
Iorhana Fernandes
✉ iorhanafernandes@hotmail.com

RECEIVED 09 April 2024
ACCEPTED 01 July 2024
PUBLISHED 12 July 2024

CITATION
Fernandes I, Zanini DS and Peixoto EM (2024) PERMA-Profiler for adolescents: validity evidence based on internal structure and related constructs.
Front. Psychol. 15:1415084.
doi: 10.3389/fpsyg.2024.1415084

PERMA-Profiler for adolescents: validity evidence based on internal structure and related constructs

Iorhana Fernandes[1*], Daniela Sacramento Zanini[1] and Evandro Morais Peixoto[2]

[1]Department of Psychology, Pontifícia Universidade Católica de Goiás, Goiânia, Brazil, [2]Department of Psychology, Universidade São Francisco, Campinas, Brazil

Introduction: The PERMA model of well-being has gained prominence in the study of well-being by the Positive Psychology movement. However, the model has been little studied regarding its applicability in different populations, such as adolescents. This study sought to evaluate the psychometric properties of the PERMA-Profiler instrument for Brazilian adolescents, as well as the measurement invariance for different age groups and gender, and investigate the relation with external variables.

Methods: Confirmatory Factor Analysis and Multigroup Confirmatory Factor Analysis were used to test the internal structure and invariance of the PERMA-Profiler. Reliability was determined with McDonald's Omega and composite reliability. A total of 1,197 adolescents between 11 and 19 years old from different regions of Brazil participated.

Results: The results of the confirmatory factor analysis indicated that the five correlated factors model was the most appropriate, presenting good factor loadings and adequate reliability. The scale proved to be invariant for adolescents of different age groups and gender. Correlations with associated variables were significant and moderate to strong, showing positive relations between positive emotions and well-being, and negative relations with negative affects and depressive and anxiety symptoms.

Discussion: These results contribute to the understanding of well-being in adolescence and highlight the importance of promoting different components of well-being for adolescents' mental health.

KEYWORDS

PERMA-Profiler, well-being, adolescence, psychometry, psychological assessment

Introduction

The PERMA model of well-being, an acronym for Positive Emotions (P), Engagement (E), Positive Relationships (R), Meaning (M), and Achievements (A), was developed by Seligman (2011) with the aim of creating a theoretical model that addresses the central aspects of well-being and the integration of eudaimonic and hedonic conceptions. This model has spread quickly around the world (Bazargan-Hejazi et al., 2021) and has been applied in different contexts with adult populations (e.g., O'Connor et al., 2016; Tansey et al., 2017; Turner and Thielking, 2019; Umucu et al., 2020; Blaire and Simmons, 2021; Tu et al., 2021). To evaluate the model, Butler and Kern (2016) developed the PERMA-Profiler instrument, which has been adopted by different researchers and translated into several languages (e.g., Cobo-Rendón et al., 2020; Giangrasso, 2021; Zewude and Hercz, 2022; Alves et al., 2023; Chaves et al., 2023; Mahamid et al., 2023).

The PERMA-Profiler scale is composed of 15 items that assess the PERMA components and eight additional items that evaluate happiness, negative emotions (sadness, anger, and anxiety), loneliness and physical health. The final measure consists of 23 items, scored on a Likert scale ranging from 0 = never to 10 = always. Regarding psychometric indicators, the authors found a factor structure of five correlated factors and three items in each factor, with good fit indices and good reliability indicators. These factors are: positive emotions ($\alpha = 0.88$), engagement ($\alpha = 0.72$), relationships ($\alpha = 0.82$), meaning ($\alpha = 0.90$), and achievements ($\alpha = 0.79$).

Although the original instrument presents good psychometric indicators and a factorial structure of five correlated factors, in subsequent adaptations some authors have found divergences regarding the internal structure of the model. Bartholomaeus et al. (2020) tested different models: unidimensional, multifactorial (five correlated factors), and hierarchical, which corresponds to the five PERMA components, comprising a general well-being factor. The unidimensional model was not accepted. The five correlated factors model and the second-order model had marginally different fit indices. The authors indicated that the second-order factorial structure might be the most appropriate to understand PERMA, due to the theoretical understanding of some authors (Goodman et al., 2017; Sellbom and Tellegen, 2019) that PERMA represents several facets of a single construct, which is well-being. However, predominantly, studies have pointed to a structure of five factors correlated with each other (Cobo-Rendón et al., 2020; Giangrasso, 2021; Wammerl et al., 2019; Zewude and Hercz, 2022; Alves et al., 2023; Chaves et al., 2023; Mahamid et al., 2023).

Such evidence corresponds to adaptations of the instrument for healthy adults, highlighting a gap in studies on different populations according to age group and gender. To date, only two studies have been found with adolescent populations using the PERMA-Profiler scale: one with Argentine adolescents (Waigel and Lemos, 2023) and the other with Indian adolescents (Singh and Raina, 2020). In both studies, the factor structure of five correlated factors had the best fit to the data, consistent with the original instrument. In Brazil, no adaptations have yet been found for this population.

Although there are few studies with the PERMA-Profiler instrument, evidence of the viability of the PERMA model applied to adolescence can be found in the literature. With regard to positive emotions, such as joy, contentment, and pleasure, Wu et al. (2022) point out that adolescents with high scores in positive emotions and low scores in negative emotions tend to have fewer mental health problems, such as depressive and anxiety symptoms, and better levels of self-esteem, gratitude, resilience, and life satisfaction.

Engagement is related to 'losing track of time' while carrying out a certain activity (Usán and Salavera, 2019). Adolescents with a higher level of engagement tend to have lower consumption of substances such as tobacco and alcohol (Pérez-Fuentes et al., 2021), a positive attitude toward authorities and higher self-esteem (Zhao et al., 2021). It is noteworthy that this component has a comprehensive nature and can be evaluated in different contexts, such as leisure activities with friends or family, activities carried out in the community, among others. Some authors have pointed out that engagement in leisure activities that involve peers predicts well-being in terms of levels of positive emotions and mental health (Asquith et al., 2022), as does engagement in sports and vacation trips or holidays enhance the perception of life satisfaction (Schmiedeberg and Schröder, 2017).

Regarding positive relationships, Carrillo et al. (2021) investigated, through a literature review, children and adolescents' understanding of well-being. Among the most reported aspects are healthy and supportive relationships, which relate to receiving adequate treatment from others and maintaining bonds with other people (adults and peers). Children and adolescents understand that relationships in which they are treated well promote well-being, as well as protection, security, experience of positive emotions and support for daily life (Carvalho et al., 2021).

Meaning concerns the significance attributed to one's existence and the feeling of belonging and connection with something or someone believed to be greater than oneself, for example God, church, family, etc. (Krok, 2018). This construct is defined as the sense that life has meaning, direction and intentionality, and that this sense plays a guiding role in establishing life goals and making decisions regarding the use of personal resources (Ribeiro et al., 2020). In adolescents, evidence points to meaning as a moderating protective factor in the relation to suicidal ideation (Zhang et al., 2017; Simões et al., 2022), and also as a factor against involvement in violent behavior or being a victim of peer violence (Zawadzka et al., 2018).

Achievement is about making progress toward a goal and being recognized internally and externally for it (Seligman, 2011). Studies show that high levels of recognition of one's own competence and achievements predict high levels of psychological well-being (Krok, 2018).

The present study

To fill this gap, this study aims to estimate the psychometric properties of the PERMA-Profiler scale for Brazilian adolescents. Specifically, the following aspects will be evaluated: (I) the internal structure of the scale, contrasting four factorial models suggested by the literature, namely, one factor, hierarchical, five correlated factors and bifactor models; (II) the scale's reliability indicators; (III) the measurement invariance in relation to gender and age group; and (IV) the correlation of PERMA with related variables.

Methods

Participants

One thousand, one hundred ninety-seven adolescents aged between 11 and 19 years old participated in this study ($M = 15.3$, $SD = 1.65$), who identified themselves as cisgender girls (54.3%), cisgender boys (41.6%), non-binary (3%) and transgender (1.1%). These participants were from all five of Brazil's regions: Central-West (42.5%), Southeast (52.8%), South (3.1%), Northeast (1.1%), and North (0.9%). Adolescents answered the questionnaires via cell phone (55.6%), computer (14.5%) and tablet (29.8%).

To participate in this research, participants must be between 10 and 19 years old. Their guardians must agree to participate by signing the Free and Informed Consent Form, and the adolescents themselves must agree to participate through the Term of Free and Informed Assent. Participants who self-report any neurological impairment that impedes their ability to read and understand the items presented will be excluded.

Instruments

Sociodemographic questionnaire

A sociodemographic questionnaire was developed containing questions regarding age, gender identification, education, housing and income.

PERMA-Profiler for adolescents

The original scale (Butler and Kern, 2016) is composed of 23 self-report items scored from 0 to 10. The measure is composed of 15 core items that assess the PERMA components and eight additional items that evaluate positive emotions, loneliness, happiness, and perception of health. The factor structure of five correlated factors showed good indicators [$\chi^2(80) = 10.61$; CFI = 0.97; TLI = 0.96 RMSEA = 0.06; SRMR = 0.03], as well as satisfactory internal consistency (Positive emotions $\alpha = 0.88$; Engagement $\alpha = 0.72$; Relationships $\alpha = 0.82$; Meaning $\alpha = 0.90$; Achievements $\alpha = 0.79$). To calculate the score for the five PERMA components, the arithmetic mean of the three items in each domain is taken. Negative emotions and health perception are calculated by the arithmetic mean of the three items, for example: (P1 + P2 + P3)/3 = Positive emotions component score. Happiness and loneliness are assessed by a single item.

Kessler K10

To evaluate psychological distress, the Kessler Psychological Distress Scale (K10) was used, adapted into Portuguese by Peixoto et al. (2021) for individuals aged 14 and over. The instrument consists of 10 self-report items that assess emotional stress in the last 30 days, scored on a 5-point Likert scale, with 1 = never and 5 = all the time. To interpret the scale, all items must be added together, and higher scores mean higher levels of anguish or psychological suffering. Due to the factorial structure of the instrument, it is also possible to know the levels of anxiety and depression separately. Anxiety is the sum of items 2, 3, 5, and 6. Depression is the sum of items 1, 4, 7, 8, 9, and 10. The instrument in original study has shown good psychometric indices and desirable reliability ($\alpha = 0.87$).

Life satisfaction

The Global Life Satisfaction Scale for Adolescents (EGSV-A) was developed by Giacomoni et al. (2012), and it is a Likert scale composed of 10 statements that are scored from 1 (not at all) to 5 (very much). The level of life satisfaction is known through the sum of all items answered. Regarding internal consistency, the scale presented good reliability indexes ($\alpha = 0.90$).

Positive and negative affects

The Positive and Negative Affects Scale for Adolescents (EAPN-A) was validated for Brazilian adolescents by Segabinazi et al. (2012), and it is a five-point Likert scale consisting of 28 items referring to positive and negative affects. Adolescents must mark how much they feel that way, ranging from 1 – "not at all" to 5 – "very much," and its application can be individual or collective. To know the result, the raw score must be calculated, which is obtained by adding the answered items for each construct (PA and NA). Scale validation studies demonstrate that the internal consistency of the scale assessed using Cronbach's alpha was 0.88 for PA and NA.

Cross-cultural adaptation procedure

With regard to the cross-cultural adaptation of the instrument and content validity, the International Test Commission (International Test Commission, 2017) translation and adaptation test checklist was used, along with the recommendations suggested by Borsa et al. (2012). The adaptation process comprised five stages: (I) Translation of the instrument; (II) Synthesis of translations; (III) Assessment by expert judges; (IV) Summary of evaluations; and (V) Evaluation by target audience. The first stage involved translation by two independent translators: a specialist in the subject of well-being and psychometrics in English, and a translator specialized in English. After the translations, a synthesis was carried out, considering semantic equivalence, idiomatic equivalence, experiential equivalence and conceptual equivalence, as proposed by Borsa et al. (2012). Based on the synthesis of translations, the language of the items was adapted for the adolescent audience (between 11 and 19 years old), and submitted to four independent judges who are doctors and experts in psychometrics, on the subject of well-being and adolescence. These judges evaluated the scale for clarity, pertinence, and relevance. After this stage, an analysis of the judges' contributions was carried out, and their suggestions were incorporated into the scale. The revised version of the instrument was then administered to 30 adolescents, who assessed the clarity and understanding of the items through structured interviews (Fernandes and Zanini, manuscript under construction). Based on the adolescents' feedback, the final version of the instrument was defined.

Data collection procedure

Participants were contacted in two ways. The first, through organic dissemination through the social media channels of the researchers involved, with the aim of reaching adolescents in different environments. In the second format, adolescents were contacted through partner schools in different regions of the country and, after consent from their guardians, the questionnaires were administered at the educational institution via computer or cell phone.

Ethical procedures

This research was approved by the Brazilian research ethics committee under number CAAE: 53061221.9.0000.0037. All participants and their guardians, respectively, signed the Free and Informed Assent Form and the Free and Informed Consent Form.

Data analysis

To evaluate the factorial structure of the scale, a Confirmatory Factor Analysis (CFA) was conducted using the Mplus software (Muthén and Muthén, 2012), with the Weighted Least Square estimation method (WLSMV). Four models were tested, namely: (I) One general factor model, in which all items compose a general

TABLE 1 Confirmatory factor analysis (CFA) results for the four contrasted models.

Model	χ^2	df	χ^2/df	CFI	TLI	RMSEA (90% CI)
One factor	2410.105	90	26.778	0.894	0.877	0.163 (0.158–0.169)
Hierarchical	1495.714	85	17.596	0.936	0.921	0.131 (0.125–0.137)
Five factors	700.287	80	8.753	0.972	0.963	0.090 (0.084–0.096)
Bifactor	1467.433	75	18.342	0.937	0.911	0.139 (0.132–0.145)

well-being factor (Ryan et al., 2019); (II) Hierarchical model, which consists of five PERMA factors at the first level and one general factor at the second level (Bartholomaeus et al., 2020); (III) Five correlated factors model, in which each PERMA factor is composed of three specific items, and all factors are correlated (Butler and Kern, 2016); and (IV) Bifactor model, which corresponds to all items loading on a general well-being factor and the items loading on each PERMA component (Wammerl et al., 2019). Unlike the hierarchical model, the bifactor model assumes that the associations of the general well-being factor with the responses to the observed items are direct and independent of the associations of the PERMA factors with the responses to the items. For all models, only the 15 central items of the instrument were used.

To interpret the results, the following fit indices were adopted: χ^2 (chi-square), df (degrees of freedom), RMSEA (Root-Mean-Square Error of Approximation), CFI (Comparative Fit Index), and TLI (Tucker–Lewis Index). The following values were used as parameters: CFI and TLI ≥ 0.90, RMSEA values ≤ 0.08 and $\chi^2/df \leq 5$ (Muthén and Muthén, 2017; Tabachnick and Fidell, 2019).

A multigroup confirmatory factor analysis (MGCFA) was performed to investigate the invariance of the PERMA-Profiler for adolescents according to gender and age group (11–15 years and 16–19 years). MGCFA evaluated measurement invariance in three models, namely: configural, metric and scalar. Model 1 (configural invariance) assessed whether the scale configuration (number of factors and items per factor) was acceptable for both groups (male and female). If the model is not supported, the factorial structure of the instrument cannot be considered equivalent for the groups evaluated. Model 2 (metric invariance) analyzed whether the factor loadings of the items could be considered equivalent between the groups. Model 3 (scalar invariance) investigated whether the level of latent trait required to endorse item categories (thresholds) was equivalent across groups (Cheung and Rensvold, 2002). Measurement invariance was assessed using the CFI difference test (ΔCFI, Cheung and Rensvold, 2002). If, when setting a parameter, a significant reduction in CFI indices is found (ΔCFI >0.01), measurement invariance cannot be accepted (Cheung and Rensvold, 2002).

To evaluate the internal consistency of the instrument, McDonald's omega was estimated, obtained from the Jamovi software (Version 2.3; The Jamovi Project, 2022), and composite reliability using the Composite Reliability Calculator (Colwell, 2016). To interpret the coefficients, values ≥0.70 were adopted as good indicators of precision (Tabachnick and Fidell, 2019). To evaluate the evidence based on the relation with external variables, Pearson correlation analyzes were carried out (r). To interpret the results, correlations with r between −0.09 and 0.09 are interpreted as null, r between 0.10 and 0.29 small, r between 0.30 and 0.49 medium or moderate, and r between 0.50 and 1.0 strong (Cohen, 1988). For this purpose, the Jamovi software was used (The Jamovi Project, 2022).

Results

Table 1 presents the fit indices of the four models tested in this study. The one factor model presented unacceptable fit indices. The bifactor and hierarchical models presented reasonable indices, although the RMSEA was higher than expected. The five correlated factors model presented the best fit to the data, with good indicators and adequate RMSEA, especially when considering the confidence interval. It is also noteworthy that models with few factors and items are expected to consequently produce a higher RMSEA (Kenny et al., 2015).

The factor loadings, as well as the reliability indicators, are presented in Table 2. The items have good factor loadings (0.319–0.884) and good reliability indicators with omega ranging from 0.77 to 0.89 and composite reliability from 0.79 to 0.90. The MGCFA results are presented in Table 3 considering gender and age group (adolescents between 11 and 15 years old and 16 and 19 years old). The results support configural, metric and scalar invariance, demonstrating that the PERMA-Profiler instrument for adolescents is an equivalent measure for men and women, and for younger (11–15 years old) and older (16–19 years old) adolescents, which allows comparison between groups.

The Table 4 presents the correlations with external variables. The PERMA scale showed a significant correlation with all instruments ($p < 0.001$). PERMA components correlate with each other, with moderate to strong magnitude. The engagement and achievement components present a moderate magnitude correlation with life satisfaction ($r = 0.39$; 0.42) and positive affects ($r = 0.48$). Likewise, relationships present a moderate magnitude correlation between meaning ($r = 0.65$) and positive affects ($r = 0.49$). The other PERMA components show a strong correlation with the indicators of life satisfaction and positive affects ($r = 0.63$–0.73).

Positive emotions correlated with negative affects ($r = 0.50$), distress ($r = 0.57$) and depression ($r = 0.59$), with a strong magnitude, just as the meaning component correlated with depression ($r = 0.51$) with strong magnitude. Positive emotions correlated with anxiety ($r = 0.50$) with moderate magnitude. Engagement showed a low correlation with all risk factors (negative affects, distress, anxiety and depression [$r = 0.15$–0.22]). Also, meaning correlated with a moderate magnitude with negative affects ($r = 0.44$), distress ($r = 0.49$) and anxiety ($r = 0.36$). The achievement component presents a significant correlation with negative affects, distress, anxiety, and depression, however with low magnitude ($r = 0.17$–0.21).

Discussion

The present study aimed to adapt and estimate the psychometric properties of the PERMA-Profiler scale for adolescents, in addition to contrasting four different models, evaluating the measurement

TABLE 2 Factor loadings of the five correlated factors model and reliability indicators.

Items	P	E	R	M	A
Item 3	0.831				
Item 13	0.848				
Item 22	0.884				
Item 2		0.685			
Item10		0.779			
Item17		0.319			
Item 8			0.696		
Item 19			0.825		
Item 21			0.729		
Item 7				0.831	
Item 9				0.871	
Item 20				0.883	
Item 1					0.742
Item 5					0.849
Item 15					0.639
McDonald's ω	0.877	0.635	0.776	0.886	0.770
Composite reliability	0.890	0.636	0.795	0.897	0.790

TABLE 3 Multigroup confirmatory factor analysis (MGCFA) for men, women and age group.

Gender	RMSEA (90% CI)	CFI	TLI	ΔCFI
Configural invariance	0.092 (0.086–0.098)	0.969	0.960	–
Metric invariance	0.085 (0.079–0.091)	0.972	0.966	+0.003
Scalar invariance	0.057 (0.052–0.062)	0.979	0.984	+0.007

Age	RMSEA (90% CI)	CFI	TLI	ΔCFI
Configural invariance	0.092 (0.085–0.098)	0.971	0.962	-
Metric invariance	0.086 (0.080–0.092)	0.973	0.966	+0.002
Scalar invariance	0.059 (0.054–0.064)	0.978	0.984	+0.005

invariance regarding age and gender, and the relation with external variables. In short, the scale presented good psychometric indicators and a factorial structure of five first-order correlated factors. The scale also proved to be suitable for assessing well-being in the PERMA model for adolescents of different age groups and in relation to gender.

The results of the confirmatory factor analysis indicate that the factor structure of five correlated factors is the most appropriate fit to understand the data. Similar results were found in the original version of the instrument (Butler and Kern, 2016) for Brazilian adults (Carvalho et al., 2021), for Indian and Argentine adolescents (Singh and Raina, 2020; Waigel and Lemos, 2023), and in several populations around the world (Cobo-Rendón et al., 2020; Giangrasso, 2021; Zewude and Hercz, 2022; Chaves et al., 2023; Mahamid et al., 2023). However, the factor structure of the PERMA model has been discussed.

Goodman et al. (2017) and Sellbom and Tellegen (2019) suggest that PERMA should be understood as a construct that is organized in a hierarchical structure, since the different facets compose well-being. This argument has been reinforced by slightly different fit indices when comparing the five first-order correlated factors model and the hierarchical model (Bartholomaeus et al., 2020). However, understanding the PERMA model applied to adolescence based on the factorial structure of five correlated factors corresponds to the essential properties indicated by Seligman (2018), which are: (I) Contribute to the construction of well-being; (II) People should seek it for itself, and not to promote another component; (III) It is defined and measured independently of another component.

Therefore, understanding PERMA from a first-order correlated model is close to the properties highlighted by Seligman (2018), and also to the instrument's original proposal, since the authors suggest the possibility of well-being profiles that would be more adaptive for each context (Butler and Kern, 2016). In other words, depending on the context, presenting high levels of a PERMA component can be a risk factor, while in other contexts it can be beneficial (e.g., a high level of work engagement can be a risk factor for burnout in a work environment with high demands and few resources). In this way, it is possible to think of an intervention that focuses only on conscious engagement in certain actions, without interventions in the other PERMA components.

TABLE 4 Correlations between PERMA, life satisfaction, positive and negative affects, and psychological distress.

	1	2	3	4	5	6	7	8	9	10
1-P	—									
2-E	0.535*	—								
3-R	0.709*	0.482*	—							
4-M	0.732*	0.599*	0.650*	—						
5-A	0.512*	0.683*	0.475*	0.637*	—					
6-SATIS	0.734*	0.390*	0.634*	0.676*	0.424*	—				
7-PA	0.657*	0.482*	0.498*	0.565*	0.481*	0.631*	—			
8-NA	−0.508*	−0.150*	−0.388*	−0.443*	−0.212*	−0.537*	−0.193*	—		
9-DIS	−0.570*	−0.194*	−0.427*	−0.495*	−0.244*	−0.598*	−0.332*	0.730*	—	
10-ANS	−0.397*	−0.119*	−0.300*	−0.360*	−0.177*	−0.426*	−0.202*	0.622*	0.874*	—
11-DEP	−0.598*	−0.222*	−0.447*	−0.510*	−0.256*	−0.631*	−0.375*	0.699*	0.944*	0.665*

*$p < 0.001$; P, positive emotions; E, engagement; R, relationships; M, meaning; A, achievements; SATIS, life satisfaction; PA, positive affects; NA, negative affects; DIS, distress; ANS, anxiety; DEP, depression.

From this perspective, the first-order correlated model privileges the independence between variables and the possibility of interventions focused on promoting just one component. On the other hand, hierarchical models favor the promotion of the general factor, that is, intervening in one of the components will (theoretically) affect the general perception of well-being, since the second-order components have dependence between them, shared in the general factor. Thus, understanding the PERMA structure as a hierarchical model (as proposed by Goodman et al., 2017) can affect the objective of the interventions, which now have as its main objective the promotion of the general factor (well-being), and to a lesser extent proportion, the promotion of each component.

Regarding reliability, all factors presented good indicators. The engagement component presents the lowest indicator ($\omega = 0.63$ CR = 0.63). This data corroborates what is found in national and international literature in studies adapting the PERMA-Profile instrument in different cultures with adolescents (Singh and Raina, 2020; Waigel and Lemos, 2023) and with adults (Wammerl et al., 2019; Carvalho et al., 2021; Alves et al., 2023). A possible explanation for this aspect is due to the specificity/amplitude of the engagement measure, since engagement behavior may refer to a specific context or activity (e.g., playing), and may not necessarily be generalizable to other contexts or activities (e.g., study).

Another factor that must be considered concerns item 17, which has the lowest factor loading (0.319), and corresponds to the flow construct indicator. Flow is understood as complete engagement, in which the individual loses track of the passage of time and there is full cognitive, emotional and psychological functioning (Csikszentmihalyi, 1998). It is noteworthy that measuring flow together with engagement and only through a single item can contribute to the lower factor loading in relation to the other items, and the consequent reliability indicator.

The need to include more items to assess engagement in adolescents is highlighted, as well as to elucidate the contexts that will be added in the assessment of the component. Although school engagement is traditionally adopted in adolescence (Krok, 2018), studies have emphasized the importance of engagement in leisure activities for building well-being (Carrillo et al., 2021). Furthermore, Fernandes and Zanini (manuscript under construction), points out the tendency of adolescents to interpret engagement as related to leisure activities with family and peers.

Regarding the assessment of scale invariance in relation to age and gender, configural, metric and scalar invariance analyzes were conducted. The results indicate measurement invariance, that is, the equivalence of the factorial structure both among younger adolescents (aged between 11 and 15 years old) and among older adolescents (aged between 16 and 19 years old), as well as among boys and girls. Similar results regarding invariance based on gender were observed in the study by Reinhardt et al. (2020), which examined the Mental Health Continuum-Short form instrument (Keyes, 2006) to assess mental well-being in adolescents. The importance of evaluating invariance for different gender identities is highlighted, and not just for gender. However, in this study, it was not feasible to carry out such an analysis, due to the small number of participants with diverse gender identities.

With regard to the relation between the PERMA components, we first point to the strong and moderate correlation between the components. Such evidence corroborates the PERMA proposal (Seligman, 2011, 2018), which theoretically proposes strong relations between the variables, however, maintaining their independence. As well as studies that show the correlation between the hedonic and eudaimonic dimensions of well-being (Kern et al., 2015; Krok, 2018; Cabrera et al., 2020; Mesurado et al., 2021). Furthermore, moderate to strong correlations were found between the PERMA components and subjective well-being in the dimensions of positive affects and life satisfaction, indicating the relation between positive aspects and the different models of well-being. Such results are repeated in several studies (Tansey et al., 2017; Lai et al., 2018; Choi, 2021; Donaldson et al., 2021).

In addition to the positive correlations between PERMA and positive variables, inverse correlations were found between PERMA and negative variables. PERMA was inversely related to negative affects, distress, depression and anxiety, with weak to moderate magnitude. Well-being in general is inversely related to negative health outcomes, such as anxiety and depression (Kern et al., 2015; Lai et al., 2018; Trigg, 2021). These associations between variables allow the interpretation that well-being variables may have a bidirectional correlation between positive and negative outcomes.

On the other hand, the engagement and achievement components present the weakest correlations with risk variables. Regarding engagement, the scope of the construct and the reliability indicators, already discussed in this study, must be considered. Such aspects can influence the correlations with the variables studied. Concerning achievement, one must consider the understanding of the construct in the adolescent population. In a study carried out by Fernandes and Zanini (manuscript under construction), the authors indicate that this component, for adolescents, is related to future goals, different from the understanding developed by adults, which is related to actions already completed (Seligman, 2011). Therefore, for the adolescent population, this component may present itself differently in relation to other variables.

In conclusion, this study offers in-depth insight into the adaptation and psychometric properties of the PERMA-Profiler scale for adolescents, highlighting the factor structure of five correlated factors as a robust model for understanding well-being in this specific group. This understanding aligns with the properties indicated by Seligman (2018), emphasizing the independent contribution of each component to the construction of well-being. This approach enables interventions focused on a specific component, recognizing the relevance of different well-being profiles in different contexts (Butler and Kern, 2016).

The results also point to the need to expand the evaluation of the engagement component, recognizing the diversity of contexts in which engagement is manifested in adolescence. While the measurement invariance was observed between different age groups and between sexes, the analysis also suggests the importance of expanding the investigation to include different gender identities. Furthermore, the correlations between the PERMA components and related variables reinforce the relevance of these elements in promoting well-being and reducing risk factors associated with mental health. However, the achievement component should be further studied in the adolescent population.

This study represents a significant advance in understanding well-being in adolescence, offering valuable insights into the structure and applicability of the PERMA model in a crucial developmental context. The findings from this study have significant practical implications, especially in the fields of education and mental health. The PERMA-Profiler scale for adolescents provide a robust tool for assessing well-being in this demographic. Educational policies can leverage this tool to monitor and enhance student well-being, implementing targeted interventions that address specific components of the PERMA model. In mental health practices, the validated PERMA-Profiler can aid in the early identification of adolescents at risk for negative outcomes such as anxiety and depression. Mental health professionals can develop tailored intervention programs that focus on enhancing particular aspects of well-being, such as positive relationships and meaning, to mitigate these risks.

However, some limitations of the research deserve to be highlighted. The limitations of this study can be indicated first in relation to the characteristics of the sample, which, although it represents part of the adolescent population in number, does not represent the diversity of gender and regions of the country. Therefore, it is suggested that future studies investigate the PERMA model in different populations in terms of sociodemographic and specific characteristics.

Furthermore, the cross-sectional research design used in this study to evaluate the correlation with related variables prevents the inference of causality between the variables, which limits their understanding. Self-reported data can be influenced by the respondent's current mood, social desirability, and recall accuracy, potentially skewing the results. Future research should consider incorporating multiple data sources, such as teacher reports, peer assessments, and behavioral observations, to triangulate findings and provide a more comprehensive understanding of adolescent well-being. Additionally, longitudinal studies would help in understanding the stability of the PERMA components over time and their impact on long-term outcomes, as well as estimating the relation with external variables.

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. Requests to access the datasets should be directed to IF, iorhanafernandes@hotmail.com.

Ethics statement

The studies involving humans were approved by Ethics Committee – Pontifícia Universidade Católica de Goiás – Approval number: 53061221.9.0000.0037. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

IF: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. DZ: Writing – original draft, Project administration, Formal analysis, Writing – review & editing, Visualization, Supervision, Methodology, Investigation, Conceptualization. EP: Validation, Writing – review & editing, Supervision, Software, Methodology, Formal analysis, Data curation.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alves, M. A., Palmer, S., and Gouveia, M. J. (2023). Psychometric properties of the Portuguese version of the PERMA-profiler. *Trends Psychol.* 2023, 1–15. doi: 10.1007/s43076-023-00261-3

Asquith, S. L., Wang, X., Quintana, D. S., and Abraham, A. (2022). The role of personality traits and leisure activities in predicting wellbeing in young people. *BMC Psychol* 10:249. doi: 10.1186/s40359-022-00954-x

Bartholomaeus, J. D., Iasiello, M. P., Jarden, A., Burke, K. J., and van Agteren, J. (2020). Evaluating the psychometric properties of PERMA profiler. *J. Well Being Assess.* 4, 163–180. doi: 10.1007/s41543-020-00031-3

Bazargan-Hejazi, S., Shirazi, A., Wang, A., Shlobin, N. A., Karunungan, K., Shulman, J., et al. (2021). Contribution of a positive psychology-based conceptual framework in reducing physician burnout and improving well-being: a systematic review. *BMC Med. Educ.* 21:593. doi: 10.1186/s12909-021-03021-y

Blaire, M., and Simmons, L. (2021). A 'PERMA' response to the pandemic: an online positive education Programme to promote wellbeing in university students. *Front. Educ.* 6:642632. doi: 10.3389/feduc.2021.642632

Borsa, J. C., Damásio, B. F., and Bandeira, D. R. (2012). Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações. *Paidéia* 22, 423–432. doi: 10.1590/S0103-863X2012000300014

Butler, J., and Kern, M. L. (2016). The PERMA-profiler: a brief multidimensional measure of flourishing. *Int. J. Wellbeing* 6, 1–48. doi: 10.5502/ijw.v6i3.526

Cabrera, M. C., Larrañaga, E., and Yubero, S. (2020). The role of emotions, moral disengagement and gender in supporting victims of bullying. *Educ. Sci.* 10:365. doi: 10.3390/educsci10120365

Carrillo, G., Alfaro, J., Aspillaga, C., Ramírez-Casas del valle, L., Inostroza, C., and Villarroel, A. (2021). (2021). Well-being from the understanding of children and adolescents: a qualitative Metasynthesis. *Child Ind. Res.* 14, 1677–1701. doi: 10.1007/s12187-021-09816-1

Carvalho, T. F., Aquino, S. D., and Natividade, J. C. (2021). Flourishing in the Brazilian context: evidence of the validity of the PERMA-profiler scale. *Curr. Psychol.* 42, 1828–1840. doi: 10.1007/s12144-021-01587-w

Chaves, C., Ballesteros-Valdés, R., Madridejos, E., and Charles-Leija, H. (2023). PERMA-profiler for the evaluation of well-being: adaptation and validation in a sample of university students and employees in the Mexican educational context. *Appl. Res. Qual. Life* 18, 1225–1247. doi: 10.1007/s11482-022-10132-1

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Choi, M. Y. (2021). Mental and physical factors influencing wellbeing among south Korean emergency workers. *Int. J. Environ. Res. Public Health* 18:70. doi: 10.3390/ijerph18010070

Cobo-Rendón, R., Pérez Villalobos, M. V., and Díaz Mujica, A. (2020). Propiedades psicométricas del PERMA-Profiler para la medición del bienestar en una muestra de estudiantes universitarios chilenos. *Rev. Cienc. Salud* 18, 119–133. doi: 10.12804/revistas.urosario.edu.co/revsalud/a.8775

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *2nd* Edn. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Colwell, S. R. (2016). The composite reliability calculator. Technical Report.

Csikszentmihalyi, M. (1998). Aprender a fluir. Barcelona, España. Kairós.

Donaldson, S. I., Heshmati, S., Lee, J. Y., and Donaldson, S. I. (2021). Examining building blocks of well-being beyond PERMA and self-report bias. *J. Posit. Psychol.* 16, 811–818. doi: 10.1080/17439760.2020.1818813

Giacomoni, C. H., Segabinazi, J. D., Zortea, M., and Zanon, C. (2012). Escala Global de Satisfação com a Vida dos Adolescentes. Porto Alegre: Relatório Técnico Não Publicado.

Giangrasso, B. (2021). Psychometric properties of the PERMA-profiler as hedonic and eudaimonic well-being measure in an Italian context. *Curr. Psychol.* 40, 1175–1184. doi: 10.1007/s12144-018-0040-3

Goodman, F., Disabato, D., Kashdan, T., and Kauffman, S. (2017). Measuring well-being: A comparison of subjective wellbeing and PERMA. *The Journal of Positive Psychology*, 1–12. doi: 10.1080/17439760.2017.1388434

International Test Commission. (2017). The ITC Guidelines for Translating and Adapting Testes (Second edition). Available at: https://www.intestcom.org/

Kenny, D. A., Kaniskan, B., and McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociol. Method Res.* 44, 486–507. doi: 10.1177/0049124114543236

Kern, M. L., Waters, L. E., Adler, A., and White, M. A. (2015). A multidimensional approach to measuring well-being in students: application of the PERMA framework. *J. Posit. Psychol.* 10, 262–271. doi: 10.1080/17439760.2014.936962

Keyes, C. L. (2006). Mental health in adolescence: is America's youth flourishing? *Am. J. Orthop.* 76, 395–402. doi: 10.1037/0002-9432.76.3.395

Krok, D. (2018). When is meaning in life most beneficial to young people? Styles of meaning in life and well-being among late adolescents. *J. Adult Dev.* 25, 96–106. doi: 10.1007/s10804-017-9280-y

Lai, M. K., Leung, C., Kwok, S. Y. C., Hui, A. N. N., Lo, H. H. M., Leung, J. T. Y., et al. (2018). A multidimensional PERMA-H positive education model, general satisfaction of school life, and character strengths use in Hong Kong senior primary school students: confirmatory factor analysis and path analysis using the APASO-II. *Front. Psychol.* 9:1090. doi: 10.3389/fpsyg.2018.01090

Mahamid, F., Veronese, G., and Bdier, D. (2023). The factorial structure and psychometric properties of the PERMA-profiler Arabic version to measure well-being within a Palestinian adult population. *Middle East. Curr. Psychiatry* 30:10. doi: 10.1186/s43045-023-00282-9

Mesurado, B., Resett, S., Tezón, M., and Vanney, C. E. (2021). Do positive emotions make you more prosocial? Direct and indirect effects of an intervention program on pro-sociality in Colombian adolescents during social isolation due to COVID-19. *Front. Psychol.* 12:710037. doi: 10.3389/fpsyg.2021.710037

Muthén, L. K., and Muthén, B. O. (2012). Mplus user's guide. *7th* Edn. Los Angeles, CA: Muthén and Muthén.

Muthén, L. K., and Muthén, B. O. (2017). Mplus: statistical analysis with latent variables: user's guide (version 8). Los Angeles, CA: Authors.

O'Connor, M., Sanson, A. V., Toumbourou, J. W., Norrish, J., and Olsson, C. A. (2016). Does positive mental health in adolescence longitudinally predict healthy transitions in young adulthood? *J. Happiness Stud.* 18, 177–198. doi: 10.1007/s10902-016-9723-3

Peixoto, E. M., Zanini, D. S., and Andrade, J. M. (2021). Cross-cultural adaptation and psychometric properties of the Kessler distress scale (K10): an application of the rating scale model. July 2021. *Psicol. Reflex. Crít.* 34:21. doi: 10.1186/s41155-021-00186-9

Pérez-Fuentes, M. C., Gázquez-Linares, J. J., Molero-Jurado, M. M., Martos-Martínez, Á., Barragán-Martín, A. B., and Simón-Márquez, M. M. (2021). Student burnout and engagement: relationship with adolescent use of alcohol and attitudes towards authority. *Int. J. Clin. Health Psychol.* 21:100225. doi: 10.1016/j.ijchp.2021.100225

Reinhardt, M., Horváth, Z., Morgan, A., and Kökönyei, G. (2020). Well-being profiles in adolescence: psychometric properties and latent profile analysis of the mental health continuum model – a methodological study. *Health Qual. Life Outcomes* 18:95. doi: 10.1186/s12955-020-01332-0

Ribeiro, C. C., Yassuda, M. S., and Neri, A. L. (2020). Propósito de vida em adultos e idosos: revisão integrativa. *Ciência Saúde Coletiva* 25, 2127–2142. doi: 10.1590/1413-81232020256.20602018

Ryan, J., Curtis, R., Olds, T., Edney, S., Vandelanotte, C., Plotnikoff, R., et al. (2019). Psychometric properties of the PERMA profiler for measuring wellbeing in Australian adults. *PLoS One* 14:e0225932. doi: 10.1371/journal.pone.0225932

Schmiedeberg, C., and Schröder, J. (2017). (2017). Leisure activities and life satisfaction: an analysis with German panel data. *Appl. Res. Qual. Life* 12, 137–151. doi: 10.1007/s11482-016-9458-7

Segabinazi, J. D., Zortea, M., Zanon, C., Bandeira, D. R., Giacomoni, C. H., and Hutz, C. S. (2012). Escala de afetos positivos e negativos para adolescentes: adaptação, normatização e evidências de validade. *Aval. Psicol.* 11, 1–12.

Seligman, M. E. (2011). Florescer: Uma nova e visionária interpretação da felicidade e do bem-estar. Rio de Janeiro: Objetiva.

Seligman, M. (2018). PERMA and the building blocks of well-being. *J. Posit. Psychol.* 13, 333–335. doi: 10.1080/17439760.2018.1437466

Sellbom, M., and Tellegen, A. (2019). Factor analysis in psychological assessment research: common pitfalls and recommendations. *Psychol. Assess.* 31, 1428–1441. doi: 10.1037/pas0000623

Simões, É. V., de Oliveira, A. M. N., de Pinho, L. B., Lourenção, L. G., de Oliveira, S. M., and de Farias, F. L. R. (2022). Reasons assigned to suicide attempts: adolescents' perceptions. *Rev. Bras. Enferm.* 75:e20210163. doi: 10.1590/0034-7167-2021-016

Singh, K., and Raina, M. (2020). Demographic correlates and validation of PERMA and WEMWBS scales in Indian adolescents. *Child Indic. Res.* 13, 1175–1186. doi: 10.1007/s12187-019-09655-1

Tabachnick, G. B., and Fidell, L. S. (2019). "Principal components and factor analysis" in Using multivariate statistics. eds. G. B. Tabachnick and L. S. Fidell (Boston: Pearson), 476–527.

Tansey, T. N., Smedema, S., Umucu, E., Iwanaga, K., Wu, J.-R., Cardoso, E. S., et al. (2017). Assessing college life adjustment of students with disabilities: application of the PERMA framework. *Rehabil. Couns. Bull.* 61, 131–142. doi: 10.1177/0034355217702136

The Jamovi Project (2022). jamovi. (Version 2.3) [Computer Software]. Available at: https://www.jamovi.org.

Trigg, J. (2021). Examining the role of pets in cancer survivors' physical and mental wellbeing. *J. Psychosoc. Oncol.* 40, 1–20. doi: 10.1080/07347332.2021.1936337

Tu, M., Wang, F., Shen, S., Wang, H., and Feng, J. (2021). Influences of psychological intervention on negative emotion, Cancer-related fatigue and level of Hope in lung Cancer chemotherapy patients based on the PERMA framework. *Iran. J. Public Health* 50, 728–736. doi: 10.18502/ijph.v50i4.5997

Turner, K., and Thielking, M. (2019). Teacher wellbeing: its effects on teaching practice and student learning. *Issues Educ. Res.* 29, 938–960.

Umucu, E., Wu, J. R., Sanchez, J., Brooks, J. M., Chiu, C. Y., Tu, W. M., et al. (2020). Psychometric validation of the PERMA-profiler as a well-being measure for student veterans. *J. Am. Coll. Heal.* 68, 271–277. doi: 10.1080/07448481.2018.1546182

Usán, P., and Salavera, C. (2019). El rendimiento escolar, la inteligencia emocional y el engagement académico en adolescentes. *Electron. J. Res. Educ. Psychol.* 17, 5–26. doi: 10.25115/ejrep.v17i47.1879

Waigel, N. C., and Lemos, V. N. (2023). Psychometric properties of PERMA profiler scale in Argentinian adolescents. *Int. J. Psychol. Res.* 16, 103–113. doi: 10.21500/20112084.5737

Wammerl, M., Jaunig, J., Mairunteregger, T., and Streit, P. (2019). The German version of the PERMA- profiler: evidence for construct and convergent validity of the PERMA theory of well-being in German speaking countries. *J. Wellbeing Assess.* 3, 75–96. doi: 10.1007/s41543-019-00021-0

Wu, Q., Zhao, J., Zhao, G., Li, X., Du, H., and Chi, P. (2022). Affective profiles and psychosocial adjustment among Chinese adolescents and adults with adverse childhood experiences: a person-centered approach. *J. Happiness Stud.* 23, 3909–3927. doi: 10.1007/s10902-022-00566-7

Zawadzka, D., Korzycka, M., and Oblacińska, A. (2018). Sense of meaning in life as a factor protecting middle school adolescents from perpetrating and experiencing peer violence. *Psychiatr. Psychol. Klin.* 18, 340–353. doi: 10.15557/PiPK.2018.0042

Zewude, G. T., and Hercz, M. (2022). Psychometric properties and measurement invariance of the PERMA profiler in an Ethiopian higher education setting. *Pedagogika* 146, 209–236. doi: 10.15823/p.2022.146.11

Zhang, R., Li, D., Chen, F., Ewalds-Kvist, B. M., and Liu, S. (2017). Interparental conflict relative to suicidal ideation in Chinese adolescents: the roles of coping strategies and meaning in life. *Front. Psychol.* 8:1010. doi: 10.3389/fpsyg.2017.01010

Zhao, Y., Zheng, Z., Pan, C., and Zhou, L. (2021). Self-esteem and academic engagement among adolescents: a moderated mediation model. *Front. Psychol.* 12:690828. doi: 10.3389/fpsyg.2021.690828

frontiers | Frontiers in Psychology

TYPE Brief Research Report
PUBLISHED 15 July 2024
DOI 10.3389/fpsyg.2024.1386235

Check for updates

OPEN ACCESS

EDITED BY
Paola Gremigni,
University of Bologna, Italy

REVIEWED BY
Thomas Salzberger,
Vienna University of Economics and Business, Austria
Tonya Hansel,
Tulane University, United States

*CORRESPONDENCE
Karina Alarcón-Castillo
✉ karina.alarcon.castillo@gmail.com

RECEIVED 14 February 2024
ACCEPTED 01 July 2024
PUBLISHED 15 July 2024

CITATION
Santibáñez-Palma JF, Ferrer-Urbina R, Sepúlveda-Páez G, Bravo de la Fuente J and Alarcón-Castillo K (2024) Development and validation of the Environmental Confinement Stressors Scale (ECSS-20). *Front. Psychol.* 15:1386235.
doi: 10.3389/fpsyg.2024.1386235

Development and validation of the Environmental Confinement Stressors Scale (ECSS-20)

J. Francisco Santibáñez-Palma, Rodrigo Ferrer-Urbina, Geraldy Sepúlveda-Páez, Josefa Bravo de la Fuente and Karina Alarcón-Castillo*

Escuela de Psicología y Filosofía, Universidad de Tarapacá, Arica, Chile

The COVID-19 pandemic has generated a global crisis with severe consequences for public health. There have been negative impacts on people's quality of life and mental health due to various stressors arising in this context, such as physical, social, economic, and psychological challenges. Noteworthy among these are the indirect effects of health measures, especially social distancing and confinement, which have significantly altered people's daily lives and social activities, producing high levels of anxiety, depression, and stress. This study proposes developing and validating a cross-sectional scale called the "Environmental Stressors Scale (ECSS-20)" to address the need to measure the impact of environmental stressors during confinement. The scale, which has been validated following ethical and methodological guidelines, consists of four dimensions: economic stressors (EE), social activities (SA), habitability (H), and exposure to virtual media (EMV). A pilot study ($n = 113$) and a main study ($n = 314$) were applied. The results showed that the instrument has a reliable and valid structure, with satisfactory internal consistency and factorial validity. Likewise, gender invariance tests supported its suitability for its applicability to women and men. Overall, the ECSS-20 is a valuable instrument for assessing the impact of confinement and improving the understanding of people's subjective experiences in this situation. Future research could further develop its applicability in different contexts and populations to better understand its usefulness and psychometric properties.

KEYWORDS
environmental stressors, COVID-19, confinement, ESEM, psychometric scales development

Introduction

The pandemic context resulting from COVID-19 has depicted a global emergency, from the health point of view, with more than 6,987,222 deaths to date [World Health Organization (WHO), 2023], and in terms of the quality of life and mental health of people, due to the set of stressors that arose in this context (i.e., physical, social, economic, and psychological) (Bavel et al., 2020; Ornell et al., 2020; Wang et al., 2020; Wetherall et al., 2022). Within these stressors are those that emerged from the indirect effects of health policies and containment efforts, specifically, the policies of confinement and social distancing (Caqueo-Urízar et al., 2020), designed to reduce personal interactions and movements (Maier and Brockmann, 2020; Mayr et al., 2020; Badenes-Plá, 2022; Yu et al., 2023), which generated changes in the social and daily activities of the population (e.g., studies, work, intimate relationships, financial management, home habitability) (Ammar et al., 2020; Gloster et al., 2020; Marroquín et al., 2020; Akbari et al., 2021; Amerio et al., 2021; Gruber et al., 2021; Lal et al., 2021; Manchia et al., 2022; Quintana, 2022; von Keyserlingk et al., 2022).

As is well known, the social environment systematically influences health, considering social, psychological, economic, demographic, local and cultural aspects (Wild, 2005; McMichael, 2011; Pettini and Mazzocco, 2022; Whipple and Evans, 2022; Eyre et al., 2023; Gudi-Mindermann et al., 2023; Ibanez and Eyre, 2023). In this sense, any variation in these aspects will impact the health status of people, being detrimental or beneficial to the population. For example, there is ample evidence on how contexts of physical and social isolation (i.e., plague, influenza, cholera, leprosy, or others) and subjective social isolation are associated with negative impacts on mental health (Cacioppo and Hawkley, 2009; Cacioppo and Cacioppo, 2014; Cacioppo et al., 2015; Bzdok and Dunbar, 2022; Gilbar et al., 2022). The growth of such literature has been exponential during and following the COVID-19 pandemic (Jakovljevic et al., 2020; Mukhtar, 2020; Wilder-Smith and Freedman, 2020; Clair et al., 2021; Liu et al., 2021; Neville et al., 2021; Takian et al., 2021), highlighting negative psychological impacts (i.e., distress, anxiety, depression, and high levels of stress) that are primarily attributed to side effects of confinement (Cuadra-Martínez et al., 2020; Gloster et al., 2020; Kujawa et al., 2020; Kokkinos et al., 2022; Nanath et al., 2022). These effects, which have been reported in all types of populations (e.g., children, adolescents, adults, pregnant women, senior citizens, among others) (Manchia et al., 2022), would be explained by the increase in environmental stressors and the variability of coping resources (Verdolini et al., 2021; McLaughlin et al., 2022; Tracy et al., 2022; Yang et al., 2022; Delhey et al., 2023). Now, stress refers to an emergent relationship between the person and the environment (Lazarus and Cohen, 1977) involving environmental stimuli, their evaluation and the organism's response (Cohen et al., 1983, 2007, 2019; Segerstrom and O'Connor, 2012). Likewise, *environmental stress* is defined as the physiological, cognitive, and emotional response that people may experience to various environmental situations, whether at the macro-level (e.g., population density in a city) or in the immediate environment (e.g., housing conditions; Gatersleben and Griffin, 2017).

A wide range of environmental stressors derived from confinement are observed in the literature (Ellen et al., 2021; Hussong et al., 2021; Kumar and Shah, 2021; Kunzler et al., 2021; Salazar et al., 2021; Sheek-Hussein et al., 2021; Szkody et al., 2021; Valdés-Florido et al., 2022; Morgado et al., 2023). Researched are: (1) economic stressors (ES), which refer to the perceived economic impact that a variation in the estimated household budget generates, and thus impacts on job insecurity and economic livelihood in households (Bazzoli et al., 2021; Friedline et al., 2021; Low and Mounts, 2022) (2) everyday activities (EA), understood as the impact on the performance of routine activities, being considered an individual facet of social practice (Rieger and Wang, 2020; Ellen et al., 2021); (3) social activities (SA), which refers to the perceived impact on social recreation activities such as interaction with others and leisure (Kunzler et al., 2021; Nielsen et al., 2021); (4) home habitability (H), referring to the operational housing conditions and comfort (i.e., conditions necessary to satisfy the physical, biological, psychological, and social well-being) of those who inhabit a dwelling (Corral-Verdugo et al., 2011; Zulaica and Oriolani, 2019); and (5) virtual media exposure (VME), understood as the level of person's exposure to, or interaction with, virtual or technological media (e.g., television viewing, computer use, use of social networks, websites and mobile applications; Dubey, 2020; Rivest-Beauregard et al., 2022).

Consequently, the frequency of environmental stressors, whether higher or lower, defines how confinement or other situations that cause stressful environmental changes in people's habitability (e.g., population displacement due to natural disasters and/or in search of shelter) will be experienced, thereby affecting their health (Choi et al., 2020; Kim et al., 2021; Bzdok and Dunbar, 2022). Thus, the evaluation of these environmental factors will provide a more complete view of how the health of the population is affected, and how to counteract this situation. In this way, it will be possible to generate preventive actions, either at the individual or collective level, that are aimed at the well-being of people.

In this line, and considering that efforts to study this phenomenon lack a psychometric instrument with evidence of validity, the present study aims to design a scale that assesses the perception of change in environmental conditions of confinement incorporating a cross-sectional approach, and thus reduce the existing gap in research on the measurement of the impact of confinement (Manchia et al., 2022). To this end, updated evidence of both reliability and validity is presented, following the guidelines of ethical and methodological standards recognized in the field of psychometric evaluation (Prieto and Delgado, 2010; American Educational Research Association, 2014; Muñiz and Fonseca-Pedrero, 2019).

Method

Procedures

The study was approved by the Ethics Committee of the Universidad de Tarapacá. An instrumental study was conducted, i.e., a battery of instruments was applied, with a cross-sectional design, i.e., applied over a period of time (Ato et al., 2013).

Initially, 68 items were profiled and evaluated by expert judges (two judges with experience in psychometrics and one judge specialized in the health area) in terms of grammatical adequacy (coherence and clarity) and construct representativeness, using a score of "−1, 0, 1" where "1" represents the grammatical adequacy and construct representativeness of the item. Means were then calculated and items with means less than or equal to 0 were eliminated; 45 items were retained from this process and applied to an online pilot study.

The pilot sample was collected through non-probability sampling strategies (Otzen and Manterola, 2017), using snowball and social network strategies (Montero and León, 2007). It consisted of 113 adults between 18 and 51 years of age ($M = 27.1$; $SD = 7.29$), 83 women (73.5%), 27 men (23.9%) and 3 (2.7%) individuals who did not identify with any of the aforementioned groups, coming from the Biobío region (45%, $n = 50$), the Arica and Parinacota region (42.3%, $n = 47$) and other regions of the country (12.7%, $n = 16$). It was surveyed online during October 2021 using a Google Form with a response procedure of 30 to 35 min.

Once the pilot sample was collected, an exploratory factor analysis (EFA) was conducted to explore the underlying structure of the data (Costello and Osborne, 2019). In addition, to provide a brief and concise scale representing the construct of interest, items with values below 0.50 on the factor loadings of each item were iteratively removed.

The results of the EFA suggested a new dimension, which included items related to exposure to virtual media (e.g., being exposed to computers or television and participating in virtual meetings). Thus, a 30-item version was obtained, which was applied to the main study sample, which like the pilot sample, was collected through non-probability sampling strategies (Otzen and Manterola, 2017), using snowball and social networking strategies (Montero and León, 2007), during January 2022 using a Google Form with a response procedure of 20 to 25 min.

Participants

The main study sample consisted of 314 adults between 18 and 79 years of age (M = 27.34; SD = 9.58), 191 women (60.8%), 123 men (39.2%), from the Biobío region (31.5%, n = 99), the Arica and Parinacota region (43.8%, n = 137), the Metropolitan region (9.6%, n = 30), and other regions of the country (15.1%, n = 48). The main study was conducted in the classrooms of the Universidad de Tarapacá during April and May 2022, using QR codes and paper-and-pencil surveys.

Instruments

The *Environmental Confinement Stressors Scale (ECSS-20)* was developed to evaluate the subjective comparison, before and during, of the most predominant environmental stressors established in periods of stress and confinement. The final version of the questionnaire consists of four dimensions of perception: (a) economic stressors (ES), (b) social activities (SA), (c) home habitability (H), and (d) exposure to virtual media (VME), with five items each, for a total of 20 items. The response options have a Likert format of 5 ranked categories (−2 = "Much less than before," 2 = "Much more than before"). In the EE and EMV dimensions, higher scores are interpreted as experiencing a significant increase in environmental stress than before confinement. In the AS and H dimensions, higher scores are interpreted as experiencing a significant decrease in environmental stress than before confinement. The statements refer to facts and behaviors associated with environmental stressors in confinement.

Perceived Stress Scale (PSS-14): It is a 14 item self-report designed to assess "the degree to which life situations are evaluated as stressful" (Cohen et al., 1983), was applied in the main study. Tapia et al. (2007) validated and adapted this inventory in Chile. In the Chilean population, this inventory has presented a Cronbach's alpha higher than 0.889 (González-Tovar and Hernández-Rodríguez, 2020). Half of the questions are positively formulated and reverse-coded. Each item is scored on a 5-point scale (0 = never, 4 = very often). Individual scores on the PSS-14 can range from 0 to 56, considering that (1) scores between 0 and 19 would be considered no stress; (2) scores ranging from 2 to 28 are considered low stress; (3) scores ranging from 29 to 38 would be considered moderate stress; and (4) scores ranging from 39 to 56 are considered high perceived stress (Tapia et al., 2007).

Data analysis

First, in the main investigation, an exploratory structural equation model (ESEM) was performed. An exploratory structural equation model is a statistical modeling technique that combines the advantages of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), allowing estimating the effects and relationships between variables in a more precise and flexible way by accounting for measurement errors in both dependent and independent variables (Morin et al., 2016; Marsh et al., 2020). Oblimin rotation (Asparouhov and Muthén, 2009) and the weighted least square mean and variance adjusted (WLSMV) estimation method were used for the ESEM, which is robust, with non-normal discrete variables from the matrix of polychoric correlations (Muthén and Asparouhov, 2011; Brown, 2015).

Second, the following cut-off points were considered for the overall model fit: values higher than 0.96 in comparative fit index (CFI) or Tucker–Lewis index (TLI) and values lower than 0.07 in root-mean-squared error of approximation (RMSEA) (Hair et al., 2019). The modification indexes and cross-loadings between various items of different dimensions were analyzed.

Using an iterative approach, three fundamental criteria were applied: selecting items with moderate or vigorous factor loadings ($\lambda > 0.50$), the elimination of redundant items, and the deletion of items with solid cross-loadings ($\lambda > 0.30$) (Muthén and Asparouhov, 2012; Xiao et al., 2019).

Third, reliability was estimated for each dimension using Cronbach's Alpha coefficient. Additionally, McDonald's hierarchical omega was provided to report more efficient reliability criteria, with values above 0.70 considered acceptable and above 0.80 adequate (Cho and Kim, 2015; McNeish, 2018).

Fourth, to assess the instrument's stability between different genders (women and men), invariance tests were performed to verify that the scores of the items have the same meaning for both groups and do not present biases (Leitgöb et al., 2023). For this purpose, the increase in RMSEA (> 0.010) and the decrease in CFI (>0.005) were considered as evidence of invariance (Cheung and Rensvold, 2002; Chen, 2007; Dimitrov, 2010).

Finally, to establish existing relationships with other variables, a SET-ESEM was performed between ECSS-20 dimensions and PSS-14 dimensions (Cohen et al., 1983); for this, the WLSMV estimation method and the polychoric correlations matrix were used. Sequential analyses were performed using Mplus (8.0) (Muthén et al., 2017) and Jamovi (2.2.5) (The Jamovi, Project, 2021) statistical software.

Results

First, based on a qualitative analysis, it was decided to discard one item of the dimension because it loaded positively on the factor despite being classified as an inverse item. Then, a 29 items ESEM model (Model 1; M1) was estimated, as shown in Table 1. This model showed an excellent statistical fit according to the parameters proposed by Hair et al. (2019), for the CFI estimator (CFI = 0.972; TLI = 0.958). However, three items of the everyday activities factor presented relevant cross-loadings (i.e., $\lambda > 0.5$) on the social activities factor (i.e., perform physical activity; obtain medical or health care easily; take walks/visits).

Consequently, after a qualitative analysis (i.e., item relevance and construct definition), the items above became part of the social activities factor. Likewise, the remaining items of the everyday activities factor (e.g., carrying out procedures typically) were discarded due to their redundancy, leaving 26 items.

TABLE 1 Fit indexes for models ESEM and CFA of ECSS-20.

Model	$\chi 2$	df	χ^2/df	RMSEA	90% CI	CFI	TLI	SRMR
M1	838.537*	271	3.09	0.082	[0.075, 0.088]	0.972	0.958	0.023
M2	730.636*	227	3.21	0.084	[0.077, 0.091]	0.972	0.960	0.024
M3	380.155*	116	3.27	0.085	[0.076, 0.095]	0.977	0.962	0.021
M3a	140.970*	164	0.85	0.069	[0.061, 0.078]	0.978	0.975	0.044

*p < .001. M1, ESEM with four factors, 29 items; M2, ESEM with four factors, 26 items; M3, ESEM with four factors, 20 items; M3a, CFA with four factors, 20 items; χ^2, Chi-square; df, degree of freedom; RMSEA, root mean square error of approximation; 90% CI, confidence interval; CFI, comparative fit index; TLI, Tucker–Lewis index; SRMR, standardized root mean square residual.

Second, a 4-factor model with 26 items was estimated (Model 2; M2), which presented better internal structure adjustments in the CFI and TLI indicators (CFI = 0.972; TLI = 0.960), compared to M1. As previously mentioned, items with lower factor loadings referring to the factor ($\lambda > 0.5$) and that presented cross-loadings ($\lambda \leq 0.3$) were iteratively eliminated to reduce the number of items in the scale. In total, 20 items were retained, estimated in Model 3 (M3).

The ESEM analysis of M3 presented better internal structure adjustments in the CFI and TLI indicators (CFI = 0.977; TLI = 0.962) than the previous models. This structure was confirmed by performing a CFA, which evidenced a satisfactory fit for the CFI, TLI, and RMSEA indicators (CFI = 0.978; TLI = 0.975; RMSEA = 0.069) (Hair et al., 2019). Finally, the ECSS-20 comprises four dimensions (i.e., ES, SA, H, and VME) and five items per dimension (i.e., 20 items in total).

Table 2 presents the factor loadings with their corresponding factorial covariances and reliability coefficients of the M3. Factorial loadings for this model proved adequate for each factor, and no relevant cross-loadings were observed. Also, structural relationships between dimensions were moderate ($r > 0.30$), mild ($r > 0.10$; Cohen, 1988), and null, and reliability estimates were adequate ($\omega > 0.89$; $\alpha > 0.89$; Cho and Kim, 2015).

Third, a multigroup CFA model was estimated between M3 men and women; the results are presented in Table 3. This model was first tested for configural invariance, i.e., a baseline model was fitted for each group separately. Compared with the configural model, the metric model showed no relevant changes in the RMSEA differential or CFI, thus confirming that the factor loadings of the items are the same in both groups. Finally, the scalar model compared with the configural model also showed no relevant changes in the RMSEA differential or CFI, which means that the intercepts of the items are the same in both groups.

Thus, strong measurement invariance is demonstrated by the existence of metric and scalar invariance, i.e., the equivalence between factor loadings and thresholds for those who identified themselves as female or male is sustained (Cheung and Rensvold, 2002; Chen, 2007; Dimitrov, 2010; Leitgöb et al., 2023).

Finally, the SET-ESEM model that estimated the association between the latent dimensions of the ECSS-20 and the PSS-14 one-dimensional showed comparative and absolute fit indices far from the recommendations [$\chi^2(469) = 1780.710$; CFI = 0.911; TLI = 0.893; RMSEA = 0.094; 90% CI = (0.090, 0.099); SRMR = 0.094]. The observed mismatches could be attributed to the factor loadings shown by the PSS-14 (see Figure 1). Finally, significant direct and inverse loadings are observed for the ES factor ($\lambda = 0.323$) and the H factor ($\lambda = -0.301$) concerning PSS-14. The details of the standardized relationships between the latent dimensions and the PSS-14 indicators are shown in the Figure 1.

Discussion

The present study focused on developing and validating the Environmental Stressors in Confinement Scale (ECSS-20), which assesses the perception of change in environmental stressors produced by confinement circumstances. Theoretical and practical contributions, limitations, and future lines of research emerging from this study are discussed below.

In first place, the final structure of the ECSS-20, composed of four dimensions: economic stressors (ES), social activities (SA), habitability (H), and exposure to virtual media (VME), proved to be robust and consistent with previous literature evidencing the existence of environmental stressors stemming from confinement (Ellen et al., 2021; Kunzler et al., 2021; Sheek-Hussein et al., 2021). The reason for this is that the observed factor loadings indicate that each dimension uniquely impacts the perception of environmental stress in confinement situations. McDonald's omega ($\omega > 0.89$) and Cronbach's alpha ($\alpha > 0.89$; Cho and Kim, 2015), internal consistency values were also shown to be satisfactory, providing evidence of the reliability of the instrument. That is, the ECSS-20 can be applied to the Chilean adult population experiencing confinement measures.

Second, the application of gender invariance tests supported the equivalence of factor loadings between women and men, suggesting that the ECSS-20 can be used in both groups without distinction. Thus, this strengthens the instrument's usefulness, as it demonstrates that it assesses the impact of environmental stressors accurately and comparatively, as has been evidenced in other studies analyzing gender invariance in confinement contexts (Prime et al., 2021).

Third, validity tests based on the association with other variables were established by demonstrating significant relationships between the ECSS-20 and PSS-14 dimensions. Specifically, direct significant relationships were observed between the economic stressors (ES) dimension and the perception of stress. Therefore, the greater the perception of variation in the estimated household budget, the greater the feeling of stress. Likewise, inverse relationships were observed between the dimension of habitability (H) and the perception of stress. In other words, the higher the perception of well-being with those who share a dwelling, the lower the feeling of stress. As a result, this supports previous research that suggests these factors impact the change in people's mental health (Kunzler et al., 2021; Sheek-Hussein et al., 2021). Thus, this demonstrates that habitability and economic conditions could function as protective and risk factors in the perception of stress during confinement situations, such as during the COVID-19 pandemic. Therefore, it is proposed to consider these subscales to assess these factors in confinement circumstances.

TABLE 2 Standardized factor loadings resulting from ESEM, factorial covariations and reliability coefficients (Cronbach's alpha and McDonald's omega) for each dimension of ECSS-20.

	Descriptive Stadistics			Factor Loadings				Reliability Statistics	
	M (SD)	S	K	ES	SA	H	VME	α if item is dropped	ω if item is dropped
Economic stressors (ES)									
1. Difficulties organizing household finances	0.08 (1.15)	−0.263	−0.574	0.804	−0.004	−0.079	0.021	0.888	0.895
2. Difficulties generating income	0.28 (1.10)	−0.315	−0.362	0.762	0.101	−0.077	0.089	0.882	0.890
3. Difficulties covering essential household services	0.04 (0.83)	−0.795	0.334	0.925	−0.048	0.074	−0.026	0.873	0.876
4. Difficulties meeting bank or retail debts	0.08 (1.00)	−0.117	0.081	0.910	0.005	0.032	−0.024	0.866	0.874
5. Difficulties paying for educational services or health services	0.07 (1.03)	−0.200	−0.037	0.796	0.018	−0.001	0.013	0.879	0.889
Social activities (SA)									
6. Participating in social gatherings	−0.26 (1.53)	0.301	−1.417	0.040	0.888	−0.018	−0.059	0.877	0.878
7. Having romantic or sexual relationships	−0.12 (1.28)	0.077	−0.895	0.046	0.786	−0.019	0.008	0.895	0.897
8. Sharing with those close to me	−0.21 (1.40)	0.215	−1.256	−0.012	0.900	−0.006	0.026	0.873	0.874
9. Physical activity	−0.21 (1.37)	0.218	−1.159	−0.031	0.759	0.054	0.036	0.897	0.899
10. Going for walks/visits	−0.24 (1.40)	0.268	−1.213	−0.013	0.860	0.052	0.023	0.873	0.876
Habitability (H)									
11. Having privacy in my home	−0.29 (1.14)	0.181	−0.400	−0.007	−0.049	0.899	−0.033	0.877	0.880
12. Finding silence in my home	−0.38 (1.15)	0.241	−0.502	0.049	0.043	0.921	−0.076	0.860	0.865
13. Having space to carry out my activities	−0.42 (1.12)	0.318	−0.385	0.008	0.111	0.822	−0.101	0.872	0.875
14. Cooking comfortably	−0.21 (0.98)	−0.100	−0.142	−0.015	−0.026	0.795	0.178	0.877	0.880
15. Feeling comfortable in the bathroom	−0.19 (0.94)	−0.138	0.709	−0.028	0.010	0.720	0.199	0.893	0.895
Virtual media exposure (VME)									
16. Using mobile applications to shop from home	0.43 (1.21)	−0.344	−0.701	0.105	−0.145	0.033	0.737	0.929	0.931
17. Being exposed to computers or television	0.68 (1.29)	−0.638	−0.600	0.005	0.040	−0.030	0.920	0.900	0.905
18. Participating in virtual meetings	0.75 (1.36)	−0.808	−0.564	−0.029	−0.004	0.041	0.867	0.911	0.915
19. Using technological devices	0.86 (1.24)	−0.819	−0.332	0.010	0.037	0.037	0.940	0.890	0.892
20. Accessing the internet	0.69 (1.23)	−0.551	−0.628	−0.003	0.018	−0.046	0.887	0.900	0.902
Correlations								**ω index**	**α index**
Economic Stressors				—	—	—	—	0.899	0.906
Social Activities				0.312*	—	—	—	0.904	0.906
Home Habitability				0.081	0.449*	—	—	0.899	0.900
Exposure to Virtual Media				0.290*	0.085	0.207*	—	0.924	0.926

*$p < 0.001$; ES, economic stressors; SA, social activities; H, home habitability; VME, exposure to virtual media; M, mean; SD, standard deviation; S, skewness standardized; K, kurtosis standardized.

In contrast, one of the limitations of this study is the mismatch in the RMSEA values of M1, M2, and M3 in the ESEM analyses, suggesting that the model structure and accuracy of fit could be improved. This discrepancy could be explained by factors such as the complexity of the interrelationships between dimensions or the presence of variables not considered in the model (Shi et al., 2020). In addition, the sample size and the absence of validity tests on a sample other than those collected in this study should be considered as a

10.3389/fpsyg.2024.1386235

TABLE 3 Fit indexes for multiple-group CFA of ECSS-20.

	χ^2	df	χ^2/df	RMSEA	90% CI	CFI	TLI	SRMR	CMs	Δ CFI	ΔRMSEA
M5	610.507	328	1.861	0.074	[0.065, 0.830]	0.937	0.927	0.055		—	—
M6	628.231*	344	1.826	0.073	[0.064, 0.081]	0.936	0.930	0.057	M6-M5	−0.001	−0.001
M7	646.164*	360	1.794	0.071	[0.062, 0.080]	0.936	0.932	0.058	M7-M6	−0.001	−0.003

M5, configural invariance; M6, metric invariance; M7, scalar invariance; * = $p < 0.001$; χ^2, chi-square; df, degree of freedom; RMSEA, root mean square error of approximation; 90% CI, confidence interval; CFI, comparative fit index; TLI, Tucker–Lewis index; SRMR, standardized root mean square residual; CMs, comparison between models; ΔCFI, change in comparative adjustment index; ΔRMSEA, change in the error of the mean square of the approximation root.

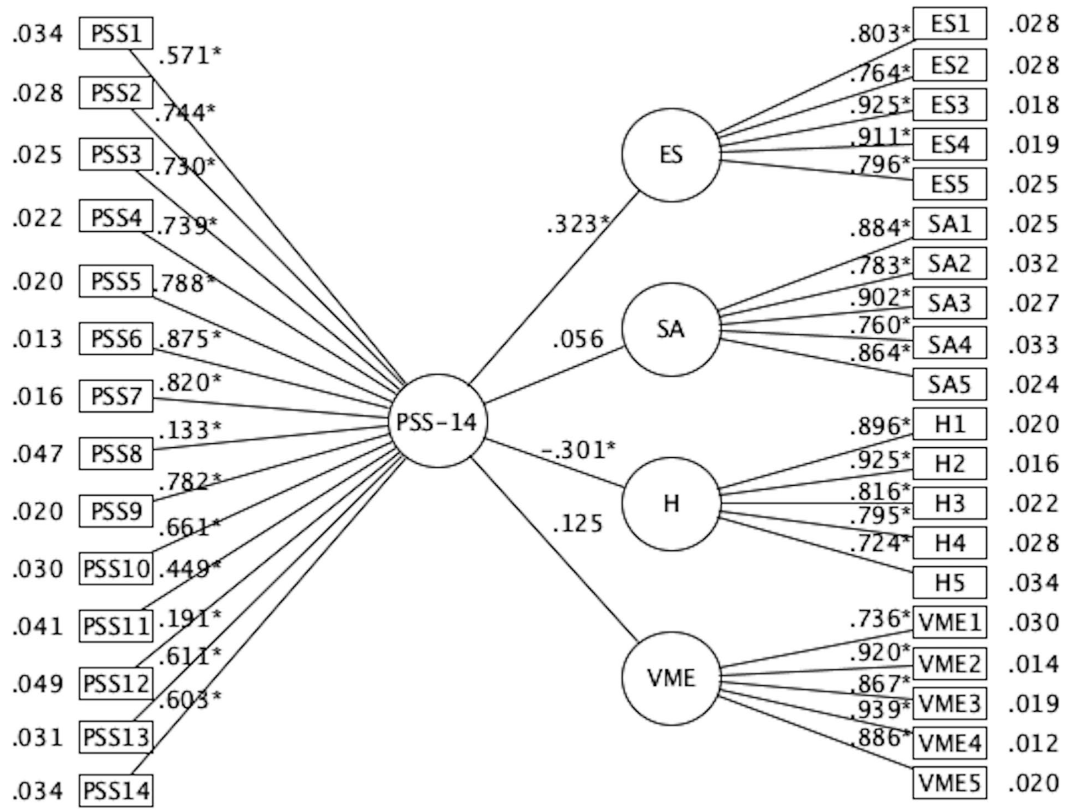

FIGURE 1
SET-ESEM ECSS-20 and PSS-14.

limitation of this study. Finally, a limitation of this research is that the survey was conducted during the medium health impact phase mandated by the Chilean government, characterized by the reduction of social interactions through measures such as social distancing or confinement, including a distance of one meter between two people and the use of a permit demonstrating current vaccination status for transit in the city (Ministerio de Salud, 2022).

This context of dynamic and sometimes irregular confinements, as evidenced in studies such as Patrono et al. (2024), demonstrates the heterogeneity of the confinement experience in the country, with differentiated impacts on the mental and physical health of the population (Duarte and Jiménez-Molina, 2022; Gutiérrez-Pérez et al., 2024). The research by Dagnino et al. (2020) underscores the significant psychological impact and the high demand for psychological support in Santiago, reflecting a critical need that could influence the structure of the ECSS-20 under different confinement intensities. Governmental policies, criticized for their improvised and unequal approach, particularly in terms of gender equity (Undurraga and López-Hornickel, 2023), and the special vulnerability of minority communities (Anandarajah et al., 2024), require careful consideration in interpreting the ECSS-20 data. Recently, Rodman et al. (2024), offered a longitudinal perspective on the deterioration of youth psychopathology due to reduced socialization, a factor that must be considered when assessing the validity of the ECSS-20 in future research. Additionally, studies on the impact of the social environment, such as that by Choi et al. (2024), highlight how specific neighborhood characteristics, such as socioeconomic deprivation and disorder, can increase the risk of dementia, mediated in part by subjective loneliness. This link underscores the importance of considering how urban environments and socialization dynamics influence mental health at all life stages (Ibanez et al., 2024; Migeot et al., 2024). The ECSS-20 ability to capture variations in the perception of environmental stress could be crucial not only for a better understanding of mental health disorders in youth but also for exploring longitudinal connections with cognitive risks in later life stages, influenced by social isolation and neighborhood conditions.

Finally, future lines of research could apply the ECSS-20 in contexts where stress is generated by confinement and/or population displacement, such as those generated by natural disasters (i.e., tsunamis, fires, landslides, extreme heat, hurricanes and tornadoes) (Sandoval-Díaz and Martínez-Labrín, 2021; Birkmann et al., 2022) or by sociopolitical situations of the countries of residence (i.e., political asylum, immigration) (Kwok and Ku, 2008; Kim et al., 2021). These events, marked by critical sociopolitical dynamics and needs for rapid adaptation, present a fertile ground to assess variations in subjective well-being, coping, and fields of spatial justice and habitability (Astudillo Pizarro and Sandoval Díaz, 2019; Sandoval-Díaz et al., 2021, 2022, 2024). Moreover, employability circumstances such as job loss or absence, or significant changes in work conditions, as well as the loss of daily social activities and hospitalizations, are critical areas where the ECSS-20 could reveal significant impacts on mental and physical health (Bahamondes-Rosado et al., 2023; Pérez-Villalobos et al., 2023). It is also advisable to explore how the ECSS-20 functions across different cultures and countries, as stress and its perception can vary considerably among different environments and populations (Tasnim et al., 2024; Tonon et al., 2024). The COVID-19 pandemic, as a natural experiment, has provided a unique context to better understand these phenomena (Gormley, 2024; Ruggeri et al., 2024). Therefore, adapting and validating the ECSS-20 in diverse cultural and environmental stress contexts could enrich our understanding of the interactions between the environment, stress, and mental health, thereby broadening the practical applications of the scale in designing targeted interventions and effective public health policies.

Conclusion

The instrument's multidimensional structure, internal consistency, gender invariance, and evidence associated with related measures support its validity and usefulness. The ECSS-20 is a valuable tool for investigating and further understanding the effects of confinement on the population's mental health. Future research could explore its applicability in different contexts and populations to strengthen understanding of its psychometric properties and utility in assessing confinement situations.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Comité ético científico Universidad de Tarapacá. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JS-P: Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. RF-U: Writing – original draft, Validation, Supervision, Software, Resources, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. GS-P: Writing – original draft, Validation, Supervision, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. JF: Writing – original draft, Investigation. KA-C: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Akbari, P., Yazdanfar, S. A., Hosseini, S. B., and Norouzian-Maleki, S. (2021). Housing and mental health during outbreak of COVID-19. *J. Build. Eng.* 43:102919. doi: 10.1016/j.jobe.2021.102919

American Educational Research Association. (2014). Standards for educational & psychological testing. Available at: https://www.aera.net/Publications/Books/Standards-for-Educational-Psychological-Testing-2014-Edition

Amerio, A., Lugo, A., Stival, C., Fanucchi, T., Gorini, G., Pacifici, R., et al. (2021). COVID-19 lockdown impact on mental health in a large representative sample of Italian adults. *J. Affect. Disord.* 292, 398–404. doi: 10.1016/j.jad.2021.05.117

Ammar, A., Trabelsi, K., Brach, M., Chtourou, H., Boukhris, O., Masmoudi, L., et al. (2020). Effects of home confinement on mental health and lifestyle behaviours during the COVID-19 outbreak: insight from the ECLB-COVID19 multicenter study. *Biol. Sport* 38, 9–21. doi: 10.5114/biolsport.2020.96857

Anandarajah, A. P., Yi, L., Anandarajah, A. A., Shelton, N., Feng, C., and Williams, E. M. (2024). Impact of COVID-19 pandemic on social isolation and loneliness among minority populations. *Am. J. Med. Sci.* 367, 21–27. doi: 10.1016/j.amjms.2023.09.020

Asparouhov, T., and Muthén, B. (2009). Exploratory structural equation modeling. *Struct. Equ. Model. Multidiscip. J.* 16, 397–438. doi: 10.1080/10705510903008204

Astudillo Pizarro, F., and Sandoval Díaz, J. (2019). Justicia espacial, desastres socionaturales y políticas del espacio: dinámicas sociopolíticas frente a los aluviones y proceso de recuperación en Copiapó, Chile. *Cuadernos de Geografía: Revista Colombiana de Geografía* 28, 303–321. doi: 10.15446/rcdg.v28n2.73520

Ato, M., López, J. J., and Benavente, A. (2013). Un sistema de clasificación de Ios diseños de investigación en psicología. *An. Psicol.* 29, 1038–1059. doi: 10.6018/analesps.29.3.178511

Badenes-Plá, N. (2022). "Changes in behaviour induced by COVID-19: Obedience to the introduced measures" in Pandemics: insurance and social protection. Springer actuarial (Cham: Springer), 143–161.

Bahamondes-Rosado, M. E., Cerdá-Suárez, L. M., Ortiz, D., de Zevallos, G. F., and Espinosa-Cristia, J. F. (2023). Technostress at work during the COVID-19 lockdown phase (2020–2021): a systematic review of the literature. *Front. Psychol.* 14:1173425. doi: 10.3389/fpsyg.2023.1173425

Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., et al. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* 4, 460–471. doi: 10.1038/s41562-020-0884-z

Bazzoli, A., Probst, T. M., and Lee, H. J. (2021). Economic stressors, COVID-19 attitudes, worry, and behaviors among U.S. working adults: a mixture analysis. *Int. J. Environ. Res. Public Health* 18:2338. doi: 10.3390/IJERPH18052338

Birkmann, J., Liwenga, E., Pandey, R., Boyd, E., Djalante, R., Gemenne, F., et al. (2022). "Poverty, livelihoods and sustainable development" in Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge, UK: Cambridge University Press), 1171–1274.

Brown, T. (2015). Confirmatory factor analysis for applied research: Guilford Publications Available at: https://books.google.cl/books?hl=es&lr=&id=tTL2BQAAQBAJ&oi=fnd&pg=PP1&dq=brown+2015+confirmatory+factor+analysis&ots=amOtuKWQaE&sig=oz--wLFHuCYpgQo73jpo-wtzIXE.

Bzdok, D., and Dunbar, R. I. (2022). Social isolation and the brain in the pandemic era. *Nature human behaviour* 6, 1333–1343. doi: 10.1038/s41562-022-01453-0

Cacioppo, J. T., and Cacioppo, S. (2014). Social relationships and health: the toxic effects of perceived social isolation. *Soc. Personal. Psychol. Compass* 8, 58–72. doi: 10.1111/spc3.12087

Cacioppo, J. T., Cacioppo, S., Capitanio, J. P., and Cole, S. W. (2015). The neuroendocrinology of social isolation. *Annu. Rev. Psychol.* 66, 733–767. doi: 10.1146/Annurev-Psych-010814-015240

Cacioppo, J. T., and Hawkley, L. C. (2009). Perceived social isolation and cognition. *Trends Cogn. Sci.* 13, 447–454. doi: 10.1016/j.tics.2009.06.005

Caqueo-Urízar, A., Urzúa, A., Aragón-Caqueo, D., Charles, C. H., El-Khatib, Z., Otu, A., et al. (2020). Mental health and the COVID-19 pandemic in Chile. *Psychol. Trauma Theory Res. Pract. Policy* 12, 521–523. doi: 10.1037/tra0000753

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14, 464–504. doi: 10.1080/10705510701301834

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Cho, E., and Kim, S. (2015). Cronbach's coefficient alpha: well known but poorly understood. *Organ. Res. Methods* 18, 207–230. doi: 10.1177/1094428114555994

Choi, E. Y., Cho, G., and Chang, V. W. (2024). Neighborhood social environment and dementia: the mediating role of social isolation. *J. Gerontol. B* 79:gbad199. doi: 10.1093/GERONB/GBAD199

Choi, Y. R., Kim, E. J., and Kim, M. K. (2020). A planning guide for temporary disaster shelters focusing on habitability. *Indoor Built Environ.* 29, 1412–1424. doi: 10.1177/1420326X19886051

Clair, R., Gordon, M., Kroon, M., and Reilly, C. (2021). The effects of social isolation on well-being and life satisfaction during pandemic. *Humanit. Soc. Sci. Commun.* 8:28. doi: 10.1057/s41599-021-00710-3

Cohen, J. (1988). Set correlation and contingency tables. *Appl. Psychol. Meas.* 12, 425–434. doi: 10.1177/014662168801200410

Cohen, S., Janicki-Deverts, D., and Miller, G. E. (2007). Psychological stress and disease. *JAMA* 298, 1685–1687. doi: 10.1001/JAMA.298.14.1685

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *J. Health Soc. Behav.* 24, 385–396. doi: 10.2307/2136404

Cohen, S., Murphy, M. L. M., and Prather, A. A. (2019). Ten surprising facts about stressful life events and disease risk. *Annu. Rev. Psychol.* 70, 577–597. doi: 10.1146/annurev-psych-010418-102857

Corral-Verdugo, V., Barrón, M., Cuen, A., and Tapia-Fonllem, C. (2011). Habitabilidad de la vivienda, estrés y violencia familiar. *PsyEcology* 2, 53–64. doi: 10.1174/217119711794394608

Costello, A. B., and Osborne, J. (2019). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* 10:7. doi: 10.7275/jyj1-4868

Cuadra-Martínez, D., Castro-Carrasco, P. J., Sandoval-Díaz, J., Pérez-Zapata, D., and Mora Dabancens, D. (2020). COVID-19 y comportamiento psicológico: revisión sistemática de los efectos psicológicos de las pandemias del siglo XXI. Revista médica de Chile, 148, 1139–1154. doi: 10.4067/S0034-98872020000801139

Dagnino, P., Anguita, V., Escobar, K., and Cifuentes, S. (2020). Psychological effects of social isolation due to quarantine in Chile: an exploratory study. *Front. Psychiatry* 11:591142. doi: 10.3389/fpsyt.2020.591142

Delhey, J., Hess, S., Boehnke, K., Deutsch, F., Eichhorn, J., Kühnen, U., et al. (2023). Life satisfaction during the COVID-19 pandemic: the role of human, economic, social, and psychological capital. *J. Happiness Stud.* 24, 2201–2222. doi: 10.1007/s10902-023-00676-w

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Meas. Eval. Couns. Dev.* 43, 121–149. doi: 10.1177/0748175610373459

Duarte, F., and Jiménez-Molina, Á. (2022). A longitudinal nationwide study of psychological distress during the COVID-19 pandemic in Chile. *Front. Psychiatry* 13:744204. doi: 10.3389/fpsyt.2022.744204

Dubey, A. D. (2020). The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. *JMIR Public Health Surveill.* 6:e19833. doi: 10.2196/19833

Ellen, C., Patricia, D. V., Miet, D. L., Peter, V., Patrick, C., Robby, D. P., et al. (2021). Meaningful activities during COVID-19 lockdown and association with mental health in Belgian adults. *BMC Public Health* 21:622. doi: 10.1186/s12889-021-10673-4

Eyre, H. A., Stirland, L. E., Jeste, D. V., Reynolds, C. F., Berk, M., Ibanez, A., et al. (2023). Life-course brain health as a determinant of late-life mental health: American Association for Geriatric Psychiatry Expert Panel recommendations. *Am. J. Geriatr. Psychiatry* 31, 1017–1031. doi: 10.1016/j.jagp.2023.09.013

Friedline, T., Chen, Z., and Morrow, S. P. (2021). Families' financial stress & well-being: the importance of the economy and economic environments. *J. Fam. Econ. Iss.* 42, 34–51. doi: 10.1007/s10834-020-09694-9

Gatersleben, B., and Griffin, I. (2017). "Environmental stress" in Handbook of environmental psychology and quality of life research. International handbooks of quality-of-life (Cham: Springer), 469–485.

Gilbar, O., Gelkopf, M., and Greene, T. (2022). Perceived stress during COVID-19: community resilience three years before the pandemic as a protective factor. *Int. J. Disaster Risk Reduct.* 82:103337. doi: 10.1016/j.ijdrr.2022.103337

Gloster, A. T., Lamnisos, D., Lubenko, J., Presti, G., Squatrito, V., Constantinou, M., et al. (2020). Impact of COVID-19 pandemic on mental health: an international study. *PLoS One* 15:e0244809. doi: 10.1371/journal.pone.0244809

Gormley, S. (2024). "Social interaction, social status, and mental health and how they were impacted by the COVID-19 pandemic as a natural experiment of social isolation" in Doctoral dissertation. Cambridge, UK.

Gruber, J., Prinstein, M. J., Clark, L. A., Rottenberg, J., Abramowitz, J. S., Albano, A. M., et al. (2021). Mental health and clinical psychological science in the time of COVID-19: challenges, opportunities, and a call to action. *Am. Psychol.* 76, 409–426. doi: 10.1037/amp0000707

Gudi-Mindermann, H., White, M., Roczen, J., Riedel, N., Dreger, S., and Bolte, G. (2023). Integrating the social environment with an equity perspective into the exposome paradigm: a new conceptual framework of the social exposome. *Environ. Res.* 233:116485. doi: 10.1016/j.envres.2023.116485

Gutiérrez-Pérez, I. A., Delgado-Floody, P., Molina-Gutiérrez, N., Campos-Jara, C., Parra-Rojas, I., Contreras-Osorio, F. H., et al. (2024). Changes in lifestyle and physical and mental health related to long-confinement due COVID-19: a study during the first and second pandemic waves in Mexico and Chile. *Psychol. Health Med.* 29, 174–190. doi: 10.1080/13548506.2023.2281295

González-Tovar, M., and Hernández-Rodríguez, S. (2020). Covid-19 y afrontamiento flexible en chilenos: Papel del estrés percibido y la ansiedad. Cuadernos de Neuropsicología. 14, 10–19. doi: 10.7714/CNPS/14.3.201

Hair, J., Black, W., Babin, B., and Anderson, R. (2019). Multivariate data analysis. *8th* Edn. Hampshire, UK: Cengage Learning, Anover, Cengage India.

Hussong, A. M., Midgette, A. J., Thomas, T. E., Coffman, J. L., and Cho, S. (2021). Coping and mental health in early adolescence during COVID-19. *Res. Child Adolesc. Psychopathol.* 49, 1113–1123. doi: 10.1007/s10802-021-00821-0

Ibanez, A., and Eyre, H. (2023). Brain capital, ecological development and sustainable environments. *BMJ Ment. Health* 26, 1–3. doi: 10.1136/bmjment-2023-300803

Ibanez, A., Melloni, L., Świeboda, P., Hynes, W., Ikiz, B., Ayadi, R., et al. (2024). Neuroecological links of the exposome and one health. *Neuron* 112, 1905–1910. doi: 10.1016/j.neuron.2024.04.016

Jakovljevic, M., Bjedov, S., Jaksic, N., and Jakovljevic, I. (2020). COVID-19 PANDEMIA and public and global mental health from the perspective of global health security. *Psychiatr. Danub.* 32, 6–14. doi: 10.24869/psyd.2020.6

Jamovi (2021). 2.2.5; The Jamovi Project. [Computer Software]. https://www.jamovi.org/

Kim, M., Kim, K., and Kim, E. (2021). Problems and implications of shelter planning focusing on habitability: a case study of a temporary disaster shelter after the Pohang earthquake in South Korea. *Int. J. Environ. Res. Public Health* 18:2868. doi: 10.3390/ijerph18062868

Kokkinos, C. M., Tsouloupas, C. N., and Voulgaridou, I. (2022). The effects of perceived psychological, educational, and financial impact of COVID-19 pandemic on Greek university students' satisfaction with life through mental health. *J. Affect. Disord.* 300, 289–295. doi: 10.1016/j.jad.2021.12.114

Kujawa, A., Green, H., Compas, B. E., Dickey, L., and Pegg, S. (2020). Exposure to COVID-19 pandemic stress: associations with depression and anxiety in emerging adults in the United States. *Depress. Anxiety* 37, 1280–1288. doi: 10.1002/da.23109

Kumar, S., and Shah, A. (2021). Revisiting food delivery apps during COVID-19 pandemic? Investigating the role of emotions. *J. Retail. Consum. Serv.* 62:102595. doi: 10.1016/j.jretconser.2021.102595

Kunzler, A. M., Röthke, N., Günthner, L., Stoffers-Winterling, J., Tüscher, O., Coenen, M., et al. (2021). Mental burden and its risk and protective factors during the early phase of the SARS-CoV-2 pandemic: systematic review and meta-analyses. *Glob. Health* 17:34. doi: 10.1186/s12992-021-00670-y

Kwok, J. Y., and Ku, H.-B. (2008). Making habitable space together with female Chinese immigrants to Hong Kong: an interdisciplinary participatory action research project. *Action Res.* 6, 261–283. doi: 10.1177/1476750308094131

Lal, A., Erondu, N. A., Heymann, D. L., Gitahi, G., and Yates, R. (2021). Fragmented health systems in COVID-19: rectifying the misalignment between global health security and universal health coverage. *Lancet* 397, 61–67. doi: 10.1016/S0140-6736(20)32228-5

Lazarus, R. S., and Cohen, J. B. (1977). "Environmental stress" in Behavior and environment (Boston, MA: Springer), 89–127.

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., et al. (2023). Measurement invariance in the social sciences: historical development, methodological challenges, state of the art, and future perspectives. *Soc. Sci. Res.* 110:102805. doi: 10.1016/j.ssresearch.2022.102805

Liu, X., Zhu, M., Zhang, R., Zhang, J., Zhang, C., Liu, P., et al. (2021). Public mental health problems during COVID-19 pandemic: a large-scale meta-analysis of the evidence. *Transl. Psychiatry* 11:384. doi: 10.1038/s41398-021-01501-9

Low, N., and Mounts, N. S. (2022). Economic stress, parenting, and adolescents' adjustment during the COVID-19 pandemic. *Fam. Relat.* 71, 90–107. doi: 10.1111/fare.12623

Maier, B. F., and Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* 368, 742–746. doi: 10.1126/science.abb4557

Manchia, M., Gathier, A. W., Yapici-Eser, H., Schmidt, M. V., de Quervain, D., van Amelsvoort, T., et al. (2022). The impact of the prolonged COVID-19 pandemic on stress resilience and mental health: a critical review across waves. *Eur. Neuropsychopharmacol.* 55, 22–83. doi: 10.1016/j.euroneuro.2021.10.864

Marroquín, B., Vine, V., and Morgan, R. (2020). Mental health during the COVID-19 pandemic: effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry Res.* 293:113419. doi: 10.1016/j.psychres.2020.113419

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., and Craven, R. G. (2020). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: optimal balance between goodness of fit and parsimony. *Multivar. Behav. Res.* 55, 102–119. doi: 10.1080/00273171.2019.1602503

Mayr, V., Nuβbaumer-Streit, B., and Gartlehner, G. (2020). Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review (review). *Gesundheitswesen* 82, 501–506. doi: 10.1055/a-1164-6611

McLaughlin, K. A., Rosen, M. L., Kasparek, S. W., and Rodman, A. M. (2022). Stress-related psychopathology during the COVID-19 pandemic. *Behav. Res. Ther.* 154:104121. doi: 10.1016/j.brat.2022.104121

McMichael, A. J. (2011). Epidemiology and the people's health: theory and context. *Int. J. Epidemiol.* 40, 1130–1132. doi: 10.1093/IJE/DYR075

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychol. Methods* 23, 412–433. doi: 10.1037/met0000144

Migeot, J., Panesso, C., Duran-Aniotz, C., Avíla-Rincón, C., Ochoa, C., Huepe, D., et al. (2024). Allostasis, health, and development in Latin America. *Neurosci. Biobehav. Rev.* 162:105697. doi: 10.1016/j.neubiorev.2024.105697

Ministerio de Salud. (2022). Plan Seguimos Cuidándonos—Ministerio de Salud—Gobierno de Chile. Available at: https://www.minsal.cl/plan-seguimos-cuidandonos-paso-a-paso/

Montero, I., and León, O. G. (2007). A guide for naming research studies in psychology. *Int. J. Clin. Health Psychol.* 7, 847–862. Available at: https://www.redalyc.org/articulo.oa?id=33770318

Morgado, A. M., Cruz, J., and Peixoto, M. M. (2023). Individual and community psychological experiences of the COVID-19 pandemic: the state of emergency in Portugal. *Curr. Psychol.* 42, 3213–3223. doi: 10.1007/s12144-021-01676-w

Morin, A. J. S., Katrin Arens, A., and Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Struct. Equ. Model. Multidiscip. J.* 23, 116–139. doi: 10.1080/10705511.2014.961800

Mukhtar, S. (2020). Psychological health during the coronavirus disease 2019 pandemic outbreak. *Int. J. Soc. Psychiatry* 66, 512–516. doi: 10.1177/0020764020925835

Muñiz, J., and Fonseca-Pedrero, E. (2019). Vista de Diez pasos para la construcción de un test. *Psichotema* 31, 7–16. doi: 10.7334/psicothema2018.291

Muthén, B., and Asparouhov, T. (2011). "Beyond multilevel regression modeling: multilevel analysis in a general latent variable framework" in Handbook of advanced multilevel analysis, 15–40. Available at: https://books.google.cl/books?hl=es&lr=&id=JlhySkysflUC&oi=fnd&pg=PA15&ots=vUDfQ1faA6&sig=AnJgXI5aLUllLCLv3VvQiNUocrk

Muthén, B., and Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychol. Methods* 17, 313–335. doi: 10.1037/a0026802

Muthén, L., Muthén, B., Asparouhov, T., and Nguyen, T.. (2017). Muthen: Mplus program (version 8.0) [computer software] (8.0). https://www.statmodel.com/

Nanath, K., Balasubramanian, S., Shukla, V., Islam, N., and Kaitheri, S. (2022). Developing a mental health index using a machine learning approach: assessing the impact of mobility and lockdown during the COVID-19 pandemic. *Technol. Forecast. Soc. Chang.* 178:121560. doi: 10.1016/j.techfore.2022.121560

Neville, F. G., Templeton, A., Smith, J. R., and Louis, W. R. (2021). Social norms, social identities and the COVID-19 pandemic: theory and recommendations. *Soc. Personal. Psychol. Compass* 15:e12596. doi: 10.1111/spc3.12596

Nielsen, L., Hinrichsen, C., Madsen, K. R., Nelausen, M. K., Meilstrup, C., Koyanagi, A., et al. (2021). Participation in social leisure activities may benefit mental health particularly among individuals that lack social connectedness at work or school. *Ment. Health Soc. Incl.* 25, 341–351. doi: 10.1108/MHSI-06-2021-0026

Ornell, F., Schuch, J. B., Sordi, A. O., and Kessler, F. H. P. (2020). "Pandemic fear" and COVID-19: mental health burden and strategies. *Braz. J. Psychiatry* 42, 232–235. doi: 10.1590/1516-4446-2020-0008

Otzen, T., and Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *Int. J. Morphol.* 35, 227–232. doi: 10.4067/S0717-95022017000100037

Patrono, A., Renzetti, S., Guerini, C., Macgowan, M., Moncada, S. M., Placidi, D., et al. (2024). Social isolation consequences: lessons from COVID-19 pandemic in a context of dynamic lock-down in Chile. *BMC Public Health* 24:599. doi: 10.1186/s12889-024-18064-1

Pérez-Villalobos, C., Ventura-Ventura, J., Spormann-Romeri, C., Paredes-Villarroel, X., Rojas-Pino, M., Jara-Reyes, C., et al. (2023). Well-being variations on students of health sciences related to their learning opportunities, resources, and daily activities in an online and on-crisis context: a survey-based study. *BMC Med. Educ.* 23:37. doi: 10.1186/s12909-023-04011-y

Pettini, G., and Mazzocco, K. (2022). What a pandemic can say about human-environment relationship. *Syst. Res. Behav. Sci.* 39, 159–162. doi: 10.1002/sres.2803

Prieto, G., and Delgado, A. R. (2010). Fiabilidad y validez. *Pap. Psicol.* 31, 67–74. Available at: https://www.researchgate.net/publication/41734046_Fiabilidad_y_Validez

Prime, H., Wade, M., May, S. S., Jenkins, J. M., and Browne, D. T. (2021). The COVID-19 family stressor scale: validation and measurement invariance in female and male caregivers. *Front. Psychiatry* 12:669106. doi: 10.3389/fpsyt.2021.669106

Quintana, G. (2022). The impact of the SARS-CoV-2 pandemic on reproduction, sexual function and behaviors: a review of the Main trends and findings. *Int. J. Sex. Health* 34, 351–365. doi: 10.1080/19317611.2022.2053921

Rieger, M. O., and Wang, M. (2020). Secret Erosion of the "lockdown"? Patterns in daily activities during the SARS-CoV-2 pandemics around the world. *Rev. Behav. Econ.* 7, 223–235. doi: 10.1561/105.00000124

Rivest-Beauregard, M., Fortin, J., Guo, C., Cipolletta, S., Sapkota, R. P., Lonergan, M., et al. (2022). Media use during the COVID-19 pandemic: cross-sectional study. *J. Med. Internet Res.* 24:e33011. doi: 10.2196/33011

Rodman, A. M., Rosen, M. L., Kasparek, S. W., Mayes, M., Lengua, L., Meltzoff, A. N., et al. (2024). Social experiences and youth psychopathology during the COVID-19 pandemic: a longitudinal study. *Dev. Psychopathol.* 36, 366–378. doi: 10.1017/S0954579422001250

Ruggeri, K., Stock, F., Haslam, S. A., Capraro, V., Boggio, P., Ellemers, N., et al. (2024). A synthesis of evidence for policy from behavioural science during COVID-19. *Nature* 625, 134–147. doi: 10.1038/s41586-023-06840-9

Salazar, A., Palomo-Osuna, J., de Sola, H., Moral-Munoz, J. A., Dueñas, M., and Failde, I. (2021). Psychological impact of the lockdown due to the COVID-19 pandemic in university workers: factors related to stress, anxiety, and depression. *Int. J. Environ. Res. Public Health* 18:4367. doi: 10.3390/ijerph18084367

Sandoval-Díaz, J., Cuadra Martínez, D., Pérez-Zapata, D., Sandoval-Díaz, J., Cuadra Martínez, D., and Pérez-Zapata, D. (2022). Del Afrontamiento Colectivo al Crecimiento Postraumático Comunitario: Análisis Mediacional del Empoderamiento ante un Desastre Climatológico. *Psykhe* 31, 1–15. doi: 10.7764/psykhe.2019.22345

Sandoval-Díaz, J., Díaz-Vargas, N., Flores-Jiménez, D., López-Salazar, C., and Bravo-Ferretti, C. (2024). Cambio climático y olas de calor sobre el bienestar subjetivo en jóvenes. *Rev. Latinoam. Cienc. Soc. Ninez Juv.* 22, 1–30. doi: 10.11600/rlcsnj.22.1.5926

Sandoval-Díaz, J., and Martínez-Labrín, S. (2021). Gestión comunitaria del riesgo de desastre: Una propuesta metodológica-reflexiva desde las metodologías participativas. *Rev. Estud. Latinoam. sobre Reducc. Riesgo Desastres* 5, 75–90. doi: 10.55467/reder.v5i2.73

Sandoval-Díaz, J., Ruiz-Rivera, N., Cuadra-Martínez, D., Sandoval-Díaz, J., Ruiz-Rivera, N., and Cuadra-Martínez, D. (2021). Experiencia y afrontamiento ante el riesgo aluvional: Un modelo mediacional múltiple. *Acta Colomb. Psicol.* 24, 130–143. doi: 10.14718/ACP.2021.24.2.12

Segerstrom, S. C., and O'Connor, D. B. (2012). Stress, health and illness: four challenges for the future. *Psychol. Health* 27, 128–140. doi: 10.1080/08870446.2012.659516

Sheek-Hussein, M., Abu-Zidan, F. M., and Stip, E. (2021). Disaster management of the psychological impact of the COVID-19 pandemic. *Int. J. Emerg. Med.* 14:19. doi: 10.1186/s12245-021-00342-z

Shi, D., Maydeu-Olivares, A., and Rosseel, Y. (2020). Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Struct. Equ. Model. Multidiscip. J.* 27, 1–15. doi: 10.1080/10705511.2019.1611434

Szkody, E., Stearns, M., Stanhope, L., and McKinney, C. (2021). Stress-buffering role of social support during COVID-19. *Fam. Process* 60, 1002–1015. doi: 10.1111/famp.12618

Takian, A., Haghighi, H., and Raoofi, A. (2021). "Challenges, opportunities, and future perspectives" in Environmental and Health Management of Novel Coronavirus Disease (COVID-19), *Londres, UK: Academic Press*. 443–477.

Tapia, D., Cruz, C., Gallardo, I., and Dasso, M. (2007). Adaptación de la Escala de Percepción Global de Estrés (EPGE) en estudiantes adultos de escasos recursos en Santiago, Chile. *Psiquiatr. Salud Ment.* 24, 109–119. Available at: https://schilesaludmental.cl/web/wp-content/uploads/2022/11/07-1y2-09.-Adaptacion-escala-percepcion-global-estres-estudiantes-adultos.pdf

Tasnim, A., Rahman, F. N., Rakhshanda, S., and Karim, A. (2024). Ageing and mental health in the context of social isolation in COVID-19 pandemic. Mymensingh Med. J., 33, 626–635. Available at: https://europepmc.org/article/med/

Tonon, A. C., de Abreu, A. C. O. V., da Silva, M. M., De S, T. P., Nishino, F., Versignassi, P., et al. (2024). Human social isolation and stress: a systematic review of different contexts and recommendations for future studies. *Trends Psychiatry Psychother.* 46:e20210452. doi: 10.47626/2237-6089-2021-0452

Tracy, E. L., Chin, B., Lehrer, H. M., Carroll, L. W., Buysse, D. J., and Hall, M. H. (2022). Coping strategies moderate the effect of perceived stress on sleep and health in older adults during the COVID-19 pandemic. *Stress. Health* 38, 708–721. doi: 10.1002/smi.3124

Undurraga, R., and López-Hornickel, N. (2023). The experience of women regarding Chilean government measures during the COVID-19 pandemic. *Bull. Lat. Am. Res.* 42, 564–577. doi: 10.1111/blar.13487

Valdés-Florido, M. J., López-Díaz, Á., Palermo-Zeballos, F. J., Garrido-Torres, N., Álvarez-Gil, P., Martínez-Molina, I., et al. (2022). Clinical characterization of brief psychotic disorders triggered by the COVID-19 pandemic: a multicenter observational study. *Eur. Arch. Psychiatry Clin. Neurosci.* 272, 5–15. doi: 10.1007/s00406-021-01256-w

Verdolini, N., Amoretti, S., Montejo, L., García-Rizo, C., Hogg, B., Mezquida, G., et al. (2021). Resilience and mental health during the COVID-19 pandemic. *J. Affect. Disord.* 283, 156–164. doi: 10.1016/j.jad.2021.01.055

von Keyserlingk, L., Yamaguchi-Pedroza, K., Arum, R., and Eccles, J. S. (2022). Stress of university students before and after campus closure in response to COVID-19. *J. Community Psychol.* 50, 285–301. doi: 10.1002/jcop.22561

Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., Ho, C. S., et al. (2020). 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* 17:1729. doi: 10.3390/IJERPH17051729

Wetherall, K., Cleare, S., McClelland, H., Melson, A. J., Niedzwiedz, C. L., O'Carroll, R. E., et al. (2022). Mental health and well-being during the second wave of COVID-19: longitudinal analyses of the UK COVID-19 mental health and wellbeing study (UK COVID-MH). *BJPsych Open* 8:e103. doi: 10.1192/bjo.2022.58

Whipple, S. S., and Evans, G. W. (2022). "The physical environment and social development" in The Wiley-Blackwell handbook of childhood social development, Wiley, New Jersey, USA. 171–188.

Wild, C. P. (2005). Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomarkers Prev.* 14, 1847–1850. doi: 10.1158/1055-9965.EPI-05-0456

Wilder-Smith, A., and Freedman, D. O. (2020). Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. *J. Travel Med.* 27:taaa020. doi: 10.1093/JTM/TAAA020

World Health Organization (WHO). (2023). WHO COVID-19 dashboard. Available at: https://covid19.who.int/?adgroupsurvey=%7Badgroupsurvey%7D&gclid=Cj0KCQiAw8OeBhCeARIsAGxWtUxR9wuVfKmK8awaDLSRrSm65bKiEJOcvWX34XpP2LX4eD_sR9ZfUqAaArNPEALw_wcB

Xiao, Y., Liu, H., and Hau, K. T. (2019). A comparison of CFA, ESEM, and BSEM in test structure analysis. *Struct. Equ. Model. Multidiscip. J.* 26, 665–677. doi: 10.1080/10705511.2018.1562928

Yang, Z., Luo, Y., Zhou, Q., Chen, F., Xu, Z., Ke, L., et al. (2022). COVID-19-related stressors and depression in Chinese adolescents: the effects of life history strategies and gender. *J. Affect. Disord.* 304, 122–127. doi: 10.1016/j.jad.2022.02.060

Yu, D., Zhang, Y., Meng, J., Wang, X., He, L., Jia, M., et al. (2023). Seeing the forest and the trees: holistic view of social distancing on the spread of COVID-19 in China. *Appl. Geogr.* 154:102941. doi: 10.1016/j.apgeog.2023.102941

Zulaica, L., and Oriolani, F. (2019). Quality of life and habitability conditions in Peri-urban areas of southern Mar del Plata, Argentina: a multimethod study. *Appl. Res. Qual. Life* 14, 659–683. doi: 10.1007/s11482-018-9618-z

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 19 July 2024
DOI 10.3389/fpsyg.2024.1392351

Check for updates

OPEN ACCESS

EDITED BY
Paola Gremigni,
University of Bologna, Italy

REVIEWED BY
Dan Wang,
University of Virginia, United States
Francisco Javier Cano-García,
Sevilla University, Spain

*CORRESPONDENCE
Yafei Tan
yafei.tan@ccnu.edu.cn

†These authors have contributed equally to this work

RECEIVED 27 February 2024
ACCEPTED 05 July 2024
PUBLISHED 19 July 2024

CITATION
Tan Y, An X, Cao M and Van den Bergh O (2024) Somatosensory Amplification Scale—Chinese version: psychometric properties and its mediating role in the relationship between alexithymia and somatization. *Front. Psychol.* 15:1392351. doi: 10.3389/fpsyg.2024.1392351

Somatosensory Amplification Scale—Chinese version: psychometric properties and its mediating role in the relationship between alexithymia and somatization

Yafei Tan[1,2,3*†], Xiaoran An[1†], Menglu Cao[4,5] and Omer Van den Bergh[6]

[1]School of Psychology, Central China Normal University, Wuhan, China, [2]Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education, Wuhan, China, [3]Key Laboratory of Human Development and Mental Health of Hubei Province, Wuhan, China, [4]Faculty of Psychology, Southwest University, Chongqing, China, [5]Center of Students' Mental Health, Sichuan Technology and Business University, Chengdu, China, [6]Health Psychology, University of Leuven, Leuven, Belgium

The Somatosensory Amplification Scale (SSAS) was designed to measure individual's tendency to experience visceral and somatic sensations as unusually intense, disturbing and alarming. In this study, we aimed to investigate the reliability and validity of the SSAS in the Chinese general population, as well as the mediating effect of somatosensory amplification in the relationship between alexithymia and somatization. A total of 386 healthy adults were enrolled in this study. Participants completed the Chinese versions of the Somatosensory Amplification Scale (SSAS-C), the somatization subscale of the Symptom Check List 90 (SCL-90 som), the Toronto Alexithymia Scale (TAS-20), and the Short form Health Anxiety Inventory (SHAI). One hundred and thirty-three participants were randomly selected to complete the SSAS-C again two weeks after the initial assessment. The reliability and validity of the SSAS-C were analyzed. Confirmatory factor analysis showed that the one-factor model achieved adequate model fits; one item was deleted due to low factor loading. The revised SSAS-C showed good internal consistency and test-retest reliability. The SSAS-C scores correlated positively with the scores of SCL-90 som, TAS-20 and the SHAI, showing good convergent validity. In addition, somatosensory amplification mediated the association between alexithymia and somatization. The Chinese version of SSAS has acceptable reliability and validity for the general population. In addition, alexithymia may increase somatization through higher somatosensory amplification.

KEYWORDS

Somatosensory Amplification Scale, psychometric properties, somatization, alexithymia, Chinese version

Introduction

The term "somatosensory amplification" refers to the tendency to experience normal unpleasant bodily sensations as abnormally intense, noxious, and disturbing (Barsky and Klerman, 1983). It involves three elements: (1) hypervigilance toward bodily sensations; (2) a tendency to focus on certain weak and infrequent somatic and visceral sensations; and (3) a tendency to appraise bodily sensations as more alarming and ominous than they actually are (Barsky et al., 1988). The concept of somatosensory amplification was first proposed as one of the risk factors to explain somatization, health anxiety, and the development of hypochondriasis (Barsky et al., 1988, 1990; Barsky, 1992). Somatization is a tendency to experience somatic distress and symptoms unexplained by organ pathology and to seek medical help for it (Lipowski, 1988; Wise and Mann, 1994; North, 2002). When the somatizing patients assume that they have medical problems, they may focus more on somatic sensations that can confirm their hypothesis (Pennebaker, 1982). Additionally, they may indeed experience more bodily sensations because of anxiety related to their concern on symptoms (Barsky and Wyshak, 1990). The symptoms of patients with somatization are far from minor or trivial, but may result in substantial distress and increased medical costs (Barsky et al., 2005).

To measure somatosensory amplification, the Somatosensory Amplification Scale (SSAS), a ten-item self-report scale, was developed (Barsky et al., 1990). Barsky et al. (1990) hypothesized that somatizing patients are overly disturbed by normal bodily sensations which are not necessarily pathological. Therefore, the scale assesses participants' sensitivity to mildly uncomfortable bodily experience that is non-pathogenic. It showed good test-retest reliability and internal consistency and close associations with hypochondriasis and somatization tendencies in both healthy and clinical populations (Barsky et al., 1990; Barsky and Wyshak, 1990; Speckens et al., 1996). For example, Takayanagi and Fujiu (2008) suggested an attention bias toward illness in individuals with high SSAS score. Furthermore, a number of studies have reported significant correlations between the SSAS score and self-reported bodily symptoms in psychiatric and psychosomatic outpatients (Spinhoven and van der Does, 1997; Muramatsu et al., 2002; Nakao et al., 2002; Nakao and Takeuchi, 2018; Barsky and Silbersweig, 2023). Based on SSAS, these studies demonstrated a role of somatosensory amplification in explaining why the relation between symptom report and real physiological dysfunction is highly variable (Barsky and Wyshak, 1990).

However, some researchers argued that SSAS was an index of negative reporting style and negative affectivity rather than a specific measurement of somatic sensitivity (Aronson et al., 2001, 2006). For example, it has been found that the SSAS score were highly related to depression and anxiety (Kosturek et al., 1998; Sayar et al., 2005). The SSAS score was also positively linked with alexithymia (Wise and Mann, 1994; Kosturek et al., 1998; Nakao et al., 2002), a personality trait characterized by difficulties in identifying feelings and distinguishing feelings from the physical sensations that accompany an emotional state (Taylor et al., 1991, 1997; Samur et al., 2013; Timoney and Holder, 2013). These arguments can be explained by a theory proposed by Van den Bergh et al. (2017), which suggests that unexplained symptom reports result from the mismatch between the prior prediction in the brain and the actual physiological input. Negative affect and attention bias to bodily signals may increase the gain of the prior. Then relatively weak bodily stimulation may be represented as stronger and more disturbing in individuals with high symptom reports.

The original English version of SSAS has been revised into several languages including French, Farsi, Spanish, Korean, Japanese, Kannada language, Italian, Turkish, and Hungarian (Jimenez Moron and Saiz Ruiz, 1996; Won and Shin, 1998; Nakao et al., 2001b; Duddu et al., 2003; De Berardis et al., 2007; Güleç and Sayar, 2007; Köteles et al., 2009; Bridou and Aguerre, 2013; Aghayousefi et al., 2015). However, there is no validated Chinese form of SSAS up to date. The present study therefore aimed to introduce the SSAS into China and examine the reliability and validity of the SSAS in the Chinese context. In the original version, the researchers proposed that somatosensory amplification is a unitary structure with some degree of stability over time. However, the French revision proposes a two-factor model, assuming that the SSAS items boil down to two factors: Exteroceptive sensitivity and Interoceptive sensitivity (Bridou and Aguerre, 2013). So, we also compared the fit of the two models to determine whether somatosensory amplification is a unitary structure representing somatosensory amplification.

In addition, a number of studies have demonstrated a close relationship between alexithymia and somatization in clinical and community samples (Cox et al., 1995; Deary et al., 1997; Bailey and Henry, 2007; Tam and Wong, 2013; Lanzara et al., 2020). For example, high-level somatization patients had significantly higher alexithymia scores than those with low levels of somatization (Joukamaa et al., 1995; Lanzara et al., 2020). Similarly, in a quantitative review of related articles, it was found that the total alexithymia score was positively correlated with different physical symptom reporting indicators, and individuals suffering from somatoform conditions were significantly more alexithymic than healthy controls, with effect sizes ranging from moderate to large (De Gucht and Heiser, 2003). Although both individuals with alexithymia and somatizing individuals have a tendency to frame psychological discomfort in physical terms (Lanzara et al., 2020), the underlying mechanism of how alexithymia may lead to somatization is still unclear. Psychopathology such as depression (Allen et al., 2011) and distress symptoms (Lanzara et al., 2020) have been proposed as mediators, but it could only partially explain the relationship between alexithymia and somatization. Some scholars have pointed out that alexithymia is a continuous and significant indicator that can predict somatization (Bach and Bach, 1995). An individual with normal emotional processing functions will be able to identify bodily sensations as part of the arousal accompanying emotions and qualify these sensations as secondary (Tyrer, 1973; Barsky and Klerman, 1983). However, alexithymic individuals fail to use emotions as signals in information processing, and instead may pay more attention to the normal physiological arousal accompanying emotions and regard them as primary phenomena and as alarming and ominous (Krystal, 1990). This somatosensory amplification may in turn lead to somatization. As stated, many studies have revealed a positive association between somatosensory amplification and alexithymia (Wise and Mann, 1994; Kosturek et al., 1998; Jones et al., 2004). There is also ample evidence showing the positive relationship between somatosensory amplification and somatization

(Güleç and Sayar, 2007; Bridou and Aguerre, 2013). Considering individual variations in both alexithymia and somatization have been associated with somatosensory amplification, the latter may act as a mediator in explaining the relationship between alexithymia and somatization.

Overall, targeting general population, the aims of this study were to (1) translate the SSAS into Chinese; (2) investigate the psychometric properties of the Chinese version of the SSAS; (3) further validate the SSAS-C as a measurement for somatosensory amplification through examining the mediating effect of somatosensory amplification in the relationship between alexithymia and somatization.

Materials and methods

Participants

Four hundred participants were recruited from the general population via online advertisements. The inclusion criteria were: age between 18 and 65 years, not pregnant, no IQ problem (i.e., able to read and understand a text), no psychotic and physical disorders. Somatosensory is immature in childhood and adolescence (Taylor et al., 2016). The elderly population may suffer sensory decline, which affect process of somatosensory (Heft and Robinson, 2017). To keep homogeneous in the present study, we only include participants aged between 18 and 65 years old. Participants were assessed using the Health History Inventory of the Body Perception Questionnaire, which asked participants to report the extent to which they are experiencing, have experienced, or have been diagnosed with psychotic and physical symptoms, such as migraine headaches, back problems, asthma, clinical depression and gastric distress or digestive problems (Porges, 1993). Participants were excluded if they reported "Usually" or "Always." Excluding incomplete responses, a final sample of 386 subjects (160 males; age = 31.59 ± 11.92 years) were used for analysis. Of these, 133 participants (55 males; age = 30.08 ± 11.84 years) were available for follow-up and completed the Chinese version of the SSAS again two weeks after the initial assessment. The study's protocol was approved by the Ethical Committee of Central China Normal University (IRB Number: CCNU-IRB-202204001). All participants provided informed consent and disclosed no conflicts of interests. The procedures in this study were performed in accordance with the Declaration of Helsinki.

Measures

Chinese version of the Somatosensory Amplification Scale (SSAS-C)

The SSAS is a 10-item self-report scale developed by Barsky et al. (1990). Respondents are asked to score each item on a scale from 1 (*Not at all true*) to 5 (*Extremely true*) according to how well each item describes themselves in general. A total amplification score is obtained by adding up the scores for each item. A higher total score indicates a greater tendency to amplify somatic sensations.

After obtaining the permission from the developer of the original English version of the SSAS, the Chinese version of the SSAS was obtained through to the standard back-translation procedures (Brislin, 1970). First, the scale was translated into Chinese by two independent bilingual authors. This preliminary version was then evaluated by three psychologists. They verified the clarity of the Chinese items and evaluated whether relevant cultural adaptations were needed, resulting in a combined forward-translation. It was then translated back into English by a professional translator. The original developer of the SSAS checked this back-translated version and confirmed that the meaning of the original version had not been changed or lost. Consequently, the final version of the SSAS in Chinese (SSAS-C) was established and used in the present study.

Somatization subscale of the Symptom Check List 90 (SCL-90 som)

This study used the Chinese somatization subscale of the Symptom Check List 90 (SCL-90 som) to measure somatization propensity (Derogatis et al., 1973; Wang, 1984). The somatization subscale consists of twelve items assessing the intensity of symptoms experienced during the previous week (e.g., nausea or upset stomach). Each item is rated on a scale ranging from 0 (*not at all distressed by the item*) to 4 (*extremely distressed by the item*). The total score was calculated across all 12 items, with a higher score reflecting a greater somatization tendency. The Chinese version of the SCL-90 som used in the current study showed a Cronbach's α of 0.89.

Toronto Alexithymia Scale (TAS-20)

The Chinese version of the 20-item Toronto Alexithymia Scale (TAS-20) developed by Yi et al. (2003) was used. The TAS-20 evaluates the individual's lack of emotional awareness (Bagby et al., 1994), consisting of three subscales: Difficulty Identifying Feelings (TAS-1), Difficulty Describing Feelings (TAS-2) and Externally-Oriented Thinking (TAS-3). The TAS-1 includes 7 items assessing individual's ability to identify their feelings and to distinguish the feelings from the bodily sensations accompanying emotions (e.g., "I have feelings that I can't quite identify"). The TAS-2 comprises 5 items measuring the ability to describe feelings (e.g., "It is difficult for me to reveal my innermost feelings, even to close friends"). The TAS-3 consists of 8 items evaluating individual's tendency to externally oriented thinking (e.g., "I prefer talking to people about their daily activities rather than their feelings"). Participants answer each item on a Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The Chinese version of the TAS-20 has shown good validity and reliability in several studies (e.g., Yang et al., 2008, 2009). The internal consistency was good in the current study, with a Cronbach's α of 0.82. The Cronbach's α of the three subscales were 0.86, 0.65 and 0.23, respectively. Consistent with previous studies, the Cronbach's α of TAS-3 was also lower than the other two subscales (Yi et al., 2003).

Short Health Anxiety Inventory (SHAI)

Health anxiety was measured using the Chinese version of the Short Health Anxiety Inventory (Zhang et al., 2013). The SHAI is a self-report inventory that contains 18 items assessing health-related anxiety not due to physical health status (Salkovskis et al., 2002).

Each item contains four different statements and participants are asked to select the one which best describes their feelings over the past six months. The score for each item is ranging from 0 to 3, resulting in a maximum total score of 54. The SHAI showed a Cronbach's α of 0.89 in the current sample.

Data analyses

The data were analyzed using the Statistical Product and Service Solutions version 27.0 (SPSS 27.0) and IBM SPSS Amos 26. We conducted a confirmatory factor analysis (CFA) to examine the factor structure of Chinese version of the SSAS. The following goodness-of-fit indices were used to evaluate model fit: relative chi-square (X^2/df), root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR) (Hu and Bentler, 1999). A range from 5.0 to 2.0 of X^2/df ratio is acceptable (Hooper et al., 2008). Values for TLI and CFI of greater than 0.90 are acceptable (Sharma et al., 2005; Hair et al., 2010). It was thought that a score for RMSEA of below 0.08 shows a good fit (MacCallum et al., 1996). A value for SRMR of below 0.05 indicate fair fit (Byrne, 1998; Diamantopoulos and Siguaw, 2000). Larger value of CFI and TLI, and smaller value of RMSEA and SRMR indicate that the model fits better. One-way ANOVA were used to make sociodemographic comparisons and item discrimination. To examine the reliability of SSAS-C, the internal consistency was assessed using item-total correlations and Cronbach's alpha coefficient. Moreover, we also reported McDonald's ω (McDonald, 1999) for the SSAS-C, which has been suggested as a better indicator of reliability (Teo and Fan, 2013). A ω value of greater than 0.70 reflects good reliability (Field, 2013). In addition, the test-retest consistency was assessed using intra-class correlations (ICC), with values below 0.50 signifying poor reliability, 0.50 to 0.75 representing moderate reliability, above 0.75 indicating good reliability (Koo and Li, 2016). The concurrent validity of the SSAS-C was examined by calculating Pearson correlations between the SSAS-C score and the scores of SCL-90 som, TAS-20 and SHAI, respectively. To examine the mediating role of somatosensory amplification in the relationship between alexithymia and somatization, we used PROCESS v4.0 prepared by Hayes in SPSS 27.0. To this end, alexithymia served as a predictor and the somatization served as the outcome. The bootstrapping method based on 5,000 bootstrap resamples was used to assess 95% bias-corrected confidence intervals of indirect effects for the mediation model. Finally, the alpha-level of significance for all statistical analyses was set to 0.05 in this study.

Results

Confirmatory factor analysis

There are two viewpoints about the structure of SSAS, respectively, the one-factor model in the original version and the two-factors model in the French revised version. The one-factor model considers that SSAS items are represented by one global dimension called somatosensory amplification. The two-factors model considers that SSAS items are represented by two factors: Exteroceptive sensitivity (items: 2, 4, 5, 7, 8, 9, 10) and Interoceptive sensitivity (items: 1, 3, 6). In this study, the results of CFA for the two models were compared and all of the fit indices for the two models are shown in Table 1. It shows that both models for the SSAS fit the sample data well. As shown in Figure 1, the factor loading of item-1 (*When someone else coughs, it makes me cough too*) was 0.33, which explains 10% latent variable variance (R^2 = 0.1). Chin (1998) describes R^2 value of 19% or below as weak. Therefore, the item-1 was deleted for further analysis. The results of CFA for the 9-item SSAS-C show improved model fit compared to the 10-item SSAS-C (Table 1). However, the estimated correlation between factor 1 and factor 2 of the two-factor model for both 10-items and 9-items SSAS-C was 1.03, indicating the two factors were not statistically distinguishable (Handscomb et al., 2016). More importantly, the one-factor model was more concise and consistent with the original English version (Barsky et al., 1990). Therefore, the one-factor model is more appropriate than the two-factor model.

TABLE 1 Goodness of fit measures for the two compared models of the Somatosensory Amplification Scale—Chinese version (*N* = 386).

		X^2/df	RMSEA	CFI	TLI	SRMR
10 items	1-factor model	3.323	0.078	0.931	0.912	0.0479
	2-factor model	3.408	0.079	0.931	0.908	0.0479
9 items	1-factor model	3.096	0.074	0.949	0.933	0.0438
	2-factor model	3.202	0.076	0.949	0.929	0.0438

RMSEA, root mean square error of approximation; CFI, comparative fit index; TLI, Tucker-Lewis index. SRMR, standardized root mean square residual.

Item discrimination

To analyze the item discrimination of the revised Chinese version of the SSAS, we quartered the group according to the total score of the scale and compared the four quartiles: 0–25%, 25–50%, 50–75%, 75–100% using one-way ANOVA and Least Significance Difference (LSD). The results showed that each item had a good discrimination among the four groups (Table 2).

Descriptive analysis

The results indicate significantly higher 9-item SSAS-C scores at retest (M = 32.60, SD = 4.31) compared to the scores at the first time (M = 29.45, SD = 5.59; t = −7.93, p < 0.001). Table 3 shows the demographic characteristics of the participants and the scores on the revised 9-item SSAS-C in different groups for both the test sample at the first time (Time 1) and the retest sample (Time 2). Women reported significantly higher scores on SSAS-C than men in this study [$t(384)$ = −3.35, p = 0.01]. In the subgroups of age and education, we found that individuals in the age range of 18 to 30 years had the highest SSAS-C scores and individuals who had a higher level of education reported higher SSAS-C scores.

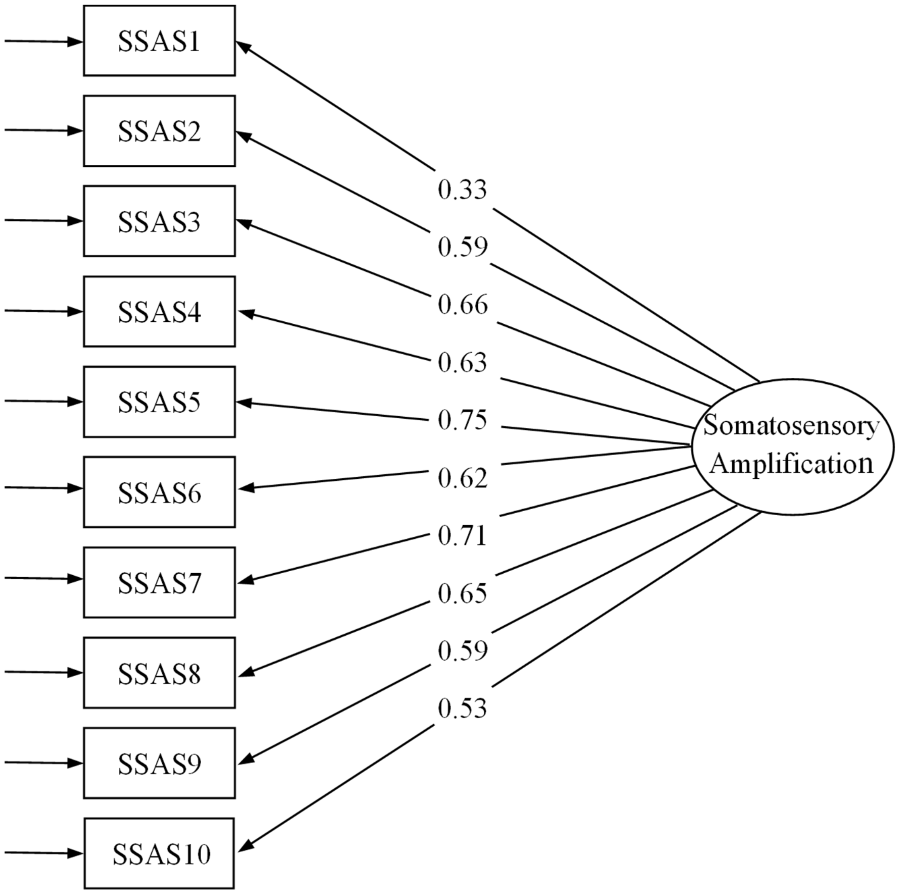

FIGURE 1
Factor loading of the Somatosensory Amplification Scale. SSAS, Somatosensory Amplification Scale.

TABLE 2 Item discrimination.

	Group M (SD)				*F*	*p*-value	LSD
	0~25% (n = 97)	25~50% (n = 96)	50~75% (n = 96)	75~100% (n = 97)			
Item 1	1.45 (0.90)	1.97 (0.95)	2.01 (0.92)	2.34 (1.10)	13.74	< 0.001	1 < 2 = 3 = 4
Item 2	2.24 (1.24)	3.35 (1.10)	3.77 (0.91)	4.28 (0.76)	69.84	< 0.001	1 < 2 < 3 < 4
Item 3	2.08 (0.99)	3.00 (0.90)	3.47 (0.69)	3.81 (0.68)	79.09	< 0.001	1 < 2 < 3 < 4
Item 4	1.87 (0.97)	2.79 (0.97)	3.33 (0.99)	3.97 (0.82)	86.54	< 0.001	1 < 2 < 3 < 4
Item 5	2.41 (1.25)	3.73 (0.78)	4.08 (0.47)	4.48 (0.50)	116.62	< 0.001	1 < 2 < 3 < 4
Item 6	2.03 (1.12)	2.95 (0.91)	3.40 (0.88)	3.87 (0.86)	65.44	< 0.001	1 < 2 < 3 < 4
Item 7	2.46 (1.24)	3.68 (0.73)	4.07 (0.63)	4.38 (0.56)	97.50	< 0.001	1 < 2 < 3 < 4
Item 8	1.82 (0.99)	3.02 (0.88)	3.43 (0.96)	3.87 (0.81)	89.42	< 0.001	1 < 2 < 3 < 4
Item 9	1.79 (0.95)	2.68 (1.00)	3.05 (0.92)	3.87 (0.88)	80.83	< 0.001	1 < 2 < 3 < 4
Item 10	1.81 (0.99)	2.66 (0.96)	2.91 (1.00)	3.68 (1.04)	57.09	< 0.001	1 < 2 = 3 < 4

Group 1: 0~25%, group 2: 25~50%, group 3: 50~75%, group 4: 75~100%; LSD, Least Significance Difference.

Reliability

The item–total correlations are shown in Table 4. It should be noted that the item-1 had a factor loading of less than 0.5 and the Cronbach's alpha of the entire scale increased after it is removed. This result suggested to delete item-1; all the following results are for the 9-item SSAS-C. The average inter-item correlation for 9 items was 0.40, with a range of 0.25–0.53. The Cronbach's alpha of the 9-item SSAS-C score was 0.86. The ω values for the 9-item SSAS-C was 0.85. These values indicate good reliability of the 9-item SSAS-C. The 2-week test-retest reliability indexed by ICC for the SSAS-C was 0.65, indicating moderate test-retest reliability.

TABLE 3 Characteristics of the participants and the Chinese version of Somatosensory Amplification Scale scores.

Characteristic	*N* (%)		SSAS-C score M (SD)		*p*-value	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Gender					< 0.001	= 0.049
Male	160 (41.45%)	55 (41.35%)	27.04 (7.89)	31.73 (4.75)		
Female	226 (58.55%)	78 (58.65%)	29.57 (6.47)	33.22 (3.88)		
Age (SD)					= 0.019	= 0.218
18–30 years	233 (60.36%)	92 (69.17%)	29.33 (6.73)	33.03 (4.06)		
30–50 years	126 (32.64%)	29 (21.80%)	27.10 (8.02)	31.52 (5.07)		
50–65 years	27 (7.00%)	12 (9.03%)	28.11 (6.07)	31.92 (3.94)		
Education					= 0.002	= 0.021
Middle school and below	46 (11.92%)	15 (11.28%)	25.19 (8.27)	29.73 (3.49)		
High school	59 (15.28%)	14 (10.53%)	28.10 (6.45)	32.57 (5.77)		
University and above	281 (72.80%)	104 (78.19%)	29.14 (7.02)	33.02 (4.07)		

SSAS-C, Chinese version of Somatosensory Amplification Scale.

TABLE 4 Somatosensory Amplification Scale internal consistency evaluation, the impact of each item on the scale and Cronbach's α values if item deleted.

SSAS item	Corrected item–total correlation	α if item deleted
1. When someone else coughs, it makes me cough too	0.45***	0.86
2. I can't stand smoke, smog, or pollutants in the air	0.65***	0.84
3. I am often aware of various things happening within my body	0.68***	0.84
4. When I bruise myself, it stays noticeable for a long time	0.69***	0.84
5. Sudden loud noises really bother me	0.76***	0.83
6. I can sometimes hear my pulse or my heartbeat throbbing in my ear	0.66***	0.84
7. I hate to be too hot or too cold	0.72***	0.83
8. I am quick to sense the hunger contractions in my stomach	0.69***	0.84
9. Even something minor, like an insect bite or a splinter, really bothers me	0.66***	0.84
10. I have a low tolerance for pain	0.61***	0.85

***$p < 0.001$. SSAS, Somatosensory Amplification Scale.

Concurrent validity

Correlations between the total scores of the SSAS-C and scores for other measurements (SCL-90 som, TAS-20, and SHAI) are shown in Table 5. As expected, the total scores of the SCL-90 som, the TAS-20 and the SHAI correlated positively with the total scores of the 9-item SSAS-C. However, the associations of SSAS-C scores with SCL-90 som ($r = 0.29$, $p < 0.001$) and SHAI scores ($r = 0.28$, $p < 0.001$) are moderate (Cohen, 1992).

Mediation analysis

As shown in Figure 2, after controlling for demographic covariates (i.e., gender, age, and education level), alexithymia significantly and positively predicted somatosensory amplification ($p < 0.001$), which in turn significantly and positively predicted somatization ($p < 0.001$). The positive direct association between alexithymia and somatization remained significant, showing that somatosensory amplification partially mediated the relationship between alexithymia and somatization. The indirect effect was 0.08 (95% CI: 0.02, 0.07). In addition, somatosensory amplification also partially mediated the associations of somatization with Difficulty Identifying Feelings (indirect effects = 0.06, 95% CI = 0.01–0.11) and Difficulty Describing Feelings (indirect effects = 0.08, 95% CI = 0.08–0.22), respectively.

Discussion

Somatosensory amplification is a major explanatory construct for various forms of somatoform disorders. The present study validated the Chinese version of the Somatosensory Amplification Scale (SSAS-C). The findings suggested that a slightly revised version of the SSAS (deleting its first item) had sound psychometric properties such as good reliability and convergent validity in a Chinese non-clinical sample. Furthermore, the scale showed a gender difference, in which women had higher levels of somatosensory amplification. Finally, the level of somatosensory amplification measured by SSAS-C showed a mediating effect on the relationship between alexithymia and somatization.

Barsky and Klerman (1983) defined somatosensory amplification as "a tendency to experience somatic and visceral sensation as unusually intense, noxious, and disturbing." It has been conceptualized as a unidimensional construct and assessed by the SSAS with a one-factor model (Barsky et al., 1990). Consistent with the results reported in most previous studies (Won and Shin, 1998; Nakao et al., 2001b; Duddu et al., 2003; De Berardis et al., 2007; Güleç and Sayar, 2007; Köteles et al., 2009;

TABLE 5 The correlations between the Somatosensory Amplification Scale and other measurements ($N = 386$).

	SCL-90 som	TAS-20	TAS-1	TAS-2	TAS-3	SHAI
SSAS	0.29***	0.37***	0.50***	0.28***	−0.01	0.28***
Mean (SD)	6.17 (5.73)	54.16 (9.63)	18.20 (5.32)	13.98 (3.31)	21.98 (3.18)	14.14 (7.54)

***$p < 0.001$. SSAS, Somatosensory Amplification Scale; SCL-90 som, Symptom Check List somatization subscale; TAS-20, Toronto Alexithymia Scale; TAS-1, Difficulty Identifying Feelings; TAS-2, Difficulty Describing Feelings; TAS-3, Externally-Oriented Thinking; SHAI, Short Health Anxiety Inventory.

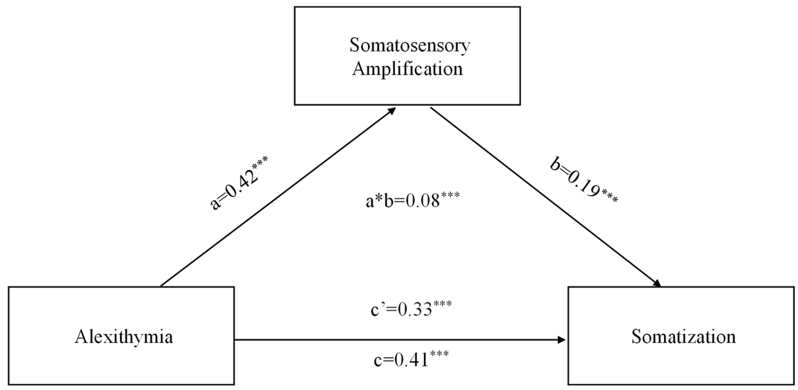

FIGURE 2
Results of Mediation analysis. a, initial variable; b, outcome; c, simple a to b effect when we do not control for the mediator; c', x to y effect when we do control for the mediator; a*b, mediation effect (indirect). The path weights in the graph were standardized and obtained after controlling for gender, age, and education level. ***$p < 0.001$.

Bridou and Aguerre, 2013), the current study showed that the one-factor model fitted the sample data well and was more appropriate than the two-factor model (Bridou and Aguerre, 2013). In addition, the confirmatory factor analysis revealed that the factor loading of item-1 "When someone else coughs, it makes me cough too" was lower than 0.5. Whereas all other items describe the perception of processes related to one's own body, item-1 deals with the impact of other people's bodily processes on one's own behavior. Other previous studies have reported similar results, showing the item-1 had relatively low factor loading (Barsky et al., 1990; Jasper et al., 2013). Another potential explanation is that this phenomenon may reflect empathy or social rules in the Chinese collectivist cultural context rather than individual somatic sensitivity. Therefore, we deleted this item, resulting in a 9-item Chinese version of SSAS. The 9-item SSAS-C showed satisfactory internal consistency (Cronbach's alpha). It would be interesting to investigate the cultural difference of somatosensory amplification using data from different culture groups in future studies. Test-retest consistency of the 9-item SSAS-C was assessed in 2 weeks indexed by ICC. The results indicate a moderate test-retest reliability, which replicates the finding of Güleç and Sayar (2007). We also found significantly higher SSAS-C scores at retest compared to the scores at the first time. This may be caused by the relatively short interval between two tests in the present study. It is possible that the test of SSAS at the first time induced them to pay more attention to somatosensory signals in anticipation of the second test, which may have increased the ratings for the items in the retest.

Moreover, in the current study, somatosensory amplification assessed by the SSAS-C showed the expected correlations with theoretically related measures. Specifically, as somatosensory amplification is a reliable indicator of somatization (Nakao and Barsky, 2007), a significantly positive correlation between the SSAS-C and the SCL-90-R somatization subscale was revealed. However, this correlation was moderate (Cohen, 1992), which replicates the finding of Bridou and Aguerre (2013). This might be due to the relatively low average and low variability in somatization scores of our sample constituting mainly university students, as well as to the difference in focus between the scales: SSAC is more about beliefs and attitudes about symptoms, whereas SCL-90 som focuses more on the symptoms themselves. Consistent with previous literature (Barsky et al., 1990), the results showed a significantly positive correlation between SSAS-C and SHAI. It suggests that a greater tendency to worry about health tends to be associated with a greater propensity to focus on, and thereby maximize, bodily sensations. On the other way around, cognitive misinterpretations of bodily sensations may cause health anxiety (Marcus et al., 2007). Additionally, we found that SSAS-C score was positively related to the level of individual alexithymia. More specifically, the SSAS was correlated with dimensions of Difficulty Identifying Feelings and Difficulty Describing Feelings, respectively, but not with dimension of Externally-Oriented Thinking. This corresponds to some previous studies indicating that the TAS-20 factor of "externally oriented thinking" was not closely correlated with the TAS-20 total score and other two factors (Fukunishi et al., 1997; Kojima et al., 2001; Leising et al., 2009), nor with the tendency to somatization (De Gucht and Heiser, 2003; Bailey and Henry, 2007; Mattila et al., 2008; Tam and Wong, 2013).

In terms of gender differences, the total score of SSAS was significantly higher in females than in males. Therefore, it would be appropriate to make gender comparisons when using the SSAS-C. This is generally in line with previous research on somatosensory amplification (Bridou and Aguerre, 2013; Nakao and Takeuchi, 2018). This finding further indicates that women are more sensitive to bodily stimuli than men, and therefore may experience more somatic distress and report a greater number of somatic symptoms (Nakao et al., 2001a).

Finally, somatic amplification was found to mediate the association between alexithymia and somatization. These findings were consistent with previous studies showing that the somatization is accompanied by alexithymia (Nakao et al., 2002; Jones et al., 2004; Güleç and Sayar, 2007). Alexithymia refers to difficulties in both identifying and expressing emotions (Nemiah and Sifneos, 1970; Taylor et al., 1997; Samur et al., 2013; Timoney and Holder, 2013). It has been defined as a possible personality risk factor for a variety of medical and psychiatric disorders involving emotional regulation problems (Taylor et al., 1991; Lumley et al., 1996; Porcelli et al., 2013; Luminet et al., 2018). Impaired emotional processing and emotion regulation capacity underlying alexithymia may lead to a focus on, amplification and misunderstanding of somatic sensations that accompany emotional arousal, thus resulting in somatization seeking medical help for medically unexplained symptoms (Pennebaker, 1982; Wise and Mann, 1994; Orrù et al., 2020). It is worthy to note that somatosensory amplification could only partially explain the relationship between alexithymia and somatization. This is consistent with previous studies showing that depression (Allen et al., 2011) and distress symptoms (Lanzara et al., 2020) mediated the alexithymia-somatization association.

Limitations and future directions

These findings of the present study must be interpreted considering a number of limitations. First, the sample in present study constituted mainly university students, which limits the application of the findings to the general population. In addition, somatic amplification was introduced as a clinical construct in relation to somatization and hypochondriasis, but whether the SSAS-C is applicable to Chinese clinical samples is still unknown. Future studies are recommended to validate the SSAS in more representative Chinese population as well as clinical sample. Second, the retrospective nature of this study design makes it difficult to draw clear conclusion on the direction of the relationship between alexithymia, somatization, and somatic amplification. Longitudinal study design is needed to disentangle the associations between somatization, alexithymia, and variations in somatic amplification. Finally, the data collected in this research is self-report. In the future, its validity as a measure of sensitivity to non-pathological bodily sensation should be established by comparing it with direct, objective measures of somatic and visceral perception.

Conclusion

The current study provides evidence for good psychometric properties of the SSAS-C in a Chinese non-clinical sample. In addition, somatosensory amplification measured by the SSAS-C was found to mediate the association between alexithymia and somatization. The SSAS is a potential tool for further investigation in the field of psychiatric disorders that are accompanied by functional somatic complaints. Moreover, the SSAS-C can be helpful in examining the cultural differences in somatosensory amplification and clarifying the relationships among somatosensory amplification, alexithymia and somatization in differing cultural contexts.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethical Committee of Central China Normal University (IRB Number: CCNU-IRB-202204001). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YT: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – original draft. XA: Conceptualization, Data curation, Methodology, Writing – original draft. MC: Conceptualization, Methodology, Supervision, Writing – review & editing. OV: Supervision, Writing – review & editing.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (grant number 32000747), and the Major Program of the National Social Science Foundation of China (grant number 22&ZD187).

Acknowledgments

We thank Dr. Arthur Barsky, MD for checking the back-translated version of the Somatosensory Amplification Scale and kind suggestions for improving it. We thank Dr. Yun Tang, Dr. Xiaojing Gu, and Dr. Xiaoqin Wang for evaluating the clarity and cultural adaptations of the Chinese items. We further thank Dr. Yun Tang for helping with the data analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aghayousefi, A., Oraki, M., Mohammadi, N., Farzad, V., and Daghaghzadeh, H. (2015). Reliability and validity of the farsi version of the somatosensory amplification scale. *Iranian J. Psychiatry Behav. Sci.* 9:e233. doi: 10.17795/ijpbs-233

Allen, L. B., Qian, L., Tsao, J. C. I., Hayes, L. P., and Zeltzer, L. K. (2011). Depression partially mediates the relationship between alexithymia and somatization in a sample of healthy children. *J. Health Psychol.* 16, 1177–1186. doi: 10.1177/1359105311402407

Aronson, K. R., Barrett, L. F., and Quigley, K. (2006). Emotional reactivity and the over report of somatic symptoms: Somatic sensitivity or negative reporting style? *J. Psychosom. Res.* 60, 521–530. doi: 10.1016/j.jpsychores.2005.09.001

Aronson, K. R., Barrett, L. F., and Quigley, K. S. (2001). Feeling your body or feeling badly: Evidence for the limited validity of the Somatosensory Amplification Scale as an index of somatic sensitivity. *J. Psychosom. Res.* 51, 387–394. doi: 10.1016/s0022-3999(01)00216-1

Bach, M., and Bach, D. (1995). Predictive value of alexithymia: A prospective study in somatizing patients. *Psychother. Psychosom.* 64, 43–48. doi: 10.1159/000288989

Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1

Bailey, P. E., and Henry, J. D. (2007). Alexithymia, somatization and negative affect in a community sample. *Psychiatry Res.* 150, 13–20. doi: 10.1016/j.psychres.2006.05.024

Barsky, A. J. (1992). Amplification, somatization, and the somatoform disorders. *Psychosomatics* 33, 28–34. doi: 10.1016/S0033-3182(92)72018-0

Barsky, A. J., and Klerman, G. L. (1983). Overview: Hypochondriasis, bodily complaints, and somatic styles. *Am. J. Psychiatry* 140, 273–283. doi: 10.1176/ajp.140.3.273

Barsky, A. J., and Silbersweig, D. A. (2023). The amplification of symptoms in the medically Ill. *J. Gen. Int. Med.* 38, 195–202. doi: 10.1007/s11606-022-07699-8

Barsky, A. J., and Wyshak, G. (1990). Hypochondriasis and somatosensory amplification. *Br. J. Psychiatry J. Ment. Sci.* 157, 404–409. doi: 10.1192/bjp.157.3.404

Barsky, A. J., Goodson, J. D., Lane, R. S., and Cleary, P. D. (1988). The amplification of somatic symptoms. *Psychosom. Med.* 50, 510–519. doi: 10.1097/00006842-198809000-00007

Barsky, A. J., Orav, E. J., and Bates, D. W. (2005). Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Arch. Gen. Psychiatry* 62, 903–910. doi: 10.1001/archpsyc.62.8.903

Barsky, A. J., Wyshak, G., and Klerman, G. L. (1990). The somatosensory amplification scale and its relationship to hypochondriasis. *J. Psychiatr. Res.* 24, 323–334. doi: 10.1016/0022-3956(90)90004-a

Bridou, M., and Aguerre, C. (2013). Validity of the French form of the somatosensory amplification scale in a non-clinical sample. *Health Psychol. Res.* 1:e11. doi: 10.4081/hpr.2013.e11

Brislin, R. W. (1970). Back-translation for cross-cultural research. *J. Cross-Cult. Psychol.* 1, 185–216. doi: 10.1177/135910457000100301

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Chin, W. W. (1998). "The partial least squares approach for structural equation modeling," in *Modern methods for business research*, ed. G. A. Marcoulides (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 295–336.

Cohen, J. (1992). *A power primer. Psychol. Bull.* 112, 155–159. doi: 10.1037//0033-2909.112.1.155

Cox, B. J., Swinson, R. P., Shulman, I. D., and Bourdeau, D. (1995). Alexithymia in panic disorder and social phobia. *Compr. Psychiatry* 36, 195–198. doi: 10.1016/0010-440X(95)90081-6

De Berardis, D., Campanella, D., Gambi, F., La Rovere, R., Sepede, G., Core, L., et al. (2007). Alexithymia, fear of bodily sensations, and somatosensory amplification in young outpatients with panic disorder. *Psychosomatics* 48, 239–246. doi: 10.1176/appi.psy.48.3.239

De Gucht, V., and Heiser, W. (2003). Alexithymia and somatisation: A quantitative review of the literature. *J. Psychosom. Res.* 54, 425–434. doi: 10.1016/S0022-3999(02)00467-1

Deary, I. J., Scott, S., and Wilson, J. A. (1997). Neuroticism, alexithymia and medically unexplained symptoms. *Pers. Individ. Diff.* 22, 551–564. doi: 10.1016/S0191-8869(96)00229-2

Derogatis, L. R., Lipman, R. S., and Covi, L. (1973). SCL-90: an outpatient psychiatric rating scale–preliminary report. *Psychopharmacol Bull.* 9, 13–28.

Diamantopoulos, A., and Siguaw, J. A. (2000). *Introducing LISREL*. Thousand Oaks, CA: SAGE Publications, Ltd, doi: 10.4135/9781849209359

Duddu, V., Chaturvedi, S. K., and Isaac, M. K. (2003). Amplification and attribution styles in somatoform and depressive disorders—A study from Bangalore, India. *Psychopathology* 36, 98–103. doi: 10.1159/000070365

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*. Thousand Oaks, CA: Sage Publications Ltd.

Fukunishi, I., Nakagawa, T., Nakamura, H., Kikuchi, M., and Takubo, M. (1997). Is alexithymia a culture-bound construct? Validity and reliability of the Japanese versions of the 20-item Toronto Alexithymia Scale and modified Beth Israel Hospital Psychosomatic Questionnaire. *Psychol. Rep.* 80(3 Pt 1), 787–799. doi: 10.2466/pr0.1997.80.3.787

Güleç, H., and Sayar, K. (2007). Reliability and validity of the Turkish form of the Somatosensory Amplification Scale. *Psychiatry Clin. Neurosci.* 61, 25–30. doi: 10.1111/j.1440-1819.2007.01606.x

Hair, J., Black, W., Babin, B., and Anderson, R. (2010). *Multivariate data analysis: A global perspective*, 7th Edn. New York, NY: Pearson.

Handscomb, L., Hall, D. A., Hoare, D. J., and Shorter, G. W. (2016). Confirmatory factor analysis of clinical outcomes in routine evaluation (CORE-OM) used as a measure of emotional distress in people with tinnitus. *Health Qual. Life Outcomes* 14:124. doi: 10.1186/s12955-016-0524-5

Heft, M. W., and Robinson, M. E. (2017). Somatosensory function in old age. *J. Oral Rehabil.* 44, 327–332. doi: 10.1111/joor.12488

Hooper, D., Coughlan, J., and Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic J. Bus. Res. Methods* 6, 53–60.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Jasper, F., Nater, U. M., Hiller, W., Ehlert, U., Fischer, S., and Witthöft, M. (2013). Rasch scalability of the somatosensory amplification scale: A mixture distribution approach. *J. Psychosom. Res.* 74, 469–478. doi: 10.1016/j.jpsychores.2013.02.006

Jimenez Moron, D., and Saiz Ruiz, J. (1996). Amplification: Somatosensory or cognitive? *Actas Luso-Espanolas Neurol. Psiquiatria Ciencias Afines* 24, 1–6.

Jones, M. P., Schettler, A., Olden, K., and Crowell, M. D. (2004). Alexithymia and somatosensory amplification in functional dyspepsia. *Psychosomatics* 45, 508–516. doi: 10.1176/appi.psy.45.6.508

Joukamaa, M., Sohlman, B., and Lehtinen, V. (1995). Alexithymia in primary health care patients. *J. Psychosom. Res.* 39, 833–842. doi: 10.1016/0022-3999(95)00020-1

Kojima, M., Frasure-Smith, N., and Lespérance, F. (2001). Alexithymia following myocardial infarction: Psychometric properties and correlates of the Toronto Alexithymia Scale. *J. Psychosom. Res.* 51, 487–495. doi: 10.1016/S0022-3999(01)00253-7

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropract. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Kosturek, A., Gregory, R. J., Sousou, A. J., and Trief, P. (1998). Alexithymia and somatic amplification in chronic pain. *Psychosomatics* 39, 399–404. doi: 10.1016/S0033-3182(98)71298-8

Köteles, F., Gémes, H., Papp, G., Túróczi, P., Pásztor, A., Freyler, A., et al. (2009). Validation of the Hungarian version of the Somatosensory Amplification Scale (SSAS). *Mentálhigiéné Pszichoszomatika* 10, 321–335. doi: 10.1556/Mental.10.2009.4.3

Krystal, H. (1990). An information processing view of object-relations. *Psychoanal. Inquiry* 10, 221–251. doi: 10.1080/07351699009533808

Lanzara, R., Conti, C., Camelio, M., Cannizzaro, P., Lalli, V., Bellomo, R. G., et al. (2020). Alexithymia and somatization in chronic pain patients: A sequential mediation model. *Front. Psychol.* 11:545881. doi: 10.3389/fpsyg.2020.545881

Leising, D., Grande, T., and Faber, R. (2009). The Toronto Alexithymia Scale (TAS-20): A measure of general psychological distress. *J. Res. Pers.* 43, 707–710. doi: 10.1016/j.jrp.2009.03.009

Lipowski, Z. J. (1988). Somatization: The concept and its clinical application. *Am. J. Psychiatry* 145, 1358–1368. doi: 10.1176/ajp.145.11.1358

Luminet, O., Taylor, G., and Bagby, R. (2018). *Alexithymia: Advances in research, theory, and clinical practice*. Cambridge: Cambridge University Press.

Lumley, M. A., Stettner, L., and Wehmer, F. (1996). How are alexithymia and physical illness linked? A review and critique of pathways. *J. Psychosom. Res.* 41, 505–518. doi: 10.1016/S0022-3999(96)00222-X

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychol. Methods* 1, 130–149. doi: 10.1037/1082-989X.1.2.130

Marcus, D. K., Gurley, J. R., Marchi, M. M., and Bauer, C. (2007). Cognitive and perceptual variables in hypochondriasis and health anxiety: A systematic review. *Clin. Psychol. Rev.* 27, 127–139. doi: 10.1016/j.cpr.2006.09.003

Mattila, A. K., Kronholm, E., Jula, A., Salminen, J. K., Koivisto, A.-M., Mielonen, R.-L., et al. (2008). Alexithymia and somatization in general population. *Psychosom. Med.* 70, 716–722. doi: 10.1097/PSY.0b013e31816ffc39

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, xi,485.

Muramatsu, K., Miyaoka, H., Muramatsu, Y., Fuse, K., Yoshimine, F., Kamijima, K., et al. (2002). The amplification of somatic symptoms in upper respiratory tract infections. *Gen. Hosp. Psychiatry* 24, 172–175. doi: 10.1016/S0163-8343(02)00177-9

Nakao, M., and Barsky, A. J. (2007). Clinical application of somatosensory amplification in psychosomatic medicine. *Biopsychosoc. Med.* 1:17. doi: 10.1186/1751-0759-1-17

Nakao, M., and Takeuchi, T. (2018). Alexithymia and somatosensory amplification link perceived psychosocial stress and somatic symptoms in outpatients with psychosomatic illness. *J. Clin. Med.* 7:112. doi: 10.3390/jcm7050112

Nakao, M., Barsky, A. J., Kumano, H., and Kuboki, T. (2002). Relationship between somatosensory amplification and alexithymia in a Japanese psychosomatic clinic. *Psychosomatics* 43, 55–60. doi: 10.1176/appi.psy.43.1.55

Nakao, M., Fricchione, G., Zuttermeister, P. C., Myers, P., Barsky, A. J., and Benson, H. (2001a). Effects of gender and marital status on somatic symptoms of patients attending a mind/body medicine clinic. *Behav. Med.* 26, 159–168. doi: 10.1080/08964280109595763

Nakao, M., Kumano, H., Kuboki, T., and Barsky, A. (2001b). Reliability and validity of the Japanese version of somatosensory amplification scale: Clinical application to psychosomatic illness. *Psychosom. Med.* 41, 539–547. doi: 10.15064/jjpm.41.7_539

Nemiah, J. C., and Sifneos, P. E. (1970). Psychosomatic illness: A problem in communication. *Psychother. Psychosom.* 18, 154–160. doi: 10.1159/000286074

North, C. S. (2002). Somatization in survivors of catastrophic trauma: A methodological review. *Environ. Health Perspect.* 110(Suppl. 4), 637–640. doi: 10.1289/ehp.02110s4637

Orrù, G., Gemignani, A., Ciacchini, R., Bazzichi, L., and Conversano, C. (2020). Machine learning increases diagnosticity in psychometric evaluation of alexithymia in fibromyalgia. *Front. Med.* 6:319. doi: 10.3389/fmed.2019.00319

Pennebaker, J. W. (1982). *The psychology of physical symptoms*. Berlin: Springer-Verlag Publishing, doi: 10.1007/978-1-4613-8196-9

Porcelli, P., Guidi, J., Sirri, L., Grandi, S., Grassi, L., Ottolini, F., et al. (2013). Alexithymia in the medically ill. Analysis of 1190 patients in gastroenterology, cardiology, oncology and dermatology. *Gen. Hosp. Psychiatry* 35, 521–527. doi: 10.1016/j.genhosppsych.2013.04.005

Porges, S. (1993). *Body perception questionnaire*. College Park, MA: University of Maryland.

Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C., and Clark, D. M. (2002). The Health Anxiety Inventory: Development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychol. Med.* 32, 843–853. doi: 10.1017/s0033291702005822

Samur, D., Tops, M., Schlinkert, C., Quirin, M., Cuijpers, P., and Koole, S. (2013). Four decades of research on alexithymia: Moving toward clinical applications. *Front. Psychol.* 4:861. doi: 10.3389/fpsyg.2013.00861

Sayar, K., Barsky, A. J., and Gulec, H. (2005). Does somatosensory amplification decrease with antidepressant treatment? *Psychosomatics* 46, 340–344. doi: 10.1176/appi.psy.46.4.340

Sharma, S., Mukherjee, S., Kumar, A., and Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *J. Bus. Res.* 58, 935–943. doi: 10.1016/j.jbusres.2003.10.007

Speckens, A. E., Spinhoven, P., Sloekers, P. P., Bolk, J. H., and van Hemert, A. M. (1996). A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *J. Psychosom. Res.* 40, 95–104. doi: 10.1016/0022-3999(95)00561-7

Spinhoven, P., and van der Does, A. J. (1997). Somatization and somatosensory amplification in psychiatric outpatients: An explorative study. *Compr. Psychiatry* 38, 93–97. doi: 10.1016/s0010-440x(97)90087-0

Takayanagi, N., and Fujiu, H. (2008). Attentional bias toward illness threat in individuals with hypochondriacal tendency and somatosensory amplification. *Japanese J. Health Psychol.* 21, 12–22. doi: 10.11560/jahp.21.2_12

Tam, B. K. H., and Wong, W. S. (2013). 2800 – Alexithymia and functional somatic symptoms among Chinese adolescents: Is somatosensory amplification a moderator or mediator? *Eur. Psychiatry* 28:1. doi: 10.1016/S0924-9338(13)77388-7

Taylor, G. J., Bagby, R. M., and Parker, J. D. (1991). The alexithymia construct. A potential paradigm for psychosomatic medicine. *Psychosomatics* 32, 153–164. doi: 10.1016/s0033-3182(91)72086-0

Taylor, G. J., Bagby, R. M., and Parker, J. D. A. (1997). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511526831

Taylor, S., McLean, B., Falkmer, T., Carey, L., Girdler, S., Elliott, C., et al. (2016). Does somatosensation change with age in children and adolescents? A systematic review. *Child Care Health Dev.* 42, 809–824. doi: 10.1111/cch.12375

Teo, T., and Fan, X. (2013). Coefficient alpha and beyond: Issues and alternatives for educational research. *Asia-Pacific Educ. Res.* 22, 209–213. doi: 10.1007/s40299-013-0075-z

Timoney, L. R., and Holder, M. D. (2013). "The importance of research on alexithymia," in *Emotional processing deficits and happiness, springer briefs in well-being and quality of life research*, (Dordrecht: Springer), 13–15. doi: 10.1007/978-94-007-7177-2_3

Tyrer, P. J. (1973). Relevance of bodily feelings in emotion. *Lancet* 1, 915–916. doi: 10.1016/s0140-6736(73)91365-2

Van den Bergh, O., Witthöft, M., Petersen, S., and Brown, R. J. (2017). Symptoms and the body: Taking the inferential leap. *Neurosci. Biobehav. Rev.* 74(Pt A), 185–203. doi: 10.1016/j.neubiorev.2017.01.015

Wang, Z. (1984). The symptom checklist 90 (SCL-90). *Shanghai Arch. Psychiatry* 2, 68–70.

Wise, T. N., and Mann, L. S. (1994). The relationship between somatosensory amplification, alexithymia, and neuroticism. *J. Psychosom. Res.* 38, 515–521. doi: 10.1016/0022-3999(94)90048-5

Won, H. T., and Shin, H. K. (1998). A study on the cognitive characteristics of somatization (I): The reliability and validity of the Korean versions of somatosensory amplification scale and symptom interpretation questionnaire. *Korean J. Clin. Psychol.* 17, 33–39.

Yang, H., Zhao, Z., Chu, Y., Liu, L., Liu, Q., and Zhou, S. (2008). Relationship between alexithymia and personality disorder of students in senior high schools and colleges. *Chinese J. Clin. Psychol.* 430, 431–439.

Yang, H., Zhou, S., Zu, L., Liu, L., and Liu, Q. (2009). Mediation of Alexithymia Between Adult Attachment and Depression. *Chinese J. Clin. Psychol.* 17, 161–163.

Yi, J., Yao, S., and Zhu, X. (2003). Reliability and validity analysis of Chinese version of TAS-20. *Chinese Ment. Health J.* 11, 763–767.

Zhang, Y., Ding, M., Xu, H., Li, P., Zhuang, D., Li, Y., et al. (2013). "Reliability and validity of the short version Health Anxiety Scale (SHAI) and its application in Chinese population," in *Compilation of papers of the 11th National Conference on Psychiatry of Chinese Medical Association and the 3rd Asian Conference on Neuropsychopharmacology*, 304.

frontiers | Frontiers in Psychology

TYPE Brief Research Report
PUBLISHED 09 September 2024
DOI 10.3389/fpsyg.2024.1440013

OPEN ACCESS

EDITED BY
Antonio de Padua Serafim,
University of São Paulo, Brazil

REVIEWED BY
Lisa Di Blas,
University of Trieste, Italy
Sara Dolan,
Baylor University, United States
Marco Tommasi,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

*CORRESPONDENCE
Timotej Glavač
✉ timotej.glavac@ff.uni-lj.si

RECEIVED 28 May 2024
ACCEPTED 15 August 2024
PUBLISHED 09 September 2024

CITATION
Glavač T, Poštuvan V, Schmeckenbecher J and Kapusta ND (2024) Slovenian validation of the Capacity to Love Inventory: associations with clinical measures and mindfulness.
Front. Psychol. 15:1440013.
doi: 10.3389/fpsyg.2024.1440013

Slovenian validation of the Capacity to Love Inventory: associations with clinical measures and mindfulness

Timotej Glavač[1]*, Vita Poštuvan[2,3], Jim Schmeckenbecher[4] and Nestor D. Kapusta[4]

[1]Department of Psychology, University of Ljubljana, Ljubljana, Slovenia, [2]Andrej Marušič Institute, Slovene Centre for Suicide Research, University of Primorska, Koper, Slovenia, [3]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia, [4]Department for Psychoanalysis and Psychotherapy, Medical University of Vienna, Vienna, Austria

Aim: The main purpose of the present study was to validate the Slovenian version of the 41- item Capacity to Love Inventory (CTL-I). Based on psychoanalytic theory, limitations to capacity to love are expected to be associated with personality dysfunction and disintegration as well as fundamental mental capacities such as self-reflection and self-awareness.

Method: To examine these assumptions, a sample of 552 Slovenian non-clinical individuals were recruited through academic networks. The construct validity of the CTL-I was assessed using a confirmatory factor analysis and convergent validity of the CTL-I and its subscales was established against IPO-16, PID-5 BF, MAAS.

Results: Our findings show that the Slovenian version of the CTL-I replicated the six-factor structure, exhibiting good model fit as well as satisfactory internal consistency of all subscales. In line with expectations, capacity to love was found to be inversely associated with dysfunctional personality traits and structural personality disturbances. Accordingly, higher dispositional mindfulness was coherently associated with all domains of CTL-I.

Conclusion: The results add to the growing evidence for the cross-cultural validity and sound psychometric properties of CTL-I, presented here in the Slovenian version. Our findings also point to the significance of dispositional mindfulness both in relation to capacity to love as well as mental health.

KEYWORDS

capacity to love inventory (CTL-I), personality, mindfulness, psychodynamic perspective, psychometrics properties

Introduction

Love stands as one of the most gratifying and fundamental of human experiences (Fletcher et al., 2015; Jankowiak and Fischer, 1992; Reis and Downey, 1999). Accordingly, research has consistently demonstrated strong associations between social relationships, particularly romantic ones, and various dimensions of well-being (Diener and Seligman, 2002; Frisch, 2005; Kansky, 2018; Mehl et al., 2010). For example, intimate relationships are linked to heightened subjective well-being (Campbell et al., 2005; Dush and Amato, 2005), while healthy romantic partnerships appear to foster improved physical health and reduced stress reactivity (Coan et al., 2013; Kiecolt-Glaser and Newton, 2001). Moreover, marriage has been shown to promote healthy behaviors among spouses over time (Homish and Leonard, 2008; Meyler et al., 2007), with research suggesting that marital satisfaction holds a stronger association with life satisfaction than domains such as job and health satisfaction (Dobrowolska et al., 2020; Heller et al., 2004). The relationship between life satisfaction and relationship satisfaction seems to be bidirectional in nature (Be et al., 2013).

The ability to form long-lasting committed intimate relationships with another person has been reflected by the concept of capacity to love (CTL) (Kernberg, 1974, 1977, 2011). Capacity to love consists of various elements, it involves the capacity to participate in, invest in, and maintain a dedicated romantic love relationship (Kernberg, 2011). More specifically, these features have been outlined based on clinical observations, and the concept is firmly rooted in psychoanalytic theory, particularly object relations theory (Kernberg, 1974, 1977, 2011; Garza-Guerrero, 2000). In order to be able to assess these constructs quantitatively, Kapusta et al. (2018) designed a 41-item questionnaire, measuring the proposed features of capacity to love. These include, Interest in the other, Basic trust Gratitude, Common ego ideal, Permanence of sexual passion, and Loss and mourning (Kapusta et al., 2018; Kernberg, 2011). Additionally, Falling in love, Forgiveness and Mature dependency, are dimensions which were also proposed in the original theoretical conceptualizations (Kernberg, 2011) but were excluded due to the unsatisfactory psychometric properties of the scales.

Interest in the other involves a deep, ongoing curiosity and interest in the emotional experience, personal history, and aspirations of the partner. This is thought to enrich one's own life and fosters mutual exploration and deeper love, contrasting sharply with the indifference often seen in narcissistic individuals. **Basic trust** in a partner's empathy and goodwill allows for openness about personal vulnerabilities and needs. This trust enables both partners to reveal themselves, fostering mutual growth and deepening their relationship. **Humility and gratitude** in mature love involve acknowledging dependency and embracing future uncertainties, as well as experiencing a genuine gratitude for the love received and responded to. **Common ego ideal** reflects a couple's commitment to their love as a lifelong project, influencing daily life and nurturing a deep interest in each other's personalities and experiences. **Permanence of sexual passion** refers to a passion, continuing throughout the relationship without diminishing after the initial phases, and involves, integrating tenderness with eroticism, supporting a deep connection between partners. Lastly, normal mourning (**Loss and mourning**) strengthens the ability to love and should not be dominated by excessive guilt or self-devaluation and can lead to positive growth after the end of a relationship. Previous validation studies have reported good psychometric properties of the proposed CTL-I model, including adequate construct validity, convergent validity and internal consistency (Kapusta et al., 2018; Margherita et al., 2018).

Derived from clinical psychoanalytic observations (Kernberg, 2011), research initially focused on examining challenges in the capacity to love among individuals with developmental deficits and psychopathology. According to the proposed psychoanalytic framework, narcissistic individuals are most often considered being at risk for experiencing deficiencies in the capacity to love. However, according to Kernberg's theoretical conceptualizations, the capacity to love is associated with various features of a healthy personality structure, and difficulties in the capacity to love may therefore be expected in various personality disorders as well as varying levels of personality organization (Kernberg, 1974). Indeed, empirical evidence consistently demonstrates that relationship distress serves as a reliable indicator of mental health problems (Whisman et al., 2000).

Difficulties in intimate relationships are one of the most common reasons individuals with acute emotional distress seek treatment (Swindle et al., 2000). Problems in intimate relationships are prevalent among individuals seeking mental health services who otherwise do not report relationship distress as their main cause for treatment (Lin et al., 1996), a finding that has been observed cross-culturally (Foran et al., 2013; Miller et al., 2013). A likely reason for this finding is that interpersonal stress has the potential to produce more intense subjective distress than non-interpersonal stress (Frans et al., 2005). Among mental health diagnoses, personality disorders (PD) may have an especially adverse effect on aspects of romantic relationships and intimate relationship satisfaction (Whisman et al., 2007; Whisman and Schonbrun, 2009). For example, in a study examining the impact of all 10 DSM-IV/DSM-5 Section II Personality disorders (American Psychiatric Association, 2013) on relationship satisfaction. South et al. (2008) found that individuals with PDs were more likely to engage in physical violence and verbal aggression as reported by their partners. PDs have also been shown to be related to higher frequencies of self-reported distress and lower relationship quality (Disney et al., 2012). PDs have also been found to decrease the probability of marriage among histrionic, avoidant and dependent individuals and increase the chances of divorce in paranoid, schizoid, antisocial, histrionic, avoidant, dependent and obsessive-compulsive individuals (Whisman et al., 2007). Similarly, Stroud et al. (2010) collected data on DSM-IV symptoms and found that self-reported paranoid, antisocial, borderline, dependent, and obsessive-compulsive symptoms significantly predicted marital dissatisfaction. A more recent study employing the Personality Inventory for DSM-5 (PID-5; Krueger et al., 2012) in a sample of 52 heterosexual couples found that the strongest effects on relationship dissatisfaction were found in the domains of detachment, negative affect, and disinhibition (DeCuyper et al., 2018). Furthermore, a study employing the DSM-5 found that nearly all the included personality constructs were negatively correlated with both participant-reported and spouse-reported marital satisfaction at each time point in the study (Humbad et al., 2010; South et al., 2008; Stroud et al., 2010).

Lastly, mindfulness, often described as the deliberate practice of focusing on the present moment's experiences—such as physical sensations, perceptions, emotions, thoughts, and imagery—in a nonjudgmental manner thereby fostering a consistent and nonreactive awareness of these experiences (Grossman et al., 2004; Kabat-Zinn, 1994; Kostanski and Hassed, 2008) has been found to be associated with various aspects of romantic relationships, such as lower relationship conflict, higher sexual satisfaction and increased emotional regulation (see Kozlowski, 2013 for a review). In relation to mental health and personality disorders, mindfulness meditation has been found to ameliorate the symptoms of various personality disorders such as antisocial, avoidant, borderline paranoid, and obsessive-compulsive personality disorders (Shorey et al., 2016; Sng and Janca, 2016; Velotti et al., 2019; Wupperman et al., 2009) as well as symptomatology related to some disorders such as psychosis

(Chadwick, 2014; Khoury et al., 2013). Indeed, dispositional mindfulness characterized as a disposition towards a higher frequency of mindful states over time (Brown and Ryan, 2003) has been shown to be inversely related to negative mental health outcomes (Tomlinson et al., 2018) and positively related to psychological well-being (Bränström et al., 2011; Hanley et al., 2015; Soysa et al., 2021).

The main purpose of our study was to expand the process of validation of the CTL-I on a Slovenian translation, as well as to further establish its psychometric properties. The concept of capacity to love shares the same theoretical and clinical basis with Kernberg's theory of personality organization (Kernberg, 1985, 2009), which convincingly explains how narcissistic and borderline individuals suffer from major difficulties in relationship functioning. We therefore employed both the Personality Inventory for DSM-5 (PID-5-BF; Krueger et al., 2012) and the Inventory of Personality Organization (IPO-16, Zimmermann et al., 2013) to measure convergent validity with the CTL-I. In contrast, the Mindful Attention Awareness Scale (MAAS, Brown and Ryan, 2003) was applied to examine the relation of CTL-I with a theoretically different, but potentially related concept of a psychological trait capacity covering one's tendency to attend to present-moment experiences in everyday activities, allowing for self-regulation of attention and the none-judgmental acceptance of one's immediate experiences (Bishop et al., 2004; Kabat-Zinn, 1994). We hypothesized that mindfulness might be beneficial to intimate relationships by enhancing one's capacity to love, given that mounting evidence has shown that mindfulness and relationship satisfaction appear to be strongly related (Kozlowski, 2013; McGill et al., 2016).

Method

Participants

Our research sample consisted of 552 Slovenian non-clinical individuals (112 men and 434 women, 6 other). The individuals who identified as "other" were excluded from the gender specific analyses. The mean age of the sample was 26.03 years (SD = 9.0). Participants were required to be at least 18 years old to participate in the study. The study protocol was approved (038–19-75/2020/4/FFUM) by the ethics committee of the Faculty of Arts, University of Maribor on 31.7.2020 and was conducted under the code of the Helsinki declaration and its subsequent amendments. Our sample was recruited through e-mail invitations sent to several faculties of all three major Slovenian public universities as well as through advertisements on social media. All the psychological measures were fitted into an online survey which the participants could access through a provided hyperlink. The first page of the survey provided a description of the study along with a consent form explaining that starting the survey (clicking a button) was equivalent to providing consent. The relationship status of our sample was as follows: 8.3% married, 45.3% in a relationship, 0.2% widowed, 2.2% divorced, 31.7% single and 34.2% living together with their partner (see Table 1).

TABLE 1 Characteristics of the sample.

Variables	Percentages (Mean ± SD)
Sex	
Men	112
Women	434
Other	6
Age	26.03
Relationship status	
Single	31.7%
Married	8.3%
Non-marital partnership	45.3%
Widowed	0.2%
Divorced	2.2%
Cohabitation	34.2%

Measures

Capacity to love inventory (CTL-I)

The CTL-I is a 41-item measure of six dimensions of the capacity to love. Items are rated on a 4-point scale. The measure consists of six subscales: Interest in the life project of the other (INT) (e.g., "*It is important to me to know the life plan of my partner*,") Basic trust (BRT) (e.g., "*My weaknesses, inner conflicts and problems are open to the other*,") Gratitude (e.g., "*I feel gratitude for the existence of my partner*") (GRT), Common ego ideal (e.g., "*We always try to work on our relationship*" (CEI)), Permanence of sexual passion (e.g., "*Sexual boredom arises in long-term relationships*" reversed) (PSP) and Loss and mourning (e.g., "*I am often unwilling to accept the end of my relationships*") (LOM). In the original version, the internal reliability of the total scale was 0.90, while the reliability scores for the specific subscales were as follows: INT—0.73, BTR—0.86, GRT—0.81, CEI—0.81, PSP—0.83 and LOM—0.75. In the present study internal reliabilities ranged from 0.67 (INT) to 0.85 (GRT). The inter-scale correlations between INT, BTR, GRT and CEI ranged from 0.57 to 0.79, while inter-scale correlations between PSP and LOM and the other scales were lower and ranged from −0.01 (between LOM and INT) and 0.31 (GRT and PSP) (shown in Table 2). A similar pattern has also been reported in previous CTL-I validation studies (Kapusta et al., 2018; Margherita et al., 2018).

Personality inventory for DSM-5 brief form

The PID-5 BF (PID-5-BF; Krueger et al., 2012) is a short 25-items version of the Personality Inventory for DSM-5 with originally 220 items rated on a 4-point Likert scale ranging from 0 (very false or often false) to 3 (very true or often true). The employed short version has shown comparable domain scores to the original version (Bach et al., 2016). The PID-5-BF assesses five key traits of dysfunctional personality proposed in Section III of the DSM-5 (American Psychiatric Association, 2013) consisting of Antagonism ("It's no big deal if I hurt other peoples' feelings"), Disinhibition ("People would describe me as reckless."), Negative affect ("I worry about almost everything."), Detachment ("I often feel like nothing I do really matters.") and Psychoticism ("My thoughts often do not make sense to others"). The psychometric properties of the PID-5 BF have been well established in several recent studies (Anderson et al., 2016; Debast et al., 2017; Hopwood et al., 2013). Internal consistencies on the present sample ranged from 0.61 (Disinhibition) to 0.79 (Psychoticism), while the internal consistency of the total scale was 0.86.

10.3389/fpsyg.2024.1440013

TABLE 2 Correlations between CTL-I subscales.

	INT	BTR	GRT	CEI	PSP	LOM
INT	-	0.57**	0.67**	0.61**	0.23**	−0.01
BTR		–	0.74**	0.68**	0.29**	0.20**
GRT			–	0.79**	0.31**	0.05
CEI				–	0.29**	0.04
PSP					–	0.07
LOM						–
CTL-I total	0.70**	0.80**	0.81**	0.77**	0.64**	0.37**
alpha	0.67	0.79	0.85	0.81	0.80	0.77

**$p \leq 0.01$, $N = 552$. INT, Interest; BTR, Basic trust; GRT, Gratitude; CEI, Common ego ideal; PSP, Permanence of sexual passion; LOM, Loss and mourning.

Inventory of personality organization (IPO-16)

The IPO-16 is a brief version of the IPO (Lenzenweger et al., 2001) measuring the severity of structural impairments in identity ("*I feel that my tastes and opinions are not really my own, but have been borrowed from other people*"), defense ("*It is hard for me to trust people because they so often turn against me or betray me*"), and reality testing ("*I act in ways that appear to others as unpredictable and erratic*") based on Kernberg's model of personality (Kernberg and Caligor, 2005). Items are rated on a 5-point Likert scale ranging from 1 = does never apply to 5 = always applies. A previous study of the German version conducted on clinical samples ($N = 1{,}300$) for the total scale showed good internal consistency at 0.85 (Zimmermann et al., 2013). The internal consistency of the total scale in our sample was 0.87.

The mindful attention awareness scale (MAAS)

The MAAS is a 15-item trait measure of one's tendency to attend to present-moment experiences in everyday activities. The items are self-rated on a scale from one ("almost always") to six ("almost never") and assess aspects of mindful experiences such as" *I could be experiencing some emotion and not be conscious of it until sometime later.*" and "*I do jobs or tasks automatically, without being aware of what I'm doing.*" The answers are scored by calculating mean performance across the 15 items. Higher scores reflect greater levels of dispositional mindfulness while lower levels reflect lower levels. The original MAAS validation study showed good internal reliability with an alpha of 0.82 (Brown and Ryan, 2003). The internal reliability of the MAAS on our sample was 0.86.

Data analysis procedure

The confirmatory factor analysis (CFA) on the Slovenian version of CTL-I was performed using the Maximum Likelihood (ML) as appropriate estimator. Descriptive statistical analyses were performed with the Statistical Package for the Social Sciences (SPSS) 27.0 (IBM Corp, 2020). Confirmatory factor analysis of the CTL-I was carried out in R 4.2.2 (R Core Team, 2022), using lavaan package (Rosseel, 2012). We tested the theory driven model developed by Kapusta et al. (2018) by means of a CFA: a six-factor model with 41 items and scales being allowed to correlate with each other, according to previous examinations (Kapusta et al., 2018; Margherita et al., 2018). Model fit was assessed by means of the fit indexes: (1) the chi-squared ($\chi 2$) statistic and its degrees of freedom; (2) the Standardized Root Mean Square Residuals (SRMR); (3) the Root Mean Square Error of Approximation (RMSEA) and its 90% confidence interval (90% CI). In line with Schermelleh-Engel et al.'s (2003) proposed criteria, the model fits the data when $\chi 2$/df equal or < 2, RMSEA equal or < 0.05 (90% CI: the lower boundary of the CI should contain zero for exact and be <0.05 for close fit) while Browne and Cudeck (1993) argued that values ranging from 0.05 to 0.08 are indicative of a good adequacy of the model. Additionally, we ran Monte Carlo simulations involving 50 simulated replications to determine if the model had adequate power levels. The model demonstrated adequate fit ($\chi^2 = 2544.183$, RMSEA = 0.065) and high statistical power for all parameter estimates. Parameter estimates were stable and reliable, with minimal bias and consistent coverage of true values. The simulations showed that power levels of the model were adequate based on the conventional cut-off criteria (Lakens, 2022).

Results

We tested the theory driven model proposed by Kapusta et al. (2018) by means of a CFA: a six-factor model with 41 items and scales being allowed to correlate with each other (Tables 2, 3). Due to our data following a non-normal distribution we used a robust standard errors optimization with a Satorra-Bentler scaled test statistic. Fit indices were: chi2/degrees of freedom = 2.71 (2070.038/763) SRMR = 0.060, RMSEA = 0.062 (90% CI 0.059–0.065). Based on the results of the $\chi 2$ statistic, a lack of overall fit was shown for the model tested ($p < 0.001$), which is likely due to the sensitivity of this statistic to large sample sizes (Hu and Bentler, 1998; Kahn, 2006). In fact, chi-square is highly sensitive to sample size: as the size of the sample increases, absolute differences become a smaller and smaller proportion of the expected value. The larger the sample, the larger and more significant will be the chi squares, even with very small discrepancies among implied and obtained covariance matrices. On the other hand, samples of reduced size may be too prone to accept poor models (Type II error).

According to other goodness-of fit indices, chi2/degrees of freedom = 3.13 (2598.870/764), SRMR = 0.060, RMSEA = 0.062 (90% CI 0.060–0.065), a good adequacy of the model was shown (Browne and Cudek, 1993; Hu and Bentler, 1998) and was comparable to previous results (Kapusta et al., 2018; Margherita et al., 2018). In order to evaluate the suitability of the proposed six-factor model compared to alternative models, we examined a five-factor model (where we excluded the Permanence of Sexual Passion dimension), as well as a four-factor model (where we also excluded the Loss and Mourning dimension). These two factors were chosen due to their lower convergent validity scores as can be seen in Table 2. The five-factor model chi2/degrees of freedom = 3.47 (2406.250/692), SRMR = 0.061, RMSEA = 0.061 as well as the four-factor model chi2/degrees of freedom = 4.25 (1818.710/428), SRMR = 0.060, RMSEA = 0.077 showed comparable fit indexes to the original model.

To investigate the potential associations among the six CTL-I subscales and other related measures, Pearson correlational coefficients were calculated. The standardized model solution is shown in Figure 1.

TABLE 3 Pearson correlations among study variables.

		Interest	Basic trust	Gratitude	Common ego ideal	Permanence of sexual passion	Loss and mourning	Total scale
MAAS		0.22**	0.33**	0.27**	0.24**	0.22**	0.30**	0.39**
PID-5	Negative affect	−0.02	−0.19**	−0.08	−0.12**	−0.08	−0.45**	−0.25**
	Detachment	−0.25**	−0.39**	−0.36**	−0.29**	−0.20**	−0.29**	−0.43**
	Antagonism	−0.21**	−0.23**	−0.22**	−0.22**	−0.12**	−0.15**	−0.27**
	Disinhibition	−0.18**	−0.25**	−0.25**	−0.22**	0.01	−0.24**	−0.25**
	Psychoticism	−0.20**	−0.29**	−0.22**	−0.16**	−0.11**	−0.32**	−0.32**
IPO-16	Identity	−0.09*	−0.23**	−0.10**	−0.11**	−0.13**	−0.46**	−0.29**
	Reality testing	−0.17**	−0.22**	−0.15**	−0.12**	−0.09*	0.25**	−0.24**
	Primitive defense	−0.17**	−0.30**	−0.24**	−0.21**	−0.11*	−0.31**	−0.32**

** indicates $p \leq 0.01$, MAAS, Mindful Attention Awareness Scale; PID 5, Personality Inventory for DSM-5; IPO-16, Inventory of Personality Organization 16.

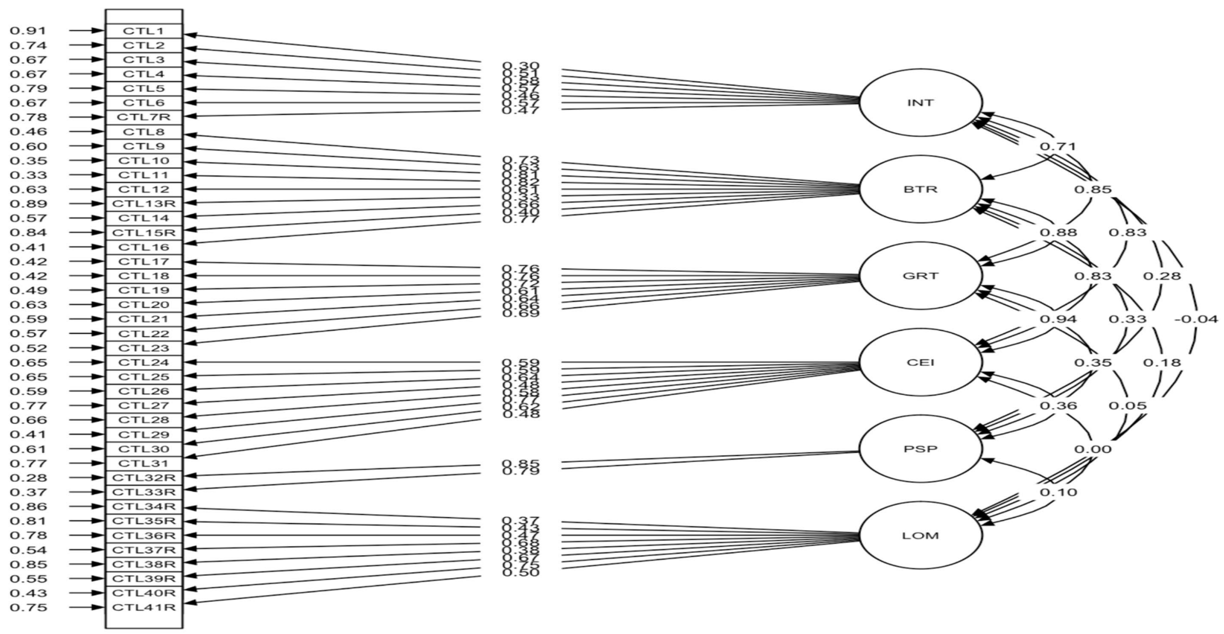

FIGURE 1
Standardized solution to the CTL model.

Table 2 shows the results of Pearson's correlation coefficient among the CTL-I subscales. Significant correlations were found among all subscales except for subscale Loss and mourning. The Loss and mourning subscale was significantly correlated with only the Basic Trust subscale and the CTL-I total.

As can be seen in Table 3, Pearson correlations between the PID-5 and CTL-I subscales showed the following results: Negative affect was significantly inversely associated with Basic trust, Common ego ideal, Loss and mourning as well as the total sum score. Detachment, Psychoticism, as well as Antagonism were significantly inversely associated with all CTL-I subscales. Similarly, Disinhibition was significantly inversely correlated with all subscales apart from the Permanence of sexual passion subscale. In examining the relationship between the three IPO subscales with the CTL-I we found that the Identity subscale was significantly inversely correlated with all the CTL-I subscales wherein the strongest association with IPO was between Identity and Loss and mourning as well as the total CTL-I score. Reality testing was also significantly inversely associated with all CTL-I subscales. Finally, the Defense subscale was similarly significantly inversely associated with all CTL-I subscales. Correlations between the Mindful Attention Awareness Scale and the CTL-I showed a strong positive relationship with all CTL-I subscales. In general, these associations fell within the small to moderate range.

To examine the unique contributions of CTL-I dimensions to the included external measures, we conducted hierarchical regression analyses controlling for gender and age in the first step, and then adding the CTL-I dimensions in the second step. The dependent variables included Identity, Primitive defense, Reality testing, Mindfulness, Antagonism, Detachment, Negative affect, Disinhibition and Psychoticism. Results of the hierarchical regressions are presented in Supplementary Tables S1, S2. Due to the covariance between some of the CTL-I scales, several associations observed in the correlations

10.3389/fpsyg.2024.1440013

(Table 2) were not observed in the regressions. Overall, these hierarchical regression analyses confirm the robust associations between Basic trust and Loss and Mourning with various external measures, consistent with the presented correlation findings. Other CTL-I dimensions, while correlated, did not uniquely predict the external measures when controlling for other factors, highlighting the unique contributions of these specific facets. The results underscore the contribution of the unique variance of Basic trust and Loss and Mourning in explaining variance in external measures. Furthermore, the results underscore that the CTL-I dimensions explain a substantial amount of the variance in the included dependent variables above what could be seen at the correlational level alone.

Additionally, to explore the potential moderating effects of gender and relationship status on the relationships between external measures and CTL dimensions, we conducted moderation analyses. Before running the moderation analyses, we also examined if homogeneity of variance assumptions were violated in relation to gender and the CTL-I dimensions. Results showed that homogeneity of variances was violated only in Basic Trust: $F\,(1{,}544) = 4.91$, $p = 0.027$. Interaction terms were created between the included external measures and the moderator variables (gender and relationship status). Overall, our findings show that in the majority of cases, gender (0—male, 1—female) was not found to be a significant moderator. Except for several cases in relation to Basic trust and the external variables such as (Primitive defense, $p = 0.026$, negative interaction, R^2 change = 0.0082), Mindfulness, ($p = 0.050$, R^2 change = 0.0062), Detachment, ($p = 0.024$, negative interaction, R^2 change = 0.0079), Disinhibition, ($p = 0.038$, negative interaction, R^2 change = 0.0074), Psychoticism, ($p = 0.036$, negative interaction, R^2 change = 0.0074) and two cases in relation to Loss and mourning with Primitive defense ($p = 0.001$, negative interaction, R^2 change = 0.0168) and Disinhibition ($p = 0.008$, negative interaction, R^2 change = 0.0119) were statistically significant at the more stringent $p \leq 01$ level. Almost none of the external variables were moderated by relationship status (coded as 0—single, 1—in a relationship), with the exception of Primitive defense in relation to Basic trust ($p = 0.050$, negative, R^2 change = 0.0058) and Primitive defense in relation to Interest in the other ($p = 0.050$, negative, R^2 change = 0.0066) as well as Psychoticism in relation to Interest in the other ($p = 0.045$, negative, R^2 change = 0.0069).

Discussion

The primary purpose of the present study was to examine the psychometric properties of the Capacity to Love-Inventory in the newly translated Slovenian version of the instrument as well as to examine its external validity with novel, converging measures. So far, the instrument has been validated in German (Kapusta et al., 2018), Polish (Kapusta et al., 2018), Italian (Margherita et al., 2018), Chinese (Li and Dong, 2021) and Slovenian (Poštuvan et al., 2022). To examine the construct validity of the CTL-I we assessed the proposed model by conducting confirmatory factor analysis. A secondary purpose of our study was to examine the relationship between capacity to love and dispositional mindfulness, which is a construct based on a distinctly non-psychoanalytical theory, thus allowing to establish trans-theoretical convergent validity.

Our results confirmed that the Slovenian version of the CTL-I has a clear factor structure and good internal consistency. The proposed six-factor model of the CTL-I demonstrates good fit across various goodness-of-fit indices and maintains theoretical robustness and was found to exhibit superior fit compared to alternative models with fewer factors. Furthermore, fit indexes of the full six-factor model in prior validation studies were very similar to those observed in the present study (Kapusta et al., 2018; Margherita et al., 2018). These results (Kapusta et al., 2018; Margherita et al., 2018) suggest comparable psychometric properties. To examine the convergent validity, we employed the PID-5-BF (Krueger et al., 2012) and the IPO-16 (Zimmermann et al., 2013), measures for assessing personality disorders and personality dysfunction, respectively. We expected both clinical measures to be negatively associated with one's capacity to love. Our analysis showed that capacity to love was significantly inversely associated to most characteristics of personality dysfunction and traits of personality disorders. These findings suggest that a limited capacity to love may indeed be associated with various forms of psychopathology. Since the capacity to love has previously been examined primarily in relation to a few clinical measures such as depression and narcissism (Kapusta et al., 2018; Margherita et al., 2018), we did not have specific assumptions about the size of these associations. Regardless the findings were in the expected directions as we expected clinical measures to be moderately negatively associated with the capacity to love as deficits in the capacity to love have been considered indicators of psychopathology (Kernberg, 1974, 1977, 2011). While limitations in the capacity to love as well as the construct's theoretical underpinnings have primarily been associated with narcissism (Kapusta et al., 2018; Kernberg, 2011), our results suggest that noticeable deficiencies of love capacity might be attributable to a wider array of personality features. These findings are in line with studies which have consistently shown personality disorders to be detrimental to relationship quality as well as relationship satisfaction (DeCuyper et al., 2018; South et al., 2020; Whisman et al., 2007; Whisman and Schonbrun, 2009). Relationship distress has been shown to consequentially and bidirectionally influence mental health and life satisfaction (Be et al., 2013). Furthermore, the experiences with significant others are laid down early in life based on interactions with caregivers and the subsequent relationships with romantic partners are likely to partially reflect the nature of these earlier caregiving relationships which are known to be crucial factors in mental health (Kernberg, 2006; Sroufe, 2000). Indeed, issues in romantic relationships as well as social relationships more broadly are commonly associated with mental health difficulties and lower well-being (Diener and Seligman, 2002; Frisch, 2005; Kouros et al., 2008; Groh et al., 2014; Kansky, 2018; Mehl et al., 2010) and our results suggest that a similar pattern might apply for the capacity to love.

We also expected capacity to love to be positively associated with mindfulness as it is a fairly well researched practice that has shown promising results in this regard (Coronado-Montoya et al., 2016: Galante et al., 2021; Kozlowski, 2013). Accordingly, the results of the present study found dispositional mindfulness to be associated with all the capacity to love subscales. In recent years an increasing number of findings show that mindfulness is beneficial to relationship quality (Kozlowski, 2013; McGill et al., 2016). Because the concept of capacity to love originates in a psychoanalytic framework, it is interesting to see mindfulness being positively related to this concept. Due to the assumption of mindfulness being both a disposition as well as a skill which can be practiced, our findings are promising as they seem to suggest that mindfulness practice may improve capacity to love and consequentially relationships satisfaction through changes on a deeper level of personality. The relationship between mindfulness and capacity to love might be a consequence of several factors, including the cultivation of empathy and empathic responding in couples (Block-Lerner et al., 2007), aiding in emotional skillfulness and the consequent benefits in emotional regulation (Wachs and Cordova, 2007) and stress response mitigation (Barnes et al., 2007). Mindfulness training has been shown to increase sexual desire and sexual arousal allowing partners to experience more sexual satisfaction in the relationship (Brotto et al., 2008). Because the present study is cross-sectional in design, causal inferences are not possible, indeed, interventional studies assessing the possible effects of mindfulness training on specific domains of CTL would be needed to test a causal association between both constructs.

Additionally, we examined the unique contribution of the CTL-I scales to the convergent measures. Results of the hierarchical regressions confirm the significant negative associations between specific CTL-I dimensions and various external measures. The findings suggest that the associations observed are not merely due to shared variance among the CTL-I dimensions but reflect the unique predictive power of the dimensions when controlling for other variables. Therein Basic trust and Loss and Mourning were observed to be especially distinct features of capacity to love. A possible explanation for the unique variance of these two dimensions is their similarity to aspects of the attachment process (Bowlby, 1969, 1980), which is known to be an important variable in relationship quality (Collins et al., 2006) as well as mental health and psychosocial functioning more broadly (Ranson and Urichuk, 2008; Zhang et al., 2022). In many ways Basic trust shares similarities with the characteristics of individuals possessing a secure attachment style, such as the capacity to tolerate one's uncertainty, that sense of security in relationship with others as well as an internalized security as a consequence of positive significant introjects (Kernberg, 2011). Secure attachment has consistently been found to be a vital aspect of relationship quality (Collins et al., 1990; Gleeson and Fitzgerald, 2014; Simpson and Rholes, 2017). The capacity for healthy mourning is a significant factor related to attachment as securely attached individuals are better able to process loss and to mourn important relationships a capacity that is influenced by earlier relationships with important others (Bowlby, 1969, 1980). It is also worth noting that in the original validation study Basic trust and Loss and Mourning had the highest (negative) correlations with pathological narcissism and narcissistic personality traits (Kapusta et al., 2018).

Apart from Basic trust and Loss and Mourning the regression analyses also showed the unique contribution of several other dimensions. For example, Gratitude was a negative predictor of Detachment and Disinhibition. Gratitude is thought to be congruent with a realistic self-regard and a genuine recognition of one's fundamental need for others in order to attain fuller enjoyment and a sense of security in life. Thus, individuals high in Detachment may struggle to form and maintain the close, emotionally fulfilling relationships that foster a sense of gratitude and Gratitude requires a certain level of emotional engagement and the ability to feel and express positive emotions towards others. Similarly, because Gratitude requires an emotional maturity and recognition of the importance of the other in one's life, individuals higher in Disinhibition, which is characterized by impulsivity might have bigger difficulties experiencing such Gratitude. Antagonism was also found to be negatively predicted by Common Ego Ideal, which is likely a consequence of altruism and empathy being key features related to the capacity for experiencing common ego ideal and that these are likely thwarted by the predisposition towards hostility in individuals higher in antagonism. Additionally, Permanence of sexual passion was found to contribute unique variance in relationship to mindfulness, a result that might be explained by the previously mentioned finding on the association between mindfulness and sexual desire (Brotto et al., 2008).

Furthermore, in order to examine the potential effects of gender and relationship status we examined additional moderation analyses across the associations. The findings of the moderations indicate that, in the majority of cases, these demographic factors do not significantly alter the relationships between the variables. However, notable exceptions were observed in specific cases. The minimal significant moderation by relationship status suggests robustness and invariance of the CTL dimensions across different relationship statuses. In relation to gender, a few moderation effects were observed in relation to Basic trust and Loss and mourning. This implies that males and females may interpret or experience these constructs slightly differently, which warrants further investigation and potential adjustments to the measurement tool to ensure fairness and accuracy across genders. While measurement invariance could not be examined on our sample, due to the considerable gender imbalance, future research should further consider these additional psychometric properties to better understand the underlying mechanisms and validate these findings across different populations and contexts.

Limitations

A limitation of our study was a disproportionate sex ratio of our sample—there was a significantly larger number of women who took part in our study. A possible explanation of this finding is that men may tend to be less interested in romantic relationships than women (Fraley et al., 2011) or are at least less interested in taking part in surveys on the topic of romantic relationships. Interestingly, previous CTL-I studies have shown similar results (Kapusta et al., 2018; Margherita et al., 2018). An additional limitation is that due to the significant gender imbalance we were not able to examine measurement invariance in relation to men and women. Furthermore, because our sample included a normative sample of students, higher ranges of psychopathology are not to be expected. Indeed, as the CTL-I validation studies have also been conducted on student samples, results show that the central tendency of the average scores is high (Kapusta et al., 2018; Margherita et al., 2018). Another limitation is the age of our sample, with most of our participants being university students they might have had less relationship experience than their older counterparts. Regardless of age however, most of our sample has had non-negligible relationship experience. Additionally, the moderation analyses showed that relationship status did not significantly influence the relationship between CTL-I dimensions and external measures. Lastly, because the present study was cross-sectional in design no causal inferences can be made on the directionality of dispositional mindfulness and capacity to love.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics committee of the Faculty of Arts, University of Maribor. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TG: Writing – review & editing, Writing – original draft, Software, Investigation, Data curation, Conceptualization. VP: Writing – review & editing, Supervision, Project administration, Investigation. JS: Writing – review & editing, Visualization, Software, Methodology, Investigation. NK: Writing – review & editing, Supervision, Resources, Project administration, Investigation, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The publication of this research was supported by a grant received by NK from the Medical University of Vienna.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1440013/full#supplementary-material

References

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (DSM-V). *5th* Edn. Washington, DC: American Psychiatric Publishing.

Anderson, J. L., Sellbom, M., and Salekin, R. T. (2016). Utility of the personality inventory for DSM-5–brief form (PID-5-BF) in the measurement of maladaptive personality and psychopathology. *Assessment* 25, 596–607. doi: 10.1177/1073191116676889

Bach, B., Maples-Keller, J. L., Bo, S., and Simonsen, E. (2016). The alternative DSM–5 personality disorder traits criterion: a comparative examination of three self-report forms in a Danish population. *Pers. Disord.* 7, 124–135. doi: 10.1037/per0000162

Barnes, S., Brown, K. W., Krusemark, E., Campbell, W. K., and Rogge, R. D. (2007). The role of mindfulness in romantic relationship satisfaction and responses to relationship stress. *J. Marital. Fam. Ther.* 33, 482–500. doi: 10.1111/j.1752-0606.2007.00033.x

Be, D., Whisman, M. A., and Uebelacker, L. A. (2013). Prospective associations between marital adjustment and life satisfaction. *Pers. Relatsh.* 20, 728–739. doi: 10.1111/pere.12011

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: a proposed operational definition. *Clin. Psychol. Sci. Pract.* 11, 230–241. doi: 10.1093/clipsy.bph077

Block-Lerner, J., Adair, C., Plumb, J. C., Rhatigan, D. L., and Orsillo, S. M. (2007). The case for mindfulness-based approaches in the cultivation of empathy: does nonjudgmental, present-moment awareness increase capacity for perspective-taking and empathic concern? *J. Marital. Fam. Ther.* 33, 501–516. doi: 10.1111/j.1752-0606.2007.00034.x

Bowlby, J. (1969). Attachment and Loss, Vol. 1: Attachment. Attachment and Loss. New York: Basic Books.

Bowlby, J. (1980). Attachment and loss. New York: Basic Books.

Bränström, R., Duncan, L. G., and Moskowitz, J. T. (2011). The association between dispositional mindfulness, psychological well-being, and perceived health in a Swedish population-based sample. *Br. J. Health Psychol.* 16, 300–316. doi: 10.1348/135910710X501683

Brotto, L. A., Basson, R., and Luria, M. (2008). A mindfulness-based group psychoeducational intervention targeting sexual arousal disorder in women. *J. Sex. Med.* 5, 1646–1659. doi: 10.1111/j.1743-6109.2008.00850.x

Brown, K. W., and Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *J. Pers. Soc. Psychol.* 84, 822–848. doi: 10.1037/0022-3514.84.4.822

Browne, M. W., and Cudeck, R. (1993). "Alternative ways of assessing model fit" in Testing structural equation models. eds. K. A. Bollen and J. S. Long (Cham: SAGE Publications), 136.

Campbell, L., Simpson, J. A., Boldry, J., and Kashy, D. A. (2005). Perceptions of conflict and support in romantic relationships: the role of attachment anxiety. *J. Pers. Soc. Psychol.* 88, 510–531. doi: 10.1037/0022-3514.88.3.510

Chadwick, P. (2014). Mindfulness for psychosis. *Br. J. Psychiatry* 204, 333–334. doi: 10.1192/bjp.bp.113.136044

Coan, J. A., Kasle, S., Jackson, A., Schaefer, H. S., and Davidson, R. J. (2013). Mutuality and the social regulation of neural threat responding. *Attach Hum. Dev.* 15, 303–315. doi: 10.1080/14616734.2013.782656

Collins, N. L., Ford, M. B., Guichard, A. C., and Allard, L. M. (2006). Working models of attachment and attribution processes in intimate relationships. *Pers. Soc. Psychol. B* 32, 201–219. doi: 10.1177/0146167205280907

Collins, N. L., and Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *J. Pers. Soc. Psychol.* 58, 644–663. doi: 10.1037/0022-3514.58.4.644

Coronado-Montoya, S., Levis, A. W., Kwakkenbos, L., Steele, R. J., Turner, E. H., and Thombs, B. D. (2016). Reporting of positive results in randomized controlled trials of mindfulness-based mental health interventions. *PLoS One* 11:e0153220. doi: 10.1371/journal.pone.0153220

Debast, I., Rossi, G., and Van Alphen, S. P. J. (2017). Construct validity of the DSM-5 section III maladaptive trait domains in older adults. *J. Personal. Disord.* 31, 671–688. doi: 10.1521/pedi_2017_31_274

DeCuyper, M., Gistelinck, F., Vergauwe, J., Pancorbo, G., and De Fruyt, F. (2018). Personality pathology and relationship satisfaction in dating and married couples. *Pers. Disord.* 9, 81–92. doi: 10.1037/per0000219

Diener, E., and Seligman, M. E. (2002). Very happy people. *Psychol. Sci.* 13, 81–84. doi: 10.1111/1467-9280.00415

Disney, K. L., Weinstein, Y., and Oltmanns, T. F. (2012). Personality disorder symptoms are differentially related to divorce frequency. *J. Fam. Psychol.* 26, 959–965. doi: 10.1037/a0030446

Dobrowolska, M., Groyecka-Bernard, A., Sorokowski, P., Randall, A. K., Hilpert, P., Ahmadi, K., et al. (2020). Global perspective on marital satisfaction. *Sustain. For.* 12:8817. doi: 10.3390/su12218817

Dush, C. M. K., and Amato, P. R. (2005). Consequences of relationship status and quality for subjective well-being. *J. Soc. Pers. Relat.* 22, 607–627. doi: 10.1177/0265407505056438

Fletcher, G. J., Simpson, J. A., Campbell, L., and Overall, N. C. (2015). Pair-bonding, romantic love, and evolution: the curious case of *Homo sapiens*. *Perspect. Psychol. Sci.* 10, 20–36. doi: 10.1177/1745691614561683

Foran, H. M., Hahlweg, K., Kliem, S., and O'Leary, K. D. (2013). Longitudinal patterns of relationship adjustment among German parents. *J. Fam. Psychol.* 27, 838–843. doi: 10.1037/a0034183

Fraley, R. C., Heffernan, M. E., Vicary, A. M., and Brumbaugh, C. C. (2011). The experiences in close relationships—relationship structures questionnaire: a method for assessing attachment orientations across relationships. *Psychol. Assess.* 23, 615–625. doi: 10.1037/a0022898

Frans, Ö., Rimmö, P. A., Åberg, L., and Fredrikson, M. (2005). Trauma exposure and post-traumatic stress disorder in the general population. *Acta Psychiatr. Scand.* 111, 291–290. doi: 10.1111/j.1600-0447.2004.00463.x

Frisch, M. B. (2005). Quality of life therapy: Applying a life satisfaction approach to positive psychology and cognitive therapy. New Jersey: John Wiley and Sons.

Galante, J., Friedrich, C., Dawson, A. F., Modrego-Alarcón, M., Gebbing, P., Delgado-Suárez, I., et al. (2021). Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: a systematic review and meta-analysis of randomised controlled trials. *PLoS Med.* 18:e1003481. doi: 10.1371/journal.pmed.1003481

Garza-Guerrero, C. (2000). Idealization and mourning in love relationships: Normal and pathological spectra. *Psychoanal. Q.* 69, 121–150. doi: 10.1002/j.2167-4086.2000.tb00557.x

Gleeson, G., and Fitzgerald, A. (2014). Exploring the association between adult attachment styles in romantic relationships, perceptions of parents from childhood and relationship satisfaction. *Health* 6, 1643–1661. doi: 10.4236/health.2014.613196

Groh, A. M., Fearon, R. P., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Steele, R. D., and Roisman, G. I. (2014). The significance of attachment security for children's social competence with peers: a meta-analytic study. *Attach Hum. Dev.* 16, 103–136. doi: 10.1080/14616734.2014.883636

Grossman, P., Niemann, L., Schmidt, S., and Walach, H. (2004). Mindfulness-based stress reduction and health benefits: a meta-analysis. *J. Psychosom. R.* 57, 35–43. doi: 10.1016/S0022-3999(03)00573-7

Hanley, A., Warner, A., and Garland, E. L. (2015). Associations between mindfulness, psychological well-being, and subjective well-being with respect to contemplative practice. *J. Happiness Stud.* 16, 1423–1436. doi: 10.1007/s10902-014-9569-5

Heller, D., Watson, D., and Ilies, R. (2004). The role of person versus situation in life satisfaction: a critical examination. *Psychol. Bull.* 130, 574–600. doi: 10.1037/0033-2909.130.4.574

Homish, G. G., and Leonard, K. E. (2008). Spousal influence on general health behaviors in a community sample. *Am. J. Health Behav.* 32, 754–763. doi: 10.5555/ajhb.2008.32.6.754

Hopwood, C. J., Wright, A. G., Krueger, R. F., Schade, N., Markon, K. E., and Morey, L. C. (2013). DSM-5 pathological personality traits and the personality assessment inventory. *Assessment* 20, 269–285. doi: 10.1177/1073191113486286

Hu, L. T., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol. Methods* 3, 424–453. doi: 10.1037/1082-989X.3.4.424

Humbad, M. N., Donnellan, M. B., Iacono, W. G., and Burt, S. A. (2010). Externalizing psychopathology and marital adjustment in long-term marriages: results from a large combined sample of married couples. *J. Abnorm. Psychol.* 119, 151–162. doi: 10.1037/a0017981

IBM Corp (2020). IBM SPSS statistics for windows, version 27.0. Armonk: IBM Corp.

Jankowiak, W. R., and Fischer, E. F. (1992). A cross-cultural perspective on romantic love. *Ethnology* 31, 149–155. doi: 10.2307/3773618

Kabat-Zinn, J. (1994). Wherever you go, there you are: Mindfulness meditation in everyday life. New York: Hyperion.

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: principles, advances, and applications. *Couns. Psychol.* 34, 684–718. doi: 10.1177/0011000006286347

Kansky, J. (2018). "What's love got to do with it? Romantic relationships and well-being" in Handbook of well-being (Salt Lake City, UT: DEF Publishers).

Kapusta, N. D., Jankowski, K. S., Wolf, V., Chéron-Le Guludec, M., Lopatka, M., Hammerer, C., et al. (2018). Measuring the capacity to love: development of the CTL-inventory. *Front. Psychol.* 9:1115. doi: 10.3389/fpsyg.2018.01115

Kernberg, O. F. (1974). Barriers to falling and remaining in love. *J. Am. Psychoanal. Assoc.* 22, 486–511. doi: 10.1177/000306517402200302

Kernberg, O. F. (1977). Boundaries and structure in love relations. *J. Am. Psychoanal. Assoc.* 25, 81–114. doi: 10.1177/000306517702500104

Kernberg, O. F. (1985). Borderline conditions and pathological narcissism. Maryland: Rowman and Littlefield.

Kernberg, O. F. (2006). Identity: recent findings and clinical implications. *Psychoanal. Q.* 75, 969–1004. doi: 10.1002/j.2167-4086.2006.tb00065.x

Kernberg, O. F. (2009). Narcissistic personality disorders: part 1. *Psychiatr. Ann.* 39, 105–167. doi: 10.3928/00485713-20090301-04

Kernberg, O. F. (2011). Limitations to the capacity to love. *Int. J. Psychoanal.* 92, 1501–1515. doi: 10.1111/j.1745-8315.2011.00456.x

Kernberg, O. F., and Caligor, E. (2005). "A psychoanalytic theory of personality disorders" in Major theories of personality disorder. eds. M. F. Lenzenweger and J. F. Clarkin. *2nd* ed (London: Guilford Press), 114–156.

Khoury, B., Lecomte, T., Gaudiano, B. A., and Paquin, K. (2013). Mindfulness interventions for psychosis: a meta-analysis. *Schizophr. Res.* 150, 176–184. doi: 10.1016/j.schres.2013.07.055

Kiecolt-Glaser, J. K., and Newton, T. L. (2001). Marriage and health: his and hers. *Psychol. Bull.* 127, 472–503. doi: 10.1037/0033-2909.127.4.472

Kostanski, M., and Hassed, C. (2008). Mindfulness as a concept and a process. *Aust. Psychol.* 43, 15–21. doi: 10.1080/00050060701593942

Kouros, C. D., Papp, L. M., and Cummings, E. M. (2008). Interrelations and moderators of longitudinal links between marital satisfaction and depressive symptoms among couples in established relationships. *J. Fam. Psychol.* 22, 667–677. doi: 10.1037/0893-3200.22.5.667

Kozlowski, A. (2013). Mindful mating: exploring the connection between mindfulness and relationship satisfaction. *Sex. Relatsh. Ther.* 28, 92–104. doi: 10.1080/14681994.2012.748889

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., and Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychol. Med.* 42, 1879–1890. doi: 10.1017/S0033291711002674

Lakens, D. (2022). Sample size justification. *Collabra: Psychol* 8:33267. doi: 10.1525/collabra.33267

Lenzenweger, M. F., Clarkin, J. F., Kernberg, O. F., and Foelsch, P. A. (2001). The inventory of personality organization: psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychol. Assess.* 13, 577–591. doi: 10.1037/1040-3590.13.4.577

Li, L., and Dong, Y. (2021). Validity and reliability of the Chinese version capacity to love inventory. *J. Southwest China Normal Univ.* 46, 31–37.

Lin, E., Goering, P., Offord, D. R., Campbell, D., and Boyle, M. H. (1996). The use of mental health services in Ontario: epidemiologic findings. *Can. J. Psychiatr.* 41, 572–577. doi: 10.1177/070674379604100905

Margherita, G., Gargiulo, A., Troisi, G., Tessitore, F., and Kapusta, N. D. (2018). Italian validation of the capacity to love inventory: preliminary results. *Front. Psychol.* 9:1434. doi: 10.3389/fpsyg.2018.01434

McGill, J., Adler-Baeder, F., and Rodriguez, P. (2016). Mindfully in love: a meta-analysis of the association between mindfulness and relationship satisfaction. *J. Hum. Sci. Ext.* 4:7. doi: 10.54718/DDCA4089

Mehl, M. R., Cazire, S., Holleran, S. E., and Clark, C. S. (2010). Eavesdropping on happiness: well-being is related to having less small talk and more substantive conversations. *Psychol. Sci.* 21, 539–541. doi: 10.1177/0956797610362675

Meyler, D., Stimpson, J. P., and Peek, M. K. (2007). Health concordance within couples: a systematic review. *Soc. Sci. Med.* 64, 2297–2310. doi: 10.1016/j.socscimed.2007.02.007

Miller, M. W., Wolf, E. J., Reardon, A. F., Harrington, K. M., Ryabchenko, K., Castillo, D., et al. (2013). PTSD and conflict behavior between veterans and their intimate partners. *J. Anxiety Disord.* 27, 240–251. doi: 10.1016/j.janxdis.2013.02.005

Poštuvan, V., Glavač, T., Schmeckenbecher, J., and Kapusta, N. D. (2022). Slovenian validation of the capacity to love inventory–a preliminary study. *Psychol. Top.* 31, 317–335. doi: 10.31820/pt.31.2.6

R Core Team (2022). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Ranson, K. E., and Urichuk, L. J. (2008). The effect of parent–child attachment relationships on child biopsychosocial outcomes: a review. *Early Child Dev. Care* 178, 129–152. doi: 10.1080/03004430600685282

Reis, H. T., and Downey, G. (1999). Social cognition in relationships: building essential bridges between two literatures. *Soc. Cogn.* 17, 97–117. doi: 10.1521/soco.1999.17.2.97

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Meth. Psychol. Res.* 8, 23–74.

Shorey, R. C., Elmquist, J., Wolford-Clevenger, C., Gawrysiak, M. J., Anderson, S., and Stuart, G. L. (2016). The relationship between dispositional mindfulness, borderline personality features, and suicidal ideation in a sample of women in residential substance use treatment. *Psychiatry Res.* 238, 122–128. doi: 10.1016/j.psychres.2016.02.040

Simpson, J. A., and Rholes, W. S. (2017). Adult attachment, stress, and romantic relationships. *Curr. Opin. Psychol.* 13, 19–24. doi: 10.1016/j.copsyc.2016.04.006

Sng, A. A., and Janca, A. (2016). Mindfulness for personality disorders. *Curr. Opin. Psychiatry* 29, 70–76. doi: 10.1097/YCO.0000000000000213

South, S. C., Boudreaux, M. J., and Oltmanns, T. F. (2020). The impact of personality disorders on longitudinal change in relationship satisfaction in long-term married couples. *J. Personal. Disord.* 34, 439–458. doi: 10.1521/pedi_2019_33_392

South, S. C., Turkheimer, E., and Oltmanns, T. F. (2008). Personality disorder symptoms and marital functioning. *J. Consult. Clin. Psychol.* 76, 769–780. doi: 10.1037/a0013346

Soysa, C. K., Zhang, F., Parmley, M., and Lahikainen, K. (2021). Dispositional mindfulness and serenity: their unique relations with stress and mental well-being. *J. Happiness Stud.* 22, 1517–1536. doi: 10.1007/s10902-020-00282-0

Sroufe, L. A. (2000). Early relationships and the development of children. *Infant Ment. Health J.* 21, 67–74. doi: 10.1002/(SICI)1097-0355(200001/04)21:1/2<67::AID-IMHJ8>3.0.CO;2-2

Stroud, C. B., Durbin, C. E., Saigal, S. D., and Knobloch-Fedders, L. M. (2010). Normal and abnormal personality traits are associated with marital satisfaction for both men and women: an actor–partner interdependence model analysis. *J. Res. Per.* 44, 466–477. doi: 10.1016/j.jrp.2010.05.011

Swindle, R. Jr., Heller, K., Pescosolido, B., and Kikuzawa, S. (2000). Responses to nervous breakdowns in America over a 40-year period: mental health policy implications. *Am. Psychol.* 55, 740–749. doi: 10.1037/0003-066X.55.7.740

Tomlinson, E. R., Yousaf, O., Vittersø, A. D., and Jones, L. (2018). Dispositional mindfulness and psychological health: a systematic review. *Mindfulness* 9, 23–43. doi: 10.1007/s12671-017-0762-6

Velotti, P., Garofalo, C., Dimaggio, G., and Fonagy, P. (2019). Mindfulness, alexithymia, and empathy moderate relations between trait aggression and antisocial personality disorder traits. *Mindfulness* 10, 1082–1090. doi: 10.1007/s12671-018-1048-3

Wachs, K., and Cordova, J. V. (2007). Mindful relating: exploring mindfulness and emotion repertoires in intimate relationships. *J. Marital. Fam. Ther.* 33, 464–481. doi: 10.1111/j.1752-0606.2007.00032.x

Whisman, M. A. (2007). Marital distress and DSM-IV psychiatric disorders in a population-based national survey. *J. Abnorm. Psychol.* 116, 638–643. doi: 10.1037/0021-843X.116.3.638

Whisman, M. A., and Schonbrun, Y. C. (2009). Social consequences of borderline personality disorder symptoms in a population-based survey: marital distress, marital violence, and marital disruption. *J. Personal. Disord.* 23, 410–415. doi: 10.1521/pedi.2009.23.4.410

Whisman, M. A., Sheldon, C. T., and Goering, P. (2000). Psychiatric disorders and dissatisfaction with social relationships: does type of relationship matter? *J. Abnorm. Psychol.* 109, 803–808. doi: 10.1037/0021-843X.109.4.803

Whisman, M. A., Tolejko, N., and Chatav, Y. (2007). Social consequences of personality disorders: probability and timing of marriage and probability of marital disruption. *J. Personal. Disord.* 21, 690–695. doi: 10.1521/pedi.2007.21.6.690

Wupperman, P., Neumann, C. S., Whitman, J. B., and Axelrod, S. R. (2009). The role of mindfulness in borderline personality disorder features. *J. Nerv. Ment.* 197, 766–771. doi: 10.1097/NMD.0b013e3181b97343

Zhang, X., Li, J., Xie, F., Chen, X., Xu, W., and Hudson, N. W. (2022). The relationship between adult attachment and mental health: a meta-analysis. *J. Pers. Soc. Psychol.* 123, 1089–1137. doi: 10.1037/pspp0000437

Zimmermann, J., Benecke, C., Horz, S., Rentrop, M., Peham, D., Bock, A., et al. (2013). Validierung einer deutschsprachigen 16-Item-Version des Inventars der Personlichkeitsorganisation (IPO-16). *Diagnostica* 59, 3–16. doi: 10.1026/0012-1924/a000076

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 10 September 2024
DOI 10.3389/fpsyg.2024.1432454

OPEN ACCESS

EDITED BY
Antonio Hernández-Mendo,
University of Malaga, Spain

REVIEWED BY
Yarisel Quiñones Rodríguez,
University of Granada, Spain
Marcin Siwek,
Jagiellonian University Medical College, Poland

*CORRESPONDENCE
Yixuan Ku
✉ kuyixuan@mail.sysu.edu.cn

†These authors have contributed equally to this work

RECEIVED 14 May 2024
ACCEPTED 23 August 2024
PUBLISHED 10 September 2024

CITATION
Ma Y, Gao Y, Yang H, Zhang Y and Ku Y (2024) Enhancing mental well-being of undergraduates: establishing cut-off values and analyzing substitutive effects of physical activity on depression regulation.
Front. Psychol. 15:1432454.
doi: 10.3389/fpsyg.2024.1432454

Enhancing mental well-being of undergraduates: establishing cut-off values and analyzing substitutive effects of physical activity on depression regulation

Yue Ma[1,2,3†], Yulin Gao[3†], Hui Yang[3], Yu Zhang[3] and Yixuan Ku[4*]

[1]Faculty of Medicine, Macau University of Science and Technology, Taipa, Macau SAR, China, [2]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, Guangzhou, China, [3]School of Nursing, Southern Medical University, Guangzhou, Guangdong, China, [4]Guangdong Provincial Key Laboratory of Brain Function and Disease, Center for Brain and Mental Well-being, Department of Psychology, Sun Yat-sen University, Guangzhou, China

Objective: This study aimed to analyze the effects of physical activity (PA), sleep quality, and sedentary behavior on subthreshold depression (StD) among undergraduates.

Methods: This study included 834 undergraduates and assessed the impact of PA time, sleep quality, and sedentary behavior on depression. The receiver operating characteristic (ROC) analysis was performed to determine cut-off values for StD risk, while the isochronous substitution analysis was performed to evaluate the effects of different activities on depression regulation.

Results: Gender, age, and academic grade had no significant influence on depression levels among undergraduates ($p > 0.05$). However, students engaging in sedentary behavior for more than 12.1 h per day or with a Pittsburgh Sleep Quality Index score above 3.5 were at an increased risk of subclinical depression. Additionally, the isochronous substitution of light-intensity physical activity for other activities (sleep, sedentary behavior, moderate and vigorous intensity physical activity) showed statistically significant effects ($p < 0.05$) in both 5-min and 10-min substitution models, demonstrating a positive effect on alleviating depression.

Conclusion: The findings indicate that specific lifestyle factors, particularly high levels of sedentary behavior and poor sleep quality, are crucial determinants of subclinical depression among undergraduates, independent of demographic variables such as gender, age, and academic grade. Notably, light-intensity PA plays a key role in StD regulation, as substituting it with more intense physical activities or improving sleep quality substantially reduces depression scores. Furthermore, the benefits such substitution became more pronounced with the increase in duration of the activity.

KEYWORDS

subclinical depression, physical activity, sedentary, cut-off value, isochronous substitution

1 Introduction

In the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5), a Beck Depression Inventory-II (BDI-II) score of 14 or higher with 2 major symptoms within 2 weeks is classified as subthreshold depression (StD) or subclinical depression (Rodríguez et al., 2012; Button et al., 2015). The prevalence of subclinical depression among undergraduates is alarmingly high, ranging from 13.2 to 39.1%, with a rapidly increasing trend (Bao et al., 2020). If left untreated, StD can lead to serious consequences, such as major depression and even suicide (Sheldon et al., 2021).

Several studies have shown that factors such as grade level, gender, and age may influence depression. For example, female students report significantly higher rates of depressive symptoms compared to male students (Acharya et al., 2018). Additionally, depressive symptoms tend to be more prevalent among older students (Bhandari et al., 2017). Other studies have identified age, grade level, and birthplace as influential factors (Wang et al., 2022). However, whether these factors influence StD remains unknown.

Undergraduates, who are in the early stages of adulthood, often experience abundant physical energy and frequent mood swings. They are also commonly confronted with challenges such as prolonged study hours, sedentary behavior (SB), and insufficient physical activity (IPA) (Li et al., 2021). SB, IPA, and sleep quality have been shown to significantly influence both survival indices and depression (Maddox et al., 2023). For example, the incidence of mild depression is four times higher in sedentary individuals compared to those who are physically active (Alotaibi and Boukelia, 2021). Engaging in more than 10 h of sedentary time per day is associated with a higher likelihood of depression (Schuch et al., 2020).

Physical activity (PA) has been proven to be an effective intervention for depression with minimal adverse events (Josefsson et al., 2014; Geneen et al., 2017). Some studies have demonstrated that engaging in PA reduces depression (Cooney et al., 2013), with 25–65 min of moderate PA (MPA) per day proving to be particularly effective (Chen et al., 2022). Supervised and group aerobic exercise programs, especially those conducted at a moderate intensity, have shown better outcomes (Heissel et al., 2023). However, few studies have analyzed the combined and quantitative effects of sleep and PA on StD in the undergraduate population.

Cut-off studies are often used to determine clinical parameter thresholds (Hajian-Tilaki, 2013; Nahm, 2022), which can help identify the PA and sleep indicators corresponding to the StD cut-off point (Negeri et al., 2021; Ghazisaeedi et al., 2022). The isochronous substitution analysis, which considers the 24-h day as a fixed component, encompassing sedentary time, moderate-to-vigorous PA (MVPA), light PA (LPA), and sleep time, reveals that an increase in one component necessitates a corresponding decrease in others (De Faria et al., 2022).

This study aims to provide a comprehensive quantitative analysis of how PA and sleep influence depression regulation among undergraduates. By defining the specific parameters that regulate depression through cut-off and substitution time analyses, this research offers valuable insights for undergraduates seeking to manage their depression and improve their sleep quality through tailored PA interventions.

2 Objects and methods

2.1 Objects

Between June 2022 and February 2024, using the PASS sample size calculation tool, this study determined the required sample size to be 513 participants; considering a 10% attrition rate, this resulted in a final total sample size of 570.

Participants were eligible for inclusion if they met the following criteria: (1) undergraduate students (not taking a break from school); (2) healthy enough to perform daily activities. The exclusion criteria are as follows (1) an existing diagnosis of a mental illness; (2) use of psychotropic medications, and (3) refusal to participate in the study. All participants provided their written informed consent after receiving a detailed explanation of the study, which was conducted in accordance with the Declaration of Helsinki.

The study was approved by the Ethics Committee of Southern Medical University[2023(NO.25)]. For the final analysis, 834 participants were randomly selected from each grade level at a university in China to complete an online questionnaire.

2.2 Questionnaires

2.2.1 International Physical Activity Questionnaire (IPAQ)

This study used the long form of the International Physical Activity Questionnaire (IPAQ), which consists of 21 questions addressing physical activities related to work, transportation, household chores and gardening, leisure sports, and sedentary time (Schuch et al., 2021). The reliability and validity of the Chinese version of the IPAQ were verified using a repeated survey method, which showed a correlation coefficient of 0.94 between the two sets of scores, indicating good test–retest reliability and construct validity (Liang et al., 2010). This study specifically focused on the time spent in various intensities of PA and the total sedentary time (Fan et al., 2014).

2.2.2 Beck Depression Inventory-II (BDI-II)

The Beck Depression Inventory comprises 21 items and is widely regarded as the gold standard for self-assessment of depression (Dozois et al., 1998). It employs a 4-point scoring system (0–3), with a total score ranging from 0 to 63 points (Kaya et al., 2021). The Cronbach's α coefficient for the Chinese version of the BDI-II is 0.94 (Wang et al., 2011).

Based on DSM-5 criteria, this study excluded individuals diagnosed with depression. For the purpose of classification, a BDI-II score greater than 14, along with the presence of two major symptoms within 2 weeks, was categorized as subclinical depression (Takagaki et al., 2014; Spek et al., 2008; Zhang et al., 2021).

2.2.3 Pittsburgh Sleep Quality Index (PSQI)

The Pittsburgh Sleep Quality Index is a self-report questionnaire that assesses sleep quality. It consists of 19 self-rated items and 5 items rated by others, categorized into seven components: A. sleep quality, B. time to fall asleep, C. sleep duration, D. sleep efficiency, E. sleep disturbances, F. use of sleeping medication, and G. daytime dysfunction. The total PSQI score ranges from 0 to 21, with higher scores indicating poorer sleep quality. The Cronbach's α coefficient for

TABLE 1 Univariate associations of demographic characteristics with BDI scores ($n = 834$).

Items		(n)	(%)	BDI-II (Means±SD)	r/t/F	p
Age	Average age	21.41 ± 1.65		8.97 ± 8.32	$r = 0.03$	0.43
Gender	Men	123	14.70	10.07 ± 10.12	$t = 1.34$	0.18
	Women	711	85.30	8.78 ± 7.69		
Grade	Freshman	197	23.62	9.76 ± 8.64	$F = 2.19$	0.09
	Sophomore	216	25.90	9.44 ± 8.80		
	Junior	223	26.74	7.84 ± 7.39		
	Senior	198	23.74	8.95 ± 8.37		

the Chinese version of the PSQI is 0.89 (Liu et al., 1996), indicating good internal consistency (Fawzy and Hamed, 2017; Liu et al., 2021).

2.3 Statistical methods

General information was analyzed using IBM SPSS 20.0 to examine potential differences in depression scores by gender and grade level. Correlation and regression analyses were conducted, incorporating ROC cutoff values (Dolle et al., 2012), to analyze the effects of physical activity (PA) and sleep indices on depression. The isochronic substitution analysis was performed using R (version 3.6.3) (Dumuid et al., 2019) by selecting four indicators—sedentary time, light physical activity (LPA) time, moderate-to-vigorous physical activity (MVPA) time, and sleep time—to explore the effects of substituting time spent in different activities on reducing depressive moods (Pasanen et al., 2022). The results are reported using 5-min and 10-min isochronous substitution models (Dumuid et al., 2019).

3 Results

3.1 Demographic analysis of depression scores

In the analysis of basic demographic variables, no statistically significant differences in depression scores were found based on gender, age, or academic grade (see Table 1).

3.2 The association between physical activity, sedentary behavior, sleep, and StD

Furthermore, the regression analysis highlighted the significant effect of PA time, sleep quality, and sedentary time on depression (Table 2). The ROC curve analysis, used to determine cutoff values, is shown in Figure 1.

Sedentary time over 12.1 h per day and a sleep quality score greater than 3.5 increased the risk of StD (see Table 3).

3.3 Substitution analysis of physical activity

Although some studies have indicated that a minimum duration of 5 min can be effective for certain health benefits (Su, 2022), 10 min is often considered the smallest unit for individual health benefits and the shortest duration at which sedentary behavior poses a health risk (Li et al., 2020). Therefore, in this study, the substitution analysis was conducted for both 5-min and 10-min intervals. Moderate and vigorous physical activities were combined as MVPA. As shown in Table 4, an increase in LPA by 10 min accompanied by a reduction in sedentary time or engaging in MVPA is associated with a significant decrease in depression.

4 Discussion

4.1 Prolonged sedentary behavior and poor sleep quality as indicators of subthreshold depression risk

This study found that a PSQI score above 3.5 and more than 12 h of sedentary behavior per day significantly increase the risk of StD. These findings align with previous research that highlights the impact of sleep quality on depression and StD (Gardani et al., 2022), emphasizing the urgent need to enhance students' awareness about the critical thresholds for maintaining sleep quality.

Regarding sedentary behavior, previous research has indicated that 8.5–10 h of sedentary time is associated with increased all-cause mortality and depression (Ku et al., 2019), which is notably lower than the threshold identified in this study. This discrepancy may be due to several factors. First, the demographic composition of this study, mainly consisting of undergraduate students, may exhibit a higher tolerance to the adverse effects of sedentary behavior compared to older adults, possibly due to their relatively faster metabolism. Additionally, existing research has suggested a dose–response relationship between PA and depression, with higher PA levels potentially exacerbating depression symptoms (Pearce et al., 2022). Furthermore, differences in sedentary time observed between this study and previous literature could result from methodological variations in activity measurement, such as the use of different instruments (Schmid et al., 2015; Ku et al., 2019).

Therefore, this study emphasizes the need for moderation in PA engagement while balancing sedentary behavior and sleep, offering nuanced insights for promoting mental well-being among undergraduates.

4.2 LPA isochronous substitution for sedentary behavior, sleep, and MVPA reduces StD risk

This study found that substitutions of 5–10 min of LPA effectively reduced the risk of StD, suggesting that undergraduates could benefit from integrating brief activity bouts during breaks between classes. The depression-reducing effects of isochronous substitution are expected to accumulate over time (Table 4). Notably, moderate to vigorous physical activity (MVPA) was associated with increased

TABLE 2 Binary logistic regression model outcomes for StD risk ($n = 834$).

Items	B	SE	*p*	*Wald*	Exp B	95% CI for Exp B	
						Lower	Upper
Constant	−5.25	1.22	<0.01	18.55	0.01		
PA time	−0.2	0.09	0.03	4.64	0.82	0.69	0.98
PSQI	0.58	0.05	<0.01	155.66	1.78	1.62	1.95
Sedentary time	0.2	0.09	0.02	5.09	1.22	1.03	1.45

TABLE 3 ROC analysis and cutoff values for StD.

Items	AUC	SE	*p*	Cut-off values	Sensitivity	Specificity	Youden index
PA time	0.36	0.02	<0.01	4.49 h	0.13	0.87	<0.01
Sedentary time	0.65	0.02	<0.01	12.10 h	0.62	0.65	0.27
PSQI	0.84	0.15	<0.01	3.50	0.82	0.70	0.52

AUC, Area under curve.

TABLE 4 Component isochronous substitution analysis of physical activity and BDI-II.

Time	(+Δ)	(−Δ)	*β*	95%CI (low, up)	*p*
5 min	(+Δ) Sleep	(−Δ) Sedentary	−0.16	(−0.23, −0.10)	<0.01
	(+Δ) Sleep	(−Δ) LPA	0.30	(0.18, 0.41)	<0.01
	(+Δ) Sleep	(−Δ) MVPA	−0.27	(−0.36, −0.17)	<0.01
	(+Δ) Sedentary	(−Δ) Sleep	0.16	(0.10, 0.23)	<0.01
	(+Δ) Sedentary	(−Δ) LPA	0.46	(0.36, 0.56)	<0.01
	(+Δ) Sedentary	(−Δ) MVPA	−0.11	(−0.19, −0.02)	0.01
	(+Δ) LPA	(−Δ) Sleep	−0.24	(−0.34, −0.14)	<0.01
	(+Δ) LPA	(−Δ) Sedentary	−0.40	(−0.48, −0.32)	<0.01
	(+Δ) LPA	(−Δ) MVPA	−0.51	(−0.63, −0.38)	<0.01
	(+Δ) MVPA	(−Δ) Sleep	0.25	(0.16, 0.34)	<0.01
	(+Δ) MVPA	(−Δ) Sedentary	0.09	(0.01, 0.16)	0.02
	(+Δ) MVPA	(−Δ) LPA	0.55	(0.42, 0.68)	<0.01
10 min	(+Δ) Sleep	(−Δ) Sedentary	−0.32	(−0.45, −0.20)	<0.01
	(+Δ) Sleep	(−Δ) LPA	0.69	(0.44, 0.94)	<0.01
	(+Δ) Sleep	(−Δ) MVPA	−0.57	(−0.77, −0.36)	<0.01
	(+Δ) Sedentary	(−Δ) Sleep	0.33	(0.20, 0.45)	<0.01
	(+Δ) Sedentary	(−Δ) LPA	1.01	(0.79, 1.23)	<0.01
	(+Δ) Sedentary	(−Δ) MVPA	−0.24	(−0.43, −0.06)	0.01
	(+Δ) LPA	(−Δ) Sleep	−0.43	(−0.62, −0.24)	<0.01
	(+Δ) LPA	(−Δ) Sedentary	−0.76	(−0.91, −0.60)	<0.01
	(+Δ) LPA	(−Δ) MVPA	−1.00	(−1.25, −0.75)	<0.01
	(+Δ) MVPA	(−Δ) Sleep	0.48	(0.32, 0.65)	<0.01
	(+Δ) MVPA	(−Δ) Sedentary	0.15	(0.01, 0.30)	0.04
	(+Δ) MVPA	(−Δ) LPA	1.16	(0.89, 1.44)	<0.01

depression scores, which contrasts with existing findings that generally favor its depression-reducing effects (Roeh et al., 2020).

There are several possible explanations for these findings. First, the average levels of PA across all intensities in this study exceeded those reported in previous research and guidelines, reflecting a U-shaped relationship between exercise and depression scores (Kim et al., 2018). This finding aligns with Pearce's findings, which suggest that excessive PA may diminish its benefits and increase uncertainty in its effects (Pearce et al., 2022). While PA generally has the potential to alleviate StD, excessive PA may exacerbate it. Therefore, LPA may

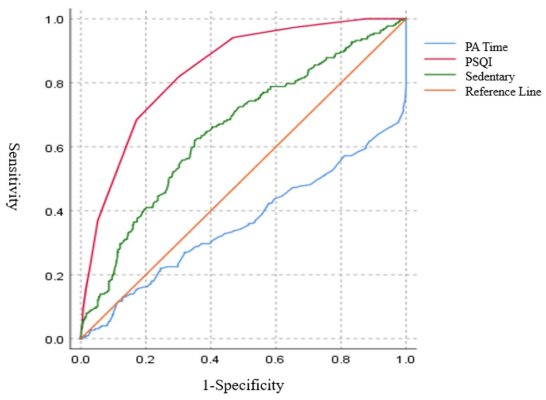

FIGURE 1
ROC curves for sleep quality, sedentary time, and PA time on StD risk.

be the optimal intensity for managing subclinical depression, as supported by recommendations from prior research (Pasanen et al., 2022). Modifications in the PA composition, particularly the LPA to MVPA ratio, could be considered in future interventions, potentially fostering sustainable PA habits among undergraduates.

Moreover, individuals who have higher levels of comorbid state anxiety along with depression may derive fewer benefits from depression-reducing exercise interventions (Blumenthal et al., 2021), indicating that supplementary strategies targeting anxiety may be necessary. Furthermore, if vigorous physical activity (VPA) exceeds the acceptable threshold for undergraduates accustomed to prolonged sedentary behavior, it could have counterproductive effects (Zahrt and Crum, 2020). Therefore, tailored interventions and PA prescriptions that consider individual characteristics and preferences are crucial for effectively addressing mental health challenges in this demographic.

5 Summary

This study comprehensively evaluated the impact of physical activity, sleep, and sedentary behavior on subthreshold depression among undergraduates, identifying critical risk factors and demonstrating the potential of activity substitution in depression regulation. Sedentary time exceeding 12 h and a PSQI score of 3.5 or higher were associated with an increased risk of subclinical depression. However, this study has certain limitations that must be acknowledged.

First, the research was conducted within a single academic institution, which may limit the generalizability of the findings. The cut-off values and the effects of isochronous substitution may vary across different populations, so future studies should include more diverse cohorts to enhance the representativeness of the results. Second, the reliance on self-reported questionnaires rather than objective measurements may introduce potential biases, including recall bias. Therefore, future research should consider employing multicenter studies or integrating evidence-based approaches to enhance the accuracy and applicability of mental health promotion and intervention strategies within educational settings.

Data availability statement

The datasets presented in this article are not readily available because the data collection is ongoing and can only be used for studies related to depression in undergraduate students. Requests to access the datasets should be directed to 1132083877@qq.com.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Southern Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YM: Formal analysis, Funding acquisition, Methodology, Writing – original draft. YG: Writing – review & editing. HY: Investigation, Software, Writing – original draft. YZ: Investigation, Software, Writing – original draft. YK: Writing – review & editing, Methodology.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Opening Project of Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education. The Research Topic of Degree and Postgraduate Education (2020MS1008). 2023 Innovative Training Program for University Students (Guangdong province) (2023142).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

References

Acharya, L., Jin, L., and Collins, W. (2018). College life is stressful today - emerging stressors and depressive symptoms in college students. *J. Am. Coll. Health.* 66, 655–664. doi: 10.1080/07448481.2018.1451869

Alotaibi, A. S., and Boukelia, B. (2021). The effect of pre-quarantine physical activity on anxiety and depressive symptoms during the covid-19 lockdown in the Kingdom of Saudi Arabia. *Int. J. Environ. Res. Public Health* 18:7771. doi: 10.3390/ijerph18157771

Bao, J., Li, H., Huang, C., Wang, Y., Hu, X., Wang, Z., et al. (2020). Anhedonia and its relationship with empathy in subclinical depressed college students. *Chin. J. Clin. Psych.* 28, 528–532. doi: 10.16128/j.cnki.1005-3611.2020.03.020

Bhandari, P. M., Neupane, D., Rijal, S., Thapa, K., Mishra, S. R., and Poudyal, A. K. (2017). Sleep quality, internet addiction and depressive symptoms among undergraduate students in Nepal. *BMC Psychiatry* 17:106. doi: 10.1186/s12888-017-1275-5

Blumenthal, J. A., Babyak, M. A., Craighead, W. E., Davidson, J., Hinderliter, A., Hoffman, B., et al. (2021). The role of comorbid anxiety in exercise and depression trials: secondary analysis of the smile-ii randomized clinical trial. *Depress. Anxiety* 38, 124–133. doi: 10.1002/da.23088

Button, K. S., Kounali, D., Thomas, L., Wiles, N. J., Peters, T. J., Welton, N. J., et al. (2015). Minimal clinically important difference on the beck depression inventory--ii according to the patient's perspective. *Psychol. Med.* 45, 3269–3279. doi: 10.1017/S0033291715001270

Chen, L. J., Hamer, M., Lai, Y. J., Huang, B. H., Ku, P. W., and Stamatakis, E. (2022). Can physical activity eliminate the mortality risk associated with poor sleep? A 15-year follow-up of 341,248 mj cohort participants. *J. Sport Health Sci.* 11, 596–604. doi: 10.1016/j.jshs.2021.03.001

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., et al. (2013). Exercise for depression. *Cochrane Database Syst. Rev.* 2013:CD4366. doi: 10.1002/14651858.CD004366.pub6

De Faria, F. R., Barbosa, D., Howe, C. A., Canabrava, K., Sasaki, J. E., and Dos, S. A. P. (2022). Time-use movement behaviors are associated with scores of depression/anxiety among adolescents: a compositional data analysis. *PLoS One* 17:e279401. doi: 10.1371/journal.pone.0279401

Dolle, K., Schulte-Körne, G., O'Leary, A. M., von Hofacker, N., Izat, Y., and Allgaier, A. K. (2012). The beck depression inventory-ii in adolescent mental health patients: cut-off scores for detecting depression and rating severity. *Psychiatry Res.* 200, 843–848. doi: 10.1016/j.psychres.2012.05.011

Dozois, D. J. A., Dobson, K. S., and Ahnberg, J. L. (1998). A psychometric evaluation of the beck depression inventory–ii. *Psychol. Assess.* 10, 83–89. doi: 10.1037/1040-3590.10.2.83

Dumuid, D., Pedišić, Z., Stanford, T. E., Martín-Fernández, J. A., Hron, K., Maher, C. A., et al. (2019). The compositional isotemporal substitution model: a method for estimating changes in a health outcome for reallocation of time between sleep, physical activity and sedentary behaviour. *Stat. Methods Med. Res.* 28, 846–857. doi: 10.1177/0962280217737805

Fan, M., Lv, J., and He, P. (2014). Calculation method of physical activity level in the international physical activity questionnaire. *Chin. J. Epidemiol.* 35, 961–964. doi: 10.3760/cma.j.issn.0254-6450.2014.08.019

Fawzy, M., and Hamed, S. A. (2017). Prevalence of psychological stress, depression and anxiety among medical students in Egypt. *Psychiatry Res.* 255, 186–194. doi: 10.1016/j.psychres.2017.05.027

Gardani, M., Bradford, D., Russell, K., Allan, S., Beattie, L., Ellis, J. G., et al. (2022). A systematic review and meta-analysis of poor sleep, insomnia symptoms and stress in undergraduate students. *Sleep Med. Rev.* 61:101565. doi: 10.1016/j.smrv.2021.101565

Geneen, L. J., Moore, R. A., Clarke, C., Martin, D., Colvin, L. A., and Smith, B. H. (2017). Physical activity and exercise for chronic pain in adults: an overview of cochrane reviews. *Cochrane Database Syst. Rev.* 4:CD11279. doi: 10.1002/14651858.CD011279.pub3

Ghazisaeedi, M., Mahmoodi, H., Arpaci, I., Mehrdar, S., and Barzegari, S. (2022). Validity, reliability, and optimal cut-off scores of the who-5, phq-9, and phq-2 to screen depression among university students in Iran. *Int. J. Mental Health Addict.* 20, 1824–1833. doi: 10.1007/s11469-021-00483-5

Hajian-Tilaki, K. (2013). Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* 4, 627–635.

Heissel, A., Heinen, D., Brokmeier, L. L., Skarabis, N., Kangas, M., Vancampfort, D., et al. (2023). Exercise as medicine for depressive symptoms? A systematic review and meta-analysis with meta-regression. *Br. J. Sports Med.* 57, 1049–1057. doi: 10.1136/bjsports-2022-106282

Josefsson, T., Lindwall, M., and Archer, T. (2014). Physical exercise intervention in depressive disorders: meta-analysis and systematic review. *Scand. J. Med. Sci. Sports* 24, 259–272. doi: 10.1111/sms.12050

Kaya, F., Bostanci, D. N., and Durar, E. (2021). Smart phone usage, sleep quality and depression in university students. *Int. J. Soc. Psychiatry* 67, 407–414. doi: 10.1177/0020764020960207

Kim, S. Y., Jeon, S. W., Shin, D. W., Oh, K. S., Shin, Y. C., and Lim, S. W. (2018). Association between physical activity and depressive symptoms in general adult populations: an analysis of the dose-response relationship. *Psychiatry Res.* 269, 258–263. doi: 10.1016/j.psychres.2018.08.076

Ku, P. W., Steptoe, A., Liao, Y., Hsueh, M. C., and Chen, L. J. (2019). A threshold of objectively-assessed daily sedentary time for all-cause mortality in older adults: a meta-regression of prospective cohort studies. *J. Clin. Med.* 8:564. doi: 10.3390/jcm8040564

Li, H., Song, J., Jiang, J., Liu, L., Zhao, Y., and Jiang, W. (2020). Analysis of the intervention effect of brisk walking on overweight and obese occupational population based on the simultaneous replacement method. *J. Beij. Sport Univ.* 43, 111–118. doi: 10.19582/j.cnki.11-3785/g8.2020.11.012

Li, J., Wei, P., Li, W., Huang, X., Liu, Y., Yin, D., et al. (2021). Analysis of the impact of physical activity status and sleep disorders on anxiety among university students in Xuzhou. *Jiangsu J. Pre. Med.* 32, 390–392. doi: 10.13668/j.issn.1006-9070.2021.04.003

Liang, Q., Wang, Y., Lin, F., Sun, B., Chen, X., and Ma, H. (2010). Reliability and validity of the Chinese version of the seven-day physical activity recall questionnaire. *Chin. J. Rehab. Med.* 25, 1078–1081. doi: 10.3969/j.issn.1001-1242.2010.11.014

Liu, C., Zhao, Y., Qin, S., Wang, X., Jiang, Y., and Wu, W. (2021). Randomized controlled trial of acupuncture for anxiety and depression in patients with chronic insomnia. *Ann. Transl. Med.* 9:1426. doi: 10.21037/atm-21-3845

Liu, X., Tang, M., Hu, L., Wang, A., Wu, H., Zhao, G., et al. (1996). Reliability and validity study of the Pittsburgh sleep quality index. *Chin. J. Psy.* 2, 103–107.

Maddox, P. A., Elahi, A., Khuram, H., Issani, A., and Hirani, R. (2023). Sleep quality and physical activity in the management of depression and anxiety. *Prev. Med.* 171:107514. doi: 10.1016/j.ypmed.2023.107514

Nahm, F. S. (2022). Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J. Anesthesiol.* 75, 25–36. doi: 10.4097/kja.21209

Negeri, Z. F., Levis, B., Sun, Y., He, C., Krishnan, A., Wu, Y., et al. (2021). Accuracy of the patient health questionnaire-9 for screening to detect major depression: updated systematic review and individual participant data meta-analysis. *BMJ* 375:n2183. doi: 10.1136/bmj.n2183

Pasanen, J., Leskinen, T., Suorsa, K., Pulakka, A., Virta, J., Auranen, K., et al. (2022). Effects of physical activity intervention on 24-h movement behaviors: a compositional data analysis. *Sci. Rep.* 12:8712. doi: 10.1038/s41598-022-12715-2

Pearce, M., Garcia, L., Abbas, A., Strain, T., Schuch, F. B., Golubic, R., et al. (2022). Association between physical activity and risk of depression: a systematic review and meta-analysis. *JAMA Psychiatry* 79, 550–559. doi: 10.1001/jamapsychiatry.2022.0609

Rodríguez, M. R., Nuevo, R., Chatterji, S., and Ayuso-Mateos, J. L. (2012). Definitions and factors associated with subthreshold depressive conditions: a systematic review. *BMC Psychiatry* 12:181. doi: 10.1186/1471-244X-12-181

Roeh, A., Lembeck, M., Papazova, I., Pross, B., Hansbauer, M., Schoenfeld, J., et al. (2020). Marathon running improves mood and negative affect. *J. Psychiatr. Res.* 130, 254–259. doi: 10.1016/j.jpsychires.2020.08.005

Schmid, D., Ricci, C., and Leitzmann, M. F. (2015). Associations of objectively assessed physical activity and sedentary time with all-cause mortality in us adults: the nhanes study. *PLoS One* 10:e119591. doi: 10.1371/journal.pone.0119591

Schuch, F. B., Bulzing, R. A., Meyer, J., Vancampfort, D., Firth, J., Stubbs, B., et al. (2020). Associations of moderate to vigorous physical activity and sedentary behavior with depressive and anxiety symptoms in self-isolating people during the covid-19 pandemic: a cross-sectional survey in Brazil. *Psychiatry Res.* 292:113339. doi: 10.1016/j.psychres.2020.113339

Schuch, F. B., Werneck, A. O., Vancampfort, D., Stubbs, B., Teychene, M., Lotufo, P. A., et al. (2021). Cross-sectional associations of leisure and transport related physical activity with depression and anxiety. *J. Psychiatr. Res.* 140, 228–234. doi: 10.1016/j.jpsychires.2021.05.053

Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., Wincott, M., et al. (2021). Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. *J. Affect. Disord.* 287, 282–292. doi: 10.1016/j.jad.2021.03.054

Spek, V., Nyklícek, I., Cuijpers, P., and Pop, V. (2008). Predictors of outcome of group and internet-based cognitive behavior therapy. *J. Affect. Disord.* 105, 137–145. doi: 10.1016/j.jad.2007.05.001

Su, Y. (2022). Study on the relationship between 24-hour activity behavior and health-related fitness and psychological health of children in Nanjing. Doctoral dissertation, China: Shanghai University of Sport.

Takagaki, K., Okamoto, Y., Jinnin, R., Mori, A., Nishiyama, Y., Yamamura, T., et al. (2014). Behavioral characteristics of subthreshold depression. *J. Affect. Disord.* 168, 472–475. doi: 10.1016/j.jad.2014.07.018

Wang, Q., Zhang, B., Zhang, S., Wei, C., Fu, D., Zhao, H., et al. (2022). Anxiety and depression and their interdependent influencing factors among medical students in Inner Mongolia: the cross-sectional survey. *BMC Med. Educ.* 22:787. doi: 10.1186/s12909-022-03839-0

Wang, Z., Yuan, C., Huang, J., Li, Z., Chen, J., Zhang, H., et al. (2011). Reliability and validity of the Chinese version of the Beck depression inventory-II in patients with depression. *Chin. J. Mental Health* 25, 476–480. doi: 10.3969/j.issn.1000-6729.2011.06.014

Zahrt, O. H., and Crum, A. J. (2020). Effects of physical activity recommendations on mindset, behavior and perceived health. *Prev. Med. Rep.* 17:101027. doi: 10.1016/j.pmedr.2019.101027

Zhang, B., Qi, S., Liu, S., Liu, X., Wei, X., and Ming, D. (2021). Altered spontaneous neural activity in the precuneus, middle and superior frontal gyri, and hippocampus in college students with subclinical depression. *BMC Psychiatry* 21:280. doi: 10.1186/s12888-021-03292-1

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 13 November 2024
DOI 10.3389/fpsyg.2024.1466426

OPEN ACCESS

EDITED BY
Victor Zaia,
Faculdade de Medicina do ABC, Brazil

REVIEWED BY
Ricardo S. S. Durães,
Methodist University of São Paulo, Brazil
Mario Alberto Trógolo,
Siglo 21 Business University, Argentina

*CORRESPONDENCE
Jian Jiang
✉ 52667428@qq.com

RECEIVED 18 July 2024
ACCEPTED 29 October 2024
PUBLISHED 13 November 2024

CITATION
Jiang J, Chen J, Lin Z, Tang X and Hu Z (2024) Validation and psychometric properties of the Depression Anxiety Stress Scale for Youth in Chinese adolescents.
Front. Psychol. 15:1466426.
doi: 10.3389/fpsyg.2024.1466426

Validation and psychometric properties of the Depression Anxiety Stress Scale for Youth in Chinese adolescents

Jian Jiang[1,2]*, Jianhua Chen[1], Zhifeng Lin[2], Xuwei Tang[2] and Zhijian Hu[2]

[1]Fujian Provincial Hospital, Fuzhou, China, [2]Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou, China

Introduction: Depression and anxiety are the most common mental health problems among adolescents. The Depression Anxiety Stress Scales for Youth (DASS-Y) is a newly developed instrument designed to assess these problems in adolescents.

Aim: The present study aims to evaluate the psychometric properties of the DASS-Y among Chinese adolescents.

Methods: A total of 326 secondary school students aged 14–18 years participated in the study. A convenience sampling method was adopted to conduct a test–retest of the DASS-Y among Chinese secondary school students. McDonald's omega, Cronbach's alpha, and intraclass correlation coefficient (ICC) along with their 95% *CI* were used to assess the internal consistency and test–retest reliability of the DASS-Y. Confirmatory Factor Analysis (CFA) evaluated the structural validity and convergent validity of the DASS-Y through the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), as well as Average Variance Extracted (AVE) and Composite Reliability (CR). Pearson correlation coefficients with the General Health Questionnaire-12 (GHQ-12), and the Pittsburgh Sleep Quality Index (PSQI) assessed criterion validity.

Results: The CFA confirmed the validity the DASS-Y three-factor model consisting of depression, anxiety, and stress. The internal consistency reliability of the DASS-Y was found to be robust, with McDonald's omega and Cronbach's alpha values exceeding 0.8 for all dimensions across two measurements. The test–retest reliability was stable. The structural validity was reasonable and effective. Additionally, convergent validity is satisfactory, while criterion validity is also satisfactory. The three-factor model consisting of depression, anxiety and stress was confirmed through CFA.

Conclusion: The DASS-Y exhibits satisfactory psychometric properties among Chinese secondary school adolescents, reliably and appropriately screening for mental health issues such as depression and anxiety within this population. Consequently, it can be employed as a standard tool for routine mental health surveillance in secondary schools.

KEYWORDS

DASS-Y, mental health, adolescent, psychometric properties, validation

Introduction

Mental health problems in adolescents have become one of the most important public health challenges of the 21st century (Williams et al., 2020). Globally, an estimated 10–20% of adolescents experience mental health conditions (Kieling et al., 2011; Feiss et al., 2019), with negative emotional disorders such as depression and anxiety being among the most prevalent psychiatric disorders (Patton et al., 2016). According to a meta-analysis on the mental health of children and adolescents in 27 countries and regions, the global prevalence of mental health issues in children and adolescents was 13.4%, with anxiety at 6.5%, depression at 2.6%, and major depression at 1.3% (Polanczyk et al., 2015). The prevalence of anxiety and depressive symptoms among Chinese children and adolescents is increasing due to rapid economic and social development (Cui et al., 2021). The latest data of 2022 show that the overall prevalence of mental health problems among Chinese school-age children and adolescents aged 6–16 is 17.5%, of which 4.7% are anxious, 3.0% are depressed, and 2.0% are severely depressed. Moreover, there is a high degree of comorbidity between depression and anxiety (Li et al., 2022). Early screening provides important opportunities for rapid identification and intervention for mental health problems such as depression and anxiety.

At present, many mental health measurement tools can be used to assess depression and anxiety symptoms in children and adolescents, such as the Patient Health Questionnaire-9 (PHQ-9) (Kroenke, 2021), the Child Anxiety and Depression Scale (CADS) (Loades et al., 2022), the Generalized Anxiety Scale-7 item (GAD-7) (Krause et al., 2021), the Beck Depression Inventory (BDI) (Smarr and Keefer, 2011), the Beck Anxiety Inventory (BAI) (Julian, 2011), However, while these measures can assess a mixture of emotional states and general distress symptoms, they have limited capacity to distinguish between depression and anxiety (Joiner et al., 1996). To compensate for these limitations, Lovibond P. F. and Lovibond (1995) and Lovibond S. H. and Lovibond (1995) developed the Depression Anxiety Stress Scales (DASS) and its short version DASS-21. Especially DASS-21, which was widely welcomed for its brevity, ease of completion and scoring (Antony et al., 1998). The DASS-21 is a simplified version of the DASS with a more streamlined test with 21 questions that screen for depression (7 questions), anxiety (7 questions), and stress (7 questions). Subjects are asked to use a 4-point combined severity/frequency scale to rate the extent to which they have experienced each item over the past week. The scale ranges from 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time). Scores for Depression, Anxiety, and Stress are calculated by summing the scores for the relevant items and multiplying by two. The DASS-21 has been extensively validated and utilized in different cultures, domains, settings, and populations around the world due to its acceptable validity, reliability, simplicity, and convenience, internal consistency coefficient between 0.73 ~ 0.91 (Ronk et al., 2013; Vignola and Tucci, 2014; Park et al., 2020). DASS-21 has also been widely applied in China, showing excellent psychometrics in different populations across the country, and Cronbach's a value ranged from 0.71 ~ 0.96 (Wang et al., 2016; Zou et al., 2018; Yu et al., 2023).

Since DASS-21 was originally developed with adults over the age of 18 years as the target population (Osman et al., 2012), it might not be appropriate to use it in adolescents under age of 18 years. Earlier studies have reported that the adult model DASS-21 cannot accurately distinguish the symptoms of depression and anxiety among adolescents in some cultural contexts (Shaw et al., 2017; Le et al., 2017). In addition, DASS-21 contains several expressions and words that may not be familiar to adolescents (Szabó, 2010), for example, questions such as "I felt that life was meaningless" and "I felt I wasn't worth much as a person" may not be able to be accurately self-reported by Chinese adolescents, who are culturally introverted and subtle with an oriental cultural background. Szabo and Lovibond (2022) developed a screening tool based on DASS and DASS-21 for depression and anxiety symptoms specifically for children and adolescents under the age of 18: the Depression Anxiety Stress Scale for Youth (DASS-Y). The DASS-Y has comprising simplified wording, appropriate terminology, and has been validated and utilized in Australia (Mackenzie et al., 2023) and Serbian (Jovanović, 2024), showing good internal consistency (a = 0.86–0.94). However, these studies have shown that there exist differences in the factor structure and stability of the DASS-Y across ethnic populations.

The DASS-Y is specifically designed to measure emotional disorders such as depression and anxiety among adolescents, and assess a subject's mental health status through specific scores, making it simpler and more practical than the current mental health measurement scales that test depression or anxiety singularly. However, there remains limited evidence regarding the application of DASS-Y among Chinese adolescent students in secondary schools, especially in terms of stability across time. Considering the differences in the expression of emotions, mental health concepts in Chinese and Western cultures and the expectations and pressures of Chinese society as well as the localization of the terminology of the survey items for processing and adaptation, the DASS-Y urgently needs to be validated in the Chinese adolescent population to confirm its applicability for screening depression and anxiety in the Chinese adolescent population. Therefore, the present study aimed to evaluate psychometric properties of DASS-Y in school adolescents in China.

Methods

Participants

A convenient random sampling method was used to select one secondary school from each of the status locations (urban and rural). Subsequently, within each selected school, one grade level was randomly chosen, and all students from the designated grade participated in the study.

We used the "Confidence Intervals for Coefficient Alpha" module of PASS 21 (NCSS, Ketchum) to determine the necessary sample size for the study. We estimate coefficient alpha with a two-sided 95% confidence interval with a width no wider than 0.1 and set examine values of $K = 21$ (scale items). From past studies, we want to use a planning estimate of 0.8 for the sample coefficient alpha. The necessary sample size for the study was determined to be 131. Considering the 10% rejection and withdrawal ratio, a total of 146 high school students were needed for this study.

No exclusion criteria were set and all students who were enrolled in the school and provided written informed consent from themselves and their parents/guardians were included in the study. Ultimately, we recruited a total of 326 students, including students from grade 9 in urban secondary schools and students from grade 10 in rural

10.3389/fpsyg.2024.1466426

secondary schools, of which 163 (50%) were girls and 139 (42.6%) were in 9th grade. The mean age of the participants was 15.72 years (SD = 0.82, range: 14–18 years), the participating students in this study met the research design requirements.

Measures

Baseline demographic characteristics

A self-administered questionnaire was used to collect basic demographic information such as gender, age (in years), grade and place of residence of the participants.

Depression Anxiety Stress Scale for Youth (DASS-Y)

The DASS-Y is a 21-tiems self-report measure containing three subscales of depression (7 items), anxiety (7 items), and stress (7 items). Subjects responded to each item by rating the frequency and severity of symptoms over the past week on a four-point Likert scale from 0 (not true) to 3 (very true). Each subscale was scored 0–21 and the total scale was scored 0–63. Higher scores indicate higher levels of symptoms in each domain and all items can also be added into a total score as a measure of general psychological distress. It is worth mentioning that, according to the use manual, the DASS-Y does not need to multiply the scores by 2 as per the DASS-21, because the DASS-Y is a standalone instrument with scores that cannot be directly compared to either of the adult questionnaires. Based on the scoring, the results are categorized into five levels: normal, mild, moderate, severe, and extremely severe (Supplementary Table S1).

Prior to this, as there is no Chinese version of the DASS-Y, the research group translated the DASS-Y into Chinese. First, the English version of the DASS-Y was translated into Chinese by two independent translators with backgrounds in studying and working in English-speaking countries, and then discussions were organized with mental health experts, epidemiologists, secondary school headmasters, class teachers and school counselors. Finally, the discussed version was translated back into English by the two researchers with bilingual backgrounds to check and reviewed by native English-speaking researchers to determine consistency with the original version.

The final Chinese version of DASS-Y was preliminarily tested among a small group of secondary school students ($n = 10$). No comprehension difficulties or ambiguities were found (data not shown in this study). The original DASS-Y can be obtained without charge from a website[1], and the final Chinese version of the questionnaire is presented in the Appendix.

General Health Questionnaire-12 (GHQ-12)

The GHQ-12 is a self-report mental health status assessment tool. It evaluates the condition of the participant's mental health in the past few weeks in terms of social relationships, physical pain, emotional state, sleep, anxiety, stress, and general feelings of well-being dimensions. The questionnaire has 12 items, of which 6 questions are positively scored and 6 questions are negatively scored, with 4 options, namely: "better than usual, same as usual, worse than usual, much worse than usual," and scored from 0 to 3 according to the Likert 3-point scale, with a total of 0–36. A score greater than 3 indicates psychological distress, with higher scores indicating more serious mental health problems and can be classified as normal and abnormal based on the score (Shek, 1987). As the GHQ-12 contains self-ratings of the dimensions of general health experience, anxiety, and stress. The GHQ-12 has been widely validated and used (Guan and Han, 2019), with internal consistency Cronbach's a of 0.721–0.903. We used the Zhang et al. (2008) supplemented GHQ-12, and its internal consistency Cronbach's a in our test–retests was 0.853 and 0.872.

Pittsburgh Sleep Quality Index (PSQI)

The PSQI consists of a 19-item self-report scale that assesses sleep quality and sleep disorders in participants over the past few weeks. These items were used to calculate seven dimensions corresponding to specific sleep domains: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disorders, use of sleep aids and daytime dysfunction. Then, these dimensional scores were summed in order to generate a single PSQI score representing overall sleep quality, with a score of more than 5 indicating the presence of a sleep disorder, and classified into four levels of very good, good, poor, and very poor based on the scores (Buysse et al., 1989). The PSQI has been widely used and validated (Raniti et al., 2018) with internal consistency Cronbach's a of 0.771–0.953. This study used the revised Chinese version of Liu et al. (1996), their internal consistency Cronbach's a was 0.853 and 0.872 in our test–retest, respectively. As sleep has been shown to be highly correlated with emotional disorders such as depression and anxiety (Beaudequin et al., 2020), the criterion-related validity of DASS-Y was validated using the PSQI scale.

Procedure

Data collection lasted for 2 months from March, 2022 to May, 2022. The ethical aspect was ensured by providing information to the adolescents and information sheet together with the consent form was sent to the family to obtain the consent of the family days before data collection.

Questionnaires were completed on school premises in a pen-and-paper written format. Following the distribution of the questionnaires, teachers read a standardized set of instructions asking the participants to read each item and select the most appropriate answer. Participants were reminded that the procedure was entirely voluntary, their responses were confidential, and there were no right or wrong answers. All students completed the same questionnaire for two times: enrollment (T_1), and 5 weeks after enrollment (T_2).

The study procedures were approved by the Ethics Committee of Fujian Provincial Hospital (K-2023-03-005).

Statistical analysis

All statistical analyses were performed in SPSS (version 26, SPSS Inc., Chicago, IL) except the confirmatory factor analysis (CFA), which was conducted using AMOS (version 26). The internal consistency reliability of the DASS-Y was examined by calculating the McDonald's omega (ω), Cronbach's alpha, and intra-class correlation coefficient (ICC) with their 95% confidence intervals (95% CI) for the

1 https://dass.psy.unsw.edu.au/DASSY.htm

test–retest data. A split-half coefficient ranging from 0.70 to 0.80 indicates good internal consistency, between 0.60 and 0.70 is considered acceptable (Campbell, 1991). The ICC was calculated using the absolute agreement type of a two-way mixed model, with ICC > 0.75 indicating good reliability (Shrout and Fleiss, 1979). A Cronbach's alpha >0.70 suggests good internal consistency and reliability, while >0.90 indicates very high internal consistency and excellent reliability (Olden et al., 2009).

The fit of the DASS-Y data was assessed through confirmatory factor analysis (CFA) using Maximum Likelihood (ML) estimation, in order to determine the appropriate fitting model for DASS-Y. The model's fit was determined by various fit indices. Calculations included chi-square (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA) with its 90% confidence interval (*CI*). Although χ^2 was reported, they were not used as a criterion for model fit evaluation due to their sensitivity to smaller sample sizes (Bentler, 2007). The model fit was assessed using the following criteria: (1) Adequate fit: RMSEA <0.08, CFI and TLI > 0.90; (2) Good fit: RMSEA <0.05, CFI and TLI > 0.95 (Bentler, 1990). Furthermore, a final fitted model diagram of DASS-Y was plotted to further illustrate the standardized factor loadings among the various items of DASS-Y.

Finally, the convergent validity of DASS-Y was assessed by calculating the Average Variance Extracted (AVE) and Composite Reliability (CR) indices. Pearson correlation coefficients (*r*) were computed between DASS-Y, GHQ-12, and PSQI to evaluate the criterion validity of DASS-Y. AVE > 0.50 and CR > 0.70 indicate good validity, AVE > 0.4 and CR > 0.80 suggest fair validity of the scale (Fornell and Larcker, 1981).

Results

The factorial structure of the DASS-Y

For DASS-Y, the initial three-factor structural model was tested using CFA. As shown in Table 1, in the test–retest analysis, the CFI values were 0.903 and 0.890, the TLI values were 0.890 and 0.875, and the RMSEA values (90% *CI*) were 0.072 (0.064 ~ 0.080) and 0.078 (0.071 ~ 0.084), respectively. The fit indices show that the original DASS-Y three-factor model does not fully meet the criteria.

Similarly, as illustrated in Figure 1, the original three-factor oblique model depicts the factor loadings among the DASS-21 and its subscales. In the test–retest, items within each subscale exhibited satisfactory factor loadings (≥0.50), with robust factor loadings were observed between the three subscales (0.85 and 0.98), except for T_2 for stress and anxiety. An anomalous value was observed between the Stress subscale and the Anxiety subscale in the T2 test (1.01).

TABLE 1 DASS-Y original three-factor model fit indices.

Time	N	χ^2	DF	CFI	TLI	RMSEA	RMSEA 90% *CI*
T_1	326	499.004	186	0.903	0.890	0.072	0.064, 0.080
T_2	326	513.782	186	0.890	0.875	0.078	0.071, 0.092

The internal consistency, test–retest reliability, and subscales-total correlation of the DASS-Y

As shown in Table 2, the McDonald's omega (ω) of the DASS-Y scale across the test–retest were 0.935 and 0.948, respectively, with 0.853 and 0.871 for the depression subscale, 0.828 and 0.854 for the anxiety subscale, and 0.844 and 0.863 for the stress subscale. The Cronbach's alpha was 0.935 and 0.948 for the total scale, with 0.853 and 0.871 for the depression subscale, 0.827 and 0.854 for the anxiety subscale, and 0.840 and 0.863 for the stress subscale. All subscales exhibited satisfactory domain-total correlations ($r > 0.90$). In addition to this, across the two measurements, the DASS-Y demonstrated good test–retest reliability, ICC = 0.901 (95% CI: 0.877–0.921), and the measurement results of each subscale were highly correlated (ICC > 0.85).

The convergent validity and criterion validity of the DASS-Y

Based on the Confirmatory Factor Analysis (CFA) model, the Average Variance Extracted (AVE) and Composite Reliability (CR) indices were calculated for each DASS-Y item on its corresponding subscale. In the test–retest, the AVE for each item on the depression subscale was 0.465 and 0.500, CR = 0.858 and 0.874, the AVE for each item on the anxiety subscale was 0.420 and 0.461, CR = 0.833 and 0.856, and the AVE for each item on the stress subscale was 0.439 and 0.476, CR = 0.842 and 0.864 (Supplementary Table S3).

As expected, the DASS-Y showed strong correlations with the GHQ-12, with Pearson correlation coefficients of 0.789 and 0.710 ($p < 0.001$) on test–retest, respectively. Depression ($r = 0.721$ and 0.694, $p < 0.001$), anxiety ($r = 0.756$ and 0.642, $p < 0.001$) and stress ($r = 0.779$ and 0.683, $p < 0.001$) were also strongly correlated with the GHQ-12. Meanwhile, the DASS-Y was also moderately correlated with the PSQI, the Pearson correlation coefficients of the DASS-Y and PSQI were 0.68 and 0.602 ($p < 0.001$) in both tests. Depression ($r = 0.609$ and 0.559, $p < 0.001$), anxiety ($r = 0.578$ and 0.575, $p < 0.001$) and stress ($r = 0.639$ and 0.577, $p < 0.001$) were also moderately correlated with the PSQI (Table 3).

Discussion

Emotional disorders such as depression and anxiety are common mental health problems in adolescents. Early screening offers an important opportunity for prompt diagnosis and timely intervention. To the best of our knowledge, this is one of two studies of DASS-Y psychometric properties in Chinese adolescent populations.

Prior to this, only Cao et al. (2023) conducted a DASS-Y study in a population of Chinese children and adolescents. In contrast to their work, we examined the internal consistency reliability, test–retest reliability, factor structure, convergent validity, and criterion validity of DASS-Y by calculating McDonald's Omega, Cronbach's Alpha, and Intraclass Correlation Coefficient (ICC) between two measurements. These indicators align with the findings reported by Cao et al. (2023), suggested that the DASS-Y exhibited high internal consistency, stable

10.3389/fpsyg.2024.1466426

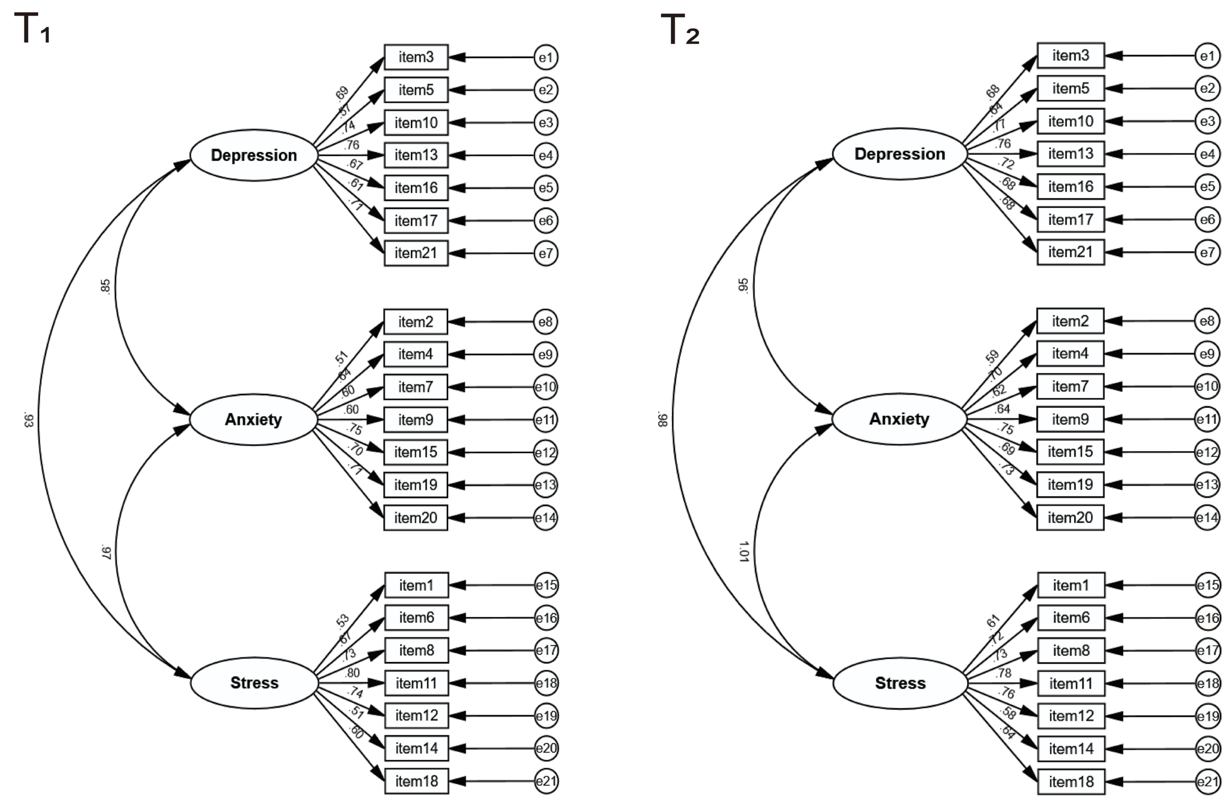

FIGURE 1
Path diagram of the 3-factor oblique DASS-Y model, with standardized parameter estimates. The lines indicate the three subscales of the DASS-Y depression, anxiety and stress and the factor structure and associations between the respective entries. The relationships between all pathways are presented as standardized pathway coefficients.

TABLE 2 Correlation coefficients for the reliability of DASS-Y and its subscales.

		ω	Cronbach's α	Domain-total correlation	ICC (95% CI)
Depression	T_1	0.853	0.853	0.910	0.897 (0.872 ~ 0.917)
	T_2	0.871	0.871	0.928	
Anxiety	T_1	0.828	0.827	0.923	0.872 (0.841 ~ 0.897)
	T_2	0.854	0.854	0.948	
Stress	T_1	0.844	0.840	0.943	0.855 (0.857 ~ 0.907)
	T_2	0.863	0.863	0.950	
DASS-Y	T_1	0.935	0.935		0.901 (0.877 ~ 0.921)
	T_2	0.948	0.948		

test–retest reliability, and cross-temporal stability in the Chinese adolescent population.

We validated the factor structure and the original three-factor model of the DASS-Y. In terms of factor structure and factor loadings, the internal components of each subscale exhibit good factor loadings, similar to the findings of studies by Szabo and Lovibond (2022), Mackenzie et al. (2023), and Jovanović (2024), indicating that each dimension can adequately reflect symptoms of depression, anxiety, and stress. However, unlike them, there is an overlap of the stress subscales, which may be related to our cultural traditions, where stress is not a psychological problem but more of an emotional expression or, like anxiety, a worry about the future.

In terms of model fit, the original three-factor structural model did not fully conform to the established criteria, which is similar to previous findings (Cao et al., 2023). Hu and Bentler (1999) argue that all fit indices are intercept values assessed based on prior models, and their results may vary under different conditions, and that, in the real world, it is necessary to also take into account factors such as the sensitivity of the responses to the test items and small sample bias, among other factors. Therefore, this may be related to the source of our sample size, where students in different regions (urban and rural) may have inconsistent understanding of emotional disorders such as depression and anxiety, leading to bias in self-reported DASS-Y, which also affects the factor structure and construct validity of our DASS-Y. Considering that DASS-Y is a new

TABLE 3 Correlations between DASS-Y and GHQ-12, and PSQI.

		Depression	Anxiety	Stress	DASS-Y
T_1	GHQ-12	0.721**	0.756**	0.779**	0.789**
	PSQI	0.609**	0.578**	0.639**	0.668**
T_2	GHQ-12	0.694**	0.642**	0.683**	0.710**
	PSQI	0.559**	0.575**	0.577**	0.602**

** $p < 0.001$.

scale with limited evidence in the Chinese population, we did not jump to conclusions after weighing the sample size ($n \leq 500$), the study context, and the purpose of the study, and further validation is needed with larger and more diverse populations in the future.

In the analysis of the convergent validity of the DASS-Y, it was observed that only the AVE of the depression subscale of T_2 reached a satisfactory level of 0.500. However, according to Fornell and Larcker (1981), CR greater than 0.6 is sufficient to establish convergent validity, even if the AVE falls below 0.5. Upon conducting tests and retests, it was found that the CR values of all subscales exceeded 0.8. Therefore, the research maintains that the DASS-Y scale possesses acceptable convergent validity and good combinatorial reliability in assessing mental health issues among Chinese adolescents. These findings further corroborate the robust convergent validity and stable test structure of the DASS-Y.

As shown in the results, the DASS-Y and the GHQ-12 exhibited strong correlations. Both instruments include core dimensions that measure negative emotional states, such as depression, anxiety, and stress, and assess mental health status through a total score. These instruments effectively identify the core symptoms of depression and anxiety, which are crucial for assessing the severity of mental health problem (Li et al., 2009). Additionally, the DASS-Y was found to have a moderate correlation with the PSQI. Previous research has demonstrated that negative emotions, including depression and anxiety, impact sleep and are strongly associated with sleep quality (Xue et al., 2022). These results confirm that the DASS-Y serves as an effective tool for screening mental health problem such as depression and anxiety among adolescents.

As a convenient, brief, and non-professionally administered instrument for assessing mental health problems, this study demonstrated that Chinese adolescent students could effectively comprehend the item content and verbal presentation of the DASS-Y. Furthermore, laypersons were able to organize and complete the questionnaire with greater ease during the testing process. Notably, the questionnaire's completion time, averaging 10 min, fell within the acceptable range for the participating students. Consequently, the DASS-Y provides an alternative testing tool for screening mental health problems among Chinese children and adolescents, thereby enriching the assessment toolbox available for this population. These findings may suggest that the DASS-Y should be included as a regular part of the mental health examination in regular student health screening in schools to routinely monitor the mental health status of students.

Limitations and future research

It is noteworthy that the present study also has some limitations. Firstly, the study specifically recruited students from grades 9 and 10, encompassing the junior high school and high school learning stages in China, with ages ranging from 14 to 18, however, this may limit the generalization of the findings to other age groups. Second, our study was conducted in only one region of China and lacked comparable data. Future studies could further enrich the dataset for cross-population, cross-region and cross-country comparisons. Thirdly, the DASS-Y is a self-report scale that reflects individual's situation during the past week. Given that negative emotional problems vary over time and according to personal feelings, and that there is as yet universally established standard for the time interval between repeated measurements, multiple sources of reporting should be adopted in the future, including parent-reported scales, teacher-reported scales, and structured interviews. Fourth, sample geographic, cultural, and linguistic contexts affect factor constructs and modeling, and future DASS-Y studies should be conducted in different regions, populations, and cultural contexts to further validate the applicability of the DASS-Y in the Chinese adolescent population. Finally, we did not discuss the discriminant validity of DASS-Y, despite its significance, as it was not the primary objective of this study, however, further in-depth discussion should be pursued in future research.

Conclusion

This study investigated the psychometric properties of the DASS-Y among Chinese adolescents. The results indicate that the DASS-Y demonstrates good internal consistency, test–retest stability, and correlational validity, has a reasonable factor structure, and can effectively differentiate depressive, anxiety, and stress symptoms among Chinese adolescents. These findings suggest that the DASS-Y can be a useful self-report measure for assessing common mental health problems (depression, anxiety) among Chinese secondary school students. We suggest that frontline teachers and school mental health counselors promote and utilize this valuable tool as an effective means of screening and assessing students' mental health issues.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by The Ethics Committee of Fujian Provincial Hospital (K-2023-03-005). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

JJ: Writing – review & editing, Writing – original draft, Resources, Project administration, Data curation, Conceptualization. JC:

Investigation, Writing – original draft, Methodology. ZL: Writing – original draft, Methodology, Formal analysis. XT: Writing – original draft, Methodology, Formal analysis. ZH: Writing – review & editing, Supervision, Project administration, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Scientific and Technological Innovation Joint Capital Projects of Fujian Province (Nos. 2020Y9018 and 2022R0088); the Natural Science Foundation of Fujian Province (Nos 2021J01726 and 2021J01733); the Central Government-Led Local Science and Technology Development Special Project (Nos. 2020L3009 and 2019L3006); and the Fujian Provincial Medical and Health System Reform Research Association Health Policy Innovation Research Project (No. 2022B07), China.

Acknowledgments

The authors would like to thank the schools, teachers, students and parents who participated in this study for their help and support and the students for completing the survey in a complete, honest and detailed manner.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1466426/full#supplementary-material

References

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., and Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychol. Assess.* 10, 176–181. doi: 10.1037/1040-3590.10.2.176

Beaudequin, D., Schwenn, P., McLoughlin, L. T., Parker, M. J., Broadhouse, K., Simcock, G., et al. (2020). Using measures of intrinsic homeostasis and extrinsic modulation to evaluate mental health in adolescents: preliminary results from the longitudinal adolescent brain study (LABS). *Psychiatry Res.* 285:112848. doi: 10.1016/j.psychres.2020.112848

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull.* 107, 238–246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Pers. Individ. Dif.* 42, 825–829. doi: 10.1016/j.paid.2006.09.024

Buysse, D. J., Reynolds, C. F. 3rd., Monk, T. H., Berman, S. R., and Kupfer, D. J. (1989). The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213. doi: 10.1016/0165-1781(89)90047-4

Campbell, M. J. (1991). Health measurement scales. A practical guide to their development and use. David L. Streiner, Geoffrey R. Norman, Oxford university press, Oxford, 1989. No. of pages: vii + 175. Price: £25 hardback, £15 paperback. *Stat. Med.* 10, 1478–1479. doi: 10.1002/sim.4780100918

Cao, C. H., Liao, X. L., Gamble, J. H., Li, L. L., Jiang, X. Y., Li, X. D., et al. (2023). Evaluating the psychometric properties of the Chinese depression anxiety stress scale for youth (DASS-Y) and DASS-21. *Child Adolesc. Psychiatry Ment. Health* 17:106. doi: 10.1186/s13034-023-00655-2

Cui, Y., Li, F., Leckman, J. F., Guo, L., Ke, X., Liu, J., et al. (2021). The prevalence of behavioral and emotional problems among Chinese school children and adolescents aged 6-16: a national survey. *Eur. Child Adolesc. Psychiatry* 30, 233–241. doi: 10.1007/s00787-020-01507-6

Feiss, R., Dolinger, S. B., Merritt, M., Reiche, E., Martin, K., Yanes, J. A., et al. (2019). A systematic review and meta-analysis of school-based stress, anxiety, and depression prevention programs for adolescents. *J. Youth Adolesc.* 48, 1668–1685. doi: 10.1007/s10964-019-01085-0

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.1177/002224378101800104

Guan, M., and Han, B. (2019). Factor structures of general health Questionnaire-12 within the number of Kins among the rural residents in China. *Front. Psychol.* 10:1774. doi: 10.3389/fpsyg.2019.01774

Hu, L.-t., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Joiner, T. E., Catanzaro, S. J., and Laurent, J. (1996). Tripartite structure of positive and negative affect, depression, and anxiety in child and adolescent psychiatric inpatients. *J. Abnorm. Psychol.* 105, 401–409. doi: 10.1037//0021-843x.105.3.401

Jovanović, V. (2024). The depression anxiety stress scales for youth (DASS-Y): evidence of validity and cross-national and gender measurement invariance. *Eur. Child Adolesc. Psychiatry*. doi: 10.1007/s00787-024-02561-0 Epub ahead of print.

Julian, L. J. (2011). Measures of anxiety: state-trait anxiety inventory (STAI), Beck anxiety inventory (BAI), and hospital anxiety and depression scale-anxiety (HADS-A). *Arthritis Care Res* 11, S467–S472. doi: 10.1002/acr.20561

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., et al. (2011). Child and adolescent mental health worldwide: evidence for action. *Lancet* 378, 1515–1525. doi: 10.1016/S0140-6736(11)60827-1

Krause, K. R., Chung, S., Adewuya, A. O., Albano, A. M., Babins-Wagner, R., Birkinshaw, L., et al. (2021). International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder. *Lancet Psychiatry* 8, 76–86. doi: 10.1016/S2215-0366(20)30356-4

Kroenke, K. (2021). PHQ-9: global uptake of a depression scale. *World Psychiatry* 20, 135–136. doi: 10.1002/wps.20821

Le, M. T. H., Tran, T. D., Holton, S., Nguyen, H. T., Wolfe, R., and Fisher, J. (2017). Reliability, convergent validity and factor structure of the DASS-21 in a sample of Vietnamese adolescents. *PLoS One* 12:e0180557. doi: 10.1371/journal.pone.0180557

Li, W. H., Chung, J. O., Chui, M. M., and Chan, P. S. (2009). Factorial structure of the Chinese version of the 12-item general health questionnaire in adolescents. *J. Clin. Nurs.* 18, 3253–3261. doi: 10.1111/j.1365-2702.2009.02905.x

Li, F., Cui, Y., Li, Y., Guo, L., Ke, X., Liu, J., et al. (2022). Prevalence of mental disorders in school children and adolescents in China: diagnostic data from detailed clinical assessments of 17,524 individuals. *J. Child Psychol. Psychiatry* 63, 34–46. doi: 10.1111/jcpp.13445

Liu, X. C., Tang, M. Q., Hu, L., Wang, A. Z., Wu, H. X., Zhao, G. F., et al. (1996). Reliability and validity of the Pittsburgh sleep quality index. *Chin J Psychiatry.* 29, 103–107.

Loades, M. E., Stallard, P., Kessler, D., and Crawley, E. (2022). Mental health screening in adolescents with CFS/ME. *Eur. Child Adolesc. Psychiatry* 31, 1003–1005. doi: 10.1007/s00787-021-01734-5

Lovibond, P. F., and Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behav. Res. Ther.* 33, 335–343. doi: 10.1016/0005-7967(94)00075-u

Lovibond, S. H., and Lovibond, P. F. (1995). Manual for the depression anxiety stress scales. *2nd* Edn. Sydney: Psychology Foundation.

Mackenzie, E., McMaugh, A., Van Bergen, P., and Parada, R. H. (2023). Online support seeking, co-rumination, and mental health in adolescent girls. *Front. Psych.* 14:1040636. doi: 10.3389/fpsyt.2023.1040636

Olden, M., Rosenfeld, B., Pessin, H., and Breitbart, W. (2009). Measuring depression at the end of life: is the Hamilton depression rating scale a valid instrument? *Assessment* 16, 43–54. doi: 10.1177/1073191108320415

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., and Lozano, G. (2012). The depression anxiety stress Scales-21 (DASS-21): further examination of dimensions, scale reliability, and correlates. *J. Clin. Psychol.* 68, 1322–1338. doi: 10.1002/jclp.21908

Park, S. H., Song, Y. J. C., Demetriou, E. A., Pepper, K. L., Thomas, E. E., Hickie, I. B., et al. (2020). Validation of the 21-item depression, anxiety, and stress scales (DASS-21) in individuals with autism spectrum disorder. *Psychiatry Res.* 291:113300. doi: 10.1016/j.psychres.2020.113300

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., et al. (2016). Our future: a lancet commission on adolescent health and wellbeing. *Lancet* 387, 2423–2478. doi: 10.1016/S0140-6736(16)00579-1

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., and Rohde, L. A. (2015). Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J. Child Psychol. Psychiatry* 56, 345–365. doi: 10.1111/jcpp.12381

Raniti, M. B., Waloszek, J. M., Schwartz, O., Allen, N. B., and Trinder, J. (2018). Factor structure and psychometric properties of the Pittsburgh sleep quality index in community-based adolescents. *Sleep* 41:66. doi: 10.1093/sleep/zsy066

Ronk, F. R., Korman, J. R., Hooke, G. R., and Page, A. C. (2013). Assessing clinical significance of treatment outcomes using the DASS-21. *Psychol. Assess.* 25, 1103–1110. doi: 10.1037/a0033100

Shaw, T., Campbell, M. A., Runions, K. C., and Zubrick, S. R. (2017). Properties of the DASS-21 in an Australian community adolescent population. *J. Clin. Psychol.* 73, 879–892. doi: 10.1002/jclp.22376

Shek, D. T. (1987). Reliability and factorial structure of the Chinese version of the general health questionnaire. *J. Clin. Psychol.* 43, 683–691. doi: 10.1002/1097-4679(198711)43:6<683::AID-JCLP2270430606>3.0.CO;2-B

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420

Smarr, K. L., and Keefer, A. L. (2011). Measures of depression and depressive symptoms: Beck depression inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), geriatric depression scale (GDS), hospital anxiety and depression scale (HADS), and patient health Questionnaire-9 (PHQ-9). *Arthritis Care Res.* 63, S454–S466. doi: 10.1002/acr.20556

Szabó, M. (2010). The short version of the depression anxiety stress scales (DASS-21): factor structure in a young adolescent sample. *J. Adolesc.* 33, 1–8. doi: 10.1016/j.adolescence.2009.05.014

Szabo, M., and Lovibond, P. F. (2022). Development and psychometric properties of the DASS-youth (DASS-Y): an extension of the depression anxiety stress scales (DASS) to adolescents and children. *Front. Psychol.* 13:766890. doi: 10.3389/fpsyg.2022.766890

Vignola, R. C., and Tucci, A. M. (2014). Adaptation and validation of the depression, anxiety and stress scale (DASS) to Brazilian Portuguese. *J. Affect. Disord.* 155, 104–109. doi: 10.1016/j.jad.2013.10.031

Wang, K., Shi, H. S., Geng, F. L., Zou, L. Q., Tan, S. P., Wang, Y., et al. (2016). Cross-cultural validation of the depression anxiety stress Scale-21 in China. *Psychol. Assess.* 28, e88–e100. doi: 10.1037/pas0000207

Williams, I., Vaisey, A., Patton, G., and Sanci, L. (2020). The effectiveness, feasibility and scalability of the school platform in adolescent mental healthcare. *Curr. Opin. Psychiatry* 33, 391–396. doi: 10.1097/YCO.0000000000000619

Xue, B., Xue, Y., Zheng, X., Shi, L., Liang, P., Dong, F., et al. (2022). Association of sleep with mental health in Chinese high school students: a cross-sectional study. *J. Sleep Res.* 31:e13697. doi: 10.1111/jsr.13697

Yu, Y., Liu, J., Skokauskas, N., Liu, F., Zhang, L., Teng, T., et al. (2023). Prevalence of depression and anxiety, and associated factors, among Chinese primary and high school students: a cross-sectional, epidemiological study. *Asia Pac. Psychiatry* 15:e12523. doi: 10.1111/appy.12523

Zhang, Y., Cui, L. J., Li, K. Q., Jiang, Q. P., Sun, X. L., Gao, L. H., et al. (2008). Supplemented edition of general health questionnaire (GHQ-12) in epidemiological survey of mental health. *Chin. Ment. Health J.* 22, 189–192. doi: 10.3321/j.issn:1000-6729.2008.03.009

Zou, P., Sun, L., Yang, W., Zeng, Y., Chen, Q., Yang, H., et al. (2018). Associations between negative life events and anxiety, depressive, and stress symptoms: a cross-sectional study among Chinese male senior college students. *Psychiatry Res.* 270, 26–33. doi: 10.1016/j.psychres.2018.09.019

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 November 2024
DOI 10.3389/fpsyg.2024.1441794

Validation of the schema coping inventory for dysfunctional coping strategies

Simone Gazzellini[1]*, Valerio Pellegrini[2], Eleonora Napoli[1], Vanessa Ventre[3], Donatella Lettori[1], Enrico Castelli[1], Barbara Basile[3] and Mauro Giacomantonio[2]

[1]Neuroscience Clinic Area, Neurorehabilitation Research Area, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy, [2]Department of Social and Developmental Psychology, 'Sapienza' University of Rome, Rome, Italy, [3]School of Cognitive Psychotherapy (SPC), Rome, Italy

Check for updates

OPEN ACCESS

EDITED BY
Victor Zaia,
Faculdade de Medicina do ABC, Brazil

REVIEWED BY
Cristian Ramos-Vera,
Cesar Vallejo University, Peru
George Tsouvelas,
National and Kapodistrian University of Athens, Greece

*CORRESPONDENCE
Simone Gazzellini
✉ simone.gazzellini@opbg.net

RECEIVED 31 May 2024
ACCEPTED 30 October 2024
PUBLISHED 27 November 2024

CITATION
Gazzellini S, Pellegrini V, Napoli E, Ventre V, Lettori D, Castelli E, Basile B and Giacomantonio M (2024) Validation of the schema coping inventory for dysfunctional coping strategies.
Front. Psychol. 15:1441794.
doi: 10.3389/fpsyg.2024.1441794

Introduction: According to the theoretical model of Schema Therapy, each human being has basic needs that require natural satisfaction from childhood. When these emotional needs are frustrated, early maladaptive schemas (EMS) develop, leading to coping styles that are strategies to manage the pain caused by activated EMS. The study validates and standardizes the Schema Coping Inventory in the Italian population and evaluates correlations between psychological variables and the SCI.

Methods: We analysed data from a community sample of 602 Italian adults, aged between 18 and 69 years, who endorsed a structured questionnaire, involving demographic information, the Italian version of the SCI, and an array of theoretically related psychological constructs.

Results and discussion: Confirmative factor analysis corroborated the tridimensional structure of the SCI (Surrender, Avoidance, Overcompensation), both in terms of the overall goodness-of-fit of the model and the single items' factor loading in the corresponding factors. The internal consistency turned out to be satisfactory. Construct validity was assessed through convergent and divergent (positive and negative) correlations with other coping style measures and psychopathological scales. Mean values and mean standard deviations are reported for the general population, for psychopathological clinical and non-pathological samples.

KEYWORDS

schema coping inventory, standardization, validation, maladaptive coping strategies, schema therapy

Introduction

Schema therapy (ST, Young et al., 2003) is an integrative therapeutic approach that combines elements of Gestalt and attachment theories alongside emotion-focused and cognitive-behavioral strategies. The main goal of ST is to identify clients' unmet core needs and help them fulfill them in healthier, more functional ways. When emotional core needs are frustrated in early childhood or adolescence (resulting from situations such as abuse, neglect, or dysfunctional parenting), early maladaptive schemas (EMSs) can develop. These schemas are associated with coping strategies that were initially adaptive for managing difficult or threatening situations during childhood but may become maladaptive in adulthood, perpetuating the schemas.

Patients adopt these coping strategies as a means of survival in threatening situations during childhood, where such responses may have been the most effective way to manage those circumstances. However, these behaviors often become maladaptive in the patient's current life, serving to reinforce the schemas.

A recent meta-analysis of 33 studies (Pilkington et al., 2021) supports the theory that childhood adversity is associated with the development of EMSs in adulthood. It further indicates that individuals who recall a lack of maternal warmth or nurturing often anticipate that their emotional needs will remain unmet. Additionally, the studies by Tsouvelas et al. (2023a, 2023b), which examined children in residential care in Greece, demonstrated that schemas related to disconnection/rejection and impaired autonomy/performance domains are predictive indicators of psychopathology. The authors concluded that assessing EMSs in children is crucial to prevent the establishment of psychopathological conditions.

Three coping styles have been identified: surrender (i.e., acting as if the schema were true), avoidance (acting as if the schema has to be avoided or escaped), and overcompensation (acting as if the opposite of the schema were true). During the child's development, such coping responses to activated schemas result in so-called 'schema modes,' which define the person's momentary emotional-cognitive-behavioral state. To make a clearer distinction, EMSs can be defined as traits. In contrast, schema modes (including coping modes) describe the momentary emotional-cognitive-behavioral state of the person, and they are considered state-based constructs (Arntz et al., 2021). Coping strategies are considered over or covert responses to EMSs, activated in the "here and now" in reaction to the activation of underlying schemata. Furthermore, in contrast to EMSs, modes include more behavioral aspects. Overall, modes refer to (1) individuals' emotional parts, the so-called child modes (i.e., feeling sad, lonely, abused, enraged, or impulsive); (2) parental introjected messages (i.e., punitive, critical, or perfectionistic inners); and (3) specific behavioral-coping responses (i.e., avoidance and overcompensation). The overcompensation coping (recently labeled as "inversion") leads to a state in which the opposite of the EMS is felt and believed, and it is characterized by thoughts, emotions, and behaviors that serve to prove the contrary of the underlying specific schemata (Arntz et al., 2021). A previous study (Van Wijk-Herbrink et al., 2018) found a significant correlation between overcompensation and externalizing behavior.

The scientific literature on coping modes is scarce. Gulder Altuner and Yararbas (2024) found that childhood violence and illicit substance use were associated with EMSs, particularly with the coping strategies of overcompensation and avoidance. Tenore et al. (2018) focused on non-clinical subjects. They found that specific schemas (mistrust/abuse, vulnerability to harm, and high standards), modes (demanding parent), and coping styles (intra-psychic avoidance) were associated with precise peculiarities for obsessive-compulsive characteristics (washing, checking, and obsessions). In a study on temperament, Mairet et al. (2014) underline that more introverted individuals use more avoidant strategies and that the impact of EMSs on coping strategies is stronger than the influence of coping strategies on such schemas.

However, a recent international working group aims to construct a cross-cultural taxonomy of modes, also extending the theory underlying ST with new insights into emotional needs (Arntz et al., 2021). Within this framework, there has also been a reconsideration of the purpose behind the different ways of coping with EMS activation, coming up with new labels for two of those: resignation instead of surrender and inversion instead of overcompensation (Zhu et al., 2023; Wang et al., 2023). At this moment, these new insights are being empirically tested and validated in over 30 countries. Within therapy, early identification of coping strategies might be particularly helpful in addressing the unhealthy ways by which the client deals with specific EMSs.

The schema coping inventory (SCI; Rijkeboer et al., 2010; Italian translation: Basile, unpublished) is a 12-item self-report questionnaire designed to investigate dysfunctional coping styles using the Schema Theory's theoretical model. The SCI has a three-factor structure, and it enables measuring the three main coping responses identified by the ST model, namely surrender, avoidance, and overcompensation. Unfortunately, a peer-reviewed formal validation of the constructs and structure of the SCI is still missing, and that represents a limit of past literature that the present study addresses and aims to overcome.

On samples of the psychopathological clinic and non-clinic adolescents, Van Wijk-Herbrink et al. (2018) tested the three-factor structure of the SCI, obtaining some empirical evidence for the model's fit. Means and standard deviations on the global and single SCI factors have been provided. Internal consistency, measured by the *Cronbach's alpha* coefficient, was higher for the clinical sample (ranging from 0.71 to 0.78) than the non-clinical sample (from 0.61 to 0.67). Significant correlations have been found between surrender and internalizing behavior problems, avoidance and internalization, overcompensation, and externalizing behavior. However, this study showed some limitations since it was carried out only on Dutch-speaking adolescent groups, a limited variety of variables were used to prove concurrent validity, and limited statistical analyses were reported to demonstrate several types of validity.

In the interest of the clinical and scientific community, we need a systematic validation study of the SCI instrument with means and standard deviation values calculated on a more general adult population, possibly linking coping strategies to psychological and psychopathological characteristics (even if a detailed theoretical dissertation of psychopathological traits of coping strategies is beyond the aims of the present study). This study was driven by the need to obtain a tool to assess coping styles in a Schema Theory framework, which could be useful both in private and public health clinics and research activities. We chose the Italian population because of the SCI-translated version that we may use with our patients, but the results of this study can be generalized to other populations (of the Western countries, at least).

To assess the presence of psychopathological distress, each participant in the present study completed the symptom checklist 90 (SCL90; Derogatis, 1994; It. transl.: Sarno et al., 2011). The aim was to provide useful statistical indexes (i.e., mean and standard deviations) for both non-pathological and pathological populations. Moreover, we hypothesize that the three coping strategies would be differently associated with psychopathological symptoms, with surrender being the most severe condition, particularly for its association with depressive status. The external validity of the SCI was determined by investigating the associations with the brief version of the COPE (Carver, 1997; It. transl.: Conti, 2010). Since avoidance is mostly employed in anxiety disorders, and particularly in social phobia (Bögels and Mansell, 2004), the Liebowitz Social Phobia Scale (LSPS; Liebowitz, 1987; It. transl.: Conti, 2010) was used to analyze the concurrent validity of the avoidance scale. We also evaluated individual optimism versus pessimism, a depressive trait linked to surrender, using the revised version of the Life Orientation Test (LOT-R; Scheier et al., 1994). The validity of the Overcompensation scale was investigated by relating it to the

State–Trait Anger Expression Inventory (STAXI; Forgays et al., 1997) and the Single Item Narcissism Scale (SINS; Konrath et al., 2014), respectively, since anger and narcissism are considered an expression of overcompensation and "inversion" of EMSs (Behary, 2013). Finally, a low internal locus of control is associated with depressive symptoms and surrender style. Therefore, we administered the locus of control (LOC; Lumpkin, 1985) scale.

Therefore, the present study aimed to standardize mean and standard deviation values for population and demographic groups and validate the SCI in an Italian adult population. For this purpose, we selected independent scales measuring the theoretical construct of the SCI. Moreover, according to the previous literature (e.g., Moritz et al., 2016; Van Wijk-Herbrink et al., 2018), we hypothesize significant correlations between dysfunctional coping strategies and psychopathological symptoms, such as the association between avoidance and anxiety/phobia and depression; overcompensation and anger and narcissism; surrender and locus of control and depression.

Method

SCI—Italian version

The 12 items comprising the SCI (Rijkeboer et al., 2010) were translated and adapted for the Italian language according to APA standards (see Supplementary material). After the initial translation, which was made by two independent English translators, followed by a discussion and agreement on a common version, backward translation and agreement by the original author of the SCI were reached (Beaton et al., 2000). Responses to the items of this final version were given on a 7-point Likert scale ranging from 1 = totally disagree to 7 = totally agree.

Participants

We analyzed data from a community sample of 602 Italian adults (305 females) aged between 18 and 69 years (*M*age = 33.7, *SD*age = 11.5). Participants were geographically distributed across the country (North = 48.5%, Centre = 26.2%, South = 16.9%, Islands = 8.3%). Regarding job positions, 33.6% declared themselves to be students, 47.6% to be engaged in full-time work, 14.3% to be unemployed, and 3.5% were retired or houseworkers. As for educational level, 3.3% had a lower secondary school diploma, 45.5% a high school diploma, 44% a degree, while 7.1% had a postgraduate qualification. Participants were enrolled through Prolific, a web platform for data recruiting. The sample size was established using an *a priori* power analysis designed for structural equation models (Moshagen and Erdfelder, 2016). Following the indication of Moshagen and Erdfelder (2016), we set a threshold of 0.05 for root mean square error of approximation (RMSEA), an *alpha* of 0.05, a power of 0.90, and 51 model degrees of freedom. Analysis indicated a minimum sample size of 295 participants to achieve the desired power.

The study conforms to the ethical principles of good clinical practice, the Helsinki Declaration, and complies with the current regulations. The Independent Ethics Committee of Bambino Gesù Children's Hospital (Protocol n. 1787/2019) approved it. The informed consent was obtained when the participants were enrolled in the study.

Measures

The following psychopathological scales have been administered in the study.

Schema Coping Inventory (Italian version): In the present research, the overall scale had a mean score of 3.47 (*SD* = 0.87), while surrender, avoidance, and overcompensation were associated with mean scores of 3.04 (*SD* = 1.24), 3.21 (*SD* = 1.24), and 4.16 (*SD* = 1.04), respectively. Descriptive statistics for the three dimensions and the overall scale across variables such as sex, age, education, parental status, clinical condition, and psychotherapy are shown in Table 1. The clinical cut-off of the SCL90 General Symptomatic Index (GSI) of 60 was used to categorize participants into "non-pathological" and "psychopathological" groups.

Brief COPE (Carver, 1997; It. transl.: Conti, 2010) is a 28-item scale, in which redundant items were dropped from the COPE (Carver et al., 1989) and only two items for each of 14 subscales were selected: Active Coping ($M = 2.83$, $SD = 0.74$, $\alpha = 0.75$); Planning ($M = 2.98$, $SD = 0.80$, $\alpha = 0.81$); Positive Reframing ($M = 2.32$, $SD = 0.86$, $\alpha = 0.82$); Acceptance ($M = 2.75$, $SD = 0.70$, $\alpha = 0.56$); Humor ($M = 1.96$, $SD = 0.75$, $\alpha = 0.66$); Religion ($M = 1.43$, $SD = 0.77$, $\alpha = 0.87$); Using Emotional Support ($M = 2.27$, $SD = 0.89$, $\alpha = 0.85$); Using Instrumental Support ($M = 2.37$, $SD = 0.87$, $\alpha = 0.84$); Self-Distraction ($M = 2.50$, $SD = 0.79$, $\alpha = 0.48$); Denial ($M = 1.35$, $SD = 0.62$, $\alpha = 0.69$); Venting ($M = 2.15$, $SD = 0.76$, $\alpha = 0.57$); Substance Use ($M = 1.24$, $SD = 0.58$, $\alpha = 0.92$); Behavioral Disengagement ($M = 1.65$, $SD = 0.73$, $\alpha = 0.78$); Self-Blame ($M = 2.76$, $SD = 0.78$, $\alpha = 0.64$).

Liebowitz Social Phobia Scale (LSPS; Liebowitz, 1987; It. transl.: Conti, 2010) is a 24-item scale employed to assess performance (13 items) and social difficulties (11 items). Fear or anxiety in the specific situation and avoidance on a 4-point scale (from 'none/ never' to 'severe/usually') were evaluated for each item separately. We computed the overall LSPS scores for each participant by summing anxiety and avoidance scores of both performance and social dimensions ($M = 92.8$, $SD = 28.3$, $\alpha = 0.97$).

The revised version of the Life Orientation Test (LOT-R; Scheier et al., 1994) is a 10-item scale. Three items measure optimism, three measure pessimism, and four need fillers. Each item is scored on a 4-point scale, from 'strongly disagree' to 'strongly agree' ($M = 30.4$, $SD = 7.25$, $\alpha = 0.85$).

State–Trait Anger Expression Inventory (STAXI; Forgays et al., 1997) is a multidimensional measure of anger. We administered the 10-item scale of the state anger dimensions, to which participants provided their rates on a 5-point Likert scale ($M = 1.80$, $SD = 0.63$, $\alpha = 0.91$).

The Single Item Narcissism Scale (SINS; Konrath et al., 2014) is a unique item measure assessing narcissistic personality ($M = 2.25$, $SD = 1.48$). It represents a useful and reliable tool for researchers in preventing the use of longer measures. Participants were asked to rate the item on a 7-point Likert scale from 1 (not very true of me) to 7 (very true of me).

Locus of Control (LOC; Lumpkin, 1985) is a 6-item scale aimed to assess the individual tendency to interpret the events of one's life as products of one's behavior or actions (I. LOC; $M = 3.36$, $SD = 0.71$, $\alpha = 0.53$), or to interpret them as external causes independent of one's will (E. LOC; $M = 3.18$, $SD = 0.71$, $\alpha = 0.54$). Participants provided their answers on a 5-point Likert scale ranging from 'strongly disagree' (1) to 'strongly agree' (5).

TABLE 1 Descriptive statistics.

Variables	Surrender	Avoidance	Overcompensation	SCI_TOT	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	*N*
Sex					
Male	2.89 (1.15)	3.20 (1.17)	4.18 (1.01)	3.42 (0.82)	297
Female	3.19 (1.30)	3.22 (1.31)	4.14 (1.06)	3.52 (0.87)	305
Age					
18–25	3.32 (1.21)	3.25 (1.21)	4.42 (1.02)	3.66 (0.80)	214
26–35	3.34 (1.14)	3.43 (1.22)	4.21 (0.94)	3.65 (0.83)	130
36–45	2.84 (1.18)	3.20 (1.32)	4.04 (1.07)	3.36 (0.88)	152
45–69	2.41 (1.24)	2.89 (1.17)	3.71 (0.99)	3.00 (0.85)	106
Education					
Secondary school	2.60 (1.10)	3.39 (0.93)	4.10 (1.14)	3.36 (0.81)	20
High school	3.17 (1.24)	3.29 (1.21)	4.18 (1.06)	3.54 (0.86)	274
Degree	3.03 (1.24)	3.16 (1.30)	4.16 (1.03)	3.45 (0.90)	265
Postgraduate	2.56 (1.03)	2.94 (1.21)	4.01 (0.92)	3.17 (0.72)	43
Parents					
No	3.18 (1.22)	3.29 (1.24)	4.24 (1.02)	3.57 (0.84)	480
Yes	2.50 (1.14)	2.90 (1.21)	3.82 (1.04)	3.07 (0.87)	122
Clinical condition					
Non-pathological[a]	2.80 (1.12)	3.03 (1.16)	4.08 (1.03)	3.30 (0.80)	467
Psychopathological	3.90 (1.24)	3.84 (1.33)	4.43 (1.02)	4.06 (0.84)	135
Psychotherapy					
No	2.91 (1.21)	3.14 (1.22)	4.12 (1.03)	3.39 (0.86)	429
Yes	3.37 (1.25)	3.39 (1.29)	4.24 (1.05)	3.67 (0.88)	173
Total	3.04 (1.24)	3.21 (1.24)	4.16 (1.04)	3.47 (0.87)	602

[a]Based on SCL90—GSI index.

Symptom Checklist 90-Revised (SCL-90-R; Derogatis, 1994; It. transl.: Sarno et al., 2011) is a 90-item self-report inventory. Each item is scored on a 5-point Likert scale ranging from 'completely disagree' to 'completely agree.' SCL-90 provides information about 10 dimensions of psychopathological distress in the last 7 days: somatization ($M=1.68$, $SD=0.63$, $\alpha=0.87$); obsessive-compulsive ($M=2.09$, $SD=0.83$, $\alpha=0.89$); interpersonal sensitivity ($M=1.87$, $SD=0.76$, $\alpha=0.87$); depression ($M=2.18$, $SD=0.89$, $\alpha=0.93$); anxiety ($M=1.74$, $SD=0.73$, $\alpha=0.89$); hostility ($M=1.63$, $SD=0.66$, $\alpha=0.82$); phobic anxiety ($M=1.37$, $SD=0.59$, $\alpha=0.83$); paranoid ideation ($M=1.79$, $SD=0.77$, $\alpha=0.81$); psychoticism ($M=1.51$, $SD=0.55$, $\alpha=0.81$); and sleep disturbance ($M=2.33$, $SD=0.93$, $\alpha=0.58$). It is also possible to obtain a Global Score Index (GSI), which provides information about the individuals' overall distress level ($M=1.81$, $SD=0.61$, $\alpha=0.98$).

Data analysis

The present study aimed to investigate the factorial structure and psychometric properties of the Italian version of the SCI. To test the goodness-of-fit of the original three-factors (i.e., surrender, avoidance, and overcompensation) structure of Rijkeboer et al. (2010), we conducted a *confirmatory factor analysis* using the diagonal weighted least squares estimator (DWLS). The model fit was evaluated following the benchmarks provided by Hu and Bentler (1999). Given the sensitivity of the *chi-square* (χ^2) statistic to sample size (Chen, 2007), we mainly based on values above 0.90 for the Comparative Fit Index (CFI) and Tucker-Lewis index (TLI) and on values below 0.08 for the root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR). The *CFA* was conducted using *lavaan* (Rosseel, 2012), an *R* package for structural equation modeling, using the *RStudio* graphical interface (Posit team, 2022).

To test the convergent and divergent validity, we computed correlations between the dimensions of the SCI and the Brief COPE's distinct dimensions of active coping, planning, positive reframing, acceptance, humor, religion, using emotional support, using instrumental support, self-distraction, denial, venting, substance use, behavioral disengagement, and self-blame. The significance level of each correlation was adjusted with the Bonferroni method, given the large number of multiple tests performed. Analyses were conducted using the Psych R package (Revelle and Revelle, 2015).

Concurrent validity was first investigated by computing correlations of SCI's dimensions with conceptually concurrent measures, such as the Liebowitz Social Phobia Scale (Liebowitz, 1987; It. transl.: Conti, 2010), Life Orientation Test (Scheier et al., 1994), state anger dimension of the State–Trait Anger Expression Inventory

(Forgays et al., 1997), Single Item Narcissism Scale (Konrath et al., 2014), Locus of Control (Lumpkin, 1985) and Symptom Checklist 90-Revised (Derogatis, 1994; It. transl.: Sarno et al., 2011). Criterion validity was further investigated using a series of multiple regression models. The models included surrender, avoidance, and overcompensation as predictors and the distinct symptoms of SCL-90-R as criteria. Regression analyses aimed to further discriminate each SCI dimension and weigh their predictive power of symptomatic outcomes. Finally, we conducted *univariate ANOVAs* considering each dimension (recoded into tertiles) of the SCI as fixed factors and the total score (GSI) of the SCL-90-R as a criterion. These analyses merely provided a descriptive overview of the overall symptomatology across distinct levels of surrender, avoidance, and overcompensation.

Results

Confirmatory factor analysis

First, we tested the original three factors and the 12-item model of the SCI (Figure 1). The analysis revealed an acceptable model fit of this factorial structure within our community sample. Specifically, besides a significant *Chi-square* ($\chi^2 = 210.83$, $df = 51$, $p < 0.001$), we found the considered incremental fit index to be over the acceptability threshold of 0.90 (CFI = 0.92). Regarding the absolute fit indices, the analyses showed a value of 0.07 for the SRMR and 0.07 for the RMSEA with a 90% confidence interval ranging from 0.06 to 0.08. These findings suggested that the model's fit to the observed data was acceptable (Table 2). Moreover, as seen in Table 3, the factor loadings of the items were moderate to high and significant, highlighting coefficients between 0.28 and 0.76 in their standardized version.

Internal consistency

To assess the internal consistency of the proposed scale, we tested its reliability within the interested community sample. We also separately assessed the reliability of the three dimensions: surrender, avoidance, and overcompensation. The reliability analysis showed a satisfactory *Cronbach's alpha* of 0.74 for the overall scale. The dimensions of surrender and avoidance yielded suitable results, with *Cronbach's alpha* of 0.71 and 0.68, respectively. However, the dimension of overcompensation showed a low reliability coefficient ($\alpha = 0.47$).

Convergent and divergent validity

We computed correlations with other theoretically convergent and divergent measures to test the convergent and divergent validity of the Italian version of the SCI. More specifically, we investigated the associations of SCI's dimensions with the distinct coping styles proposed by the Brief COPE. As shown in Table 4, we found corroboration in favor of the construct validity of SCI. The dimension of surrender highlighted convergent associations with the maladaptive coping styles of denial ($r = 0.25$, $p < 0.001$), behavioral disengagement ($r = 0.52$, $p < 0.001$), substance use ($r = 0.19$, $p < 0.001$), and self-blame ($r = 0.31$, $p < 0.001$). The avoidance dimension showed significant positive associations with denial ($r = 0.18$, $p < 0.01$), behavioral disengagement ($r = 0.36$, $p < 0.001$), and self-blame ($r = 0.16$, $p < 0.05$). Overcompensation was not significantly related to any maladaptive coping styles of the brief COPE.

Surrender also showed significant divergent associations with 'functional' coping dimensions of positive reframing ($r = -0.29$, $p < 0.001$), active coping ($r = -0.40$, $p < 0.001$), acceptance ($r = -0.27$, $p < 0.001$), and planning ($r = -0.41$, $p < 0.001$). Avoidance highlighted significant and negative relations with positive reframing ($r = -0.23$, $p < 0.001$), using instrumental support ($r = -0.25$, $p < 0.001$), active coping ($r = -0.28$, $p < 0.001$), acceptance ($r = -0.18$, $p < 0.01$), and planning ($r = -0.26$, $p < 0.001$). Avoidance was the unique SCI dimension significantly associated with using emotional support ($r = -0.23$, $p < 0.01$). Overcompensation was not significantly related to functional coping styles.

Concurrent validity

We computed correlations among each dimension of the SCI and several variables that may be framed as their potential external criteria. Then, following the indication of Eid et al. (2011) for comparing correlations from dependent samples, we also contrasted the emerged coefficients. As shown in Table 4, the Liebowitz Social Phobia Scale (LSPS) is positively related to Surrender and avoidance but not to overcompensation. The association between Surrender and LSPS was stronger than that with Avoidance ($z = 7.41$, $p < 0.001$). The three dimensions of SCI also correlated positively with anger (STAXI; see Table 4). In this case, the only relationships that differed were those with surrender and avoidance ($z = 2.14$, $p = 0.02$). As for narcissism, the analyses showed an inverse pattern: overcompensation was the unique SCI dimension to be significantly related to it. Surrender, avoidance, and overcompensation correlated negatively with the LOT. The relationship with Surrender was stronger compared to that with Avoidance ($z = -7.11$, $p < 0.001$) and Overcompensation ($z = -10.35$, $p < 0.001$). In turn, the association between LOT and Avoidance was significantly stronger than that with overcompensation ($z = -4.77$, $p < 0.001$). A similar pattern was repeated with the Internal Locus of control. Both surrender and avoidance showed a significant correlation with the Internal Locus of control, with Surrender showing a significantly larger coefficient than avoidance ($z = -4.18$, $p < 0.001$), while overcompensation was not significantly related to the internal locus of control. As for the external locus of control, all the dimensions of SCI showed a positive relationship. However, surrender showed a higher coefficient than the other two ($z = 1.92$, $p = 0.03$; $z = 3.24$, $p < 0.001$, respectively, for avoidance and overcompensation), which were also different from each other ($z = 1.80$, $p = 0.04$). Finally, each dimension of SCI correlated positively with psychopathological symptomatology measured with SCL-90-R. Surrender had a stronger association with psychopathological symptoms than avoidance ($z = 5.72$, $p < 0.001$) and overcompensation ($z = 7.47$, $p < 0.001$). In turn, avoidance had a significantly stronger relation to symptomatology than overcompensation ($z = 3.07$, $p < 0.001$). To provide an overview of the psychopathological symptomatology for different levels of the three dysfunctional coping styles, we also implemented a series of *ANOVAs* (see Supplementary material).

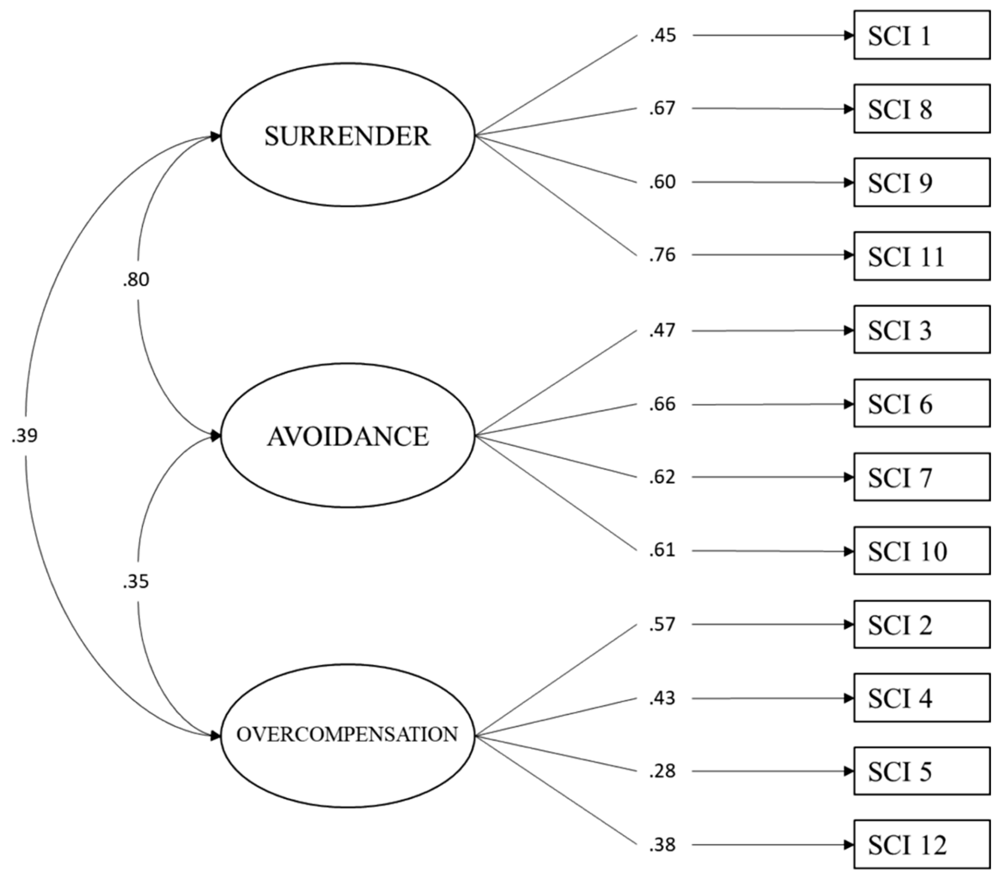

FIGURE 1
Factor structure of the Italian version of the schema coping inventory.

TABLE 2 Goodness-of-fit indicators for the 3-factor and 12-item model of the Italian version of schema coping inventory.

Models	χ^2	*df*	CFI	SRMR	RMSEA
Sample (*N* = 602)	210.83	51	0.92	0.07	0.07 (0.062, 0.082)

Multiple regression analyses

To further examine the criterion validity of SCI, we implemented a multiple regression model for each symptom dimension of the SCL-90-R (considered as a criterion) and the surrender, avoidance, and overcompensation dimensions as predictors. For the symptoms dimension of somatization, we found surrender as a unique significant predictor ($\beta = 0.33$, $se = 0.046$, $z = 7.06$, $p < 0.001$, 95%CI = 0.234, 0.415). Both surrender ($\beta = 0.52$, $se = 0.041$, $z = 12.8$, $p < 0.001$, 95%CI = 0.440, 0.600) and overcompensation ($\beta = 0.11$, $se = 0.035$, $z = 3.17$, $p = 0.002$, 95%CI = 0.042, 0.117) turned out to be positively associated with obsessive-compulsive symptoms. As for interpersonal sensitivity, all three of the SCI's dimensions of Surrender ($\beta = 0.44$, $se = 0.041$, $z = 10.9$, $p < 0.001$, 95%CI = 0.362, 0.522), avoidance ($\beta = 0.14$, $se = 0.041$, $z = 3.53$, $p < 0.001$, 95%CI = 0.064, 0.223), and overcompensation ($\beta = 0.06$, $se = 0.035$, $z = 3.17$, $p = 0.002$, 95%CI = 0.042, 0.117) simultaneously resulted as significant predictors. Depression symptoms was positively related only with surrender ($\beta = 0.53$, $se = 0.040$, $z = 13.0$, $p < 0.001$, 95%CI = 0.446, 0.605). The anxiety dimension of SCL90-R was predicted by both surrender ($\beta = 0.47$, $se = 0.043$, $z = 11.0$, $p < 0.001$, 95%CI = 0.385, 0.553) and overcompensation ($\beta = 0.08$, $se = 0.036$, $z = 2.17$, $p = 0.031$, 95%CI = 0.007, 0.150), as well as Hostility ($\beta = 0.25$, $se = 0.046$, $z = 5.40$, $p < 0.001$, 95%CI = 0.158, 0.338; $\beta = 0.22$, $se = 0.039$, $z = 5.72$, $p < 0.001$, 95%CI = 0.147, 0.300; respectively, for surrender and overcompensation). Phobic anxiety symptomatology was instead predicted by the dimensions of surrender ($\beta = 0.35$, $se = 0.045$, $z = 7.82$, $p < 0.001$, 95%CI = 0.147, 0.325) and avoidance ($\beta = 0.11$, $se = 0.045$, $z = 2.42$, $p = 0.016$, 95%CI = 0.020, 0.196). Paranoid ideation and psychoticism were both positively associated with all three dimensions of SCI (see Table 5). Finally, surrender only predicted sleep disturbance ($\beta = 0.29$, $se = 0.047$, $z = 6.15$, $p < 0.001$, 95%CI = 0.198, 0.385).

Discussion

In the present research, using a large community sample ($N = 602$), we provide empirical corroboration in favor of the factorial structure and psychometric properties of the Schema Coping Inventory (SCI-Italian version; Rijkeboer et al., 2010). SCI measures dysfunctional strategies to cope with maladaptive early schemas. Such coping mechanisms refer to how a person deals with an internal experience (schema activation) rather than an external circumstance (Arntz et al., 2021). Means and standard deviations (for SCI and all the psychopathological questionnaires), confirmatory factor analysis,

TABLE 3 Items' factor loading for the 3-factor and 12-item model of the Italian version of the schema coping inventory.

						95%CI	
Factor	Item	*β*	*se*	*z*	*p*	Lower	Upper
Surrender	SCI_1	0.45	0.030	15.2	< 0.001	0.395	0.512
	SCI_8	0.67	0.038	17.6	< 0.001	0.592	0.741
	SCI_9	0.60	0.031	19.3	< 0.001	0.540	0.663
	SCI_11	0.76	0.037	20.6	< 0.001	0.683	0.827
Avoidance	SCI_3	0.47	0.031	15.1	< 0.001	0.409	0.531
	SCI_6	0.66	0.035	19.2	< 0.001	0.594	0.729
	SCI_7	0.62	0.035	17.7	< 0.001	0.550	0.687
	SCI_10	0.61	0.033	18.1	< 0.001	0.540	0.670
Overcompensation	SCI_2	0.57	0.056	10.1	< 0.001	0.456	0.676
	SCI_4	0.43	0.048	8.89	< 0.001	0.333	0.521
	SCI_5	0.28	0.042	6.66	< 0.001	0.197	0.361
	SCI_12	0.38	0.045	8.52	< 0.001	0.294	0.470

TABLE 4 Intercorrelation of the schema coping inventory's dimensions of surrender, avoidance, and overcompensation with brief cope's dimensions, the Liebowitz Social Phobia Scale (LSPS), the Single Item Narcissism Scale (SINS), life orientation test-revised (LOT-R), trait anger (STAXI), internal (I. LOC) and external (E. LOC) locus of control.

	Surrender	Avoidance	Overcompensation
P. Reframing	−0.29***	−0.23***	0.06
S. Distraction	0.11	0.12	0.07
Venting	0.12	−0.12	0.12
U. I. support	0.05	−0.25***	−0.02
Active coping	−0.40***	−0.28***	−0.13
Denial	0.25***	0.18**	0.03
Religion	−0.11	−0.11	−0.06
Humor	−0.03	−0.001	0.03
B. Diseng.	0.52***	0.36***	0.15
U. E. Support	0.09	−0.23***	0.001
Substance Use	0.19***	0.06	0.05
Acceptance	−0.27***	−0.18**	−0.12
Planning	−0.41***	−0.26***	−0.11
Self-blame	0.31***	0.16*	0.12
SCL-GSI	0.55***	0.36***	0.21***
LSPS	0.60***	0.48***	0.13
SINS	0.10	0.12	0.30***
LOT-R	−0.64**	−0.42***	−0.19***
STAXI	0.27***	0.19***	0.22***
I. LOC	−0.41***	−0.26***	−0.09
E. LOC	0.37***	0.27***	0.18**

p is computed according to Bonferroni multiple tests adjustment; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

internal consistency, construct validity, and criterion validity analyses have been reported.

The factor analysis corroborated the tridimensional structure of the SCI, i.e., Surrender, Avoidance, and Overcompensation, both in terms of the overall goodness-of-fit of the model and the single items' factor loading in the corresponding factors. The *internal consistency* of the whole scale turned out to be satisfactory. Regarding the reliability of each of the three dimensions, we found adequate coefficients associated with the dimensions of Surrender and Avoidance, while the dimension of Overcompensation showed a low reliability coefficient. Unfortunately, comparing this data with previous studies is impossible (Rijkeboer et al., 2010; Van

TABLE 5 Multiple regression of surrender, avoidance, and overcompensation dimensions on SCL-90 symptomologies.

						95%CI	
Criterion	Predictor	*β*	*se*	*z*	*p*	Lower	Upper
SOM	Surrender	0.33	0.046	7.06	0.001	0.234	0.415
	Avoidance	0.04	0.046	0.76	0.45	−0.056	0.125
	Overcompensation	0.05	0.039	1.37	0.17	−0.023	0.130
O-C	Surrender	0.52	0.041	12.8	0.001	0.440	0.600
	Avoidance	0.03	0.041	0.63	0.53	−0.054	0.105
	Overcompensation	0.11	0.035	3.17	0.002	0.042	0.177
INT	Surrender	0.44	0.041	10.9	0.001	0.362	0.522
	Avoidance	0.14	0.041	3.53	0.001	0.064	0.223
	Overcompensation	0.10	0.035	2.97	0.003	0.035	0.171
DEP	Surrender	0.53	0.040	13.0	0.001	0.446	0.605
	Avoidance	0.06	0.040	1.43	0.15	−0.022	0.137
	Overcompensation	0.06	0.034	1.07	0.08	−0.009	0.126
ANX	Surrender	0.47	0.043	11.0	0.001	0.385	0.553
	Avoidance	0.01	0.043	0.29	0.77	−0.071	0.096
	Overcompensation	0.08	0.036	2.17	0.031	0.007	0.150
HOS	Surrender	0.25	0.046	5.40	0.001	0.158	0.338
	Avoidance	−0.003	0.046	−0.07	0.94	−0.093	0.087
	Overcompensation	0.22	0.039	5.72	0.001	0.147	0.300
PHOB	Surrender	0.35	0.045	7.82	0.001	0.262	0.438
	Avoidance	0.11	0.045	2.42	0.016	0.020	0.196
	Overcompensation	−0.009	0.038	−0.23	0.82	−0.083	0.066
PAR	Surrender	0.24	0.045	5.20	0.001	0.147	0.325
	Avoidance	014	0.045	3.15	0.002	0.054	0.232
	Overcompensation	0.14	0.039	3.67	0.001	0.066	0.217
PSY	Surrender	0.35	0.043	8.19	0.001	0.269	0.439
	Avoidance	0.15	0.043	3.47	0.001	0.065	0.235
	Overcompensation	0.08	0.037	2.14	0.033	0.006	0.151
SLEEP	Surrender	0.29	0.047	6.15	0.001	0.198	0.385
	Avoidance	−0.06	0.047	−1.17	0.24	−0.148	0.038
	Overcompensation	0.01	0.040	0.35	0.72	−0.065	0.093

Summary of explained variance for each multiple regression model: Somatization (SOM): $R^2 = 0.13$; Obsessive-Compulsive (O-C): $R^2 = 0.32$; Interpersonal Sensitivity (INT): $R^2 = 0.32$; Depression (DEP): $R^2 = 0.33$; Anxiety (ANX): $R^2 = 0.25$; Hostility (HOS): $R^2 = 0.13$; Phobic Anxiety (PHOB): $R^2 = 0.18$; Paranoid Ideation(PAR): $R^2 = 0.15$; Psychoticism(PSY): $R^2 = 0.23$; Sleep Disturbance (SLEEP): $R^2 = 0.07$.

Wijk-Herbrink et al., 2018). Convergent and divergent validity have been assessed by correlating the SCI scales with other measures indexing coping styles and psychopathological features.

The dysfunctional Surrender coping style of the SCI (SCI-surrender) significantly and positively correlated with other Brief COPE dimensions associated with depressive and surrender styles, i.e., denial, behavioral disengagement, self-blame, self-distraction, and substance use. In addition, SCI-surrender was negatively related to the 'functional' coping dimensions of positive reframing, active coping, acceptance, and planning (Carver, 1997). These data may be interpreted as evidence of multiple components in surrender, representing disengagement strategies (denial, self-distraction, and substance use) and/or responses to self-blame (Janovsky et al., 2023). All of them would work by blocking functional, active coping, such as planning, positive reframing, and, finally, hindering acceptance mechanisms. Notably, self-blame, together with behavioral disengagement and denial, represents significant components associated with avoidance behavior as well. SCI-avoidance showed moderately significant and negative correlations with use of instrumental support and use of emotive support, suggesting a specific reduction of social support searching in the case of avoidance (Polman et al., 2010; Forbes et al., 2020), whereas their correlations with surrender and overcompensation were null or weak. This evidence suggests that the use of an avoidance coping strategy is associated with a reduced use and expectation of different types of support: emotive and instrumental. Several significant correlations emerged between the three SCI factors and other psychopathological dimensions (Janovsky et al., 2023). By contrasting correlation coefficients of interest, we found that surrender showed a significantly higher correlation with LSPS (Derogatis, 1994), measuring fear and anxiety in performance and social exposure than avoidance and overcompensation, which are significantly associated with social phobia (Beidel and Randall, 1994; Gonzalez Diez et al., 2012). Surrender-LSPS association would sound contrary to expectations and suggests a possible pathological generalization of the surrender strategy even to anxiety disorders. Surrender and avoidance were associated with anger, whereas overcompensation showed the strongest correlation with narcissism, as hypothesized (Rosenthal and Hooley, 2010; Behary, 2013; Young et al., 2003). As was found in earlier research (Van Wijk-Herbrink et al., 2018), overcompensation was mainly related to externalizing tendencies, as it was associated with, for example, narcissism and hostility. However, since we did not find significant correlations between overcompensation and other brief COPE factors, we suggest that a more defined operationalization and theoretical agreement on the construct should be achieved.

As predicted, the values on the LOT-R scale, assessing aptitude for optimism and pessimism, especially individual pessimism (Scheier et al., 1994), were more strongly correlated with the surrender style than the other two styles. We found the same pattern for the correlations with locus of control (see also Reinert, 1997). These results might be related to the belief people can have in perceiving themselves as less capable and skilled in facing the schema, feeling as if they are "a child" in a world defined by the representation of their underlying schema (Arntz et al., 2021). Regarding the overall correlations with SCL-90 GSI, surrender showed the strongest correlation with the psychopathological global index, followed by avoidance and overcompensation, respectively. The evidence confirms the initial hypothesis of differential psychopathological sequela for the three coping strategies. Surrender was the dysfunctional coping style that was more strongly associated with the presence of psychopathological mechanisms, i.e., SCL-90 psychopathological global index and specific measures, such as social phobia, anger, pessimism, and polarization of internal and external locus of control attribution. Surrendering as a coping strategy tends to develop earlier in childhood compared to overcompensation, particularly in highly dysfunctional contexts. As a result, it may lead to more severe and pervasive early maladaptive schemas (EMSs).

As the *multiple regression analysis* showed, SCI criterion validity has been investigated using the psychopathological dimensions of the SCL-90-R. The surrender strategy proved to be exclusive in explaining the variability of the dimensions of depression, somatization, and sleep disturbance. Therefore, surrender is attested to be the coping strategy associated with mood depression and its related symptoms. Surrender and overcompensation were significantly related to the 'anxiety' dimensions of obsessive-compulsive, hostility, and anxiety (see also Tenore et al., 2018). Surrender and avoidance significantly explained phobic symptomatology, which is typically maintained by emotional and behavioral avoidance.

On the other hand, all three coping strategies were significantly related to the dimensions of paranoid ideation, psychoticism, and interpersonal sensitivity. As we can observe, the surrender factor is a good predictor for each psychopathological dimension of the SCL-90 scale (see also Starcevic et al., 2015). The SCI validity is supported by the significant correlations with the COPE brief factors: once again, surrender showed strong correlations with dysfunctional coping strategies. Therefore, surrender coping is giving in to EMS, which will lead to vulnerable child modes, and it is more related to psychopathological features (i.e., depression, anxiety, OCD, and so on) with respect to avoidance and overcompensation.

In the schema therapy theory, Surrender is hypothesized to lead to, for example, the vulnerable child modes in which intense distress is experienced (Arntz et al., 2021; Sójta and Strzelecki, 2023). As expected, Surrender had the strongest associations with psychopathological symptoms as measured by the SCL-90, such as depression, anxiety, obsessive-compulsive symptoms, somatization, and sleep problems. The hypothesized function of Avoidance is to escape the pain of the EMSs. In terms of overcompensation to deny the pain by assuming the opposite, in line with these functions, hardly any associations between avoidance or overcompensation and the symptoms indexed by the SCL-90 LSPS were found.

The results of our study are in line with the findings by Van Wijk-Herbrink et al. (2018), that is, there is a significant association between dysfunctional coping styles and psychopathological traits. However, we overcome the limitations of data analysis and generalization found in the study by Van Wijk-Herbrink et al. (2018), which was conducted solely on a Dutch adolescent group consisting of both community and clinical samples with externalizing or personality disorders.

In our opinion, a clinically relevant limitation of the SCI questionnaire is the absence of a dimensional measure of a positive and functional coping strategy, which would allow classifying individuals' behaviors even in an 'adaptive functional' profile, in line with previous tools, such as COPE Brief. Previous studies demonstrated that the use of positive coping strategies is correlated with a better quality of life and better capacity for resilience in people with psychiatric disorders (Chi et al., 2022; Tsouvelas et al., 2023a, 2023b). On the contrary, the notable strength of the SCI is the reduced number of items and factors, which is useful in both research and clinical application.

In our opinion, SCI classification in surrender, avoidance, and overcompensation catches and summarizes the variety of dysfunctional coping behaviors in response to schema activation. In particular, the construct of 'overcompensation' highlights the importance of the 'quantity' dimension of a class of behaviors that may be functional and adaptive until a certain quantitative threshold but become dysfunctional beyond such a threshold. Such a hypothesis should be verified in future research. However, overcompensation, as defined right now, may be subject to several theoretical interpretations, and a more detailed operationalization would be helpful.

Limitations and conclusion

Some limitations need to be acknowledged. A first limitation could be found in the use of a paid platform for the recruitment of participants. While Prolific is a highly effective participant recruitment platform, it has some limitations that can impact the quality of the collected data. One major issue is uncertainty about participant identity. Despite Prolific's verification processes, it is not always possible to ensure that participants do not misrepresent their demographic data or update it inconsistently over time. Another potential issue is the lack of control over participants' environments. It is difficult to control for variables such as distractions, the type of device used, or the environment the participant is in while completing the questionnaire. This is a disadvantage compared to procedures carried out in controlled laboratory environments, where researchers can supervise participants in real time. However, Prolific stands out as a valuable alternative to other crowd-working platforms due to its clear guidelines for participants and researchers. One of its key advantages is transparency. Participants are fully informed that they are being recruited for research, and they understand the details regarding payment, treatment, and their rights and responsibilities within this context. For researchers, Prolific offers greater transparency about the participant pool than other platforms, allowing for more thorough screening based on various criteria before selecting participants for their studies (Palan and Schitter, 2018).

The sample may not fully represent the broader Italian population due to platform recruitment, particularly regarding socioeconomic or cultural diversity. Future research could address this limitation with a more diverse or clinically representative sample. However, data collection was implemented with accurate attention checks, which, if failed, led to the automatic exclusion of the participants and the non-registration of the related data. Thus, we obtained reliable data on which to base our analysis and the related evidence.

Another limitation might be traced to the low-reliability value of the overcompensation dimension of SCI, which mirrors the need for a more detailed theoretical definition of the construct and for the development of a new psychometrically robust instrument to measure dysfunctional coping styles derived from schema therapy-associated theory.

We can conclude that having measures describing the dysfunctional coping styles for psychological pain and stress allows us to (1) increase the patient's awareness of his/her own coping style; (2) increase the patient's awareness of other possible coping styles for stress and pain; (3) plan therapeutic actions addressing the dysfunctional style, aiming at modifying it, enlarging strategies' variability, and hence working on cognitive flexibility, suggesting more functional strategies; and (4) increase awareness and acceptance of different coping styles into the framework of interpersonal systems, i.e., couple, family, clinical or rehabilitative pediatric team, professional team.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

Ethics statement

The studies involving humans were approved by the Independent Ethics Committee of Bambino Gesù Children's Hospital (Protocol n. 1787/2019), which gave its approval. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SG: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. VP: Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. EN: Data curation, Methodology, Writing – original draft. VV: Data curation, Investigation, Writing – original draft. DL: Project administration, Writing – original draft. EC: Funding acquisition, Project administration, Supervision, Writing – original draft. BB: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. MG: Investigation, Writing – original draft, Writing – review & editing, Conceptualization, Methodology.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported also by the Italian Ministry of Health with "Current Research Funds".

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1441794/full#supplementary-material

References

Arntz, A., Rijkeboer, M., Chan, E., Fassbinder, E., Karaosmanoglu, A., Lee, C. W., et al. (2021). Towards a reformulated theory underlying schema therapy: position paper of an international workgroup. *Cognit. Ther. Res.* 45, 1007–1020. doi: 10.1007/s10608-021-10209-5

Beaton, D. E., Bombardier, C., Guillemin, F., and Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 25, 3186–3191. doi: 10.1097/00007632-200012150-00014

Behary, W. T. (2013). Disarming the narcissist: Surviving and thriving with the self-absorbed. Oakland (US): New Harbinger Publications.

Beidel, D. C., and Randall, J. (1994). "Social phobia" in International handbook of phobic and anxiety disorders in children and adolescents. eds. T. H. Ollendick, N. J. King and W. Yule (New York (US): Springer, Plenum Press), 111–129.

Bögels, S. M., and Mansell, W. (2004). Attention processes in the maintenance and treatment of social phobia: hypervigilance, avoidance and self-focused attention. *Clin. Psychol. Rev.* 24, 827–856. doi: 10.1016/j.cpr.2004.06.005

Carver, C. S. (1997). You want to measure coping but your protocol's too long: consider the brief cope. *Int. J. Behav. Med.* 4, 92–100. doi: 10.1207/s15327558ijbm0401_6

Carver, C. S., Scheier, M. F., and Weintraub, J. K. (1989). Assessing coping strategies: a theoretically based approach. *J. Pers. Soc. Psychol.* 56, 267–283. doi: 10.1037/0022-3514.56.2.267

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* 14, 464–504. doi: 10.1080/10705510701301834

Chi, D., Zhong, H., Wang, Y., Ma, H., Zhang, Y., and Du, X. (2022). Relationships between positive schemas and life satisfaction in psychiatric inpatients. *Front. Psychol.* 13:1061516. doi: 10.3389/fpsyg.2022.1061516

Conti, L. (2010). Repertorio delle scale di valutazione in psichiatria. Firenze: SEE Editrice.

Derogatis, L. R. (1994). SCL-90-R: Administration, Scoring and Procedures Manual. *3rd* Edn. Minneapolis, MN: National Computer Systems.

Eid, M., Gollwitzer, M., and Schmitt, M. (2011). Statistik und Forschungsmethoden Lehrbuch. Weinheim, Germany: Beltz.

Forbes, C. N., Tull, M. T., Xie, H., Christ, N. M., Brickman, K., Mattin, M., et al. (2020). Emotional avoidance and social support interact to predict depression symptom severity one year after traumatic exposure. *Psychiatry Res.* 284:112746. doi: 10.1016/j.psychres.2020.112746

Forgays, D. G., Forgays, D. K., and Spielberger, C. D. (1997). Factor structure of the state–trait anger expression inventory. *J. Pers. Assess.* 69, 497–507. doi: 10.1207/s15327752jpa6903_5

Gonzalez Diez, Z., Calvete, Z. E., and Orue, S. I. (2012). P-129 – early Maladaptative schemas and social anxiety: the moderating effect of avoidant vs. overcompensation coping. *Eur. Psychiatry* 27:1. doi: 10.1016/S0924-9338(12)74296-7

Gulder Altuner, O., and Yararbas, G. (2024). Investigation of early maladaptive schemas and coping strategies in Turkish adults under probation. *Health Soc. Care Community* 2024:8953611. doi: 10.1155/2024/8953611

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Janovsky, T., Rock, A. J., Thorsteinsson, E. B., Clark, G. I., Polad, V., and Cosh, S. (2023). Assessing the relationship between early maladaptive schemas and interpersonal problems using interpersonal scenarios depicting rejection. *PLoS One* 18:e0288543. doi: 10.1371/journal.pone.0288543

Konrath, S., Meier, B. P., and Bushman, B. J. (2014). Development and validation of the single item narcissism scale (SINS). *PLoS One* 9:e0103469. doi: 10.1371/journal.pone.0103469

Liebowitz, M. R. (1987). Social phobia. *Mod. Probl. Pharmacopsychiatry* 22, 141–173. doi: 10.1159/000414022

Lumpkin, J. R. (1985). Validity of a brief locus of control scale for survey research. *Psychol. Rep.* 57, 655–659. doi: 10.2466/pr0.1985.57.2.655

Mairet, K., Boag, S., and Warburton, W. (2014). How important is temperament? The relationship between coping styles, early maladaptive schemas and social anxiety. *Int. J. Psychol. Psychol. Ther.* 14, 171–189.

Moritz, S., Jahns, A. K., Schröder, J., Berger, T., Lincoln, T. M., Klein, J. P., et al. (2016). More adaptive versus less maladaptive coping: what is more predictive of symptom severity? Development of a new scale to investigate coping profiles across different psychopathological syndromes. *J. Affect. Disord.* 191, 300–307. doi: 10.1016/j.jad.2015.11.027

Moshagen, M., and Erdfelder, E. (2016). A new strategy for testing structural equation models. *Struct. Equ. Model.* 23, 54–60. doi: 10.1080/10705511.2014.950896

Palan, S., and Schitter, C. (2018). Prolific. Ac—a subject pool for online experiments. *J. Behav. Exp. Finance* 17, 22–27. doi: 10.1016/j.jbef.2017.12.004

Pilkington, P. D., Bishop, A., and Younan, R. (2021). Adverse childhood experiences and early maladaptive schemas in adulthood: a systematic review and meta-analysis. *Clinic. Psych. Psychoth.* 28, 569–584. doi: 10.1002/cpp.2533

Polman, R., Borkoles, E., and Nicholls, A. R. (2010). Type D personality, stress, and symptoms of burnout: the influence of avoidance coping and social support. *Br. J. Health Psychol.* 15, 681–696. doi: 10.1348/135910709X479069

Posit team (2022). RStudio: Integrated development environment for R. Boston, MA: RStudio, PBC.

Reinert, D. F. (1997). The surrender scale: reliability, factor structure, and validity. *Alcohol. Treat. Q.* 15, 15–32. doi: 10.1300/J020v15n03_02

Revelle, W., and Revelle, M. W. (2015). Package 'psych'. *The Comprehensive R Archive Network*, pp. 337–338. Available at: https://personality-project.org/r/psych/.

Rijkeboer, M. M., Lobbestael, J., Arntz, A., and Genderen, H.Van (2010). Schema Coping Vragenlijst. Utrecht: Uitgave in eigen beheer.

Rosenthal, S. A., and Hooley, J. M. (2010). Narcissism assessment in social–personality research: does the association between narcissism and psychological health result from a confound with self-esteem? *J. Res. Pers.* 44, 453–465. doi: 10.1016/j.jrp.2010.05.008

Rosseel, Y. (2012). Lavaan: an r package for structural equation modeling and more. Version 0.5–12 (beta). *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Sarno, I., Preti, E., Prunas, A., and Madeddu, F. (2011). SCL-90-R symptom Checklist-90-R Adattamento italiano. Firenze (IT): Giunti Organizzazioni Speciali.

Scheier, M. F., Carver, C. S., and Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the life orientation test. *J. Pers. Soc. Psychol.* 67, 1063–1078. doi: 10.1037/0022-3514.67.6.1063

Sójta, K., and Strzelecki, D. (2023). Early maladaptive schemas and their impact on parenting: do dysfunctional schemas pass generationally: a systematic review. *J. Clin. Med.* 12:1263. doi: 10.3390/jcm12041263

Starcevic, V., Berle, D., Viswasam, K., Hannan, A., Milicevic, D., Brakouias, V., et al. (2015). Specificity of the relationships between dysphoria and related constructs in an outpatient sample. *Psychiatry Q.* 86, 459–469. doi: 10.1007/s11126-015-9344-8

Tenore, K., Mancini, F., and Basile, B. (2018). Schemas, modes and coping strategies in obsessive-compulsive like symptoms. *Clin. Neuropsych.* 15, 384–392.

Tsouvelas, G., Chondrokouki, M., Antoniou, X., Athanasiadou, Y., and Nikolaidis, G. (2023a). Life satisfaction and early maladaptive schemas in children in residential care. *Dev. Adolescent Health* 2, 28–36. doi: 10.54088/98jui

Tsouvelas, G., Chondrokouki, M., Antoniou, X., and Nikolaidis, G. (2023b). Early maladaptive schemas and symptoms of psychopathology in children in residential care. *Psychiatriki* 34, 301–311. doi: 10.22365/jpsych.2023.007

Van Wijk-Herbrink, M. F., Roelofs, J., Broers, N. J., Rijkeboer, M. M., Arntz, A., and Bernstein, D. P. (2018). Validation of schema coping inventory and schema mode inventory in adolescents. *J. Personal. Disord.* 32, 220–241. doi: 10.1521/pedi_2017_31_295

Wang, T., Kong, L. N., Yao, Y., and Li, L. (2023). Self-efficacy, coping strategies and quality of life among patients with chronic hepatitis B. *West. J. Nur. Res.* 45, 800–806. doi: 10.1177/01939459231184714

Young, J. E., Klosko, J. S., and Weishaar, M. E. (2003). Schema therapy: A Practitioner's guide. New York: Guilford Press.

Zhu, F., Liu, C., Zhang, W., Qiang, W., Yin, X., and Lu, Q. (2023). The mediating effect of coping styles between self-compassion and body image disturbance in young breast cancer survivors: a cross-sectional study. *BMC Nurs.* 22:178. doi: 10.1186/s12912-023-01342-5

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 06 January 2025
DOI 10.3389/fmed.2024.1451979

Check for updates

OPEN ACCESS

EDITED BY
Giulia Casu,
University of Bologna, Italy

REVIEWED BY
Marcos A. Halty,
Universidad de Tarapacá, Chile
Matteo Orsoni,
University of Bologna, Italy

*CORRESPONDENCE
Chao Xu
honghuixuchao@163.com
Yuanchi Huang
huangychell@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 20 June 2024
ACCEPTED 10 December 2024
PUBLISHED 06 January 2025

CITATION
Tai Z, Wan D, Zan Q, Huang Y and Xu C (2025) Comparison of the EQ-5D-3L and EQ-5D-5L instruments in patients undergoing unicompartmental knee arthroplasty. *Front. Med.* 11:1451979.
doi: 10.3389/fmed.2024.1451979

Comparison of the EQ-5D-3L and EQ-5D-5L instruments in patients undergoing unicompartmental knee arthroplasty

Zhipeng Tai[1,2†], Dongping Wan[3†], Qiang Zan[2], Yuanchi Huang[1]* and Chao Xu[1]*

[1]Department of Knee Joint Surgery, Honghui Hospital, Xi'an Jiaotong University, Xi'an, China, [2]The First Clinical Medical College, Shaanxi University of Chinese Medicine, Xianyang, China, [3]Guangxi University of Chinese Medicine, Nanning, China

Purpose: The purpose of this investigation is to assess and contrast the effectiveness of the two EuroQol five dimensions questionnaire (EQ-5D) versions—EQ-5D-3L and EQ-5D-5L—in assessing one-year quality of life outcomes for patients with knee osteoarthritis (KOA) undergoing unicompartmental knee arthroplasty (UKA).

Material and method: From the medical records at the Honghui Hospital, Xi'an Jiaotong University, 402 individuals aged 50 and above, who were one-year post-operation, were selected to fill out survey questionnaires during their return hospital visits. Of these, 231 respondents (57.5%) completed the questionnaire; 228 completed both versions, and 56 completed the EQ-5D retest questionnaire. The assessment included missing data, ceiling effects, informativity and discriminatory power, as well as response consistency, redistribution properties, and inconsistency. Reliability and validity were also evaluated.

Results: The results indicate that the EQ-5D-5L surpasses the EQ-5D-3L in construct validity, informativity, detection precision, and discriminatory power. Consistency reliability is also better in the EQ-5D-5L than in the EQ-5D-3L. Both instrument versions maintained reliable levels of test–retest reliability.

Conclusion: In patients with KOA undergoing UKA, the EQ-5D-5L has proven superior in measurement capabilities when compared with the EQ-5D-3L one-year post-operation. Thus, it is advised to utilize the EQ-5D-5L for ongoing assessments of quality of life in this specific group of patients.

KEYWORDS

EQ-5D-3L, EQ-5D-5L, comparing, patient-reported outcome measures, unicompartmental knee arthroplasty

1 Introduction

The EuroQol five dimensions questionnaire (EQ-5D), created by the EuroQol Group, incorporates five dimensions and is globally recognized as one of the foremost patient-reported outcome measures (PROMs) (1), extensively used in various countries. It employs distinct scoring algorithms across more than 20 nations (2, 3). This tool is essential for research in clinical and health services, economic analyses that compute cost per quality-adjusted life year, and has lately expanded its applications (4). Additionally, it provides an overarching evaluation of health-related quality of life and complements specific instruments for knee evaluations (5). The EQ-5D was originally available only in the EQ-5D-3L, which includes five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), with each dimension having three functional levels (no problems, some problems, extreme problems). Building upon the EQ-5D-3L, the EQ-5D-5L was subsequently developed (6). The EQ-5D-5L includes five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), with each dimension having five functional levels (no problems, slight problems, moderate problems, severe problems, extreme problems) (7, 8).

PROMs are increasingly popular in clinical assessments. To reduce patient burden while providing useful clinical information, clinicians should ensure the scales used are sensitive enough to changes and are also brief. The EQ-5D-3L has raised increasing doubts regarding its sensitivity to changes (5, 7, 9–12). Recent studies have indicated that the EQ-5D-5L, in comparison to the EQ-5D-3L, has shown improved sensitivity across various patient populations with different medical conditions. Zhu and colleagues (13) in the evaluation of EQ-5D across patients with breast cancer, colorectal cancer, or lung cancer, found that the EQ-5D-5L displayed improved performance in ceiling effects, reliability, and validity when compared to the EQ-5D-3L. In a study by Kontodimopoulos and colleagues (14), involving patients with rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis, and osteopenia/osteoporosis, the EQ-5D-5L demonstrated greater accuracy, consistency, and reliability compared to the EQ-5D-3L. For patients undergoing total knee replacement due to KOA, similar conclusions have been drawn regarding the responsiveness in preoperative and postoperative assessments (5, 10, 15–18), comparative studies on the construct validity of the EQ-5D-5L and EQ-5D-3L in this population have also shown that the EQ-5D-5L outperforms the EQ-5D-3L, for example, Greene and colleagues (10) along with Conner-Spady and colleagues (5) have established through their research that for individuals who have received total hip or knee replacements, the EQ-5D-5L offers improved construct validity compared to the EQ-5D-3L.

For patients undergoing UKA for unicompartmental osteoarthritis, the procedure is less invasive, with reduced bone loss and minimal blood loss. Compared to patients undergoing total knee arthroplasty, undergoing UKA patients experience a faster initial recovery and achieve better outcomes in functional results, pain assessment, revision rates, and complication incidence (19–21). Therefore, higher precision in assessment PROMs is required for evaluating this patient group. The measurement properties of both the EQ-5D-3L and EQ-5D-5L in patients undergoing UKA have already been demonstrated. However, only studies on the validity and responsiveness of individual versions exist. For example, Baryeh and colleagues (22) compared the relationship between patient satisfaction of UKA patients and preoperative and postoperative changes in EQ-5D-5L. Farrow and colleagues (23) compared the relationship between preoperative EQ-5D-3L scores and the extent of postoperative improvements in quality of life and joint function among patients who received total hip or knee replacements and those who underwent UKA.

Currently, no studies have compared the measurement properties of the two EQ-5D versions in patients undergoing UKA, and there is insufficient evidence to recommend either the EQ-5D-3L or EQ-5D-5L for assessing health status in this population. Therefore, this study aims to assess the measurement properties of both EQ-5D versions in evaluating one-year quality of life outcomes in patients post-UKA surgery.

Abbreviations: EQ-5D, EuroQol EQ-5D; BMI, Body Mass Index; SEM, Standard error of measurement; SDC, Smallest detectable change; ICC, Intraclass correlation coefficient; KOA, knee osteoarthritis; UKA, unicompartmental knee arthroplasty; PROMs, patient-reported outcome measures; SF-6D, Short Form-6 Dimension; WOMAC, Western Ontario and McMaster Universities Arthritis Index; VAS, visual analog scale.

2 Materials and methods

2.1 Data collection

In the medical records of the knee joint surgery department at the Honghui Hospital, Xi'an Jiaotong University, 402 patients aged 50 years or older who had undergone UKA and were one-year post-operation were selected. They were invited to complete questionnaires that incorporated both versions of EQ-5D and EQ-VAS, alongside general health-related measures like the Short Form-6 Dimension (SF-6D), and knee-specific PROMs, such as the Western Ontario and McMaster Universities Arthritis Index (WOMAC). Prior to questionnaire administration, the research objectives were clarified to the patients, and informed consent was secured. The sequence in which the two EQ-5D versions were presented was randomized, as was the order of domains within each EQ-5D in a subsequent retest. Exclusion criteria included illiteracy, lack of cognitive recognition, presence of severe chronic conditions such as malignant tumors, and cardiovascular or cerebrovascular diseases, along with psychiatric disorders. The study focused on data from patients who completed both EQ-5D versions for statistical purposes. Of the subjects approached, 231 individuals (57.5%) completed the questionnaires; 228 completed both versions. 96 respondents were randomly informed to complete a retest questionnaire 2 weeks later, and 56 (58%) completed the EQ-5D retest questionnaire.

The study adheres to the principles of the Declaration of Helsinki, and ethics approval was obtained from the medical ethics committee of Honghui Hospital, Xi'an Jiaotong University (No.202201039).

2.2 Instruments

The EQ-5D instrument encompasses five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The EQ-5D-3L articulates each dimension across three levels, depicting 243 distinct health states, usually denoted as health carriers from 11,111 (indicating perfect health) to 33,333 (representing the poorest health condition) (7, 9, 24). Conversely, the EQ-5D-5L details each dimension through five levels (7, 9), thereby delineating 3,125 distinct health states, conventionally reported from 11,111 (perfect health) to 55,555 (poorest health) (25). In addition to the EQ-5D-3L and EQ-5D-5L descriptive systems, the EQ-5D tool also includes a visual analog scale (VAS), EQ-VAS (2, 8), which is a separate component where health is self-rated on a 20-cm vertical scale, ranging from "Best imaginable health state" (100) to "Worst imaginable health state" (0) (8). Both the EQ-5D-3L and EQ-5D-5L are designed to be completed within 2 to 5 min (10). The Chinese value sets for both EQ-5D-3L and EQ-5D-5L have been developed, with the value range for EQ-5D-3L being [−0.149, 1.000] and for EQ-5D-5L being [−0.391, 1.000] (where 1 represents a state of perfect health) (26, 27).

The SF-6D is a universal PROMs, defined by six hierarchical health dimensions: physical functioning, role limitations, social functioning, pain, mental health, and vitality. Each dimension is divided into 4 to 6 levels (9). A six-digit code represents an SF-6D health state, derived by selecting a level from each dimension, commencing with physical functioning and culminating in vitality (9). The code 111,111 denotes optimal health, while 645,655 represents the least favorable health condition (28).

The WOMAC is a 24-item knee joint disease-specific functional measure that encompasses three domains: pain, stiffness, and physical function (29–31). Each of the 24 questions is rated using either a Likert 5-point scale or a 100-mm VAS, with scores ranging from "none or 0" to "extreme or 100." In this investigation, the VAS from the Fajardo version was employed. Patients are instructed on VAS usage before answering the questions (30, 31). Section A scores range from 0 to 500, Section B from 0 to 200, and Section C from 0 to 1700 (30, 31). Higher scores reflect more pain, stiffness, and poorer physical function (30, 31).

2.3 Statistical analysis

To assess missing data, the collected dataset is analyzed to compare the two versions of EQ-5D, focusing only on samples with complete data for both versions in the statistical analysis. The evaluation includes examining the five items of EQ-5D, the index values, and the EQ-VAS. Shannon's indices (H') and evenness index (J') are used to evaluate the informativity and discriminatory power of the two EQ-5D versions (32). Shannon index (H') expresses the absolute amount of informativity captured in a system. In this formula, H' is calculated as follows:

$$H' = -\sum_{i=1}^{L} p_i \log_2 p_i$$

Where L denotes the number of response levels in a given dimension, and p_i represents the proportion of observations at the i -th level ($i = 1, \ldots, L$). As H' increases, more information is captured.

J' is the evenness index, representing the ratio of the observed H' to the maximum possible $H_{\max}$. The formula for J' is:

$$J' = \frac{H_{\max}}{H'}$$

The maximum amount of information is obtained when the responses are uniformly distributed across all levels. Under these conditions, H' reaches its maximum value $H_{\max}$ (1.58 for EQ-5D-3L and 2.32 for EQ-5D-5L) (16).

J' is used to assess the uniformity of information distribution, ranging from 0 to 1, with values closer to 1 indicating a more even distribution of information (16).

Previous studies have demonstrated variations within specific dimensions of EQ-5D, where responses from the EQ-5D-3L deviate by two or more levels compared to the EQ-5D-5L. The 'New' response classification signifies a one-level difference between the questionnaire versions, whereas an 'Inconsistent' response is characterized by alterations spanning two or more severity levels between the assessments (28, 33). Thus, the magnitude of inconsistency spans from 1 (representing a discrepancy of two levels) to 3 (indicating a discrepancy of four levels) (11). Calculations are performed to determine both the proportion of consistent versus inconsistent response pairs between the EQ-5D-3L and EQ-5D-5L and the average magnitude of these inconsistencies (11).

Compare the SDC for groups and individuals between the two versions of EQ-5D. The SDC for groups represents the minimum change necessary to identify a true difference within a population, whereas the SDC for individuals reflects the smallest detectable change for an individual. By comparing the two versions of EQ-5D, this approach enables assessment of the sensitivity of each version in detecting clinically meaningful changes. The SDC_{groups} and $SDC_{individuals}$, representing the smallest change in value for a scale that can be considered an actual change rather than a measurement error, are calculated using the following formula: $SDC_{individuals} = 1.96 \cdot \sqrt{2} \cdot SEM$, and $SDC_{group} = SDC_{individuals} / \sqrt{n}$. Where $SEM = SD \cdot \sqrt{(1-R)}$, SD = the sample standard deviation, R = the calculated Intraclass correlation coefficient (ICC), and n is the sample size (34).

To assess the test–retest reliability of the EQ-5D tools, ICC are commonly employed, as they gauge the level of agreement or score stability across time. Test–retest reliability examines the consistency with which the EQ-5D-3L and EQ-5D-5L instruments produce similar outcomes when administered to the same patient group at two distinct time points (35). An ICC value of 0.70 or above is considered indicative of excellent test–retest reliability (35).

The internal consistency of the EQ-5D-3L and EQ-5D-5L versions is independently evaluated for each version through the application of Cronbach's alpha, which serves as a standard metric for assessing internal reliability. This coefficient measures the extent to which the items within each version align as a cohesive group, thereby representing the proportion of variance in scores that can be attributed to differences among individuals rather than inconsistencies within the instrument. A comparative analysis of the Cronbach's alpha values between the two versions enables an assessment of their respective internal consistency (36). A Cronbach's alpha value above 0.7 denotes acceptable consistency, whereas values below 0.7 suggest poor consistency (1).

According to the Consensus-based Standards for the Selection of Health Measurement Instruments (35), examine the construct validity of both iterations of EQ-5D against the EQ-VAS, SF-6D, and WOMAC by means of hypothesis testing, juxtaposing the scores [assessing hypotheses via Spearman's correlation (Rho)] (16). The magnitudes of these coefficients are interpreted as small correlation (Rho < 0.20), moderate correlation (0.20 ≤ Rho < 0.50), or high correlation (Rho ≥ 0.50). A comparison of the ceiling effects of the two versions is conducted. Given the focus on the postoperative recovery period in this survey, it is expected that there will be more ceiling effects and fewer floor effects. However, comparing the proportion of reduction in ceiling effects remains meaningful.

All statistical analyses were conducted using IBM SPSS Statistics version R26.0.0.0.

TABLE 1 Patient characteristics.

Variable	%
Age ≥ 60	178 (78.1%)
Female	164 (71.9%)
Male	64 (28.1%)
Years disease duration	
Years ≤ 5	66 (28.9%)
5 < Years ≤ 10	102 (44.8%)
Years > 10	60 (26.3%)
Surgery joint side	
Right	107 (46.9%)
Left	121 (53.1%)
Educational level	
Primary or secondary	210 (92.1%)
Tertiary	18 (7.9%)
Living status	
Alone	6 (2.6%)
With spouse	144 (63.2%)
With children	73 (32.2%)
BMI (kg/m²)	
BMI ≤ 23.9 (underweight or healthy)	64 (28.1%)
23.9 < BMI ≤ 27.9 (overweight)	107 (45.0%)
BMI > 27.9 (obese)	57 (25.0%)
Occupation	
Heavy manual worker	17 (7.4%)
Light manual worker	17 (7.4%)
Agricultural	24 (10.5%)
Retired	170 (74.6%)

3 Results

3.1 Data quality

Of the 402 patients who received invitations, 231 respondents (57.5%) completed the survey. Table 1 presents the characteristics of the 228 respondents who completed both versions of EQ-5D. According to widely accepted and recommended measurement standards (37), the sample size is adequate, meeting the Consensus-based Standards for the Selection of Health Measurement Instruments guidelines, which recommend at least 100 participants for studies assessing construct validity. Of the 96 patients who were randomly informed, 56 (58.3%) completed the retest questionnaire for both versions of EQ-5D. Table 1 presents the patient characteristics.

3.2 Informativity and discriminatory power

Figure 1 shows that Shannon's H' ranges from 0.429 (Anxiety/depression) to 0.773 (Pain/discomfort) for the EQ-5D-3L and from 0.752 (Anxiety/depression) to 1.898 (Usual activity) for the EQ-5D-5L. The J' ranges from 0.271 (Anxiety/depression) to 0.489 (Pain/discomfort) for EQ-5D-3L and from 0.475 (Anxiety/depression) to 0.676 (Usual activity) for EQ-5D-5L. As shown in Figure 1, overall, in terms of informativity and discriminatory power across all dimensions, the EQ-5D-5L consistently demonstrates higher H' and J' values compared to the EQ-5D-3L, indicating that the EQ-5D-5L provides greater informativity and discriminatory power than the EQ-5D-3L.

3.3 Re-allocation of attributes and inconsistency

Table 2 shows the redistribution properties of responses from EQ-5D-3L to EQ-5D-5L, along with the number of consistent and inconsistent response pairs. The table provides a cross-tabulation of dimension scores. In the reallocation characteristics of responses from the EQ-5D-3L to the EQ-5D-5L, there were 228 evaluations of both EQ-5D-3L and EQ-5D-5L, producing 1,128 response pairs, among which there were 12 inconsistent responses appearing in 12 out of the 228 evaluations (5.3%). Two respondents selected "1" indicating no problems in the EQ-5D-3L Usual activities dimension, but selected "3" indicating moderate problems in the EQ-5D-5L Usual activities. Six respondents chose "2" indicating some problems in the EQ-5D-3L Mobility dimension, but selected "1" indicating no problems in the EQ-5D-5L Mobility. Two respondents selected "2" in the EQ-5D-3L Anxiety/depression dimension, but selected "1" indicating no problems in the EQ-5D-5L Anxiety/depression. Two respondents chose "3" indicating extreme problems in the EQ-5D-3L Pain/discomfort dimension, but selected "2" indicating moderate problems in the EQ-5D-5L Pain/discomfort. The rate of inconsistency in responses from EQ-5D-3L to EQ-5D-5L: Mobility and Self-care had an inconsistency rate of 2.6%. Usual activities, Pain/discomfort, and Anxiety/depression had an inconsistency rate of 0.9% (Figure 2).

3.4 Reliability

Figure 3 shows that the SDC at the individual level is 0.191 for EQ-5D-3L and 0.133 for EQ-5D-5L, while at the group level, it is 0.013 for EQ-5D-3L and 0.009 for EQ-5D-5L. The measurement accuracy of EQ-5D-5L is higher than that of EQ-5D-3L. Figure 4 shows the test–retest reliability of EQ-5D, evaluated according to recommended and widely used measurement standards (ICC > 0.7) (37), indicating good test–retest reliability. The Cronbach's alpha for within-group correlation is 0.881 for EQ-5D-5L and 0.710 for EQ-5D-3L.

3.5 Validity

Figure 5 shows the correlations between EQ-5D-3L and EQ-5D-5L dimensions, index scores, WOMAC, SF-6D, and EQ-VAS scores. All correlations were statistically significant with $p \le 0.01$. The Spearman correlation coefficients between the EQ-5D-3L dimensions and EQ-VAS range from 0.26 to 0.49, and for the EQ-5D-5L dimensions, from 0.41 to 0.72. Correspondingly, the correlation between the EQ-5D index and EQ-VAS are 0.63 and 0.77, respectively.

10.3389/fmed.2024.1451979

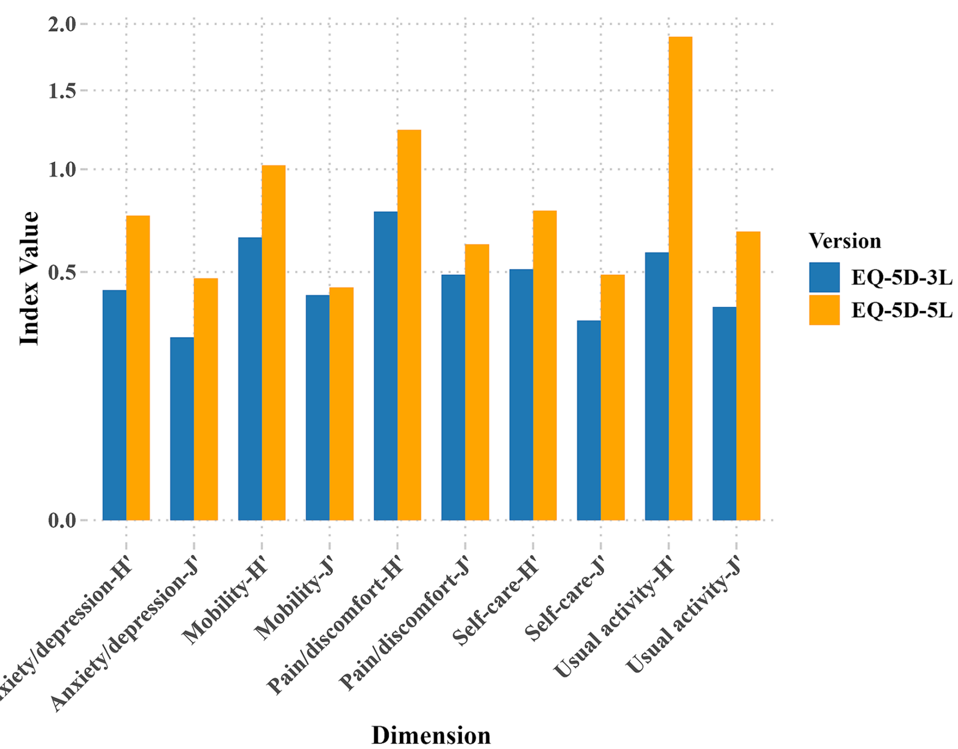

FIGURE 1
Informativity and discriminatory power of EQ-5D: results for H' and J'.

Compared to the EQ-5D-3L, all dimensions and index scores of the EQ-5D-5L show a high correlation with EQ-VAS. In terms of Spearman correlation coefficients with SF-6D, the EQ-5D-5L index and dimensions generally have higher correlations than the EQ-5D-3L, ranging from 0.30 to 0.72 for EQ-5D-5L, compared to 0.22 to 0.55 for EQ-5D-3L. In the correlation analysis with WOMAC, the EQ-5D-5L dimensions generally showed higher correlations than the EQ-5D-3L dimensions. The EQ-5D-5L index score correlated more strongly with WOMAC than the EQ-5D-3L index, which showed moderate correlations with stiffness and function (0.48 and 0.49 respectively), and was highly correlated with pain and the overall score (0.50 and 0.53 respectively). The EQ-5D-5L index was highly correlated with all WOMAC dimensions (from 0.62 to 0.76). Overall, the EQ-5D-5L demonstrates higher correlations with both the overall and specific domain scores of WOMAC and SF-6D compared to the EQ-5D-3L. The same is true for their correlation with EQ-VAS. This indicates that the EQ-5D-5L has better construct validity. Compared to the EQ-5D-3L index, the EQ-5D-5L index demonstrates a 24.6% reduction in the ceiling effect.

4 Discussion

This study compares the psychometric properties of the EQ-5D-5L and EQ-5D-3L in KOA undergoing UKA. Compared to the EQ-5D-3L, the EQ-5D-5L showed superior performance across all dimensions in terms of absolute informativity (H') and the uniformity of information distribution (J'). Consequently, the EQ-5D-5L exhibited greater informativity and discriminatory power. Both the EQ-5D-3L and EQ-5D-5L demonstrated good reliability. In terms of validity, the EQ-5D-5L also outperformed the EQ-5D-3L. Our study suggests that the EQ-5D-5L is recommended for assessing quality of life in assessing one-year quality of life outcomes for patients with KOA undergoing UKA.

Current research indicates that the Shannon's H' index for the dimensions of EQ-5D-3L and EQ-5D-5L ranges from 0.429 (Anxiety/Depression) to 0.773 (Pain/Discomfort) and from 0.752 (Anxiety/Depression) to 1.898 (Usual Activities), respectively. The J' index for the EQ-5D-3L and EQ-5D-5L dimensions ranges from 0.271 (Anxiety/Depression) to 0.489 (Pain/Discomfort) and from 0.475 (Anxiety/Depression) to 0.676 (Usual Activities). Overall, the EQ-5D-5L version exhibits greater Informativity and discriminatory power than the EQ-5D-3L. Compared to the EQ-5D-3L, the EQ-5D-5L adds two levels per dimension, expanding its sensitivity and applicability. This enhancement results in greater informativity and improved discriminatory power. This finding showing a strong consistency with the findings reported in other studies. Conner-Spady and colleagues (5) examined the informativity and discriminatory power of the EQ-5D in a cohort of 176 patients undergoing hip or knee replacement surgery due to osteoarthritis. Their findings showed that the EQ-5D-5L demonstrated greater informativity and improved discriminatory power compared to the EQ-5D-3L. Similarly, Michalowsky and colleagues (11) compared the psychometric properties of the EQ-5D-3L and EQ-5D-5L based on proxy ratings by informal caregivers and health professionals for patients with dementia. Their results also indicated that the EQ-5D-5L exhibited greater informativity and superior discriminatory power relative to the EQ-5D-3L.

TABLE 2 Redistribution properties from EQ-5D-3L to EQ-5D-5L responses.

EQ-5D-3L Dimensions		EQ-5D-5L 1	2	3	4	5	Consistent n	%	Inconsistent n	%
1	Mobility	168	22	**0**	**0**	**0**	190	100%	0	0%
	Self-care	192	10	**0**	**0**	**0**	202	100%	0	0%
	Usual activities	170	28	**2**	**0**	**0**	198	99.0%	2	1.0%
	Pain/discomfort	96	90	**0**	**0**	**0**	186	100%	0	0%
	Anxiety/depression	186	22	**0**	**0**	**0**	208	100%	0	0%
2	Mobility	**6**	18	12	2	**0**	32	84.2%	6	15.8%
	Self-care	**0**	20	6	0	**0**	26	100%	0	0%
	Usual activities	**0**	20	6	0	**0**	26	100%	0	0%
	Pain/discomfort	**0**	32	6	0	**0**	38	100%	0	0%
	Anxiety/depression	**2**	14	4	0	**0**	18	90.0%	2	10.0%
3	Mobility	**0**	**0**	**0**	0	0	0	0%	0	0%
	Self-care	**0**	**0**	**0**	0	0	0	0%	0	0%
	Usual activities	**0**	**0**	**0**	2	0	2	100%	0	0%
	Pain/discomfort	**0**	**0**	**2**	2	0	2	50.0%	2	50.0%
	Anxiety/depression	**0**	**0**	**0**	0	0	0	0%	0	0%
Consistent response pair	n	812	276	34	6	0	1,128	98.9%	–	–
	%	71.2%	24.2%	3.0%	0.5%	0%				
Inconsistent response pair	n	8	0	4	0	0	–	–	12	1.1%
	%	0.7%	0%	0.4%	0%	0%				
Total response pair	n	820	276	38	6	0	1,140 (100%)			

Inconsistencies (marked in bold) are defined as differences of more than one between EQ-5D-3L and EQ-5D-5L response categories.
Redistribution properties from EQ-5D-3L to EQ-5D-5L responses of $n = 228$ EQ-5D-3L and EQ-5D-5L assessments, generating 1,128 response pairs 12 inconsistent pairs arise out of 12 of 228 assessments (5.3%).

According to the recommended and widely applied measurement standards (37), both versions of the EQ-5D have good test–retest reliability. No indications of consistent disparities between the initial and subsequent assessment outcomes are observed. The Cronbach's alpha for within-group correlation is 0.881 for EQ-5D-5L and 0.710 for EQ-5D-3L, indicating good internal consistency for both versions of EQ-5D (Cronbach's alpha >0.7). The SDC at the group level for the EQ-5D-3L and EQ-5D-5L are 0.133 and 0.009, respectively, and at the individual level, they are 0.191 and 0.013, respectively. The EQ-5D-5L's detection accuracy is higher than that of the EQ-5D-3L. These results are consistent with previous research by Garratt and colleagues (16) who examined patients undergoing surgical fixation for closed ankle fractures. In their study, the SDC values for the EQ-5D-5L index were reported as 0.02 at the group level and 0.20 at the individual level, underscoring the EQ-5D-5L's greater sensitivity for detecting subtle changes in health status across both levels. Both studies demonstrate the sensitivity advantage of the EQ-5D-5L across different health states and patient populations. This finding further validates the reliability and applicability of the EQ-5D-5L in detecting subtle health changes across diverse populations.

Our findings indicated that the EQ-5D-5L demonstrated stronger construct validity coefficients with other health-related measures compared to the EQ-5D-3L, highlighting its enhanced sensitivity in evaluating postoperative quality of life. Similarly, Conner-Spady and colleagues (5) assessed the construct validity of the EQ-5D among 176 patients undergoing hip or knee replacement for osteoarthritis and found that the EQ-5D-5L showed better alignment with related health measures than the EQ-5D-3L, consistent with our findings. Greene and colleagues (10) further examined a cohort of total hip arthroplasty patients, observing that the EQ-5D-5L provided a more refined assessment of health-related quality of life compared to the EQ-5D-3L. Eneqvist and colleagues (38) in their study of Swedish total hip replacement patients, also confirmed that the EQ-5D-5L demonstrated superior construct validity coefficients.

Given the survey focus on one-year follow-up, it is expected that a significant portion of participants will reach optimal levels of health status. In this study, the ceiling effect observed for the EQ-5D-5L index decreased by 24.6% compared to the EQ-5D-3L index. Although systematic reviews comparing the EQ-5D-3L and EQ-5D-5L do not include populations with one-year follow-up, they have identified more substantial differences in ceiling effects, with the highest disparities reaching up to 30% (32). Recent findings have indicated a noteworthy decrease in the ceiling effect of the EQ-5D-5L when compared to the EQ-5D-3L in multiple recent investigations (5, 10, 11, 18, 39).

This study's focus on the postoperative recovery phase likely contributed to an elevated ceiling effect compared to other stages of disease progression. A key limitation is the absence of responsiveness assessments; future studies should include these to

10.3389/fmed.2024.1451979

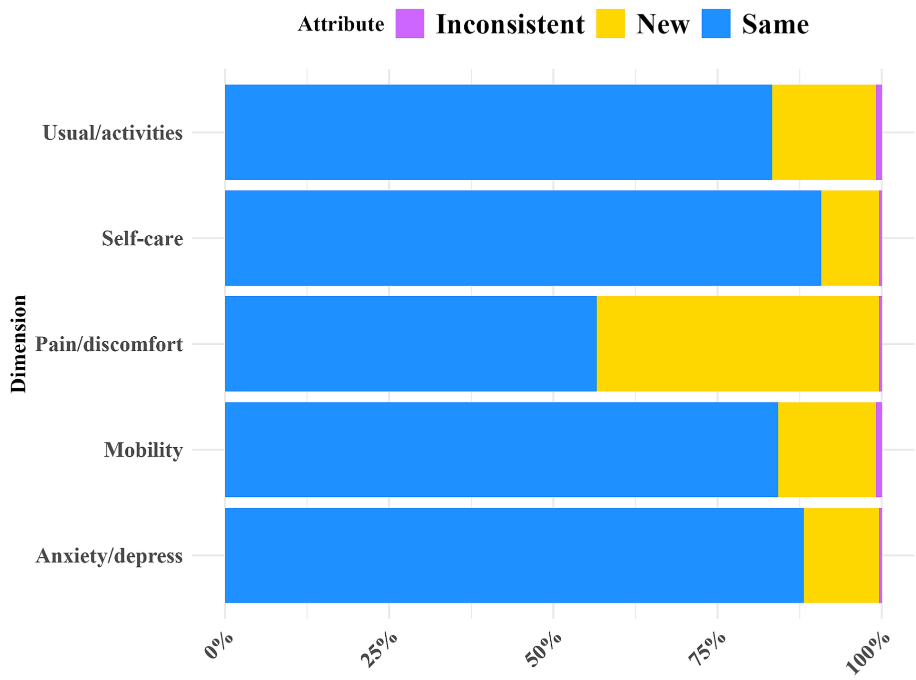

FIGURE 2
Redistribution properties from EQ-5D-3L to EQ-5D-5L responses.

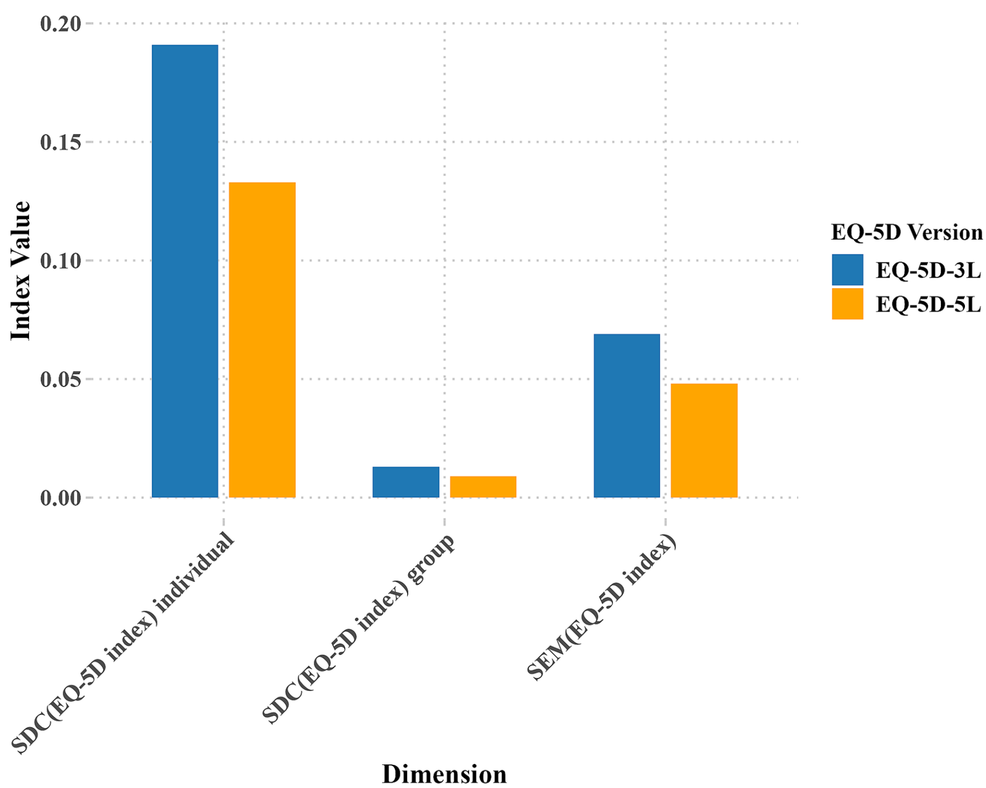

FIGURE 3
Comparison of EQ-5D index SDC and SEM results.

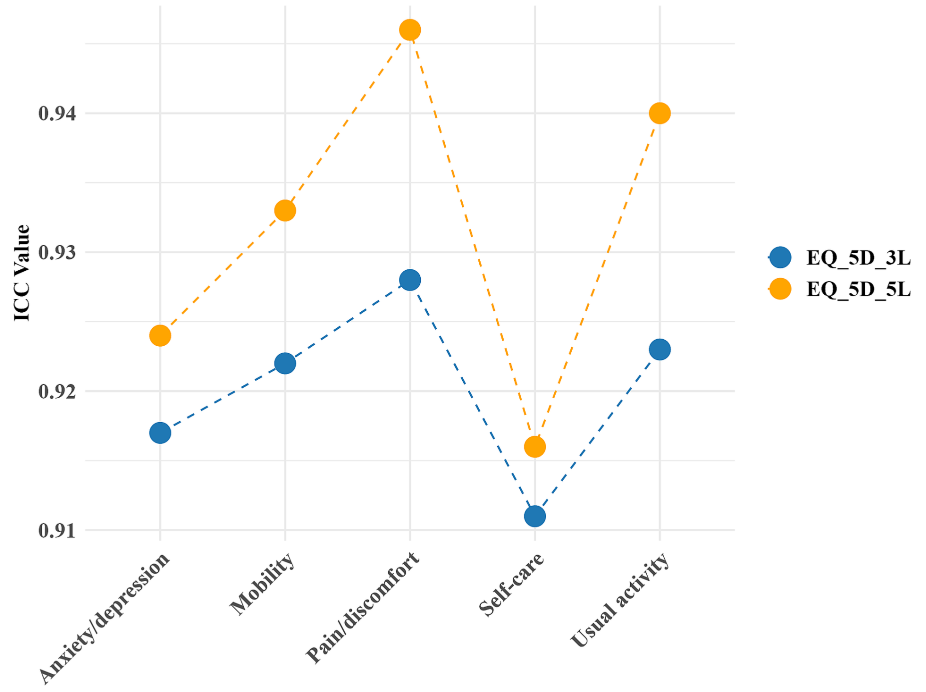

FIGURE 4
The ICC for test–retest reliability of EQ-5D.

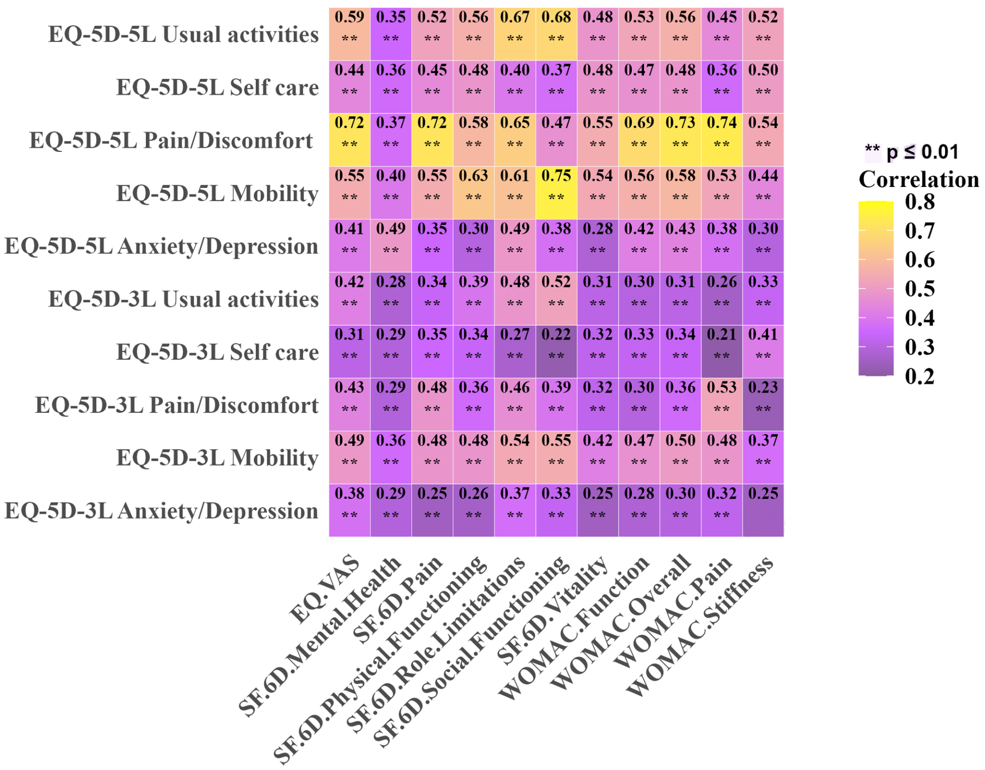

FIGURE 5
Spearman's correlation for EQ-5D at the dimension level with WOMAC, SF-6D, and EQ-VAS. ***p* ≤ 0.01.

assess the instruments' sensitivity to change over time. As a single-center study with exclusively Chinese participants, however, the generalizability of these findings may be limited. Conducting research across multiple centers and diverse ethnic populations could provide a more comprehensive view of EQ-5D-3L and EQ-5D-5L performance in varied healthcare contexts. Additionally, extending follow-up beyond one year could further elucidate the EQ-5D's capacity to reflect one-year quality of life changes. Finally, including other PROMs alongside EQ-5D would allow for a multidimensional assessment of health-related quality of life, and could help validate EQ-5D's construct validity through comparisons with other established health-related quality of life instruments.

5 Conclusion

Overall, in patients with KOA undergoing UKA, one-year post-operation, both versions were within acceptable performance ranges, with the EQ-5D-5L demonstrating superior performance. Therefore, the EQ-5D-5L is recommended for assessing quality of life in this patient group.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Medical Ethics Committee of Honghui Hospital, Xi'an Jiaotong University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

ZT: Formal analysis, Methodology, Visualization, Writing – original draft, Data curation. DW: Formal analysis, Writing – original draft. QZ: Writing – review & editing. YH: Formal analysis, Writing – original draft. CX: Formal analysis, Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Key Research and Development Program of Shaanxi Province (Grant no. 2023-YBSF-464); General Research Plan of Xi'an Health Commission (Grant no. 2023yb30); and the National Natural Science Foundation of China (Grant no. 82173627).

Acknowledgments

We sincerely thank Jianbing Ma for his contributions to the statistical design and writing approach of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1451979/full#supplementary-material

References

1. Xu RH, Keetharuth AD, Wang LL, Cheung AW, Wong EL. Measuring health-related quality of life and well-being: a head-to-head psychometric comparison of the EQ-5D-5L, ReQoL-UI and ICECAP-a. *Eur J Health Econ*. (2022) 23:165–76. doi: 10.1007/s10198-021-01359-0

2. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. *Appl Health Econ Health Policy*. (2017) 15:127–37. doi: 10.1007/s40258-017-0310-5

3. Garratt AM, Hansen TM, Augestad LA, Rand K, Stavem K. Norwegian population norms for the EQ-5D-5L: results from a general population survey. *Qual Life Res*. (2022) 31:517–26. doi: 10.1007/s11136-021-02938-7

4. Solberg TK, Olsen JA, Ingebrigtsen T, Hofoss D, Nygaard OP. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J*. (2005) 14:1000–7. doi: 10.1007/s00586-005-0898-2

5. Conner-Spady BL, Marshall DA, Bohm E, Dunbar MJ, Loucks L, Al KA, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. (2015) 24:1775–84. doi: 10.1007/s11136-014-0910-6

6. Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. *Value Health*. (2014) 17:70–6. doi: 10.1016/j.jval.2013.10.006

7. Garratt AM, Furunes H, Hellum C, Solberg T, Brox JI, Storheim K, et al. Evaluation of the EQ-5D-3L and 5L versions in low back pain patients. *Health Qual Life Outcomes*. (2021) 19:155. doi: 10.1186/s12955-021-01792-y

8. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. (2011) 20:1727–36. doi: 10.1007/s11136-011-9903-x

9. Olariu E, Caplescu R, Vale L, Niculescu-Aron IG, Oluboyede Y, Paveliu MS. Population norms for the EQ-5D-3L and EQ-5D-5L in Romania. *Health Qual Life Outcomes*. (2023) 21:80. doi: 10.1186/s12955-023-02144-8

10. Greene ME, Rader KA, Garellick G, Malchau H, Freiberg AA, Rolfson O. The EQ-5D-5L improves on the EQ-5D-3L for health-related quality-of-life assessment in patients undergoing total hip arthroplasty. *Clin Orthop Relat Res*. (2015) 473:3383–90. doi: 10.1007/s11999-014-4091-y

11. Michalowsky B, Hoffmann W, Mohr W, Radke A, Xie F. Comparing the psychometric properties of EQ-5D-3L and EQ-5D-5L proxy ratings by informal caregivers and a health professional for people with dementia. *Health Qual Life Outcomes*. (2022) 20:140. doi: 10.1186/s12955-022-02049-y

12. Scalone L, Ciampichini R, Fagiuoli S, Gardini I, Fusco F, Gaeta L, et al. Comparing the performance of the standard EQ-5D 3L with the new version EQ-5D 5L in patients with chronic hepatic diseases. *Qual Life Res*. (2013) 22:1707–16. doi: 10.1007/s11136-012-0318-0

13. Zhu J, Yan XX, Liu CC, Wang H, Wang L, Cao SM, et al. Comparing EQ-5D-3L and EQ-5D-5L performance in common cancers: suggestions for instrument choosing. *Qual Life Res*. (2021) 30:841–54. doi: 10.1007/s11136-020-02636-w

14. Kontodimopoulos N, Stamatopoulou E, Gazi S, Moschou D, Krikelis M, Talias MA. A comparison of EQ-5D-3L, EQ-5D-5L, and SF-6D utilities of patients with musculoskeletal disorders of different severity: a health-related quality of life approach. *J Clin Med*. (2022) 11:4097. doi: 10.3390/jcm11144097

15. Bilbao A, Garcia-Perez L, Arenaza JC, Garcia I, Ariza-Cardiel G, Trujillo-Martin E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res*. (2018) 27:2897–908. doi: 10.1007/s11136-018-1929-x

16. Garratt A, Stavem K. Comparison of EQ-5D-3L and 5L versions following operative fixation of closed ankle fractures. *Qual Life Res*. (2022) 31:2071–82. doi: 10.1007/s11136-022-03105-2

17. Jin X, Al SF, Ohinmaa A, Marshall DA, Johnson JA. Responsiveness of the EQ-5D-3L and EQ-5D-5L in patients following total hip or knee replacement. *Qual Life Res*. (2019) 28:2409–17. doi: 10.1007/s11136-019-02200-1

18. Garcia-Perez L, Ramos-Garcia V, Serrano-Aguilar P, Pais-Brito JL, Aciego DMM, Martin-Fernandez J, et al. EQ-5D-5L utilities per health states in Spanish population with knee or hip osteoarthritis. *Health Qual Life Outcomes*. (2019) 17:164. doi: 10.1186/s12955-019-1230-x

19. Hao P, Yang L, He R, Chen H, Sun M, Liang S. Analysis of gait and effectiveness after unicompartmental knee arthroplasty. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi*. (2020) 34:1369–75. doi: 10.7507/1002-1892.202004076

20. Radhakrishnan GT, Magan A, Kayani B, Asokan A, Ronca F, Haddad FS. Return to sport after unicompartmental knee arthroplasty: a systematic review and meta-analysis. *Orthop J Sports Med*. (2022) 10:951684341. doi: 10.1177/23259671221079285

21. Papalia R, Zampogna B, Torre G, Diaz BL, Vasta S, Papalia G, et al. Return to sport activity in the elderly patients after unicompartmental knee arthroplasty: a systematic review and Meta-analysis. *J Clin Med*. (2020) 9:1756. doi: 10.3390/jcm9061756

22. Baryeh K, Maillot C, Gummaraju A, Riviere C. Disappointing relationship between functional performance and patient satisfaction of UKA patients: a cross sectional study. *Orthop Traumatol Surg Res*. (2021) 107:102865. doi: 10.1016/j.otsr.2021.102865

23. Farrow L, Redmore J, Talukdar P, Clement N, Ashcroft GP. Prioritisation of patients awaiting hip and knee arthroplasty: lower pre-operative EQ-5D is associated with greater improvement in quality of life and joint function. *Musculoskeletal Care*. (2022) 20:892–8. doi: 10.1002/msc.1645

24. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med*. (2001) 33:337–43. doi: 10.3109/07853890109002087

25. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. (2012) 15:708–15. doi: 10.1016/j.jval.2012.02.008

26. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. *Value Health*. (2017) 20:662–9. doi: 10.1016/j.jval.2016.11.016

27. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. *Value Health*. (2014) 17:597–604. doi: 10.1016/j.jval.2014.05.007

28. Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. *Value Health*. (2008) 11:275–84. doi: 10.1111/j.1524-4733.2007.00230.x

29. Feinstein A, Shen L, Rose J, Cayer C, Bockus C, Meza C, et al. A French version of a voice recognition symbol digit modalities test analog. *Can J Neurol Sci*. (2023) 50:925–8. doi: 10.1017/cjn.2022.343

30. Faucher M, Poiraudeau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Assessment of the test-retest reliability and construct validity of a modified WOMAC index in knee osteoarthritis. *Joint Bone Spine*. (2004) 71:121–7. doi: 10.1016/S1297-319X(03)00112-X

31. Xie F, Li SC, Goeree R, Tarride JE, O'Reilly D, Lo NN, et al. Validation of Chinese Western Ontario and McMaster universities osteoarthritis index (WOMAC) in patients scheduled for total knee replacement. *Qual Life Res*. (2008) 17:595–601. doi: 10.1007/s11136-008-9340-7

32. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *PharmacoEconomics*. (2018) 36:645–61. doi: 10.1007/s40273-018-0642-5

33. Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. *Med Care*. (2007) 45:259–63. doi: 10.1097/01.mlr.0000254515.63841.81

34. van Kampen DA, Willems WJ, van Beers LW, Castelein RM, Scholtes VA, Terwee CB. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). *J Orthop Surg Res*. (2013) 8:40. doi: 10.1186/1749-799X-8-40

35. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. (2010) 19:539–49. doi: 10.1007/s11136-010-9606-8

36. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. (2011) 2:53–5. doi: 10.5116/ijme.4dfb.8dfd

37. Terwee CB, Prinsen CA, Ricci GM, Suman A, de Vet HC, Mokkink LB. The quality of systematic reviews of health-related outcome measurement instruments. *Qual Life Res*. (2016) 25:767–79. doi: 10.1007/s11136-015-1122-4

38. Eneqvist T, Nemes S, Karrholm J, Burstrom K, Rolfson O. How do EQ-5D-3L and EQ-5D-5L compare in a Swedish total hip replacement population? *Acta Orthop*. (2020) 91:272–8. doi: 10.1080/17453674.2020.1746124

39. Bhadhuri A, Kind P, Salari P, Jungo KT, Boland B, Byrne S, et al. Measurement properties of EQ-5D-3L and EQ-5D-5L in recording self-reported health status in older patients with substantial multimorbidity and polypharmacy. *Health Qual Life Outcomes*. (2020) 18:317. doi: 10.1186/s12955-020-01564-0

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 16 January 2025
DOI 10.3389/fpsyg.2024.1497422

Check for updates

OPEN ACCESS

EDITED BY
Paola Gremigni,
University of Bologna, Italy

REVIEWED BY
Mario Alberto Trógolo,
Siglo 21 Business University, Argentina
Biao Zeng,
Beijing Normal University, China

*CORRESPONDENCE
Jian Chen
✉ clegendj@aliyun.com

RECEIVED 17 September 2024
ACCEPTED 29 November 2024
PUBLISHED 16 January 2025

CITATION
Chen J, Chen XY and Miao C (2025) Validation of existential fulfillment scale in Chinese university students.
Front. Psychol. 15:1497422.
doi: 10.3389/fpsyg.2024.1497422

Validation of existential fulfillment scale in Chinese university students

Jian Chen*, Xiaoyan Chen and Chen Miao

School of Psychology, Fujian Normal University, Fuzhou, Fujian, China

Introduction: This study aimed to assess the psychometric properties of the Existential Fulfillment Scale (EFS) in a Chinese university student sample, emphasizing the cultural fit of the scale.

Methods: A cohort of 1,600 undergraduate students from six universities in Fujian Province completed questionnaires including the EFS, Meaning of Life Questionnaire (MLQ), Index of Well-Being (IWB), and Self-Depression Scale (SDS). We conducted item analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and assessments of criterion-related validity, internal consistency, and test-retest reliability.

Results: The Chinese EFS consists of two dimensions—self-acceptance and self-breakthrough—across 14 items, reflecting cultural distinctions from the original model by combining the dimensions of self-actualization and self-transcendence. This revised structure aligns with Chinese cultural perspectives on individual growth, where self-actualization often integrates aspects of self-transcendence. The scale showed positive associations with the MLQ and IWB and a negative association with the SDS, supporting the scale's criterion-related validity. Internal consistency ranged from 0.87 to 0.97, and test-retest reliability ranged from 0.75 to 0.83.

Discussion: These findings indicate that the Chinese EFS is a reliable tool for assessing existential fulfillment among Chinese university students.

KEYWORDS

existential fulfillment, reliability, validity, Chinese university students, confirmatory analysis (CFA)

1 Introduction

1.1 Conception of existential fulfillment

Existential fulfillment is a central construct in existential psychology and stands as a foundational element of Viktor Frankl's theory of the "Existential Vacuum" and logotherapy (Loonstra et al., 2007). Frankl (1963) described the existential vacuum as a condition marked by an absence of meaning and purpose, which leads individuals to experience profound unease, emotional detachment, and a pervasive sense of life's meaninglessness (Frankl, 1963; Yalom, 1980). This existential void is often accompanied by feelings of insecurity and desolation, as individuals struggle to find direction.

In contrast to the existential vacuum, existential fulfillment embodies a life enriched with significance and purpose (Längle et al., 2003; Loonstra et al., 2007). Frankl (1962) argued that the search for meaning is the most profound motivational force for humans, achieved through the pursuit and embodiment of life's inherent values. This drive, termed the "will to meaning," reflects humanity's inherent quest to realize their potential and discover value in their existence (Frankl, 1962). Expanding on this, Rollo May introduced the concept of a "sense of existence" or "sense of being," emphasizing the experiential dimension of existential fulfillment. May's approach highlighted key characteristics of existential experience: self-centering, self-affirmation, engagement, awareness, self-consciousness, and anxiety (May et al., 1958; Yang, 1999).

Clinical and empirical research in psychology consistently highlights the central importance of existential fulfillment, recognizing it not only as a core aspect of personality but also as a crucial determinant of mental health. A wealth of studies have demonstrated strong correlations between existential fulfillment and various indicators of psychological well-being, including emotional stress levels, neuroticism, burnout, work engagement, internet addiction, vitality and self-awareness of health (Tomic and Tomic, 2008; Mausch and Rys, 2020; Längle, 2011; Längle et al., 2003; Tomic and Tomic, 2011; Loonstra et al., 2010; Levchenko, 2024; Mikhaylova and Dmitrieva, 2021; Riethof and Bob, 2019). Individuals with a strong sense of existential fulfillment exhibit an enhanced capacity for value-based decision-making, personal goal setting, and extracting meaning from life (Yang, 1999). Conversely, those experiencing an existential vacuum face a heightened risk of mental health challenges, such as depression, anxiety, and despair, which can lead to a loss of creativity, freedom, and willpower, and, in severe cases, contribute to disorders such as schizophrenia (Längle et al., 2003; Frankl, 1962; Ratcliffe, 2009). Echoing these sentiments, Rollo May posits that the erosion of existential fulfillment is central to mental illness, advocating for its restoration as a primary aim in psychotherapy (May et al., 1958; Yang, 1999).

1.2 Measurement of existential fulfillment

Drawing upon the foundational principles of Viktor Frankl's logotherapy and the concept of existential vacuum, Längle et al. (2003) initially defined existential fulfillment as a multidimensional construct, encompassing fundamental elements of human reality: self-distance (perception), self-transcendence (value recognition), freedom (decision-making), and responsibility (acting). The Existence Scale (ES), developed by Längle et al. (2003), was designed to assess these existential competencies, evaluating individuals' capacity to engage meaningfully with themselves and the world.

The ES comprises 46 items across four subscales—self-distance, self-transcendence, freedom, and responsibility—and utilizes a six-point Likert scale ranging from "fully disagree" to "fully agree." Although its theoretical foundation was robust, confirmatory factor analysis (CFA) did not confirm the ES's construct validity, as Brouwers and Tomic (2012) observed. This limitation sparked renewed interest in refining the conceptualization and measurement of existential fulfillment, prompting further development of assessment tools (Loonstra et al., 2007).

Prominent among these scholarly pursuits, Loonstra et al. (2007) made a significant contribution to existential psychology by refining and clarifying the concept of existential fulfillment, which they identified as an essential aspect of a meaningful and purpose-driven lifestyle. Their conceptualization, deeply rooted in humanistic-existential psychology, integrates the philosophical principles of Viktor Frankl's logotherapy, Irvin Yalom's existential psychotherapy, the humanistic ideals of Abraham Maslow, Carl Rogers, and the existential humanism of Erich Fromm. This unified theoretical framework defines existential fulfillment as a life goal that fully honors the nature of human existence, acknowledging three fundamental limitations: the finality imposed by mortality, the boundaries of individual potential, and the constraints shaped by relationships with others and the larger world. To navigate these intrinsic limitations and achieve fulfillment, individuals are tasked with meeting three core existential needs: self-acceptance, self-actualization, and self-transcendence.

The fulfillment of these three tasks is thought to culminate in existential fulfillment. According to the authors, self-acceptance involves a wholehearted embrace of one's reality, including the acceptance of mortality, personal limits, and the recognition of one's small role in a vast reality. Self-actualization is the process through which an individual realizes their intrinsic value and authentic potential. Self-transcendence, central to mental health, represents the cognitive ability to engage meaningfully with both one's inner self and the external world (Loonstra et al., 2007).

To effectively measure existential fulfillment, a scale must meet two criteria. First, it should assess the degree to which people live with purpose and meaning, reflecting the nature of existential fulfillment as a life goal. Second, the scale should differentiate between various qualitative types of values, goals, and life meanings, distinguishing between extrinsic and intrinsic, as well as self-centered and self-transcending, aspirations (Loonstra et al., 2007). Existing scales, such as the Purpose in Life Test (PLT; Crumbaugh, 1968), Life Regard Inventory (LRI; Battista and Almond, 1973), and Meaning in Life Scale (MLS; Steger et al., 2006), satisfy the first criterion by assessing life's purpose and meaning. However, they fall short of distinguishing between self-alienated, self-actualizing, and self-transcending values, thereby failing to meet the second criterion (Loonstra et al., 2007). Hence, a new scale was deemed necessary.

Recognizing these considerations and drawing on an advanced, holistic understanding of existential fulfillment, Loonstra et al. (2007) developed the Existential Fulfillment Scale (EFS). The EFS encapsulates three core constructs of existentialism—self-acceptance, self-actualization, and self-transcendence—each representing distinct dimensions of existential fulfillment. Comprising 15 items, with five items for each dimension, the EFS is well-structured to assess these aspects. The CFA showed that, compared to four alternative models, the three-factor oblique model yielded the best fit, with a Root Mean Square Residual (RMR) of 0.08, Goodness of Fit Index (GFI) of 0.91, Adjusted Goodness of Fit Index (AGFI) of 0.88, Normed Comparative Fit Index (CFI) of 0.89, and Parsimony Normed Comparative Fit Index (PCFI) of 0.73. Additionally, significant correlations emerged between self-acceptance and self-actualization ($r = 0.41$, $p < 0.001$) and between self-actualization and self-transcendence ($r = 0.40$, $p < 0.001$), though no significant correlation was observed between self-acceptance and self-transcendence ($r = 0.06$, $p > 0.05$). These results have been corroborated by subsequent studies (Loonstra et al., 2010; Tomic and Tomic, 2011; Bekenkamp et al., 2014; Kay, 2014, 2019). Overall, the study supports the presence of three distinct but related dimensions within existential fulfillment, with internal consistency coefficients for self-acceptance, self-actualization, and self-transcendence at 0.74, 0.71, and 0.88, respectively.

However, the study did not examine several critical indicators of validity, including convergent, discriminant, and criterion-related validity, nor did it assess stability through test–retest reliability, which the authors acknowledge as a limitation. Furthermore, they recommend that future studies theoretically explore and empirically investigate the interrelationships among the three constructs of existential fulfillment in greater depth (Loonstra et al., 2007).

Despite these limitations, the EFS has become widely accepted in empirical research on existential fulfillment, effectively replacing the Existence Scale (ES) in many studies (Loonstra et al., 2010; Tomic and Tomic, 2011; Bekenkamp et al., 2014; Kay, 2014, 2019; Tomic, 2016). The scale has consistently demonstrated strong reliability and validity across multiple studies and has been validated and applied in various cross-cultural contexts, underscoring its broad acceptance and esteem within the research community.

1.3 Existential fulfillment, well-being and depression

Existential fulfillment plays a crucial role in psychological health and life satisfaction, forming an integral part of an individual's overall well-being (Tomic and Tomic, 2011). Frankl's (2004) logotherapy underscores the centrality of the quest for meaning in human life, suggesting that achieving existential goals can foster fulfillment and enhance well-being. Tomic and Tomic's (2011) findings support this, identifying a positive correlation between self-transcendence—a dimension of existential fulfillment—and well-being, particularly in professions with high interpersonal demands, such as nursing. Similarly, Bekenkamp et al. (2014) observed that elevated levels of self-actualization and self-acceptance contribute to better perceived physical health, a significant correlate of overall well-being. Collectively, these studies suggest that existential fulfillment serves as a protective factor against psychological distress and is a vital contributor to well-being.

Conversely, existential fulfillment is inversely related to depression, with a lack of meaning and purpose often preceding depressive symptoms and burnout (Shek et al., 2022; Alfuqaha et al., 2021). Mascaro and Rosen (2008) found that a sense of existential meaning could predict lower depressive symptoms over time, while Längle et al. (2003) and Loonstra et al. (2010) demonstrated that higher levels of existential fulfillment were associated with reduced depression, especially in high-stress fields such as teaching. Thakur and Basu (2010) also identified a significant correlation between depression and existential meaning, suggesting that promoting fulfillment may be an effective strategy for both preventing and alleviating depressive symptoms. These findings underscore the value of existential dimensions in both the understanding and treatment of depression.

1.4 The present study

China's rapidly evolving socio-economic landscape has led to transformative shifts, moving beyond material scarcity and prompting heightened spiritual and existential inquiries among its people. This trend has given rise to various psychological challenges, including feelings of emptiness, a lack of purpose, and diminished spiritual direction (Wu, 2018; Zhang, 2018), thus intensifying the relevance of existential concerns in contemporary mental health discourse. Given this context, investigating the existential psychology of Chinese individuals has taken on both an urgent and a practical importance. However, empirical research into these existential challenges in China remains sparse, with few studies specifically addressing existential fulfillment. To address this gap, we propose that adopting and adapting a relevant existential fulfillment scale could be a critical first step in advancing research in this field. Considering the dual priorities of reliability and validity, the EFS emerges as a well-suited instrument for revision to meet the specific demands of this research. Thus, this study aims to assess the psychometric properties of the EFS and evaluate its applicability within a sample of Chinese college students, who serve as a representative demographic for this research.

Our objectives are twofold: first, to validate the factorial structure of the EFS through item analysis and exploratory factor analysis (EFA); second, to establish the factorial validity of the EFS in alignment with previous research, confirm its criterion-related validity, and assess its reliability through internal consistency and test–retest stability.

2 Materials and methods

2.1 Participants

Samples 1 and 2 were utilized for comprehensibility testing. Sample 1 included 25 undergraduate students from a university in Fujian Province, with 6 freshmen, 6 sophomores, 5 juniors, and 8 seniors. This sample comprised 12 male and 13 female students, with a mean age of 20.18 years ($SD = 1.98$). Sample 2 matched this composition and was drawn from the same institution, consisting of 7 freshmen, 8 sophomores, 5 juniors, and 5 seniors, including 11 male and 14 female students, with a mean age of 20.02 years ($SD = 1.81$).

Sample 3 was designated for Item Analysis and EFA. This group comprised undergraduates from six universities in Fujian Province, recruited via a stratified cluster sampling method based on academic disciplines, gender, and academic level. Of the 380 questionnaires distributed, 324 valid responses were returned, yielding a response rate of 85.3%. This sample included 141 males and 183 females, with a mean age of 20.31 years (SD = 2.11). Academic standings included 95 freshmen, 83 sophomores, 75 juniors, and 71 seniors, representing diverse fields: 111 students from the Arts, 101 from Science and Technology, and 112 from other disciplines.

Sample 4 was used for CFA and internal consistency test. This sample, drawn again from undergraduates across the same six universities, was collected via stratified cluster sampling to reflect academic major, gender, and year. Of the 800 questionnaires distributed, 727 valid responses were collected, achieving a response rate of 90.9%. The sample included 321 males and 406 females, with a mean age of 20.67 years ($SD = 1.98$), and covered 171 freshmen, 173 sophomores, 195 juniors, and 188 seniors across various disciplines: 241 students from the Arts, 276 from Science and Technology, and 210 from other fields.

Sample 5 and Sample 6 were used to assess criterion-related validity. Sample 5 included 317 students (131 male, 186 female, mean age = 21.03 years, $SD = 1.62$) from a university in Fujian Province, who completed the EFS, Meaning of Life Questionnaire (MLQ), and Index of Well-Being (IWB). Sample 6 comprised 33 students (15 male, 18 female, mean age = 20.84 years, $SD = 1.25$), who completed the EFS and the Self-Depression Scale (SDS).

Sample 7 was used to evaluate test–retest reliability, including 149 undergraduates (61 male, 88 female, mean age = 20.67 years, $SD = 1.24$) from a university in Fujian Province. Participants completed the questionnaire twice over a two-week period.

Each of these samples was independently recruited, with no overlap among participants. Samples 1–7 were initially utilized for the original analyses as described previously. However, to maximize the utility of these samples and provide additional evidence of validity and reliability, cross-validation analyses were subsequently performed as follows:

- CFA was first conducted on Sample 4 as the primary analysis, followed by CFA on Sample 5 to provide additional structural validity evidence.
- Internal consistency reliability analysis was initially conducted on Sample 4 and was later extended to Samples 5 and 7 to enhance the robustness of reliability evidence.

2.2 Measures

2.2.1 Existential fulfillment scale (EFS)

The EFS, developed by Loonstra et al. (2007), includes 15 items designed to measure key dimensions of existential fulfillment across three factors: self-acceptance, self-actualization, and self-transcendence. Responses are captured on a 5-point Likert scale ranging from 1 ("not at all") to 5 ("fully"), with higher scores indicating greater existential fulfillment.

2.2.2 Meaning of life questionnaire (MLQ)

The MLQ, initially developed by Steger et al. (2006) and subsequently translated and revised by Wang and Dai (2008), was employed in this study. Comprising 10 items, the MLQ captures two core dimensions: the pursuit of meaning (seeking) and the presence of meaning in life (experiencing). In this study, the internal consistency reliability coefficients for the total scale, as well as the two subscales for seeking and experiencing meaning, were 0.84, 0.68, and 0.69, respectively, indicating a satisfactory level of reliability.

Since the EFS draws on Viktor Frankl's logotherapy framework, as elaborated by Loonstra et al. (2007), it was hypothesized that existential fulfillment would conceptually overlap with the sense of meaning in life. Consequently, the MLQ scores were included as a criterion-related validity indicator, with the expectation of a positive correlation between the MLQ and EFS scores.

2.2.3 Index of well-being (IWB)

The IWB, originally created by Campbell et al. (1976) and adapted by Wang et al. (1999), was also utilized in this study. This scale consists of two components: a general affective index with 8 items and a life satisfaction index comprising a single item. The overall score is computed by combining the average score of the Overall Emotionality Index with the Life Satisfaction Index score, which is weighted at 1.1 to reflect its relative importance. Previous studies report a test–retest reliability of 0.849 for the IWB (Wang et al., 1999). In this study, the internal consistency of the total scale and the general affective index were 0.82 and 0.87, respectively, indicating strong reliability.

Clinical and empirical research in psychotherapy consistently links existential fulfillment with subjective well-being (Tomic and Tomic, 2008; Mausch and Rys, 2020). Thus, IWB scores were incorporated to assess the criterion-related validity of the EFS, with a hypothesized positive correlation between these two scales.

2.2.4 Self-depression scale (SDS)

The SDS, originally developed by Zung and cited in Wang et al. (1999), was used in this study as a measure of depressive symptoms. This 20-item scale is organized into four dimensions: psychotic-affective symptoms (2 items), somatic disorders (8 items), psychomotor disorders (2 items), and psychological disorders of depression (8 items). The SDS evaluates the severity of depressive symptoms experienced by participants over the preceding week. Internal consistency reliability for the total scale and its four subscales in this study were 0.79, 0.77, 0.72, 0.74, and 0.67, respectively.

In line with Frankl's (1962) logotherapy, which positions the "existential vacuum" as an antithesis to existential fulfillment and associates it with depressive symptoms, the SDS was included as a criterion-related validity measure for the EFS. It was hypothesized that EFS scores would negatively correlate with SDS scores, reflecting the inverse relationship between existential fulfillment and depressive symptoms.

2.3 Procedure

2.3.1 Data collection

The study protocol was approved by the Ethical Review Board of the School of Psychology at Fujian Normal University (Protocol No. Psy240051). Samples 1 and 2 were collected for comprehensibility testing by recruiting students from the researcher's institution to complete an online questionnaire. For Samples 3 through 8, participants were independently recruited from six universities across Fujian Province through collaborations with faculty members who facilitated student recruitment in their respective classes. Each sample was recruited separately, ensuring there was no overlap among participants.

To strengthen the validity and generalizability of the findings, independent samples were designated for each stage of analysis, including item analysis, EFA, CFA, criterion-related validity and reliability testing. This approach reduced potential biases from overlapping samples and enhanced the reliability of the scale's validation across different datasets. By using distinct samples for structural validation and criterion-related validity assessments and reliability testing, each analysis phase could contribute independently to a comprehensive evaluation of the scale's psychometric properties.

Drawing on prior experience, the research team anticipated that including too many items in a single questionnaire could lead to fatigue or irritability among Chinese college students, potentially affecting response quality. Therefore, data collection was conducted in phases, aligning with each analytical objective to optimize data validity. For instance, during the CFA phase, participants were asked only to complete the EFS, minimizing respondent fatigue and enhancing data accuracy for structural examination.

Initially, the MLQ and IWB were selected as criterion measures to assess the criterion-related validity of the EFS. However, after analyzing data from these scales, the strong relationship between existential fulfillment and mental health became apparent. To further reinforce the criterion-related validity of the EFS, the SDS was included as an additional criterion, prompting the recruitment of Sample 6 specifically for this purpose.

2.3.2 Translation of instrument

Dr. Welko Tomic, one of the original authors of the EFS, authorized the adaptation of the scale for use in Chinese. The translation process followed rigorous standards to ensure fidelity to the original English version, as outlined by the International Test Commission (ITC) guidelines[1]. In this study, the EFS was translated into Simplified Chinese and then back-translated into English through a multi-stage process designed to ensure conceptual, semantic, and cultural equivalence. Initially, a bilingual psychology researcher with a year of study experience in the United States translated the English version into Simplified Chinese, carefully preserving the meaning of the original items. Another bilingual psychology researcher, holding a Ph.D. from a U.S. university, then back-translated the Chinese version into English. Both translators were native Chinese speakers proficient in English, ensuring linguistic accuracy and cultural sensitivity. To assess conceptual equivalence, a panel of psychology experts familiar with existential fulfillment constructs in both Chinese and Western contexts reviewed the translated items. This review ensured that the items retained the underlying constructs of the original scale. Semantic equivalence was further tested with a pretest involving 25 bilingual participants, who rated the clarity and similarity of meaning between the English and Chinese items. Discrepancies were resolved through consultation with the translators. Cultural relevance was validated by consulting Chinese psychologists experienced in cross-cultural research, who reviewed each item for appropriateness in the Chinese university student context. The English back-translation was then sent to Dr. Tomic, who confirmed its accuracy, endorsing the adapted scale for use in the Chinese context.

An initial comprehensibility assessment was subsequently conducted. A sample of 25 university students from various academic disciplines (Sample 1) was asked to rate the clarity of each translated item on a 5-point Likert scale, where 1 indicated "difficult to comprehend" and 5 represented "fully comprehensible." Participants were encouraged to explain any item rated below 4 to clarify areas needing improvement. Results revealed that items 1, 10, and 13 scored below 4, with comments indicating that these items were "not easy to understand" or "too broad." Based on this feedback, the research team implemented targeted revisions to improve clarity.

A second comprehensibility test was conducted after these revisions. A new cohort of 25 college students, representing a range of academic majors, genders, and grade levels (Sample 2), assessed the scale's clarity using the same 5-point Likert scale. This time, no complaints were raised regarding the phrasing, and all 15 items scored 4 or above, indicating high understandability.

Finally, the finalized Chinese version of the EFS, with response options identical to the original scale, was established for use in this study.

2.4 Data analysis

Initially, item analysis for the EFS was conducted using corrected item-total correlations, following the methodological approach outlined by Wu (2010).

Subsequently, EFA of the EFS was performed using SPSS software, version 19.0. The factor structure of the EFS items was determined through Principal Axis Factoring analysis, followed by Promax rotation to enhance factor interpretability.

To assess construct validity, CFA of the Chinese EFS was conducted using Mplus software, version 8.3. This analysis aimed to confirm that the indicators measured corresponded to three underlying latent factors. Established criteria for an adequate model fit were applied, targeting a CFI and Tucker–Lewis Index (TLI) of 0.90 or above, and a Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) of 0.08 or below, in line with methodological guidelines provided by Byrne (1994) and Hu and Bentler (1995).

The criterion-related validity of the Chinese EFS was subsequently examined using the Chinese versions of the MLQ, IWB, and SDS.

Finally, the internal consistency of the Chinese EFS was evaluated by calculating the composite reliability, with higher values indicating greater internal consistency. Additionally, the test–retest reliability of the EFS was assessed with a two-week interval to confirm temporal stability.

3 Results

3.1 Item analysis

An item analysis was conducted on data from 324 participants in Sample 3. The analysis indicated that Item 8 had a total correlation coefficient of 0.35, which fell below the threshold of 0.40. Consequently, in accordance with the screening criteria outlined by Wu (2010), this item was excluded. The remaining 14 items showed correlation coefficients above 0.40, ranging from 0.41 to 0.76, and were statistically significant.

Next, participants were subsequently divided into high and low groups based on the upper and lower 27% of their total scores. A Multivariate Analysis of Variance (MANOVA) with a Bonferroni correction was applied to compare two groups. The results indicated a significant multivariate effect for group, with Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root all yielding a significant F-value of 57.99, df = 14, 186, $p < 0.001$, and a large effect size (Partial Eta Squared = 0.81). This suggests that the group variable significantly influenced the scores on the items.

Subsequent pairwise comparisons, adjusted for multiple comparisons using the Bonferroni method, revealed significant mean differences for all 14 items between the two groups. Each item demonstrated a statistically significant difference at the 0.05 level, with all p-values being less than 0.000, indicating a strong and consistent pattern of group differences across all items. The mean differences ranged from 0.78 to 1.44, with standard errors varying between 0.10 and 0.14, and all 95% confidence intervals excluding zero, confirming the significance of the observed differences.

In summary, the MANOVA and Bonferroni-adjusted pairwise comparisons provided robust evidence that the two groups significantly differed on all 14 items, affirming each item's discriminative capability.

3.2 Exploratory factor analysis (EFA)

Exploratory factor analysis was conducted using data from Sample 3 (Table 1). The Kaiser–Meyer–Olkin (KMO) measure was 0.840, and Bartlett's test of sphericity yielded a chi-square value of 1258.319

1 https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf

TABLE 1 Standardized factor loadings of the Chinese EFS.

Item	Item content	Factor 1	Factor 2	Factor 3
1	I often feel uncertain about the impression I make on other people	**0.49**	0.22	0.35
3	I do a lot of things that I would actually rather not do	**0.67**	0.31	0.16
11	I find it very hard to accept myself	**0.48**	0.43	0.35
12	I often do things because I have to, not because I really want to do them	**0.73**	0.30	0.16
4	I feel incorporated in a larger meaningful entity.	0.27	**0.60**	0.32
5	Deep inside I feel free.	0.34	**0.60**	0.43
6	I think I am part of a meaningful entity.	0.40	**0.79**	0.45
7	Even in busy times I experience feelings of inner calmness.	0.33	**0.54**	0.43
2	I'll remain motivated to carry on even in times of bad luck.	0.25	**0.50**	0.50
9	It is my opinion that my life is meaningful.	0.43	0.55	**0.74**
10	I have experienced that there is more in life than I can perceive with my senses.	−0.01	0.17	**0.49**
13	I think my life has such a deep meaning that it surpasses my personal interests.	0.02	0.34	**0.63**
14	I completely approve of the things that I do	0.33	0.49	**0.62**
15	My ideals inspire me.	0.25	0.54	**0.73**

Factor 1 = self-acceptance; Factor 2 = self-actualization; Factor 3 = self-transcendence. Bolded values in the table indicate the factor loadings for items that are ultimately assigned to each of the three factors after the EFA.

($p < 0.001$), confirming the data's suitability for factor analysis. A Principal Axis Factoring analysis with Promax rotation identified three factors with eigenvalues greater than 1 (1.13, 1.89, and 4.43), explaining 53.22% of the variance, which is close to the acceptable threshold of 50% (Wu, 2010). Factor loadings ranged from 0.48 to 0.79, with all items meeting the communality criterion (≥0.4). Inter-factor correlations were 0.53 between Factors 1 and 2, 0.32 between Factors 1 and 3, and 0.64 between Factors 2 and 3.

Two notable discrepancies emerged when comparing the adapted scale to the original. First, the self-acceptance dimension was reduced to four items after the exclusion of Item 8. Second, some items were reassigned: Item 4 ("I feel incorporated in a larger meaningful entity") and Item 6 ("I think I am part of a meaningful entity"), originally under self-transcendence, were moved to self-actualization, while Item 14 ("I completely approve of the things that I do") and Item 15 ("My ideals inspire me"), initially categorized as self-actualization, were shifted to self-transcendence. Despite these adjustments, the factors retained their original names, aligning with the theoretical basis of the scale. These adjustments are discussed further in the discussion section.

Additionally, several items in Factors 2 and 3, such as Items 2, 5, 6, 7, 9, 14, and 15, had loadings above 0.4 on both factors. Item 5 showed equal loadings of 0.50 on both, raising the question of whether these could be considered a single factor, which will be explored in the CFA.

3.3 Confirmatory factor analysis (CFA)

The CFA was conducted using Mplus version 8.3 on a dataset of 727 valid responses from Sample 4, with parameters estimated via robust maximum likelihood (MLR) estimation. Following the original structural framework of the EFS, a three-factor model (M3) was created to represent the dimensions of self-acceptance, self-actualization, and self-transcendence. To enable comparative analysis, two alternative models, M1 and M2, were also tested.

M1 was a single-factor model where all items load onto a single construct. This model was tested to explore whether existential fulfillment could be validated as a unified psychological entity, aligning with one of the original objectives set by Loonstra et al. (2007).

M2 was a two-factor model in which self-acceptance forms one factor while self-actualization and self-transcendence combine into a second factor. This model addresses discrepancies in item allocation noted between the original EFS and its Chinese adaptation during EFA and is theoretically supported by Frankl's (1962) assertion that self-actualization is often contingent upon self-transcendence, suggesting a potential linkage.

Model fit indices, summarized in Table 2, indicated that the three-factor model achieved the best fit, though the two-factor model was also acceptable, with both models meeting established criteria for goodness-of-fit.

To further evaluate the structural validity of the three-factor model compared to the two-factor model, we analyzed the total scores, individual factor scores, and inter-factor correlations for each model. Findings are as follows:

In the three-factor model, the total scale score showed significant associations with self-acceptance ($r = 0.62$, $p < 0.001$), self-actualization ($r = 0.89$, $p < 0.001$), and self-transcendence ($r = 0.85$, $p < 0.001$). Self-acceptance correlated moderately with both self-actualization ($r = 0.30$, $p < 0.001$) and self-transcendence ($r = 0.22$, $p < 0.001$), while self-actualization demonstrated a strong correlation with self-transcendence ($r = 0.76$, $p < 0.001$).

As observed in the EFA, many items within the self-actualization and self-transcendence dimensions had high loadings on both factors, and the CFA fit indices further indicated that a two-factor model—where self-actualization and self-transcendence are combined into a single factor—was also viable. Additionally, correlation analyses among the three factors revealed a significant and high correlation between self-actualization and self-transcendence. Collectively, these findings suggest that, in the Chinese context, self-actualization and

10.3389/fpsyg.2024.1497422

TABLE 2 Confirmatory factor analysis fit index.

Model	χ/df	*CFI*	*TLI*	*RMSEA* (90 Percent C.I)	*SRMR*
1-factor (M1)	10.64	0.80	0.76	0.12 (0.11,0.12)	0.09
2-factor (M2)	4.82	0.92	0.90	0.07 (0.06,0.07)	0.06
3-factor (M3)	4.63	0.93	0.91	0.07 (0.06,0.08)	0.06

self-transcendence may function as a single factor. To further investigate, a bifactor model analysis was conducted for these two dimensions on Sample 4.

3.4 Evaluation of the bifactor model

To evaluate the bifactor model, we utilized two indices: explained common variance (ECV) and percent uncontaminated correlations (PUC) (Rodriguez et al., 2016a, 2016b). ECV measures the proportion of common variance attributed to both general and domain-specific factors. In this study, the general factor explained 85.2% of the common variance, with domain-specific factors accounting for the remaining 14.8%. This distribution suggests that the common variance is predominantly represented by the general factor. PUC, which measures the extent of EFS item correlations influenced by both general and domain-specific factors, yielded a value of 0.60. The high ECV combined with a moderate PUC indicates that many item correlations are primarily influenced by the general factor. This finding supports the notion that self-actualization and self-transcendence within the EFS may be treated as a unified dimension.

As previously noted, the constructs of self-actualization and self-transcendence reflect a fundamental human drive to explore and realize inherent potential while engaging meaningfully with the realities beyond oneself (Loonstra et al., 2007). This suggests that individuals achieving existential fulfillment continually strive to exceed their current capacities and self-concept. To encapsulate this ongoing developmental process within the Chinese context, we introduce the term "self-breakthrough" to represent this key factor in the Chinese EFS.

To provide more robust evidence for the construct validity of the scale, we conducted a CFA of the two-factor model, as depicted in Figure 1, using Sample 5. The results indicated that the CFI was 0.91, the TLI was 0.88, the RMSEA was 0.06 with a 90% confidence interval of 0.05–0.07, and the SRMR was 0.07. While the TLI fell slightly below the commonly accepted threshold of 0.90 for a good fit, all other fit indices were within ranges indicative of a well-fitting model. The results further reinforce the stability and structural validity of the two-factor Chinese version of the EFS.

In conclusion, the findings from these analyses indicate that the two-factor model of the Chinese EFS, comprising self-acceptance and Self-Breakthrough, provides the most robust structural validity.

3.5 Criterion-related validity test

To evaluate the criterion-related validity of the Chinese EFS, correlations were examined between the Chinese EFS and the MLQ, IWB, and SDS using data from Samples 5 and 6. The Chinese EFS showed a moderate positive correlation with the MLQ ($r = 0.49$, $p < 0.001$) and the IWB ($r = 0.63$, $p < 0.001$), and a moderate negative correlation with the SDS ($r = -0.54$, $p < 0.001$), confirming the scale's criterion validity. See Table 3 for detailed correlations.

3.6 Reliability test

To assess the internal consistency of the scale, we first examined item-total correlations on Sample 4. Results showed that correlations between individual items and the total score ranged from 0.41 to 0.76, all of which exceeded the minimum criterion of 0.40 (Wu, 2010), confirming the homogeneity of the items across the scale.

Next, we employed CFA to assess the composite reliability (coefficient omega (ω)) of the total scale and subscales. Coefficient ω represents the proportion of variance in the total score attributable to all sources of common variance modeled within the scale. For the total scale, ω was 0.97. Regarding the two domain-specific factors, namely self-acceptance and self-breakthrough, ω was 0.89 and 0.96, respectively.

To further evaluate the reliability of the scale, we also assessed its internal consistency across Sample 5 and Sample 7. For Sample 5, the coefficient ω values for the total scale and the subscales of self-acceptance and self-breakthrough were 0.92, 0.87, and 0.91, respectively. For the initial test data of Sample 7, these coefficients were 0.97, 0.92, and 0.96, respectively. For the retest data of Sample 7, the coefficients were 0.96, 0.89, and 0.93, respectively. The cross-analysis of these samples suggests robust internal consistency reliability for both the overall scale and its subscales.

Additionally, test–retest reliability was assessed with a two-week interval on Sample 7, yielding values between 0.72 ($p < 0.001$) and 0.86 ($p < 0.001$) for both the total scale and its subscales, highlighting the scale's strong temporal stability.

A comprehensive summary of these reliability results is presented in Table 4.

4 Discussion

In this research, the Chinese version of the EFS underwent its initial revision and cross-cultural validation, focusing specifically on Chinese college students. Through EFA, comparative analysis of multiple competing structural models using CFA, and examination of a two-factor model, this study found that, within the Chinese cultural context, the structure of the Chinese EFS diverges from the original three-dimensional model, reducing instead to two dimensions. Specifically, the dimensions of self-actualization and self-transcendence from the original scale have been combined into a single factor, which we have termed "self-breakthrough." We propose that this cross-cultural difference in scale structure is related both to humanistic-existential psychology's nuanced understanding of the

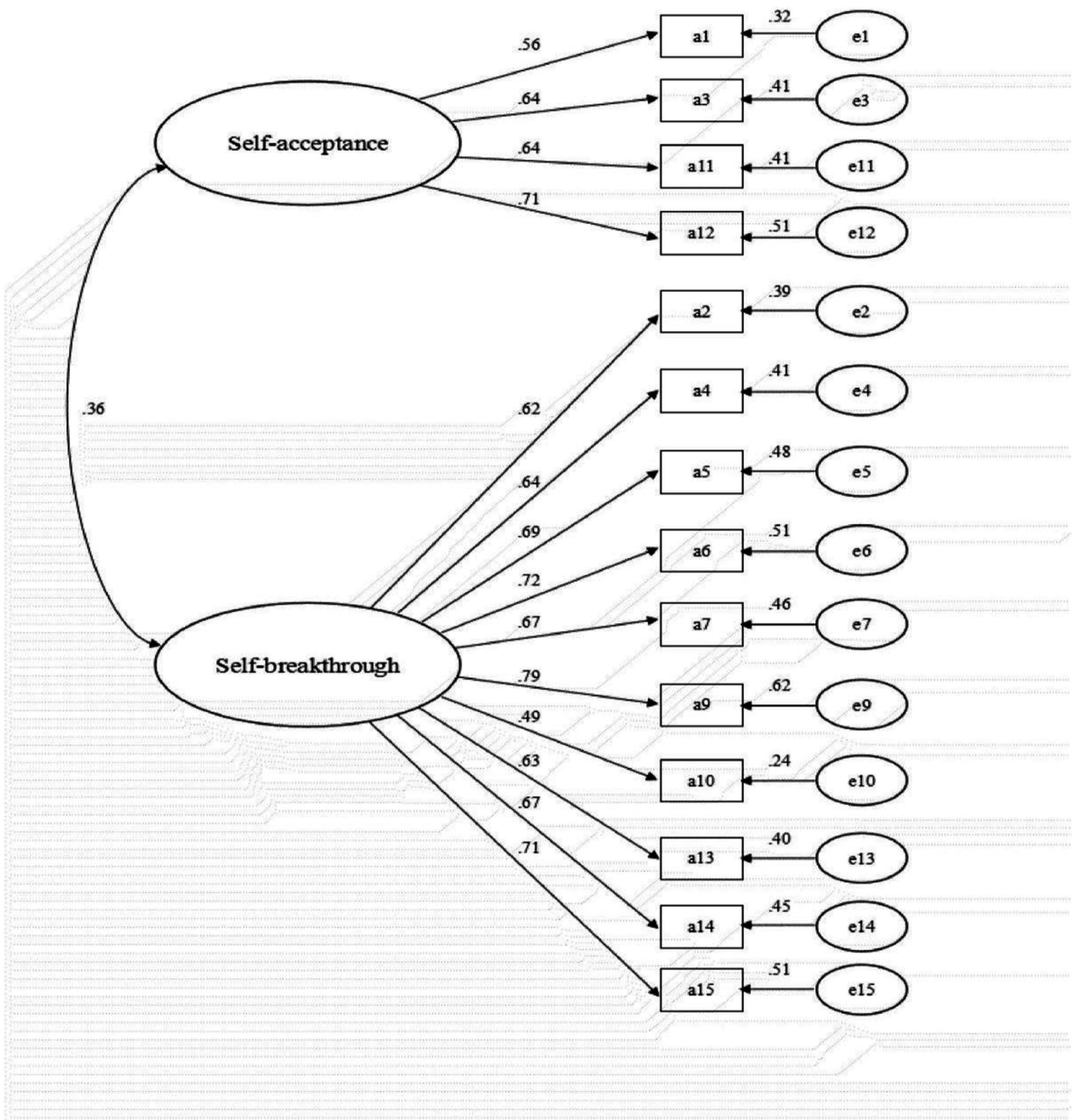

FIGURE 1
Confirmatory factor analysis for the Chinese EFS. Q1–Q15 represent items 1–15 in Table 1.

relationship between self-actualization and self-transcendence and to the cultural perceptions of individual success and fulfillment in China.

One key rationale for this restructuring stems from divergent theoretical perspectives on self-transcendence. Humanistic-existential psychologists, such as Frankl (1962, 1963) and Maslow (1971), have noted that while self-actualization and self-transcendence can be conceptually distinct, they also overlap significantly. Frankl, for instance, suggested that the pursuit of meaning involves both reaching one's potential (self-actualization) and transcending personal limitations (self-transcendence). Maslow (1971) similarly posited that self-actualization leads individuals to a deeper perception of reality and a capacity for profound interpersonal relationships, suggesting that self-actualized individuals often experience transcendence as they extend beyond themselves in meaningful relationships with others. Given this theoretical background, it is understandable that the distinction between self-actualization and self-transcendence is not always clear-cut, supporting their combination into a single factor within the Chinese context.

The second explanation for these structural adjustments is cultural context. In Eastern societies, particularly within Chinese culture, self-actualization is viewed as a socially oriented process, emphasizing the individual's role responsibility and balance between personal and collective needs (Xu et al., 2017). This contrasts with the Western, more individualistic interpretation of self-actualization, which often emphasizes personal responsibility and the pursuit of individual goals and ideals (Maslow, 1971; Maehr, 1974). In Chinese culture, where collectivist ideologies are prevalent, self-actualization is considered

TABLE 3 Criterion-related validity of the Chinese version of EFS.

	Existential fulfillment	Self-acceptance	Self-breakthrough
MLQ (n = 317)	**0.49****	**0.15****	**0.55****
Seeking of meaning	0.58**	0.24**	0.62**
Experiencing of meaning in life	0.28**	0.03**	0.34**
IWB (n = 317)	**0.63****	**0.54****	**0.54****
General affective index	0.68**	0.55**	0.59**
Life satisfaction index	0.38**	0.16**	0.40**
SDS (n = 33)	**−0.54****	**−0.68****	**−0.34**
Psychotic-affective symptoms	−0.48**	−0.59**	−0.31
Somatic disorders	−0.34**	−0.60**	−0.13
Psychomotor disorders	−0.02	−0.07	−0.01
Psychological disorders of depression	−0.64**	−0.59**	−0.49**

The bolded numbers indicate the correlation coefficients between the total scores of the three criterion-related scales and the Chinese version of the EFS. **$p < 0.01$.

TABLE 4 Internal consistency reliability and retest reliability of the Chinese version of EFS.

	Existential fulfillment	Self-acceptance	Self-breakthrough
Internal consistency reliability (Sample 4)	0.97	0.89	0.96
Internal consistency reliability (Sample 5)	0.92	0.87	0.91
Internal consistency reliability (Sample 7, initial test)	0.97	0.92	0.96
Internal consistency reliability (Sample 7, retest)	0.96	0.89	0.93
Retest reliability (Sample 7, Two-week interval)	0.76	0.75	0.83

incomplete without the support of the group, and self-transcendence is associated with individuals' pursuit of future goals and ideals that often extend beyond their immediate realities (Yu and Yang, 1987; Yu, 2008). Thus, the structural differentiation in the Chinese EFS reflects both cultural specificity and cross-cultural theoretical perspectives. This study therefore indicates that a two-factor model is the optimal structure for measuring existential fulfillment in Chinese cultural settings.

Furthermore, this study found that the Chinese EFS outperforms the original scale regarding CFA fit indices. The original scale's three-factor model yielded a CFI of 0.89 and an RMR of 0.08 (Loonstra et al., 2007), whereas our two-factor Chinese EFS achieved a CFI of 0.92 and an SRMR of 0.06. This improvement may be attributed to differences in the study populations. The original scale was tested on a more complex and heterogeneous sample of psychology graduates with an average age of 42.6, whereas our sample consists of Chinese undergraduate students, a relatively homogenous and less complex demographic. This difference in sample composition likely contributes to the observed variance in fit indices, supporting the suitability of the revised Chinese EFS for use in a younger, undergraduate context.

Consistent with our expectations, the Chinese EFS demonstrated moderate positive correlations with MLQ and IWB, supporting its criterion-related validity, along with a moderate negative correlation with SDS. These results suggest that higher EFS scores are associated with reduced depressive symptoms and increased levels of meaning and well-being, underscoring the scale's robust criterion-related validity.

In terms of reliability, the Chinese EFS achieved an omega coefficient (ω) ranging from 0.87 to 0.97, indicating strong internal consistency. The test–retest reliability over a two-week interval was 0.76, with subscale values of 0.75 and 0.83. This temporal stability supports that the revised Chinese EFS meets high psychometric standards.

In summary, the Chinese EFS is a concise, effective tool for measuring existential fulfillment among Chinese college students, offering a reduced item set (14 items) that maintains clarity and minimizes respondent fatigue. Given its solid psychometric properties and cultural adaptability, the Chinese EFS stands to make a meaningful contribution to existential studies within Chinese populations.

4.1 Limitation and future directions

It is essential to acknowledge certain limitations in this research that require careful consideration. First, the study's sample was drawn exclusively from universities in Fujian Province, potentially limiting the generalizability of the findings to the broader university-aged population in China. Fujian, located in southeastern China, has a moderately developed economy and a unique, relatively open yet conservative culture, which may affect the applicability of the study's findings in other regions.

Secondly, despite efforts to recruit from various universities in Fujian, the limited number of top-tier institutions in the sample may not fully represent the diversity of the Chinese university landscape, potentially impacting the external validity of the results.

Thirdly, while the study provides initial evidence regarding the structure of existential fulfillment within a Chinese cultural context, it does not fully address the ongoing debate surrounding the specific connotations and structure of this construct. Although cultural influences were found to shape existential fulfillment, more in-depth exploration and validation of these cultural factors were beyond the scope of this study.

A significant limitation of the current study is the incomplete assessment of certain psychometric properties of the EFS. While this study offers preliminary evidence for the EFS's use in a Chinese context, it did not rigorously evaluate key psychometric aspects such as discriminant validity and measurement invariance. Discriminant validity, which assesses the degree to which the EFS differs from related constructs, is crucial for establishing the unique role of existential fulfillment in psychological research. Similarly, measurement invariance, which tests whether the scale measures the same construct across different groups (e.g., gender or regional subgroups), is necessary to ensure the EFS's applicability across diverse populations within China. Other aspects such as convergent validity, predictive validity, and known-group validity were also not thoroughly examined, which are important for establishing the EFS's robustness.

Future research should aim to address these limitations by developing a locally adapted version of the EFS that reflects the unique connotations of existential fulfillment within Chinese culture. This process might involve conducting in-depth interviews to explore these cultural meanings, followed by coding techniques to inform the scale's structure. A rigorous examination of measurement invariance and discriminant validity should be incorporated in future studies to confirm that the EFS provides consistent and distinct measures across various subgroups. Additionally, by employing a bottom-up approach in scale development, future research could yield a more nuanced and culturally sensitive tool, enhancing our understanding of existential fulfillment across China's diverse populations. Such comprehensive psychometric evaluation would strengthen the EFS's performance across varied contexts, thereby reinforcing its applicability in psychological research and assessment.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the School of Psychology, Fujian Normal University, China. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JC: Writing – review & editing. XC: Writing – original draft. CM: Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the general project of National Social Science Foundation for Education in China, Research on the Influence Mechanism and Intervention of Youth's Sense of Existence, Project no. BBA200038.

Acknowledgments

The authors gratefully acknowledge Dr. Welko Tomic from the Open University, Heerlen, the Netherlands, for generously granting permission to adapt and utilize the Existential Fulfillment Scale (EFS) for this study. Our sincere thanks are also extended to Dr. Tomic for the invaluable insights and expert guidance provided throughout the scale adaptation process. The authors wishes to extend their heartfelt thanks to Dr. Yin Bin for his invaluable guidance and specific assistance during the revision of the scale and the manuscript. His expertise and meticulous attention to detail have been instrumental in enhancing the quality of this research. The authors would like to express their sincere gratitude to Dr. Lin Rongmao, for his substantial contributions to the statistical analysis of the data. His expertise and guidance were invaluable to the completion of this research. The author would like to extend sincere gratitude to Ying Yang for her invaluable contributions to the revision of this manuscript. The authors would like to thank the editor and reviewers for their invaluable assistance in improving the quality of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alfuqaha, O. A., Al-Olaimat, Y., Abdelfattah, A. S., Jarrar, R. J., Almudallal, B. M., and Abu Ajamieh, Z. I. (2021). Existential vacuum and external locus of control as predictors of burnout among nurses. *Nurs. Reports* 11, 558–567. doi: 10.3390/nursrep11030053

Battista, J., and Almond, R. (1973). The development of meaning in life. *Psychiatry* 36, 409–427. doi: 10.1080/00332747.1973.11023774

Bekenkamp, J., Groothof, H. A. K., Bloemers, W., and Tomic, W. (2014). The relationship between physical health and meaning in life among parents of special needs children. *Eur. J. Psychol.* 10, 67–78. doi: 10.5964/ejop.v10i1.674

Brouwers, A., and Tomic, W. (2012). Factorial structure of the existence scale. *J. Articles Support Null Hypoth.* 8, 21–30. doi: 10.4103/0973-14329.103602

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/windows: basic concepts, applications, and programming, vol. *159*. Thousand Oaks, CA: Sage, 345.

Campbell, A., Converse, P. E., and Rodgers, W. L. (1976). The quality of American life: perceptions, evaluations, and satisfactions. New York, NY, U.S.: Russell Sage Foundation.

Crumbaugh, J. C. (1968). Cross-validation of purpose in life test based on Frankl's concepts. *J. Individ. Psychol.* 24, 74–81. doi: 10.1037/h0025929

Frankl, V. E. (1962). Psychiatry and man's quest for meaning. *J. Relig. Health* 1, 93–103. doi: 10.1007/BF01532076

Frankl, V. E. (1963). Man's search for meaning. New York, NY: Simon & Schuster.

Frankl, V. E. (2004). On the theory and therapy of mental disorders: an introduction to logotherapy and existential analysis. New York, NY: Brunner-Routledge. doi: 10.4324/9781003401896

Hu, L. T., and Bentler, P. M. (1995). "Evaluating model fit" in Structural equation modeling: concepts, issues, and applications. ed. R. H. Hoyle, vol. *6* (Thousand Oaks, CA: Sage), 1–55.

Kay, J. (2014). Existential fulfillment, work engagement and job burnout. *Kabarak J. Res. Innov.* 2, 89–101. doi: 10.4172/2155-9899.1000125

Kay, J. (2019). Influence of existential fulfillment on job burnout among Kenya universities professional counsellors. *J. Psychol. Guid. Couns.* 1, 1–9. doi: 10.51317/ecjpgc.v1i1.84

Längle, A. (2011). Emotionality: an existential-analytical understanding and practice. *Reconstruct. Emot. Spaces Exper. Regulat.* 2011, 41–62.

Längle, A., Orgler, C., and Kundi, M. (2003). The existence scale: a new approach to assess the ability to find personal meaning in life and to reach existential fulfillment. *Eur. Psychother.* 4, 135–151. doi: 10.1027/1016-9040.4.1.135

Levchenko, L. S. (2024). Correlation of various aspects of internet addiction with characteristics of existential fulfillment. In 2024 IEEE 25th international conference of young professionals in Electron devices and materials (EDM) (2290–2293). IEEE.

Loonstra, B., Brouwers, A., and Tomic, W. (2007). Conceptualization, construction and validation of the existential fulfillment scale. *Eur. Psychother.* 8, 62–77. doi: 10.1558/ruth.v8i1.62

Loonstra, B., Brouwers, A., and Tomic, W. (2010). Feelings of existential fulfilment and burnout among secondary school teachers. *Teach. Teach. Educ.* 25, 752–757. doi: 10.1016/j.tate.2009.01.002

Maehr, M. L. (1974). Culture and achievement motivation. *Am. Psychol.* 29, 887–896. doi: 10.1037/h0037521

Mascaro, N., and Rosen, D. H. (2008). Assessment of existential meaning and its longitudinal relations with depressive symptoms. *J. Soc. Clin. Psychol.* 27, 576–599. doi: 10.1521/jscp.2008.27.6.576

Maslow, A. H. (1971). The farther reaches of human nature. New York, NY: Viking Press.

Mausch, K., and Rys, E. (2020). Existential fulfillment, religiousness, values and state of mental health: results of empirical research among students. *Exist. Anal.* 31, 142–161.

May, R., Angel, E., and Ellenberger, H. F. (1958). Existence: a new dimension in psychiatry and psychology. New York, NY: Basic Books.

Mikhaylova, M. V., and Dmitrieva, V. A. (2021). "The interdependence of humanistic values and the level of existential fulfillment of an intellectual leader" in EPJ web of conferences, vol. *248*. ed. R. Frühwirth (Les Ulis, France: EDP Sciences), 03005.

Ratcliffe, M. (2009). Existential feeling and psychopathology. *Philos. Psychiatry Psychol.* 16, 179–194. doi: 10.1353/ppp.0.0229

Riethof, N., and Bob, P. (2019). Burnout syndrome and logotherapy: Logotherapy as useful conceptual framework for explanation and prevention of burnout. *Front. Psych.* 10:382. doi: 10.3389/fpsyt.2019.00382

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *J. Pers. Assess.* 98, 223–237. doi: 10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2016b). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychol. Methods* 21, 137–150. doi: 10.1037/met0000045

Shek, D. T., Chai, W., and Tan, L. (2022). The relationship between anxiety and depression under the pandemic: the role of life meaning. *Front. Psychol.* 13:1059330. doi: 10.3389/fpsyg.2022.1059330

Steger, M. F., Frazier, P., Oishi, S., and Kaler, M. (2006). The meaning in life questionnaire: assessing the presence of and search for meaning in life. *J. Couns. Psychol.* 53, 80–93. doi: 10.1037/0022-0167.53.1.80

Thakur, K., and Basu, S. (2010). A probe of existential meaning in depression. *SIS J. Project. Psychol. Mental Health* 17:228.

Tomic, M. (2016). "Workload, existential fulfillment, and work engagement among city council members" in Logotherapy and existential analysis: proceedings of the Viktor Frankl institute Vienna, vol. *1*. eds. A. Batthyány and C. L. McLafferty, Jr (Cham, Switzerland: Springer International Publishing), 219–235.

Tomic, W., and Tomic, E. (2008). Existential fulfillment and burnout among principals and teachers. *J. Beliefs Values* 29, 11–27. doi: 10.1080/13617670801928191

Tomic, M., and Tomic, E. (2011). Existential fulfilment, workload and work engagement among nurses. *J. Res. Nurs.* 16, 468–479. doi: 10.1177/1744987110383353

Wang, M. C., and Dai, X. Y. (2008). Chinese meaning in life questionnaire revised in college students and its reliability and validity test. *Chin. J. Clin. Psychol.* 5, 459–461.

Wang, X. L., Wang, X. L., and Ma, H. (1999). Handbook of mental health rating scales. Beijing, China: Chinese Mental Health Journal Press.

Wu, M. L. (2010). Practice of statistical analysis of questionnaires. Chongqing, China: Chongqing University Press.

Wu, L. (2018). The diagnosis and treatment of "emptiness syndrome" among Chinese youths from the perspective of modernity. *Contemp. Youth Res.* 1, 79–84. doi: 10.3969/j.issn.1006-1789.2018.01.013

Xu, X. B., Sun, C., and Wang, F. Y. (2017). Spiritual well-being: concept, measurement, related variables and intervention. *Adv. Psychol. Sci.* 25, 275–289. doi: 10.3724/SP.J.1042.2017.00275

Yalom, I. D. (1980). Existential psychotherapy. New York, NY: Basic Books.

Yang, S. G. (1999). Searching for the true meaning of existence. Wuhan: Hubei Education Press.

Yu, A. B. (2008). "Achievement motivation and achievement concepts: an exploration of Chinese cultural psychology" in Chinese indigenous psychology. eds. Y. G. Yang, G. G. Huang and Z. F. Yang (Chongqing: Chongqing University Press), 643–687.

Yu, A. B., and Yang, G. S. (1987). Social-oriented achievement motivation versus individual-oriented achievement motivation: concepts and empirical research. *Bull. Inst. Ethnol. Acad. Sin.* 64, 51–98. doi: 10.1016/S0921-4556(08)62112-X

Zhang, Q. (2018). The guidance of ideological and political education for "Buddhist-like young college students" by college counselors. *Qual. Educ. West* 12, 35–36. doi: 10.16681/j.cnki.wcqe.201812017

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 March 2025
DOI 10.3389/fpsyg.2025.1417443

Check for updates

OPEN ACCESS

EDITED BY
Paola Gremigni,
University of Bologna, Italy

REVIEWED BY
Pierluigi Diotaiuti,
University of Cassino, Italy
Richard Xu,
Hong Kong Polytechnic University,
Hong Kong SAR, China

*CORRESPONDENCE
Changming Duan
✉ duanc@ku.edu

†PRESENT ADDRESS
Kehan Shen,
Department of Psychology, California State University, Sacramento, Sacramento, CA, United States

RECEIVED 15 April 2024
ACCEPTED 03 March 2025
PUBLISHED 24 March 2025

CITATION
Lei Y, Duan C and Shen K (2025) Development and validation of the Chinese mental health value scale: a tool for culturally-informed psychological assessment.
Front. Psychol. 16:1417443.
doi: 10.3389/fpsyg.2025.1417443

Development and validation of the Chinese mental health value scale: a tool for culturally-informed psychological assessment

Yujia Lei[1], Changming Duan[2]* and Kehan Shen[3†]

[1]Center for Counseling & Psychological Services, Washington University in St. Louis, St. Louis, MO, United States, [2]Department of Educational Psychology, University of Kansas, Lawrence, KS, United States, [3]Counseling and Psychological Services, University of California, Berkeley, Berkeley, CA, United States

Background: Mental health values play a significant role in defining, promoting, and intervening in mental health. As part of an individual's value system, these values are inherently cultural. However, they remain underexplored in Chinese culture context, particularly among the large young adult population of university students. A culturally informed tool to assess mental health values is much needed to support China's efforts in promoting mental health among college students.

Methods: Four steps were taken to complete the Chinese Mental Health Value Scale (CMHVS) development, namely, item pool construction including expert reviews, a pilot study for item revision and selection, data collection for explorative and confirmatory factor analyses, and validity testing.

Results: A 35-item, seven-factor model was identified with high reliability (Cronbach's alpha of 0.96), and evidence for convergent validity. The seven factors were *Expected Self, Relating to Others, Life Principles, Family, Purpose and Meaning, Achievement*, and *Communication*.

Conclusion: The CMHVS provides a culturally grounded method for assessing mental health values in Chinese university students. It has potential applications in research and clinical settings, improving culturally sensitive mental health promotion and intervention.

KEYWORDS

mental health, values, Chinese culture, college students, scale development, wellbeing

Introduction

The recent paradigm shift from the traditional healthcare model to population health perspective promoted by American Psychological Association (APA) as an organizing framework (Evans, 2021) repositions psychology's roles in health care and health equity enhancement. This shift also revalidates the call for focusing on mental health (vs. mental illness only) in health care (Haeyen et al., 2018). This was consistent with the two continua model of mental health and illness (Keyes, 2014), which highlighted the distinctive features of mental illness and mental health. The model called for more investigation of mental health (as opposed to mental illness), and provided a refreshing perspective on enhancing preventive and interventive efforts to address the global mental health crisis (Patel et al., 2018). The understanding that promoting mental health may decrease the negative impact of mental illness is consistent with the Eastern *Taiji* philosophy of *Yin* and *Yang* (Scheid, 2016), which underlies Chinese traditional medicine and views mental health and illness as two opposite as well as complementary energies. Strengthening mental health is an essential aspect of reducing mental illness. Understanding mental health values is a critical step toward fostering value-consistent healthy behaviors, mindsets, and overall wellbeing.

Unfortunately, mental health and mental illness are often used as exchangeable terms in both the day-to-day language and professional communications in the western societies (Benton, 2018), which unavoidably leads to decreased attention to mental health in face of mental illness challenges. The sobering fact is, however, that the de facto clinical treatment of mental illness by and large has not yielded the desired outcome of eliminating or even reducing mental illness (Keyes, 2014). Thus, it is time that we reexamine what mental health is and how it can be promoted as a part of the effort to remedy the crisis due to increased mental illness. It becomes concerning if we consider the degree to which mental illness theories and intervention methods based on western medical models have been transported to many parts of the world (Duan, 2019). For instance, western models have strongly influenced mental health education and clinical practice in China (Jia, 2016). Diagnostic labels for mental illnesses have become widely publicized and have entered public awareness, particularly among young adults. This trend led to the intensified debate in recent years regarding the cultural fit of western models and the consequence of overlooking core Chinese values like family, relational harmony, and implicit communication. Thus, it is imperative that the awareness be raised that the effectiveness of mental illness interventions for specific populations rely on the success in mental health promotion, which must be grounded in an understanding the mental health values specific to the cultural contexts.

In this study, we examined and evaluated the mental health values of Chinese college students, a population estimated to comprise of 30.32 million undergraduate and 7.6 million graduate students in 2019 in China (Textor, 2021) and 703,500 overseas (Ministry of Education, 2020). Represent a crucial demographic in China, their mental health is a growing public concern, particularly given academic stress, family and cultural expectations regarding achievement. Presently, mental health interventions for this population both in China and overseas are largely prescribed by western theories and systems (Duan, 2019). The issue of cultural fit is a serious one as mental illnesses "are every bit as shaped and influenced by cultural beliefs and expectations" (Watters, 2011, p. 4). Scholars and practitioners should be reminded that effective counseling is to facilitate psychological growth and bring positive therapeutic changes defined in a given cultural context (Wampold, 2001). Understanding Chinese college students' mental health values in a Chinese cultural context can thus inform more effective, culturally sensitive interventions. Moreover, most current mental health scales on Chinese students are based on psychopathology, while this study aimed to assess from mental health, a different perspective that offered a full picture based on the two continua model (Keyes, 2014).

Mental health, values, and culture

Defining mental health has not been easy, partly due to the tangle between mental health and mental illness conceptions, and partly the cultural nature of the construct. The World Health Organization (WHO) provides a broad definition containing "a state of well-being in which an individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and is able to make a contribution to his or her community" (World Health Organization, 2018, p. 1). This definition challenges western medical model based clinical practice and draws attention to both mental illness reduction and mental health enhancement. Moreover, the WHO's conceptualization makes it clear that mental health is a culture-determined phenomenon (Bains, 2005).

Cultural values determine personality and pathology (Marsella and Yamada, 2000) and influence how mental health is viewed, assessed, and promoted (Jensen and Bergin, 1988; Tyler et al., 1983). Understanding cultural specific mental health values "as orienting beliefs about what is good or bad for clients and how that good can be achieved" (Jensen and Bergin, 1988; p. 290) is critical in combatting mental illness and promoting health. In a global context, mental health services cannot ignore the differences in values across countries and cultures as they reflect individuals' interaction with their cultural contexts (World Health Organization, 2004).

Chinese cultural context

Traditional cultural values rooted in Confucianism, Taoism, and Buddhism are built into people's deep structure of mind and shape psychology (Hwang, 2012). Distinctive cultural features include a supreme emphasis on family (*jia*, 家), relationship (*guanxi*, 关系), and harmony (*he*, 和), which leads the culture being viewed as collectivistic and relational in nature (Gold et al., 2002). The Chinese *wo* (我—I or Me in English), the source of the idea of the "self," for instance, is defined in relational terms and tied to the family (Lin, 2001). Thus, Filial Piety (*xiao*, 孝), an ancient Chinese philosophical concept, has remained an important moral tenet that expects children to offer love, respect, support, and deference to their parents and other elders in the family. Moreover, due to the paramount importance placed in relationship, family, and harmony, the society values the individual self as being field-dependent, contextual, relationship/family/others-oriented, holistic, interdependent, social-centric, and authoritarian-oriented, which sharply contrasts the individualistic definition of healthy self as field-independent, separated, unique, self-sufficient, egocentric/auto-centric, self-absorbed, and egalitarian (Yang, 2003). Li et al. (2019) found empirical evidence that Chinese tended to have an interdependent self-construal that was positively associated with relational self-esteem (vs. personal self-esteem). Liu and Goto (2007) revealed a significant relationship between self-construal and mental distress and family relations among Asian American adolescents.

Chinese culture is viewed being high in context, placing high importance of context in interpersonal relationship and communication (Hall, 1976). People tent to see themselves as part of networks of friends and family, and are socialized to be observant of others' reactions to the self. In communication, speakers rely on context information other than mere words to convey meaning they intend to deliver (Kim et al., 1998), and show interpersonal sensitivity and use feelings to guide behavior (Gudykunst, 2001). As the result, implicit communication (*hanxu*, 含蓄), listening-centered communication (*ting hua*, 听话), polite communication (*ke qi*, 客气), and insider/in group-communication (*zi ji ren*, 自己人) are highly valued. As Gao (1998) stated, "the primary functions of [Chinese] communication are to maintain existing relationships among individuals, reinforce role and status differences, and to preserve harmony within the group" (p. 168).

Empirical research on mental health values

Only a few early studies are available in the English literature on mental health values, not including those focusing on specific religious/spiritual values. The Mental Health Values Questionnaire (MHVQ) developed by Tyler et al. (1983) measured a person's conception of good mental health in self-acceptance, negative traits, achievement, affective control, good interpersonal relationships, untrustworthiness, religious commitment, and receptivity to unconventional experiences. Tyler et al. (1989) showed that the degree of agreement between counselor and patient on several MHVQ subscales had a positive influence in therapeutic outcomes in treatment of inpatient chemical dependency clients. To examine its cross-cultural utility, Tyler and Suan (1990) discovered that Japanese American undergraduate students scored higher on good interpersonal relations, trustworthiness, and absence of negative personal traits than Caucasian American students.

Jensen and Bergin (1988) surveyed a large group of psychologists, therapists, social workers, and psychiatrists (94% Caucasian, 60% male), using a questionnaire developed for the study. They found a consensus that certain basic values are important for defining mentally healthy lifestyles and for guiding and evaluating psychotherapy. These values include, in order of importance assigned: competent perception and expression of feelings, freedom/autonomy/responsibility, integration/coping ability, self-awareness/growth, human relatedness/interpersonal and family commitment, self-maintenance and physical fitness, mature values, forgiveness, regulated sexual fulfillment, and spirituality/religiosity. A few years later, Kelly (1995) developed their Mental Health Value Survey for members of American Counseling Association. They found seven highly endorsed values, namely, positive human relatedness, compassionate responsiveness, responsible self-expression, forgiveness, autonomy, purposeful personal development, and sexual acceptance. They also found differences associated with age, gender, and theoretical orientations on some of the values as well.

The present study

The present study sought to highlight the importance and cultural nature of mental health values by: (1) developing a Mental Health Value Scale for use with Chinese college students; and (2) validating the scale by investigating the relationships between mental health values and broad cultural values (individualism–collectivism and Asian cultural values) and psychological wellbeing (life satisfaction and psychological stress).

Methods

Scale development

Following the scale development procedures outlined by DeVellis (2011), we took the following four steps to complete the Chinese Mental Health Value Scale (CMHVS) development. Informed consent was appropriately obtained from participants in each step. We used expert review (Chinese psychologists working and training in the U.S. and visiting scholars from China) at each step to minimize the interference of researcher biases.

Step 1. To build an item pool, we first did a thorough review of the literature on Chinese cultural values. Then we conducted a brief survey through email to students ($N = 96$) from a university in Beijing. The students were asked to describe characteristics of people they perceive as mentally healthy or unhealthy. The first author and another Chinese doctoral candidate in Educational Psychology reviewed and coded the responses and obtained seven domains with 16 sub-domains to reflect Chinese mental health values. Next, we conducted two focus groups, one with three Chinese professionals who had at least 1 year of clinical experience working with college students, and one with five college students. Participants shared their views and perceptions regarding "what matters and what is important?" in mental health. The first author transcribed and coded the transcripts and obtained 154 items in eight domains with 19 sub-domains. Lastly, a panel of four Chinese expert reviewers reviewed the items independently before a group decision was made to use the selected 102 items describing eight domains in the pool. The eight domains were Balance, Relationship, Emotion, Capacity, Following Norms, Self, Philosophy of Life, and Modernization.

Step 2. To evaluate and improve prospective items, we conducted a pilot study by asking participants to rate the 102 items through a Qualtrics survey. Sixty-seven Chinese college students took the survey and rated each item in both the importance to their mental health and to that of others. They were also asked to provide comments on the items if they had any. The result showed that all but eight items were perceived as important or very important to themselves and others. Then six Chinese doctoral students in counseling/clinical psychology did another round of review and edited the language based on participants' comments on items. This process resulted in an improved version of CMHVS with 94 items (including 33 reversed items).

Step 3. We administered the 94 items along with a demographic questionnaire (and four scales for validity testing. See details in Step 4) to the targeted population. To establish the factor structure of CMHVS, an exploratory factor analysis (EFA) was conducted and followed by a confirmatory factor analysis (CFA). The data collection was done by sending the link of a Qualtrics survey either through social media *WeChat* or through collaborating colleagues in China for distribution.

Data were collected from 1,597 participants (Table 1). We deleted 223 cases that had more than 75% of the total responses missing and 289 cases for failure on the validity-check item ("Please choose 'not important to me at all.'). The remaining missing values were replaced by the value of −999. To check for univariate outliers, z scores were examined (i.e., above 3.29) for each of the scale totals (Tabachnick and Fidell, 2012). Twelve cases above 3.29 were identified and treated as missing values and replaced with the Exception-Maximization technique. In addition, Mahalanobis distances among the variables

TABLE 1 Participant demographics.

Characteristics	EFA		CFA	
	n	%	*n*	%
Gender				
Female	343	64.84%	365	69.00%
Male	186	35.16%	164	31.00%
Age				
18–22	351	66.35%	352	66.54%
23–26	152	28.73%	152	28.73%
27–30	16	3.02%	14	2.65%
30–40	10	1.89%	11	2.08%
Degree				
Associate degree	52	9.8%	57	10.78%
Undergraduate	298	56.33%	294	55.58%
Master's	159	30.06%	165	31.19%
Doctoral	20	3.78%	14	2.65%
Grade				
Freshman	107	20.23%	112	21.17%
Sophomore	98	18.53%	90	17.13%
Junior	75	14.18%	69	13.04%
Senior	70	13.23%	80	15.12%
Department/Program				
Literal Arts	27	5.10%	30	5.67%
Science	69	13.04%	59	11.15%
Engineering	106	20.04%	109	20.60%
Philosophy	4	0.76%	7	01.32%
Medical	46	8.70%	51	9.64%
Law	42	7.94%	38	7.18%
Agriculture	2	0.38%	1	0.19%
Management	40	7.56%	67	12.67%
Economics	40	7.56%	35	6.62%
Education	66	12.48%	72	13.61%
Art	13	2.46%	14	2.65%
Others	45	8.51%	44	8.32%
Ethnicity/Race				
Han	491	92.82%	484	91.49%
Zhuang	10	1.89%	13	2.46%
Hui	5	0.95%	4	0.76
Man	4	0.76%	7	1.32%
Miao	3	0.57%	1	0.19%
Others	16	3.02%	20	3.78%
Religious beliefs				
No	476	89.98	490	92.63%
Yes	49	9.26%	37	6.99%
Missing	4	0.76%	3	0.57%

(Continued)

TABLE 1 (Continued)

Characteristics	EFA		CFA	
	n	%	*n*	%
Only child				
Yes	226	42.72%	223	42.16%
No	303	57.28%	306	57.85%
Relationship status				
Single	319	60.30%	332	62.76%
Dating	194	36.68%	181	34.21%
Married	16	03.02%	16	03.02%

EFA, Exploratory factor analysis; CFA, confirmatory factor analysis.

were used to examine multivariate outliers, which resulted in five cases being deleted (probability of Mahalanobis D^2 is less than 0.001). The remaining 1,058 cases were then randomly split into two data sets, Sample A ($n = 529$; Mage = 21.6, SDage = 3.0) and Sample B ($n = 529$; Mage = 21.6, SDage = 3.1). See Table 1 for detailed demographics of the samples. Then EFA was performed using Sample A and CFA Sample B. We randomly divided the sample to ensure that EFA and CFA used different samples.

Step 4. To obtain evidence of convergent validity, we conducted correlational analyses of CMHVS with four possible correlates, namely, individualism–collectivism orientation, Asian culture values, life satisfaction, and psychological stress. The following instruments were included in the data collection.

Individualism and Collectivism Scale (ICS; Triandis and Gelfand, 1998). This 16-item instrument was to measure individualism and collectivism values. It contains four subscales representing Horizontal Individualism (HI), Vertical Individualism (VI), Horizontal Collectivism (HC), and Vertical Collectivism (VC). Participants' responses were recorded on a 9-point Likert scale, ranging from 1 (never or definitely no) to 9 (always or definitely yes). The scale showed good reliability in a Chinese student sample and the test–retest reliability is 0.75 (Huang et al., 2006).

Asian Values Scale (AVS; Kim et al., 1999) was designed to measure client adherence to Asian cultural values. AVS contains 24 statements reflecting Asian cultural values such as collectivism, conformity to norms, emotional self-control, family recognition through achievement, filial piety, and humility. Based on a Confirmative Factor Analysis we conducted in this study, only two subscales (Conformity to Norms and Humility) appeared to have adequate reliability (0.688 and 0.731). Participant responses were recorded by a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). The scale showed good test–retest reliability at 0.83 (Kim et al., 1999).

Satisfaction with Life Scale (SWLS; Diener et al., 1985) measures people's general degree of satisfaction with their lives. It contains five items and uses a 7-point Likert scale, ranging from 1 (strongly agree) to 7 (strongly disagree). The internal consistency reported by Pavot and Diener (1993) was 0.85.

The Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000) was designed to measure psychological distress including depression, anxiety, and somatic symptoms. It consists of 18 symptoms experienced within the past 7 days and uses a 5-point Likert scale to record responses, ranging from 1 (not at all) to 5 (extremely). Scores can be obtained for anxiety, depression, and somatization individually in addition to the Global Severity Index (GSI). The BSI-18 has shown good internal consistency (0.92) for the GSI in Chinese samples (Wang and Mallinckrodt, 2006).

Results

Exploratory factor analysis

Exploratory factor analysis (EFA) with Sample A ($n = 524$) was conducted by the maximum likelihood estimation method in Mplus program (6th Version). A parallel analysis with all 94 items showed that two factors should be retained, with one being occupied by 32 out of 33 reversed items. Research suggested that reversed items often can create an artificial factor due to its wording (Kam, 2023). Given that the reverse items were randomly selected from the original item pool without any theoretical reasons, the 2-factor model was more likely to reflect a response pattern instead of item content. Additionally, a careful review showed that the contents of the reversed items were well represented by other non-reversed items. Therefore, we dropped the 33 reverse items and used the rest 61 items in the following analysis.

Table 2 presents indices of seven models, representing 1-, 2-, 3-, 4-, 5-, 6-, 7-factor models as Mplus stopped estimating after seven factors were extracted. The underlying dimensions of each model were carefully examined and compared to each other under the scrutiny of the expert panel (the same panel for the pilot study). The panel agreed that the 7-factor model could be adequately explained by the theoretical framework recognized in the literature review. The fit statistics supported that the 7-factor model as more parsimonious than other models as well.

In making item retention decisions, we used three criteria: (a) loading at least 0.40 on one factor, only allowing items with theoretical significance when the loading is slightly lower, (b) cross-loadings not exceeding 0.30, and (c) retaining factors that each had at least three items (Kahn, 2006). Based on the 7-factor model, seven items with low loadings (less than 0.40 on any factor) were deleted. A repeated EFA with the remaining items (Kahn, 2006) resulted in three additional items with loadings less than 0.40 and two items with cross-loadings higher than 0.30 deleted. Subsequently, the scale was modified

TABLE 2 Fit statistics for EFA models with initial 61 items.

Model	χ^2	df	CFI	TLI	RMSEA	SRMR	AIC	BIC
Unstructured	5,819.05**	1,769	0.77	0.76	0.07	0.05	92,897.38	93,678.97
2 Factors	5,120.83**	1,709	0.81	0.79	0.06	0.05	92,319.16	93,357.01
3 Factors	4,445.93**	1,650	0.84	0.82	0.06	0.04	91,762.25	93,052.09
4 Factors	3,830.55**	1,592	0.87	0.85	0.05	0.03	91,262.87	92,800.43
5 Factors	3,523.46**	1,535	0.89	0.87	0.05	0.03	91,069.79	92,850.79
6 Factors	3,201.58**	1,479	0.90	0.88	0.05	0.03	90,859.91	92,880.09
7 Factors	2,902.64**	1,424	0.92	0.89	0.04	0.03	90,670.97	92,926.05

**$p < 0.001$. CFI, Comparative Fit Index; TLI, Tucker-Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; AIC, Akaike; BIC, Bayesian.

through the same iterative process of deleting the five weakest items and three cross loaded items. In addition, the meaning of each item was examined, and the six items that were deemed unfit to the belonging factor were removed. It was worth noting that all the 26 items were removed one at a time and the EFAs were run multiple times until the final model is reached.

Good model fit indices were achieved for the final scale that contains 35 items loaded on seven factors, χ^2 (371) = 709.12, CFI = 0.96, TLI = 0.94, SRMR = 0.02, RMSEA = 0.04, 90% confidence interval (0.04, 0.05). Our expert panel also deemed the seven factors as conceptually distinctive and theoretically sound. Although loadings of four items (#17, #18, #29, #35) were slightly below 0.40, they were retained in the scale for two reasons: (1) the theoretical framework strongly supported their existence; (2) their loadings were distinctively higher on the factor than other factors. Table 3 presents the seven factors and their respective items with factor loadings.

Factor 1 (11 items) was labeled *Expected Self*. It reflects the values of emotional regulation, daily life functioning, mental strength and resilience, balance between mental and physical wellbeing, and consistency between private and social life. The highest loading items were, "being aware of one's own negative emotions" and "adjusting to changes at different life and developmental stages."

Factor 2 (7 items) was labeled *Relating to Others*. It describes the value of connecting and maintaining harmonious relationships with others. All items emphasize what an individual self can contribute to others' welfare, and how an individual self can relate to others, such as "being capable of loving others" and "feeling gratitude." The highest loading items were, "bringing positive energy to others" and "helping people in need."

Factor 3 (3 items) was labeled *Life Principle*. This factor reflects the Confucian value of being modest and prudent. Modesty is one of the core values in Chinese culture, which not only teaches unpretentiousness and avoiding bragging but also promotes humbleness in relating to others. The highest loading items were "being modest" and "being prudent."

Factor 4 (4 items) was labeled *Family*. This factor demonstrates the value of one's family, especially having a close relationship with parents. The highest loading items were, "having healthy parents and family members" and "having a harmonious family."

Factor 5 (4 items) was labeled *Purpose and Meaning*. This factor reflects the value of pursuing purpose and meaning in life. The highest loading items were "having clear goals in life" and "being self-aware."

Factor 6 (3 items) was labeled *Achievement*. This factor reflects the value of achieving success in career, academic and social status. The highest loading items were "seeking career success" and "seeking academic success."

Factor 7 (3 items) was labeled *Communication Style*. This factor reflects value of a high-context communication style (i.e., implicit and indirect) and behaviors (i.e., accepting criticism and not deviating from the norm set by the mainstream). The highest loading items were "being able to implicitly express one's opinions and emotions when in disagreement with others" and "being capable of expressing one's emotions and opinions appropriately."

Gender differences were calculated for each item and each factor using Sample B (n = 529, 365 females, 164 males), which resulted in significant difference in six items spread across 6 factors (items 3,15, 19, 28, 30, 33). Meanwhile, two genders differ in the overall score of Factor 6 (Achievement), showing that male students placed higher values (M = 4.8, SD = 1.0) on achievement than female students (M = 4.6, SD = 1.1).

Confirmative factor analysis

A confirmatory factor analysis, using the maximum likelihood estimation method, was conducted on the 35-item Chinese Mental Health Value Scale for Sample B (n = 529). The seven-factor model indicated the following fit indices: χ^2(539) = 1,179.12, CFI = 0.91, TLI = 0.90, SRMR = 0.05, RMSEA = 0.05, 90% confidence interval (0.04, 0.05). Due to the low indices of TLI, modification provided by Mplus output was considered to improve the indices based on theoretical and statistical justifications (Muthén and Muthén, 2007). Under careful examination, four modifications were included in the final CFA model because of correlations between four pairs: items 2 and 1 (factor 1), items 38 and 39 (factor 3), items 33 and 49 (factor 4), and items 64 and 65 (factor 2). As a result, the final 7-factor model with modifications indicated good fit indices, χ^2 (535) = 1,009.08, CFI = 0.93, TLI = 0.92, SRMR = 0.04, RMSEA = 0.05, 90% confidence interval (0.04, 0.05).

Reliability

Reliability estimates were calculated for Sample A (n = 524) and Sample B (n = 529) respectively. The internal consistency coefficients

TABLE 3 Items, factor loadings, means and standard deviations.

Factors and items	1	2	3	4	5	6	7	M	SD
Factor 1: Expected self									
1. Being aware of one's own negative emotions	**0.64**	−0.01	−0.02	−0.02	−0.01	0.02	−0.01	5.2	1.2
2. Adjusting to changes at different life and developmental stages	**0.61**	−0.00	−0.05	−0.01	0.02	0.08	−0.01	5.5	1.2
3. Taking responsibility for the consequences of one's choices	**0.54**	0.13	0.04	0.04	−0.03	−0.06	0.07	5.6	1.2
4. Balancing physical and mental wellbeing	**0.53**	−0.01	0.12	−0.06	0.09	−0.05	0.03	5.6	1.2
5. Being able to see things from others' perspectives	**0.51**	0.06	0.06	0.09	−0.06	0.03	0.04	5.6	1.2
6. Getting along well with others	**0.49**	−0.01	−0.01	0.16	0.02	0.20	0.03	5.7	1.1
7. Accepting life's adversities	**0.47**	0.10	0.09	−0.09	0.01	−0.07	0.02	5.6	1.2
8. Maintaining normal functioning in daily life activities such as sleeping, eating, working, and studying	**0.45**	−0.04	−0.02	0.19	0.04	0.12	0.01	5.5	1.3
9. Being able to control one's negative emotions	**0.44**	0.16	0.04	−0.05	0.17	−0.08	0.12	5.4	1.2
10. Maintaining harmony in interpersonal relationships	**0.42**	0.09	0.04	0.24	−0.03	0.06	0.03	5.5	1.2
11. Balancing one's private life and social life	**0.40**	0.00	0.03	0.10	0.11	0.28	−0.03	5.0	1.3
Factor 2: Relating to others									
12. Bringing positive energy to others	−0.04	**0.76**	0.02	−0.08	0.01	0.2	−0.00	5.5	1.2
13. Helping people in need	0.01	**0.63**	0.13	0.04	−0.04	0.07	0.05	5.3	1.1
14. Having a peaceful mind	0.15	**0.55**	−0.01	0.08	0.16	0.01	−0.00	5.7	1.1
15. Being capable of loving others	0.18	**0.47**	−0.06	0.14	0.08	−0.02	−0.01	5.6	1.2
16. Feeling gratitude	−0.01	**0.45**	0.16	0.28	0.02	−0.09	0.09	5.8	1.1
17. Getting along well with people from different backgrounds (i.e., geographical, religious, ethnic, etc.)	0.28	**0.38**	0.03	0.06	0.05	0.07	−0.04	5.3	1.2
18. Having dialectical views of people and events	0.15	**0.35**	0.02	0.05	0.01	0.18	0.11	5.2	1.3
Factor 3: Life principles									
19. Being modest	−0.02	0.06	**0.90**	−0.00	0.02	−0.00	−0.02	5.2	1.3
20. Being prudent	0.06	0.04	**0.66**	0.02	0.02	0.12	0.06	5.2	1.3
21. Being neither arrogant about winning nor discouraged after losing	0.10	−0.07	**0.43**	0.06	0.28	0.08	0.04	5.2	1.3
Factor 4: Family									
22. Having healthy parents and family members	0.00	−0.03	0.14	**0.73**	0.05	−0.03	0.05	6.3	1.1
23. Having a harmonious family	−0.02	0.17	0.00	**0.60**	0.23	−0.04	−0.06	6.2	1.1
24. Showing filial piety to one's parents	0.01	0.02	0.20	**0.56**	−0.05	0.19	−0.03	6.2	1.1
25. Having a sense of belonging in one's family	0.07	0.20	−0.07	**0.50**	0.02	0.11	0.10	5.8	1.2
Factor 5: Purpose and meaning									
26. Having clear goals in life	−0.06	−0.01	−0.00	0.05	**0.80**	0.18	0.03	5.6	1.3

(Continued)

10.3389/fpsyg.2025.1417443

TABLE 3 (Continued)

Factors and items	1	2	3	4	5	6	7	M	SD
27. Being self-aware	0.08	0.09	0.03	−0.01	**0.75**	0.00	0.03	5.7	1.2
28. Feeling satisfied	0.12	0.24	0.07	0.01	**0.43**	−0.05	−0.04	5.3	1.3
29. Looking for meaning in life	0.07	0.23	−0.02	−0.02	**0.36**	0.15	0.13	5.5	1.2
Factor 6: Achievement									
30. Seeking career success	0.116	0.033	0.14	−0.03	−0.02	**0.60**	0.01	4.9	1.3
31. Seeking academic success	−0.103	0.16	0.02	0.02	0.20	**0.55**	0.11	5.0	1.3
32. Having a respectable social status	0.106	−0.006	0.05	−0.01	0.04	**0.53**	0.00	4.2	1.5
Factor 7: Communication style									
33. Being able to implicitly express one's opinions and emotions when in disagreement with others	−0.021	0.005	0.00	−0.01	0.02	−0.02	**0.95**	5.1	1.2
34. Being capable of expressing one's emotions and opinions appropriately	0.09	0.004	0.03	0.07	0.01	0.05	**0.67**	5.3	1.2
35. Being able to avoid interpersonal conflict with others	0.046	0.18	0.05	−0.02	0.04	0.16	**0.35**	5.1	1.3

N = 529. Bold values represent factor loadings exceeding 0.35. Using the varimax for orthogonal rotations.

for the subscales of the CMHVS and the whole scale were acceptable to very good. For Sample A, internal reliability (Cronbach's alpha) was 0.96 for the whole CMHVS and 87, 0.88, 0.84, 0.83, 0.85, 0.72, 0.81 for its seven subscales, respectively. For Sample B, the internal reliability was 0.95 for the total scale and 0.87, 0.87, 0.81, 0.80, 0.82, 0.70, 0.77 for its seven subscales, respectively. The seven latent factors were strongly correlated with each other, ranging from 0.51 to 0.85. Correlations between the factors were presented in Table 4.

Convergent validity

Validity of CMHVS was tested by examining its correlation with AVS, SWLS, COS, and BSI-18. Results were shown in Tables 5**,** 6. As expected, the total score and seven subscale scores of CMHVS were all positively correlated with AVS and SWLS. Furthermore, the seven subscales of CMHVS were positively and significantly correlated with all four subscales (HI, VI, HC, and VC) of COS, the magnitude of the correlations was consistently larger for Collectivism-related than Individualism-related subscales, and the peak correlation appeared between Horizontal Collectivism and CMHVS. There was no significant correlation between the BSI-18 and CMHVS.

Discussion

As expected, the study results offered knowledge and insight regarding Chinese college students' mental health values and produced an evaluation scale with sound psychometric qualities. The CMHVS shows high reliability with Cronbach's alpha of 0.96, and convergent validity with its high positive correlations with AVS, SWLS, and COS. From a cultural perspective, this scale tightens the connection between mental health and cultural values. Further, it offers support to practice that aims at promoting culturally supported and functional mental health for Chinese college students.

Mental health as a culturally informed phenomenon

Consistent with the principle of cultural psychology, "One mind, many mentalities" (Shweder et al., 1998), CMHVS sheds light on both the universal mind of human beings and the culturally specific mentality regarding mental health for Chinese colleges students. CMHVS shares some commonalities with the available mental health value scales in the U.S. (Jensen and Bergin, 1988; Kelly, 1995; Tyler et al., 1983). The most obvious is that all scales contain largely positive concepts and desirable behaviors, implying that mental health is a positive construct. For instance, shared subscales include interpersonal relationship, emotional control/regulation, and personal success/achievement. Although there are differences in the specific content or scope/depth of each area across measures, these subscales nonetheless reflect the common and basic dimensions of human mental experience.

Differences between CMHVS and western scales are notable as well. First, CMHVS appeared to reflect collectivistic/relational values more and individualistic values less than the existing western scales. Expectedly Chinese students failed to show high value of the western cultural ideals such as independence, autonomy, self-reliance, and

10.3389/fpsyg.2025.1417443

TABLE 4 Correlations among seven factors of CMHVS.

Factor	1	2	3	4	5	6	7
1. Expected self	—						
2. Relating to others	0.77**	—					
3. Life principles	0.77**	0.85**	—				
4. Family	0.62**	0.85**	0.72**	—			
5. Purpose and meaning	0.65**	0.78**	0.72**	0.68**	—		
6. Achievement	0.61**	0.60**	0.76**	0.51**	0.58**	—	
7. Communication	0.68**	0.74**	0.70**	0.61**	0.66**	0.61**	—
M	5.45	5.47	5.22	6.10	5.53	4.67	5.15
SD	0.78	0.90	1.10	0.89	1.00	1.10	0.99

CMHVS, Chinese Mental Health Value Scale.
**Correlation is significant at the 0.001 level.

explicit and direct communication, which are the foundations of western psychology. Because Chinese are more likely to define self-in-relations, and a collectivistic/relational orientation is more discernable in CMHVS than in other scales. The psychological anthropologist Alan Page Fiske famously categorized human interactions into four types, communal sharing, authority ranking, equality matching, and market pricing (Fiske, 1991). The values in CMHVS seem to reflect more communal sharing (e.g., "bringing positive energy to others") and authority ranking (e.g., "showing filial piety to parents") and less equality matching and market pricing than existing western scales. The Confucian social relationism is shown throughout the scale especially in the subscales of *Expected Self*, relation to others, and communication. The others-focus mentality is clear and salient in how one expects the self to take other's perspective and stay in harmony with and being a giver/helper to others.

Secondly, the highly valued emotional expression and taking active actions for change in western cultures are weak in CMHVS. Instead, a status of *being* is salient in "having a peaceful mind," "being modest," "feeling gratitude," "feeling satisfied," or "being modest." The Chinese folk wisdom values one's capacity to control negative emotions and accept life adversities as a way of dealing with stress. Additionally, values based in dialecticism emphasize the balance and harmonious states in physical, emotional, interpersonal and spiritual life, such as "being neither arrogant about winning nor discouraged after losing," etc. Such passive ways of being and relating to the world may not fit dominant western values but certainly contribute to peaceful mind maintenance and conflict avoidance for Chinese.

Indeed, CMHVS portrays that it is the culture that determines what traits, emotions, and behaviors would be experienced as healthy or unhealthy. The *guanxi* (relationship) orientation, *he* (harmony) mentality, and *jia* (family) centrality are distinctively reflected in multiple subscales. Rather than focusing on independence and autonomy as valued by western cultures, the *Expected Self* includes being able to take others' perspective, building harmonious relationship with others and managing all relationships in life to maintain harmony. Being altruistic and grateful to and loving others as well as bringing positive energy to others are valued as ways of relating to others. These values reflect the self-in-relation framework as a key expression of Chinese relationism (Hwang, 2012).

Cultural values, mental health values, and psychological wellbeing

The study result showed that participants who held higher Asian values endorsed CMHVS more, meaning mental health values are consistent with general cultural values. Moreover, the seven subscales of CMHVS were positively correlated with all four subscales of Collectivism–Individualism Scale, showing that individualism and collectivism are not dichotomous but co-existent for contemporary Chinese college students, although participants reported higher collectivism than individualism. Notably, the peak correlations exist between Horizontal Collectivism (emphasis on sociability and interdependence) and almost all subscales of CMHVS, which shows that Chinese college students highly value relationships. The high correlations between both Hierarchical and Vertical individualism scales and *Achievement* (Factor 5) of CMHVS, however, seemed to show that participants were aware that they need to work hard and be competitive to succeed or excel in today's environment where academic stress is extremely high.

As expected, CMHVS was positively correlated with life satisfaction, which means that Chinese college students' mental health values are functional to them in increasing life satisfaction. This relationship becomes even more significant when considering the lack of correlation between CMHVS and psychological symptoms measured by BSI-18 except for the negative correlation between Somatization and Depression and *Expected Self* (Factor 1). It is worth noting that many items of *Expected Self* describe personal strengths and resilience, such as being aware of one's negative emotions, controlling one's negative emotions, and accepting adversities in life. Likely, such personal strengths may prevent depression and somatization symptoms. Compared to Western scales like the Mental Health Continuum–Short Form (MHC-SF; Keyes, 2005), which emphasizes autonomy and self-acceptance, the CMHVS highlights relational and familial dimensions of wellbeing. While both scales capture psychological flourishing, CMHVS uniquely integrates Confucian ideals of self-cultivation and collective harmony, providing a more culturally relevant measure for Chinese students.

10.3389/fpsyg.2025.1417443

TABLE 5 Intercorrelations among subscales of CMHVS and five referent measures.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Expected self	—																
2	Relating to others	0.79**	—															
3	Life principles	0.68**	0.74**	—														
4	Family	0.65**	0.80**	0.65**	—													
5	Achievement	0.65**	0.71**	0.72**	0.62**	—												
6	Purpose and meaning	0.69**	0.80**	0.60**	0.67**	0.67**	—											
7	Communication	0.68**	0.73**	0.62**	0.57**	0.68**	0.67**	—										
8	AVS_norm	0.21**	0.23**	0.18**	0.31**	0.19**	0.18**	0.14**	—									
9	AVS_humility	0.35**	0.45**	0.53**	0.37**	0.31**	0.31**	0.33**	0.28**	—								
10	SWLS	0.18**	0.22**	0.15**	0.11**	0.15**	0.12**	0.15**	−0.01	0.18**	—							
11	COS_HI	0.33**	0.33**	0.31**	0.22**	0.40**	0.35**	0.33**	−0.17**	0.31**	0.17**	—						
12	COS_VI	0.32**	0.31**	0.34**	0.26**	0.57**	0.29**	0.31**	0.08	0.30**	0.07	0.84**	—					
13	COS_HC	0.49**	0.55**	0.44**	0.44**	0.47**	0.41**	0.43**	0.21**	0.52**	0.33**	0.45**	0.44**	—				
14	COS_VC	0.29**	0.38**	0.36**	0.47**	0.42**	0.27**	0.29**	0.27**	0.49**	0.19**	0.43**	0.49**	0.71**	—			
15	BSI_Anxiety	−0.07	−0.02	−0.02	−0.03	0.04	0.01	−0.04	−0.13*	−0.12*	−0.34**	0.01	0.09	−0.19**	−0.11**	—		
16	BSI_Soma	−0.12**	−0.01	−0.02	−0.07	0.04	−0.02	−0.03	−0.14*	−0.19**	−0.31**	0.01	−0.01	−0.16**	−0.13**	0.82**	—	
17	BSI_Depre	−0.08*	−0.03	−0.04	−0.05	0.03	0.00	−0.05	−0.12*	−0.16**	−0.43**	0.04	0.13**	−0.21**	−0.11*	0.95**	0.79**	—

CMHVS, Chinese Mental Health Value Scale; 1–7, seven factors of CMHVS; AVS, Asian Value Scale; AVS_norm, AVS_Conformity to Norm; SWLS, Satisfaction with Life Scale; COS, Cultural Orientation Scale; HI, Hierarchical Individualism; VI, Vertical Individualism; HC, Hierarchical Collectivism; VC, Vertical Collectivism; BSI, Brief Symptom Inventory. *$p < 0.05$, **$p < 0.01$.

TABLE 6 Correlations among total scores of CMHVS and referent measures.

	CMHVS	AVS	SWLS	COS_Indi	COS_Coll	BSI
CMHVS	—					
AVS	0.29**	—				
SWLS	0.14**	−0.03	—			
COS_Indi	0.32**	0.04	0.08*	—		
COS_Coll	0.45**	0.28**	0.25**	0.34**	—	
BSI	−0.05	−0.14*	−0.31**	0.07*	−0.15**	—

CMHVS, Chinese Mental Health Value Scale; AVS, Asian Value Scale; SWLS, Satisfaction with Life Scale; COS, Cultural Orientation Scale; COS_Indi, COS_Individualism; COS_Coll, COS_Collectivism; BSI, Brief Symptom Inventory.
*$p < 0.05$, **$p < 0.01$.

Characteristics of college students' mental health values

China is believed to have experienced large and transformative cultural changes in recent decades, including rising importance of materialism, individualism, freedom, democracy (Xu and Hamamura, 2014). Interestingly, however, the values captured by CMHVS generally do not show clear support to this perceived trend but back the enduring theme of family and relationship. Students' adherence to traditional cultural values in viewing mental health is apparent. The values such as relationship, harmony, filial piety, and context-driven communication are underlying their mental health values in *Expected Self*, *Relating to others*, *Family*, and *Communication* subscales.

However, indirect support for modernization of values can be seen through the positive correlations between the individualism subscales of COS and seven factors of CMHVS. Perhaps in college students' mind independence, autonomy, self-reliance, and other individualistic concepts are important for them if they want to be successful. The item "taking responsibility for the consequences of one's choices" fits the traditional value of being a responsible and accountable person on one hand and is consistent with the individualistic notion of being self-reliant on the other. *Achievement* is an example of Chinese college students, especially male students, operating in the mist of both the old and new value systems.

Implications for practice and research

The CMHVS established a solid connection between Chinese college students' mental health values and their cultural contexts. Using it as an assessment tool or as a reference, researchers and clinicians may be reminded that mental health promotion for Chinese students needs to be examined with full consideration of their values.

For clinicians trained with western theories in China or elsewhere, this study can inform them that conceptualization of clients should not be dictated by the western theories. Some valued behaviors may not appear healthy in a western cultural context or by western theoretical perceptions, but needed to be validated and confirmed because they are correlated with wellbeing for Chinese students. When Chinese students present themselves as weak in ego strength or low in self-esteem, clinicians needs to be reminded that the self-in-relation or others-focused relationship is the essence of being for those students. The importance of emotional balance in academic settings has been emphasized in previous research (Diotaiuti et al., 2021), which found that procrastination and emotional self-regulation significantly impact students' performance. Our findings align with this perspective, highlighting the role of mental health values in promoting psychological wellbeing and academic success. CMHVS may be helpful to gaining an understanding their mental health values to avoid premature judgment. Furthermore, practitioners need to be aware of their own values, the values embedded in their theoretical orientations, and their student clients' values to avoid imposing irrelevant values or overriding values of their clients. CMHVS may provide a tool for clinicians to take into consideration of "balancing physical and mental wellbeing" and "balancing one's private life and social life" for their Chinese clients.

CMHVS also provides a research tool for understanding Chinese college students' mental health needs and exploring effective educational/preventive/interventional methods for this population. Specific to counseling, future research may exam the relationship between mental health values and Chinese client in-session behaviors, including openness to various therapeutic interventions. Additionally, it may be meaningful to explore whether and to what extent Chinese students would use those values to monitor their behaviors and guide their daily life. Due to the fast-increasing number of Chinese international students in the United States or other countries, researchers may also use CMHVS to examine its cross-sample validity and its association to these students' acculturation and cultural identity.

Limitations

First, although the sample size of the study was large enough, there was a methodological limitation that the data were randomly split into two samples for EFA and CFA. Thus, the two samples may reflect similar sampling bias, measurement bias, or experimenter effects. Second, the current study did not directly measure western mental health values among Chinese students and compare them with Chinese ones, thus future research can incorporate mental health value scales developed based on western culture. Third, although most reverse items were proportionately represented by other non-reverse items, it was a loss of information to remove them in EFA due to the responding patterns. It is likely that when measuring values, reverse items may cause confusion about how to give precise answers. Fourth, the current research only used CFA in scale validation, while other psychometric methods can also be included for future research such as classical test theory, item response theory, and Rasch model theory (see Xu et al., 2020, for an example). Fifth, given the emphasis on relational and moral values, some participants may have responded in ways that reflect social expectations rather than personal beliefs. Future studies could incorporate social desirability scales to control for this bias. Finally, while this study establishes the CMHVS as a reliable research tool, its clinical utility remains to be tested. Future studies should examine its application in therapy settings to assess its impact on treatment outcomes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

Ethics statement

The studies involving humans were approved by Yujia Lei, University of Kansas IRB. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. CD: Conceptualization, Writing – review & editing. KS: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bains, J. (2005). Race, culture and psychiatry: a history of transcultural psychiatry. *Hist. Psychiatry* 16, 139–154. doi: 10.1177/0957154X05046167

Benton, S. (2018). The difference between mental health and mental illness. Available online at: https://www.psychologytoday.com/us/blog/reaching-across-the-divide/201804/the-difference-between-mental-health-and-mental-illness (Accessed March 15, 2024).

Derogatis, L. R. (2000). The brief symptom inventory–18 (BSI-18): Administration, scoring and procedures manual. Eden Prairie, MN: National Computer Systems.

DeVellis, R. F. (2011). Scale development: theory and applications. Newbury Park, CA: SAGE Publications.

Diener, E., Emmons, R. A., Larsem, R. J., and Griffin, S. (1985). The satisfaction with life scale. *J. Pers. Assess.* 49, 71–75. doi: 10.1207/s15327752jpa4901_13

Diotaiuti, P., Valente, G., Mancone, S., and Bellizzi, F. (2021). A mediating model of emotional balance and procrastination on academic performance. *Front. Psychol.* 12:665196. doi: 10.3389/fpsyg.2021.665196

Duan, C. (2019). Culture and mental health counseling: a reflective view based on observations in China. *Athens J. Soc. Sci.* 6, 71–88. doi: 10.30958/ajss.6-1-5

Evans, A. C. (2021). Population health: utilizing the full breadth of psychology to promote health. A presentation made at American Psychological Association Fall Consolidated Meeting, Washington DC.

Fiske, A. P. (1991). Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing. New York, NY: Free Press.

Gao, G. (1998). "Don't take my word for it."—understanding Chinese speaking practices. *Int. J. Intercult. Relat.* 22, 163–186. doi: 10.1016/S0147-1767(98)00003-0

Gold, T., Guthrie, D., and Wank, D. (2002). Social connections in China: Institutions, culture, and the changing nature of guanxi. Cambridge, MA: Cambridge University Press.

Gudykunst, W. B. (2001). Asian American ethnicity and communication. Thousand Oaks, CA: Sage.

Haeyen, S., van Hooren, S., van der Veld, W. M., and Hutschemaekers, G. (2018). Promoting mental health versus reducing mental illness in art therapy with patients with personality disorders: a quantitative study. *Arts Psychother.* 58, 11–16. doi: 10.1016/j.aip.2017.12.009

Hall (1976). Beyond culture. *1st* Edn: Palatine, IL: Anchor Press.

Huang, R. Z., Yao, S.-Q., and Tao, Z. (2006). Reliability and validity of individualism and collectivism scale in Chinese students. *Chin. J. Clin. Psych.* 114, 563–565.

Hwang, G. (2012). Foundations of Chinese psychology: Confucian social relations. New York, NY: Springer.

Jensen, J. P., and Bergin, A. E. (1988). Mental health values of professional therapists: a national interdisciplinary survey. *Prof. Psychol. Res. Pract.* 19, 290–297. doi: 10.1037/0735-7028.19.3.290

Jia, X. (2016). Psychoanalytic training in China: cultural colonialism or acculturation. *J. Neurosci. Mental Health* 16, 377–382. doi: 10.3969/j.issn.1009-6574.2016.04.002

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: principles, advances, and applications. *Couns. Psychol.* 34, 684–718. doi: 10.1177/0011000006286347

Kam, C. C. S. (2023). Why do regular and reversed items load on separate factors? Response difficulty vs. item extremity. *Educ. Psychol. Meas.* 83, 1085–1112. doi: 10.1177/00131644221143972

Kelly, E. W. (1995). Counselor values: a national survey. *J. Couns. Dev.* 73, 648–653. doi: 10.1002/j.1556-6676.1995.tb01810.x

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

Keyes, C. L. M. (2014). "Mental health as a complete state: how the salutogenic perspective completes the picture" in Bridging occupational, organizational and public health: a transdisciplinary approach. eds. G. F. Bauer and O. Hämmig (Dordrecht: Springer Science Business+Media), 179–192.

Kim, B. S. K., Atkinson, D. R., and Yang, P. H. (1999). The Asian values scale: development, factor analysis, validation, and reliability. *J. Couns. Psychol.* 46, 342–352. doi: 10.1037/0022-0167.46.3.342

Kim, D., Pan, Y., and Park, H. S. (1998). High-versus low-context culture: a comparison of Chinese, Korean, and American cultures. *Psychol. Mark.* 15, 507–521.

Li, J., Liu, M., Peng, M., Jiang, K., Chen, H., and Yang, J. (2019). Positive representation of relational self-esteem versus personal self-esteem in Chinese with interdependent self-construal. *Neuropsychologia* 134:107195. doi: 10.1016/j.neuropsychologia.2019.107195

Lin, N. (2001). Guanxi: A conceptual analysis. In So, A. Y., Lin, N. & Poston, D. (Eds.). The Chinese triangle of Mainland China, Taiwan, and Hong Kong: comparative institutional analyses (p. 301). Greenwood Press, Westport, Conn.

Liu, F., and Goto, S. G. (2007). Self-construal, mental distress, and family relations: a mediated moderation analysis with Asian American adolescents. *Cult. Divers. Ethn. Minor. Psychol.* 13, 134–142. doi: 10.1037/1099-9809.13.2.134

Marsella, A. J., and Yamada, A. M. (2000). "Culture and mental health: an introduction and overview of foundations, concepts, and issues" in Handbook of multicultural mental health (Cambridge, MA: Academic Press), 3–24.

Ministry of Education. (2020). Statistics on Chinese learners studying overseas in 2019. Available online at: http://en.moe.gov.cn/news/press_releases/202012/t20201224_507474.html (Accessed March 15, 2024).

Muthén, L. K., and Muthén, B. O. (2007). Mplus user's guide, vol. *7*. *5th* Edn. Los Angeles: Muthén & Muthén.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., et al. (2018). The lancet commission on global mental health and sustainable development. *Lancet* 392, 1553–1598. doi: 10.1016/S0140-6736(18)31612-X

Pavot, W., and Diener, E. (1993). Review of the satisfaction with life scale. *Psychol. Assess.* 5, 164–172. doi: 10.1037/1040-3590.5.2.164

Scheid, V. (2016). Holism, Chinese medicine and systems ideologies: Rewriting the past to imagine the future. In Whitehead, A., Woods, A., Atkinson, S., et al. (Eds.) The Edinburgh Companion to the Critical Medical Humanities. Edinburgh University Press. Edinburgh, UK.

Shweder, R. A., Goodnow, J., Hatano, G., LeVine, R. A., Markus, H., and Miller, P. (1998). "The cultural psychology of development: one mind, many mentalities" in Handbook of child psychology: theoretical models of human development, vol. *1*. *5th* ed (Hoboken, NJ: John Wiley & Sons Inc.), 865–937.

Tabachnick, B. G., and Fidell, L. S. (2012). Using Multivariate Statistics 6th Ed. Person Education, Boston.

Textor, C. (2021). Number of undergraduate students enrolled at public colleges and universities in China from 2009 to 2019. Available online at: https://www.statista.com/statistics/227028/number-of-students-at-universities-in-china/ (Accessed March 15, 2024).

Triandis, H. C., and Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *J. Pers. Soc. Psychol.* 74, 118–128. doi: 10.1037/0022-3514.74.1.118

Tyler, J. D., Clark, J. A., Olson, D., Klapp, D. A., and Cheloha, R. S. (1983). Measuring mental health values. *Couns. Values* 28, 20–30. doi: 10.1163/2161007X-02801003

Tyler, J. D., Clark, J. A., and Wittenstorm, R. C. (1989). Mental health values and response to alcoholism treatment. *Couns. Values* 33, 204–216. doi: 10.1163/2161007X-03303007

Tyler, J., and Suan, L. (1990). Mental health values differences between native American and Caucasian American college students. *J. Rural Commun. Psychol.* 11, 204–216.

Wampold, B. E. (2001). Contextualizing psychotherapy as a healing practice: culture, history, and methods. *Appl. Prev. Psychol.* 10, 69–86. doi: 10.1017/S0962-1849(02)01001-6

Wang, C.-C. D., and Mallinckrodt, B. (2006). Acculturation, attachment, and psychosocial adjustment of Chinese/Taiwanese international students. *J. Couns. Psychol.* 53, 422–433. doi: 10.1037/0022-0167.53.4.422

Watters, E. (2011). Crazy like us: the globalization of the American psyche. New York, NY: Free Press.

World Health Organization. (2004). Promoting mental health: concepts, emerging evidence, practice. Available online at: https://www.who.int/mental_health/evidence/en/promoting_mhh.pdf (Accessed March 15, 2024).

World Health Organization. (2018). Mental health: strengthening our response. Available online at: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response (Accessed March 15, 2024).

Xu, Y., and Hamamura, T. (2014). Folk beliefs of cultural changes in China. *Front. Psychol.* 5:1066. doi: 10.3389/fpsyg.2014.01066

Xu, R. H., Wong, E. L. Y., Lu, S. Y. J., Zhou, L. M., Chang, J. H., and Wang, D. (2020). Validation of the Toronto empathy questionnaire (TEQ) among medical students in China: analyses using three psychometric methods. *Front. Psychol.* 11:810. doi: 10.3389/fpsyg.2020.00810

Yang, K. S. (2003). Methodological and theoretical issues on psychological traditionality and modernity research in an Asian society: in response to Kwang-Kuo Hwang and beyond. In. *Asian J. Soc. Psychol.* 6, 263–285. doi: 10.1046/j.1467-839X.2003.00126.x

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 20 May 2025
DOI 10.3389/fpsyg.2025.1441174

Check for updates

OPEN ACCESS

EDITED BY
Giulia Casu,
University of Bologna, Italy

REVIEWED BY
Daniela Traficante,
Catholic University of the Sacred Heart, Italy
Su Woan Wo,
Sunway University, Malaysia

*CORRESPONDENCE
Lucia Ráczová
✉ lraczova@ukf.sk

RECEIVED 30 May 2024
ACCEPTED 01 May 2025
PUBLISHED 20 May 2025

CITATION
Ráczová L, Urbánek T, Jurišová E, Popelková M and Sollár T (2025) Assessing early childhood developmental functioning in the parental screening tool: an application of the Rasch model. *Front. Psychol.* 16:1441174. doi: 10.3389/fpsyg.2025.1441174

Assessing early childhood developmental functioning in the parental screening tool: an application of the Rasch model

Lucia Ráczová[1]*, Tomáš Urbánek[2], Erika Jurišová[1], Marta Popelková[1] and Tomáš Sollár[1]

[1]Department of Psychological Sciences, Faculty of Social Sciences and Health Care, Constantine The Philosopher University in Nitra, Nitra, Slovakia, [2]Institute of Applied Psychology, Faculty of Social Sciences and Health Care, Constantine The Philosopher University in Nitra, Nitra, Slovakia

Objectives: Ensuring rapid and efficient detection of developmental difficulties in early childhood necessitates aligning screening tools with the timing of preventive examinations in each country, emphasizing the need for quick and effective unidimensional screening methods.

Aim: This study aims to assess the scalability and unidimensionality of developmental functioning in two- and three-year-old children using Guttman scaling and the Rasch model.

Methods: The anonymized data from 1,640 children aged 26 to 44 months, whose caregivers completed the S-PMV11 method, were gathered from routine pediatric preventive check-ups via the National Database.

Results: The findings indicate the effective scalability of developmental functioning using the Guttman approach, thus enabling the early identification of children at risk. Additionally, the Rasch model confirms the unidimensionality of developmental functioning, highlighting the importance of early intervention addressing.

Conclusion: Despite the instrument's construct validity, significant concerns arise regarding its ability to capture children with borderline developmental functioning. These concerns could be addressed and improved by simply organizing the items by difficulty level in the pediatric response sheet, allowing the pediatrician to effectively identify any early signs of developmental issues.

KEYWORDS

developmental functioning, S-PMV11, Rasch model, Guttman scaling, developmental difficulties, parental screening tool, early childhood

1 Introduction

Developmental difficulties and neurodevelopmental disorders identified in the early years of life are significantly less numerous than their actual prevalence, not only in the at-risk population but also in the general population of children (Olusanya et al., 2023; Boyle et al., 2011). It is estimated that only about 30% of children with these difficulties are identified before entering school (Palfrey et al., 1987). However, recent systematic review (e.g., Francés et al., 2022) report much wider variability in prevalence rates of developmental difficulties and neurodevelopmental disorders, fluctuating globally between 4.70 and 88.50%, influenced by methodological differences, sociocultural factors, and disparities in diagnostic practices and professional training. As a result, early detection programs are increasingly recommended to recognize early warning signs of developmental difficulties and to inform the development of policies aimed at supporting the most vulnerable segments of the population (Francés et al., 2023; Cibralic et al., 2023). The call for early identification of children at risk for developmental difficulties has resulted in the use of psychomotor development screening scales during pediatric preventive check-ups (Yunilda et al., 2023; Prieto et al., 2022; Lipkin et al., 2020; Oberklaid and Efron, 2005).

Developmental difficulty in children manifests through delays in reaching developmental milestones or atypical, limited expression of abilities and skills typical of the child's age (Institute of Medicine, 2001). Developmental milestones reflect a child's optimal developmental functioning and are manifested in the child's skills engaged in everyday interactions with the environment (Misirliyan et al. 2023). Developmental functioning can be defined as an umbrella term for optimal psychomotor development and daily functioning of the child, thus representing a counterpart to disability and impairment. It is understood as the capacity for adequate interaction between an individual's bodily functions, activities, and participation in environmental conditions (World Health Organization, 2007). The inadequate developmental functioning of the child is thought to elicit increased attention from pediatricians and early childhood development specialists (World Health Organization, 2007; Coelho and Pinto, 2018).

Widespread use of parental screening tools can capture child manifestations that may not be observed by the pediatrician during an examination in the clinic (Cibralic et al., 2023; Vitrikas et al., 2017). Parental screenings are a recommended method to refine clinical assessment of a child's developmental functioning (Glascoe, 2005), as it has been shown that up to 80% of caregivers correctly identify abnormalities in a child's development at the age of 2 years (Chawarska et al., 2007). Parental screening tools are therefore inexpensive and effective means of identifying children with developmental difficulties across the population (Prieto et al., 2022). Methodologies such as the Parent's Evaluation of Developmental Status (PEDS) and its complementary part Developmental Milestones (PEDS: DM) and Ages and Stages Questionnaires (ASQ) are recommended by the American Academy of Pediatrics. However, the schedule of preventive examinations during which the screening can be performed and interpreted varies from country to country. Therefore, it is crucial that the screening tool corresponds to the schedule of preventive examinations in a particular country (Cibralic et al., 2023). A mismatch between the dates when screening is required and the schedule of checkups may present a barrier to its use (Lipkin et al., 2020). To align with the pediatric health care system in Slovakia, the Method for Monitoring and Screening for Developmental Difficulties (S-PMV) was developed. This tool is designed to screen for psychomotor development during the preventive checkups in general pediatric care (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021; Ministry of Health, 2021). The compulsory use of S-PMV11 during the preventive examination at the age of 2 or 3 years allows professionals to catch developmental difficulties before the child enters kindergarten. The tool has already established its structural validity and underwent norm standardization in 2013. The current study aims to further examine its structural validity using new, practice-based data, providing an updated perspective based on more recent findings.

Developmental screening tools should be optimized to be fit-for-purpose and sensitive to capturing development as a whole or discrete developmental domains (van Buuren and Eekhout, 2023). In the context of the S-PMV11 instrument, developmental functionality is viewed as unidimensional. A unidimensional screening instrument can produce a scale based on a hierarchy of dichotomous items that progress from the easiest to the most challenging, or from the less to the most extremely delineated (Bond and Fox, 2007). A simple scale meeting Guttman's requirements would allow the pediatrician to verify the accuracy of the parental assessment during the preventive exam by directly administering the selected item that most closely matches the child's current level of ability. Since the S-PMV11 is used at 2 or 3 years of age, it can be assumed that items will be more challenging for two-year-olds than for older children. Analyses will therefore be conducted separately for the two age groups. The first aim is to investigate the scalability of developmental milestones for two- and three-year-olds in the S-PMV11 instrument, which implies the research question RQ1: *Is the scalability of developmental milestones for two- and three-year-olds possible when using Guttman scaling?*

The process of detecting errors in response patterns is also known as scalogram analysis and forms the basis of the Rasch model (Engelhard, 2008). The Rasch model, which is based on the item response theory approach, considers a scale to be a set of homogeneous items (Jelínek et al., 2007). The constructed model then displays the probability of an individual with a certain level of latent ability responding to a particular item on a uniform scale (Embretson and Reise, 2000). Thus, the use of the Rasch model can contribute to validating the accuracy and sensitivity of capturing developmental functioning. Analyses will again be conducted separately for the two age groups. Research question 2 therefore focuses on the use of the Rasch model, RQ2: *Can developmental functionality be considered unidimensional when using the Rasch model in two-year-olds and three-year-olds?*

2 Methods

2.1 Design and data collection

The research design is quantitative, cross-sectional, and oriented toward examining and validating psychometric indicators of the developmental functionality domain in S-PMV11. Data from S-PMV11 were sourced from the National Database of Selected Developmental Indicators (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021). The study sample included anonymized data from children whose primary caregivers completed the screening between April 1, 2021, and December 31, 2021. The entire sample (N = 1,640) consisted of children aged 26 to 44 months. The children were further divided into younger children (< 35 months) and older children (≥ 36 months), according to the age criteria outlined in the S-PMV11 Manual (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021).

2.2 Measurement and instruments

The method whose psychometric properties were validated is called the Developmental monitoring and screening method for the 11th check-up in primary care (shortly S-PMV11). S-PMV11 is one of 10 screening tools that form a single series designed for monitoring the development of psychomotor functions and screening for developmental difficulties based on parental assessment in primary pediatric care. S-PMV11 specifically targets the detection of potential developmental risks in children aged 2 to 3 years. It is obligatory used in pediatric practice in Slovakia. The data were collected online during the 11th preventive check-up in general care for children and adolescents carried out at 2 and 3 years of age (Ministry of Health, 2021). The method was completed by the child's caregiver.

S-PMV11 is a standardized screening tool designed for the early identification of deviations from typical development and behavior in children aged 2 and 3 years. Deviations identified indicate potential developmental risks, affecting the child's functional abilities. This method comprises two versions: one tailored for younger children (aged 26–35 months) and one for older children (aged 36–40 months). Both versions are identical, differing solely in the percentile bands for norms, borderline cases, and risks. In the screening record, the pediatrician classifies children into three categories based on their developmental score percentiles. Developmental functioning sum score consists of 20 items mentioned below. A score above the 26th percentile falls within the normative range, indicating no significant deviations from the population norm. Scores between the 11th and 25th percentiles are classified as borderline, suggesting mild deviations requiring increased monitoring, while scores below the 10th percentile indicate elevated risk, warranting further evaluation and possible specialized care. According to standard pediatric procedures (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021), younger children scoring 17 to 20 points fall within the normative range, those scoring 14 to 16 points are classified in the borderline range, and those scoring 13 points or lower fall into the risk category. For older children, scores of 18 to 20 points indicate the normative range, scores of 16 to 17 points correspond to the borderline range, and scores of 15 points or lower place a child in the risk category.

The developmental functioning contains 20 items with dichotomous "*yes*"-"*not yet*" responses, focusing on the areas of motor skills (4 items, example of item: "*M3 Can draw a circle.*"); social behavior [3 items, example of item: *"SS2 Shows interest in other children (watches what they are doing, approaches them, joins in play)."*]; cognition [3 items, example of item: *"K2 Can assemble a picture made of at least 3 pieces (puzzle)."*]; speech understanding (3 items, example of item: *"PR3 Can answer the question 'How old are you?' or show the number using fingers."*); speech production (3 items, example of item: *"H1 Speaks in sentences of three or more words."*); self-care (3 items, example item: *"SE2 Can eat independently using a spoon."*), and kindergarten readiness (1 item: *"Z1 Do you think your child is mature and ready for kindergarten?"*). The maximum possible score for developmental functioning is 20 points. The items in the pediatric response sheet denote various aspects of developmental functioning: M for motor skills, SS for social behavior, K for cognition, PR for speech understanding, H for speaking, SE for self-care, and Z for kindergarten readiness. Within each category, items are labeled with abbreviations such as M1, M2, M3, and M4 for motor skills. However, these numerical labels do not indicate the difficulty level of the items; they are simply identifiers, with the numbers not reflecting an increasing level of difficulty.

2.3 Data analysis method

To verify the scalability of developmental functionality from a deterministic conception of the Guttman scale (Aim 1), the achieved values of the coefficient of reproducibility (coefficient of reproducibility = CR) were investigated. The coefficient takes values from 0 to 1, with values higher than CR = 0.9 indicating acceptable unidimensional use of the scale. Examination of the scalability of the developmental functioning items by analyzing Guttman response patterns was conducted in R, specifically in the guttman package, version 0.1.0 (Libura, 2024).

The premise of the validated Rasch model (Aim 2) was the latent continuous capacity for child developmental functioning. After logarithmizing the data, we determined the difficulty of items in the population of children using an item difficulty index (denoted as b). The fit of the data to the model was assessed with a discriminant value fixed at 1 ($\alpha = 1$) with the variance freely estimated. The model was estimated using the maximum likelihood method. Parameters such as χ^2, *p*-value, RMSEA, SRMR, TLI, and CFI were examined to evaluate the fit of the data to the model. Interpretation of the Rasch model is similar to confirmatory factor analysis, where, as noted by Markland (2007), the significance of χ^2 should not be the sole criterion for determining data fit due to conservatism, particularly in large research sets. Therefore, we also utilized the comparative fit index (CFI) and the Tucker-Lewis index (TLI), which are expected to be 0.9 or higher, as well as the standardized root mean square residual (SRMR) and the root mean square error of approximation (RMSEA), which should be less than 0.08 (Bond and Fox, 2013). An item fit index was computed for the items, providing a direct comparison of modeled and observed frequencies of correct and incorrect responses (Orlando and Thissen, 2000). Standardized and unstandardized item infit and outfit indices were also added, with too low values indicating overfitting typical of Guttman response patterns, indicating possible local dependence of the items. High values of fit, in turn, indicate unpredictability of responses (underfit), which may be due to item weakness or guessing on the part of the respondent. Standardized infit and outfit values in the range of −2 to +2 are considered ideal (Bond and Fox, 2013). Empirical reliability values were also investigated. The item characteristic functions, item information functions, and screening information function with measurement error were graphed. R software (R Core Team, 2024) and the packages readr (Wickham et al., 2024), tidyverse (Wickham et al., 2019), mirt (Chalmers, 2012), ggplot2 (Wickham, 2016), psych (Revelle, 2024), cowplot (Wilke, 2024) were used.

3 Results

In this section, we aim to achieve two objectives: Aim 1 focuses on assessing the scalability of developmental milestones in children aged 2 and 3 years old using Guttman scaling. Aim 2 involves verifying the S-PMV11 instrument using the Rasch model, examining both the developmental milestones and the overall model of developmental functioning in younger and older children. Through these aims, we seek to deepen our understanding of developmental milestones and functioning in young children.

3.1 Sample characteristics

The mean age for the younger children was $M = 30.51$ months ($SD = 2.78$; Mdn = 31; $min = 26$, $max = 35$), and for the older children, it was $M = 36.8$ months ($SD = 1.17$; $Mdn = 36$; $min = 36$, $max = 44$). The basic information of the sample, such as gender and information about the person who completed screening are displayed in Table 1.

3.2 Internal consistency

In our research, good values of internal consistency of developmental functionality were found for younger children $\alpha = 0.84$ and $\omega = 0.87$ and for older children $\alpha = 0.86$ and $\omega = 0.88$.

3.3 Scalability of developmental milestones using Guttman scaling

After calculating the scalability of developmental functioning in younger children ($n = 1,336$), the unidimensionality of the S-PMV11 instrument can be highlighted. In the S-PMV11, item abbreviations reflect aspects of functioning that are measured, e.g., K1 means the focus of the item on cognition. The items in the text denote various content of developmental functioning: M for motor skills, SS for social behavior, K for cognition, PR for speech understanding, H for speaking, SE for self-care, and Z for Kindergarten readiness. Items were ranked in order of difficulty from easiest to most difficult as follows: M4, SS1, PR1, SS2, K1, SE2, M1, M2, SS3, SE1, K3, PR2, PR3, K2, M3, SE3, Z1, H2, H1, H3 (for the names of the items, see Table 2). For a total of 26,720 data (1,336 respondents x 20 items), a reproducibility coefficient index of CR = 0.90 was calculated, which corresponds to the established Gutmman scaling criterion reported as 0.90.

For older children, the items were ranked in a different order of difficulty, with item Z1 being the most challenging. The order from easiest to most difficult item was as follows: SS1, SS2, M4, PR1, SE2, M1, SS3, K1, M2, PR3, K3, PR2, SE1, M3, K2, SE3, H1, H2, H3, Z1 (for the names of the items, see Table 2). For 6,080 data (304 respondents x 20 items), the index of coefficient of reproducibility was at CR = 0.90, thus the S-PMV11 instrument for 3-year-old children meets the requirements of a unidimensional scale.

TABLE 1 Sample characteristics.

Sample characteristics	Younger children ($n = 1,336$)		Older children ($n = 304$)	
	n	%	n	%
Gender				
Girls	622	46.56	139	45.72
Boys	714	53.44	165	54.28
Screening completed by				
Mother	1,101	82.41	248	81.58
Father	46	3.44	10	3.29
Both parents	181	13.55	42	13.82
Another person	8	0.60	4	1.32

Another person = Another person refers to someone other than the parent who takes care of the child, such as a grandparent or another caregiver.

3.4 Rasch model of developmental functioning

In the overall developmental functioning scores, younger children achieved mean values of $M = 17.61$, $SD = 3.14$, $Mdn = 19$, $skew = -1.78$, $kurt = 3.41$, $min = 2$, $max = 20$. Up to 529 younger children out of 1,336 achieved full scores, which corresponds to approximately 39.6%. For developmental functioning, older children scored $M = 17.74$, $SD = 3.15$, $Mdn = 19$, $skew = -1.93$, $kurt = 3.87$, $min = 4$, $max = 20$. For older children, 128 older children out of 304 achieved a full score, accounting for 42.11% of the responses. Together with the skewness and kurtosis indices, the descriptive statistics highlight the ceiling effect of the S-PMV11 instrument. According to the normative ranges defined in standard procedures for primary pediatrics, 74.0% ($n = 984$) of the younger children in the research sample fell within the normative range, 15.7% ($n = 209$) were in the borderline range, and 10.3% ($n = 137$) were in the risk range. Among the older children, 68.6% ($n = 208$) were classified within the normative range, 13.5% ($n = 41$) in the borderline range, and 17.8% ($n = 54$) in the risk range.

3.4.1 Results of the Rasch model for younger children

The fit of the younger children's data (< 35 months) to the Rasch model is acceptable and shows indices: $\chi^2 = 1,309$ (189, n = 1,336), $p < 0.001$; RMSEA = 0.067 [CI 95% = 0.063–0.070]; SRMSR = 0.098; TLI = 0.924; CFI = 0.924. The significance of chi-square statistics may be biased by the sample size, and therefore, when describing the results, we mainly focus on the RMSEA and TLI indices, which showed a satisfactory fit of the model to the data. The empirical reliability was at 0.730 and we consider it acceptable. The results for the items are shown in Table 2. The items showing the greatest variability in the SX^2 indicator were the items focusing on drawing a circle (M3), walking up stairs (M2), and talking (H1, H2, and H3). While the item on joining the pieces of a building block (M4) is minimally challenging, the items focusing on speaking are among the most challenging indicators. Based on the parameter b, as well as looking at the item characteristic functions, item information functions, and test information function shown in Figure 1, it can be concluded that the items measure too low an estimated developmental functional ability. In the S-PMV11, the theta indicator provides the most information about ability at the −3 level, which corresponds to approximately three standard deviations from the norm.

3.4.2 Results of the Rasch model for older children

The data fit of older children (≥ 36 months) with the Rasch model is acceptable, with the following indices: $\chi^2 = 464.225$ (df = 189, $n = 304$), $p < 0.001$; RMSEA = 0.069 [CI 95% = 0.061–0.077]; SRMSR = 0.112; TLI = 0.924; CFI = 0.925. The RMSEA and TLI indices demonstrated satisfactory model fit. The empirical reliability was at the level of 0.721, which we consider acceptable. Results for individual items with abbreviated content are displayed in Table 3, and the visualization of curves is provided in Figure 2. Concerning the capture of age differences in the construct of developmental functionality, highlighting the SX^2 indicators as well as to the infit and outfit indicators for speech items, which are less variable among older children compared to younger ones. While these captured differences provide evidence of

TABLE 2 Item scores for the Rasch model in younger children.

Item	b	SE(b)	SX^2	RMSEA SX^2	df	*p*	Infit	Infit t_u	Outfit	Outfit t_v
M4: connects the parts of the blocks	−6.318	0.312	10.659	0.035	4	0.031	0.583	−2.184	0.193	−3.845
SS1: shares joy	−5.933	0.268	15.727	0.024	9	0.073	0.636	−2.109	0.190	−3.626
PR1: fetches the ball when asked	−5.540	0.232	29.743	0.038	10	0.001	0.611	−2.642	0.083	−4.340
SS2: is interested in other children	−4.965	0.191	42.573	0.041	13	0.000	0.937	−0.444	1.384	1.279
K1: inserts shapes correctly	−4.817	0.183	15.450	0.012	13	0.280	0.927	−0.567	0.663	−1.337
SE2: eats by spoon	−4.683	0.175	17.778	0.017	13	0.166	0.807	−1.742	0.395	−3.091
M1: jumps forward with both feet	−4.046	0.147	14.945	0.014	12	0.244	0.898	−1.158	0.619	−2.386
M2: walks up the stairs	−3.917	0.143	52.172	0.050	12	0.000	1.042	0.518	1.051	0.355
SS3: plays with dolls	−3.871	0.141	18.821	0.021	12	0.093	0.822	−2.268	0.726	−1.779
SE1: dresses up	−3.503	0.130	44.251	0.045	12	0.000	0.928	−1.015	0.938	−0.417
K3: assigns colors	−2.969	0.117	20.298	0.023	12	0.062	0.929	−1.238	0.768	−2.443
PR2: completes the storyline	−2.841	0.115	26.816	0.030	12	0.008	0.807	−3.742	0.662	−4.014
PR3: answers about his/her age	−2.630	0.111	7.765	0.000	12	0.803	0.814	−3.885	0.610	−5.351
K2: composes a picture	−2.480	0.108	28.490	0.035	11	0.003	0.905	−2.005	0.749	−3.502
M3: draws a circle	−2.368	0.107	62.608	0.059	11	0.000	0.981	−0.404	0.860	−1.972
SE3: communicates a need in a timely manner	−2.136	0.103	9.952	0.000	11	0.535	0.801	−4.931	0.664	−5.886
Z1: is ready for Kindergarten	−2.109	0.103	10.381	0.000	11	0.496	0.758	−6.145	0.612	−7.064
H2: uses prepositions	−1.717	0.098	53.949	0.057	10	0.000	0.607	−12.071	0.488	−12.074
H1: says 3 or more words	−1.687	0.098	61.151	0.066	9	0.000	0.595	−12.611	0.476	−12.599
H3: uses the past tense	−1.576	0.097	48.102	0.057	9	0.000	0.628	−11.795	0.544	−11.023

b, difficulty; SE(b), standard error of estimate for difficulty; SX^2, item fit index (Orlando and Thissen, 2000); RMSEA, root mean square error of approximation for SX^2; df, degrees of freedom; *p*, significance of item fit (<0.05); infit/outfit, item fit; infit tu, standardized infit; outfit tv, standardized outfit; content of items: M, motor; SS, social behavior; K, cognition; PR, speech understanding; H, speaking; SE, selfcare; Z, Kindergarten readiness. The items from the S-PMV11 are presented in an abbreviated version.

the tool's construct validity, we highlight the test information function, indicating that in older children, screening most accurately identifies children at a level of approximately −3.5 latent ability.

4 Discussion

The aim of this study was to investigate the scalability and unidimensionality of developmental milestones in 2 and 3 year old children by using Guttman scaling and also to verify the S-PMV11 instrument using the Rasch model, both for developmental milestones and for the whole model of developmental functioning. Such approaches offer a more relevant interpretation of achievement scores compared to classical test theory (Tesio et al., 2024). Although it is rather rare to obtain data that fit a Guttman formula (Abdi and Williams, 2010), in the case of the S-PMV11 instrument, it has been shown that developmental milestones can be scaled in terms of the Guttman approach, for both age groups of children. The data for both

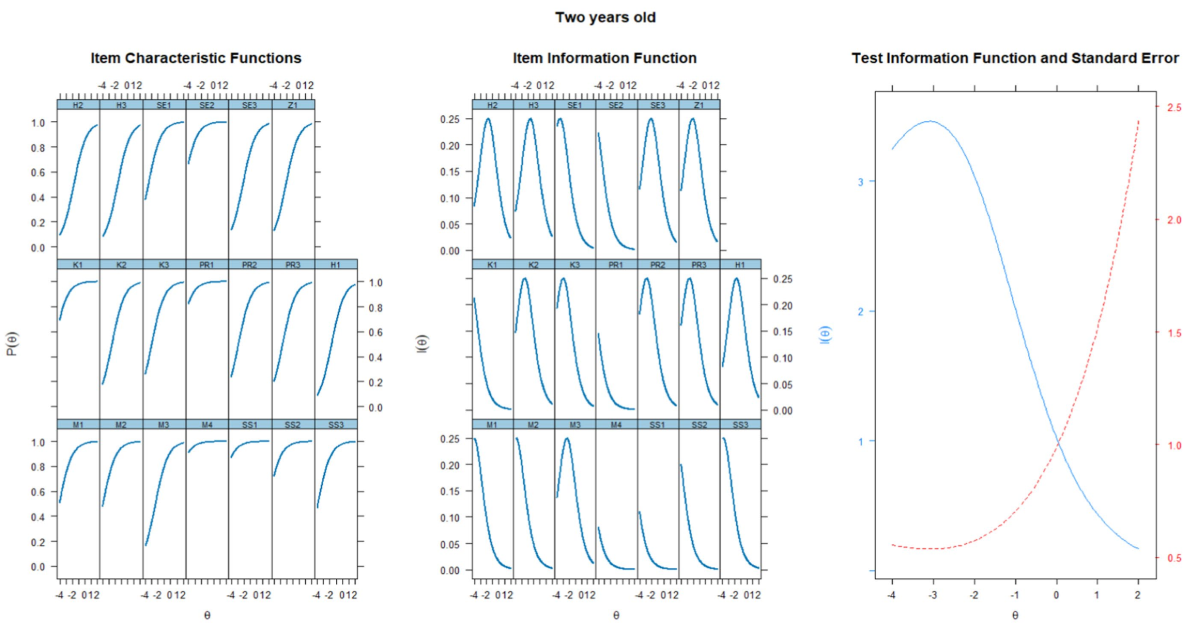

FIGURE 1
Characteristic curves of items, information functions of items, and test information function for a group of two-year-old children.

age groups were also shown to fit the Rasch model, however, the test information function achieved at approximately −3 latent ability is questionable. Our findings need to be put in the context of the practical application of screening at the population level in order to search for potential developmental difficulties. The practical purpose of the tool is to capture children at approximately −1 latent ability level of developmental functioning as stated in the tool manual (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021). Despite the identified unidimensionality and support for construct validity of S-PMV11, our findings yielded significant information about the very low level of measured ability at the −3 deviation level. We question the clinical application of screening in capturing children at possible borderline risk at around −1 ability level of developmental functioning. Parents of children should be educated early on about options for supporting developmental functioning, as late identification of difficulties does not allow professionals to develop further developmentally supportive interventions that are most effective the earlier they are initiated (Guthrie et al., 2023; Lipkin et al., 2020). We are concerned that the tool is not sensitive enough to capture children with borderline levels of developmental functional ability and therefore children may be systematically overlooked despite the need for additional population-wide care. Interpretation of our findings is therefore a syncretization of the psychometric as well as the practical evaluation of the instrument.

It is possible to find certain common features in psychomotor development that are particularly useful for clinical practice, as they lead to the rapid recognition of possible developmental difficulties. In their own research on Guttman errors in response patterns, the authors Blanchin et al. (2016) outlined the use of these errors for detecting subgroups of respondents and subgroups of items; in our interpretation, we will focus on the items measuring speaking aspect, which also proved to be the most challenging in Rasch's model. Interpretation through Guttman scaling combined with the Rasch model allows us to gain theoretical as well as practical qualitative insight into the items. The interpretation of the Rasch model needs to take into account the clinical interpretation (Tesio et al., 2024) of whether the fit indices for the items match the theoretical and practical applications. Based on our findings, it can be concluded that the data in both younger and older children demonstrated fit parameters for the Rasch model that are consistent with the interpretation of a unidimensional construct with increasing developmental difficulty when it comes to speaking expressions.

The items measuring speaking aspect in our case were the most homogeneous in content, therefore they were the most challenging items with the highest error rates and can register children who, although meeting the easier milestones, do not have sufficiently developed speaking skills. Thus, the parent perceives the child as ready for kindergarten despite not completing the speech items, as they may believe that this is a natural variability in speech (Holm et al., 2023). This interpretation is also consistent with the Kindergarten readiness item, which is most difficult only in the group of older children, whereas in younger children it is located before the items measuring speaking aspect. Thus, thanks to the Guttman formula, as well as to the infit and outfit values of the items in the Rasch model, it is possible to identify a heterogeneous group of children called "late talkers," who show less frequent verbal imitation and reduced lexical production (Zuccarini et al., 2023). The research on phonological skills requires further investigation due to the interindividual variability of this population, the significant prevalence of which has been pointed out by other authors (Zuccarini et al., 2023; Suttora et al., 2021; Desmarais et al., 2008). At the same time, talking can be understood as the result of the integration of social interaction, child talk, or cognition, but also as a manifestation of symbolic thinking, with the onset of egocentrism as indicators of this period (Heo et al., 2011). Despite being the most challenging milestones among the other items, their difficulty corresponded to the low level of developmental functioning

TABLE 3 Item scores for the Rasch model in older children.

Item	b	SE(b)	SX^2	RMSEA SX^2	df	*p*	Infit	Infit t_u	Outfit	Outfit t_v
SS1: shares joy	−5.703	0.507	NaN	NaN	0	NaN	0.801	−0.474	0.389	−0.746
SS2: is interested in other children	−5.312	0.443	0.743	0.000	1	0.389	0.908	−0.214	0.469	−0.749
M4: connects the parts of the blocks	−5.152	0.421	3.622	0.052	2	0.163	0.763	−0.824	0.186	−1.830
PR1: fetches the ball when asked	−5.152	0.421	5.021	0.071	2	0.081	0.666	−1.257	0.161	−1.957
SE2: eats by spoon	−4.879	0.387	2.183	0.017	2	0.336	0.772	−0.905	0.382	−1.314
M1: jumps forward with both feet	−4.360	0.335	10.376	0.049	6	0.110	1.051	0.314	0.829	−0.293
SS3: plays with dolls	−4.115	0.315	5.923	0.000	8	0.656	0.854	−0.760	0.465	−1.699
K1: inserts shapes correctly	−4.115	0.315	9.496	0.025	8	0.302	0.747	−1.430	0.380	−2.105
M2: walks up the stairs	−3.772	0.291	11.830	0.040	8	0.159	1.066	0.448	0.865	−0.353
PR3: answers about his/her age	−3.332	0.266	6.566	0.000	8	0.584	0.815	−1.380	0.562	−2.061
K3: assigns colors	−3.236	0.261	5.582	0.000	8	0.694	0.760	−1.913	0.564	−2.163
PR2: completes the storyline	−3.236	0.261	12.301	0.042	8	0.138	1.001	0.049	0.696	−1.393
SE1: dresses up	−3.236	0.261	6.322	0.000	8	0.611	0.964	−0.226	0.737	−1.171
M3: draws a circle	−2.851	0.244	9.841	0.028	8	0.276	0.829	−1.520	0.711	−1.656
K2: composes a picture	−2.325	0.225	3.557	0.000	7	0.829	0.778	−2.466	0.590	−3.412
SE3: communicates a need in a timely manner	−2.291	0.224	14.447	0.059	7	0.044	0.853	−1.589	0.799	−1.524
H1: says 3 or more words	−1.963	0.215	5.256	0.000	7	0.629	0.665	−4.472	0.526	−4.984
H2: uses prepositions	−1.906	0.214	12.685	0.052	7	0.080	0.597	−5.662	0.449	−6.236
H3: uses the past tense	−1.823	0.212	15.024	0.062	7	0.036	0.641	−5.085	0.510	−5.582
Z1: is ready for Kindergarten	−1.796	0.211	5.597	0.000	7	0.588	0.772	−3.074	0.673	−3.463

b, difficulty; SE(b), standard error of estimate for difficulty; SX^2, item fit index (Orlando and Thissen, 2000); RMSEA, root mean square error of approximation for SX^2; df, degrees of freedom; *p*, significance of item fit (<0.05); infit/outfit, item fit; infit tu, standardized infit; outfit tv, standardized outfit; content of items: M, motor; SS, social behavior; K, cognition; PR, speech understanding; H, speaking; SE, selfcare; Z, Kindergarten readiness. The items from the S-PMV11 are presented in an abbreviated version.

ability in both groups of children. Therefore, non-performance of these items in a population of children should be a significant indicator of possible developmental difficulty.

For two-year-old children, the easiest milestone in the Guttman scaling M4 (Builds blocks, can connect the pieces of the building block) was found to be item SS1 (Shares joy) and SS2, which saturate the content of engaging in play with other children (Köngäs et al., 2022) and are the basis for the development of other indicators of children's functionality. Failure to meet these developmental milestones may have clinically significant interpretive value because they are milestones that are manifest in socialization, in the formation of self-esteem, or in the child's globally perceived well-being. The child's social behavior and engagement in play with other children can be termed social visual engagement (SVE), which is manifested by learning to interact socially with others while observing other people, playing with others, and listening to the conversations of others. If a child's establishment of these fundamental expressions is absent between the ages of 24 and 36 months, they may even

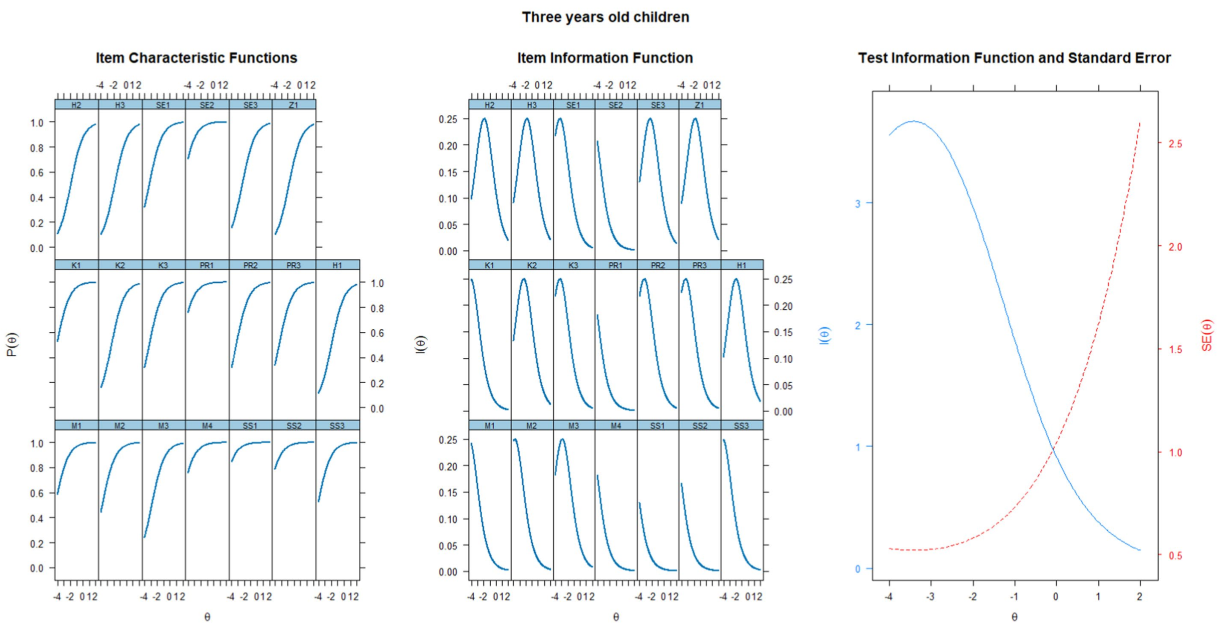

FIGURE 2
Characteristic curves of items, information functions of items, and test information function for a group of three-year-old children.

be significant clinical indicators of potential autism spectrum disorders (Klin, 2023).

From a clinical perspective, the distribution of data across a large sample shows that the developmental function bands for both younger and older children align closely with those defined in standard primary pediatric guidelines. For children within the normative range (negative results), it is important to communicate item completion and offer guidance to promote motor, communication, and play skills. Phrases such as "no significant differences from same-age peers" or "the score is similar to other children of the same age" can indicate low risk. For children in the borderline or risk range (positive results), the focus should shift to addressing unmet items and supporting family competencies. Communication may include statements like "slight differences from peers" or "the score is lower than typical, but we will verify this in a follow-up." Follow-up evaluations or specialist referrals may be necessary. As a screening tool, it is important to note that while many items may appear easy, failure to complete these easier items, even in children classified within the normative range, can serve as a significant indicator that the child is not developing typically. Beyond scores, the difficulty of the items a child fails to complete is critical for accurate assessment. Misclassification can occur if a child fails to complete easier items, which may indicate a risk of uneven development. The pediatric response sheet currently groups items by broad developmental areas, such as motor skills or social behaviors, which does not account for the varying levels of difficulty among them. It would be more beneficial to organize the items based on their difficulty level. This approach would allow clinicians to more effectively identify patterns in the child's responses and recognize where they struggle the most. For example, a pediatrician would be able to see more clearly if a child is unable to perform the three most challenging tasks, as opposed to simply noting that the child struggles with some tasks without understanding the level of difficulty involved. Such a system would provide more insight into the child's developmental needs and help guide decisions on further care or interventions. This would allow clinicians to more easily identify patterns in the child's responses based on the complexity of the items, making it easier to determine if additional care or interventions are needed. Our research demonstrates that organizing the items by difficulty allows for a more nuanced and accurate assessment. The Rasch approach in our study confirms that the instrument captures significant developmental concerns at the population level. Therefore, we recommend that the tool be used with expert caution when items are unmet. To aid pediatricians, we suggest ranking items by difficulty in the scoring sheet, which would streamline result interpretation and improve alignment with population trends. Although all items in the screening tool are important as potential "red flags," our focus is not solely on categorizing children as "normal" or "at risk." Rather, we aim to emphasize the importance of understanding certain items that may be naturally easier for most children. When a child fails to complete these tasks, it can serve as an important indicator, even if the child's overall score falls within the normal range. This is one of the key implications of our study — some items may be considered developmentally simple for the general population. We believe that it is crucial for pediatricians, particularly in the fast-paced clinical setting, to have the ability to effectively identify potential issues. This approach would allow for more nuanced decision-making, rather than automatically assuming that a child is within the normative range based on the total score. Individual analysis of a parent's response pattern can help identify critical developmental milestones where interventions may be needed. Thanks to Rasch's approach, we have shown that the instrument captures at the population level those children who already have significant developmental problems. However, it is equally important to identify children who may benefit from more thorough monitoring and a more comprehensive view of their overall development. This would also support raising parental awareness about the importance of completing all items listed in the tool, ensuring that no potential developmental concerns are overlooked.

4.1 Limits and recommendations for further research

We consider the most serious limitation of this study to be the lack of demographic and other descriptive information of children and caregivers, which may be important intervening covariates. These include, for example, the child's previous health problems or previously diagnosed illnesses. On the caregiver's side, information such as caregiver age and educational attainment was absent, which may be related to inadequate assessment of the child's functioning in the S-PMV11. This limit is caused by the effect of the data collection managed at the national level with no possibility to change the dataset. Missing information is available to the pediatrician in direct contact with the caregiver and through the child's medical record. Therefore, our interpreted findings are also set in the context of population-wide use of the instrument, with the conclusion to point out the possibility of systematically overlooked children with possible developmental difficulties at a level not so obvious as −3 latent ability. We believe that providing more detailed information about the parent and child in the Database could reveal additional associations with respect to other covariates such as parent education or child health status. There is an urgent need in creating and validating tools that are used to capture children with potential developmental difficulties early in life. Other screening methods, such as the Ages and Stages Questionnaire (Squires and Bricker, 2009), use three response options based on whether the child is performing the activity regularly, sometimes, or not yet. The instructions indicate that parents may choose not to answer if they are unsure about the response. While it would be possible to verify this procedure in the context of the screening, the current version of the screening tool follows the guidelines outlined in the manual (©S-PMV, Prof. K. Matulay Endowment Fund (EF), 2016–2021), which advises parents to select "Not yet" if they are uncertain, and for the pediatrician to clarify the item during the preventive check-up. If the parent still cannot assess whether the behavior is present, it is recommended that the pediatrician treat the response as "not yet." Even a parent's inability to evaluate a particular behavior—whether due to lack of observation or other reasons—can provide valuable insight into the child's psychosocial developmental context. Another potential limitation and recommendation for future improvement is to further explore the process of handling uncertain responses in the screening, ensuring that the tool remains effective and accurate in real-world practice. The consequence of this study lies primarily in drawing attention to the unidimensionality of the tool and highlighting the possibility of arranging items in the pediatric response sheet according to their difficulty level, rather than by subdomains such as motor skills, social behavior, language, etc., since these areas are interrelated and form a common construct. Additionally, we would consider incorporating more challenging items, especially for three-year-old children.

Data availability statement

The datasets presented in this article are not readily available because the data originate from the National Database and are not publicly available, but they can be shown upon request. Requests to access the datasets should be directed to lraczova@ukf.sk.

Ethics statement

The studies involving humans were approved by Ethics Committee of the National Institute of Children's Diseases, Limbová 1, 83,340 Bratislava, Slovakia under the number EK2/12/2024. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

LR: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. TU: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. EJ: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. MP: Conceptualization, Data curation, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. TS: Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The study was conducted with support from project VEGA 1/0534/22: Psychometric properties and validity of the screening of psychomotor development (S-PMV) as a monitoring method for early detection of developmental difficulties in children; and UGA IX/5/2024: Verification of Selected Psychometric Properties of Developmental Screenings from the S-PMV Set for Early Age.

Acknowledgments

We express our gratitude to the FOND K. Matulay n.f., represented by PaedDr. Oľga Matušková.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abdi, H., and Williams, L. J. (2010). Principal component analysis. *Wiley Interdis. Rev: Computational Statistics.* 2, 433–459. doi: 10.1002/wics.101

Blanchin, M., Sébille, V., Guilleux, A., and Hardouin, J. B. (2016). The Guttman errors as a tool for response shift detection at subgroup and item levels. *Qual. Life Res.* 25, 1385–1393. doi: 10.1007/s11136-016-1268-8

Bond, T. G., and Fox, C. M. (2007). Applying the Rasch model: Fundamental measurement in the human sciences. *2nd* Edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Bond, T. G., and Fox, C. M. (2013). Applying the Rasch model: Fundamental measurement in the human sciences. New York, NY: Psychology Press.

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., et al. (2011). Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics* 127, 1034–1042. doi: 10.1542/peds.2010-2989

Chalmers, R. P. (2012). Mirt: a multidimensional item response theory package for the R environment. *J. Stat. Softw.* 48, 1–29. doi: 10.18637/jss.v048.i06

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L. E., and Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *J. Autism Dev. Disord.* 37, 62–72. doi: 10.1007/s10803-006-0330-8

Cibralic, S., Hawker, P., Khan, F., Lucien, A., Diaz, A. M., Woolfenden, S., et al. (2023). Developmental screening tools for identification of children with developmental difficulties in high-income countries: a systematic review. *Front. Child Adolescent Psychiatry* 2, 1–16. doi: 10.3389/frcha.2023.1074004

Coelho, V., and Pinto, A. I. (2018). The relationship between Children's developmental functioning and participation in social activities in Portuguese inclusive preschool settings. *Front. Educ.* 3, 1–13. doi: 10.3389/feduc.2018.00016

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., and Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *Int. J. Lang. Commun. Disorders* 43, 361–389. doi: 10.1080/13682820701546854

Embretson, S. E., and Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum.

Engelhard, G. (2008). Historical perspectives on invariant measurement: Guttman, Rasch, and Mokken. *Measurement: Interdis. Res. Perspectives* 6, 155–189. doi: 10.1080/15366360802197792

Francés, L., Quintero, J., Fernández, A., Ruiz, A., Caules, J., Fillon, G., et al. (2022). Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the DSM-5: a systematic review in accordance with the PRISMA criteria. *Child Adolesc. Psychiatry Ment. Health* 16:27. doi: 10.1186/s13034-022-00462-1

Francés, L., Ruiz, A., Soler, C. V., Francés, J., Caules, J., Hervás, A., et al. (2023). Prevalence, comorbidities, and profiles of neurodevelopmental disorders according to the DSM-5-TR in children aged 6 years old in a European region. *Front. Psychol.* 14:1260747. doi: 10.3389/fpsyt.2023.1260747

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Ment. Retard. Dev. Disabil. Res. Rev.* 11, 173–179. doi: 10.1002/mrdd.20068

Guthrie, W., Wetherby, A. M., Woods, J., Schatschneider, C., Holland, R. D., Morgan, L., et al. (2023). The earlier the better: an RCT of treatment timing effects for toddlers on the autism spectrum. *Autism* 27, 136236132311591–136236132312309. doi: 10.1177/13623613231159153

Heo, J., Han, S., Koch, C., and Aydin, H. (2011). Piaget's egocentrism and language learning: language egocentrism (LE) and language differentiation (LD). *J. Lang. Teach. Res.* 2:4. doi: 10.4304/jltr.2.4.733-739

Holm, A., van Reyk, O., Crosbie, S., de Bono, S., Morgan, A., and Dodd, B. (2023). Preschool children's consistency of word production. *Clinical Linguistics and Phonetics* 37, 223–241. doi: 10.1080/02699206.2022.2041099

Institute of Medicine (2001). Neurological, psychiatric, and developmental disorders: Meeting the challenge in the developing world. Washington, DC: The National Academies Press.

Jelínek, M., Květoň, P., and Vobořil, D. (2007). Testování v psychologii. Teorie odpovědi na položku a počítačové adaptivní testování. Brno: Masaryk Univerzity.

Klin, A. (2023). Translating advances in developmental social neuroscience into greater access to early diagnosis in autism spectrum disorder. *Medicina* 83, 32–36

Köngäs, M., Määttä, K., and Uusiautti, S. (2022). Participation in play activities in the children's peer culture. *Early Child Dev. Care* 192, 1533–1546. doi: 10.1080/03004430.2021.1912743

Libura, M. (2024). _guttman: Analysis of reliability of Guttman scales_. R package version 0.1.0.

Lipkin, P. H., Macias, M. M., Norwood, K. W., Brei, T. J., Davidson, L. F., Davis, B. E., et al. (2020). Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics* 145:1. doi: 10.1542/peds.2019-3449

Markland, D. (2007). The golden rule is that there are no golden rules: a commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personal. Individ. Differ.* 42, 851–858. doi: 10.1016/j.paid.2006.09.023

Ministry of Health. (2021). Štandard vyšetrení psychomotorického vývinu detí pri 2.-11. preventívnej prehliadke v primárnej starostlivosti - 1. revízia. Available online at: https://www.standardnepostupy.sk/standardy-primarna-pediatria/ (Accessed March 15, 2023).

Misirliyan, S. S., Boehning, A. P., and Shah, M. (2023). "Development milestones" in StatPearls (Treasure Island (FL): StatPearls Publishing).

Oberklaid, F., and Efron, D. (2005). Developmental delay: identification and management. *Aust. Fam. Physician* 34:9.

Olusanya, B. O., Smythe, T., Ogbo, F. A., Nair, M. K. C., Scher, M., and Davis, A. C. (2023). Global prevalence of developmental disabilities in children and adolescents: a systematic umbrella review. *Front. Public Health* 11:22009. doi: 10.3389/fpubh.2023.1122009

Orlando, M., and Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Appl. Psychol. Meas.* 24, 50–64. doi: 10.1177/01466216000241003

Palfrey, J. S., Singer, J. D., Walker, D. K., and Butler, J. A. (1987). Early identification of children's special needs: a study in five metropolitan communities. *J. Pediatr.* 111, 651–659.

Prieto, J. A., Cueto, S., Carballo-Fazanes, A., and Abelairas-Gómez, C. (2022). Psychomotor development disorders in apparently healthy children and considerations of family evaluation. *J. Human Sport Exercise* 17, 518–527. doi: 10.14198/jhse.2022.173.04

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Revelle, W. (2024). _psych: Procedures for psychological, psychometric and personality Research_. Evanston, Illinois: Northwestern University.

©S-PMV, Prof. K. Matulay Endowment Fund (EF). (2016–2021). S-PMV 2–11. Available online at: https://www.zdraviedietata.sk/vyvinove-dotazniky (Accessed January 5, 2022).

Squires, J., and Bricker, D. (2009) in Ages & stages questionnaires®, third edition (ASQ®-3): A parent-completed child monitoring system. ed. H. Paul (Baltimore: Brookes Publishing Co., Inc).

Suttora, C., Zuccarini, M., Aceti, A., Corvaglia, L., Guarini, A., and Sansavini, A. (2021). The effects of a parent-implemented language intervention on late-talkers' expressive skills: the mediational role of parental speech contingency and dialogic Reading abilities. *Front. Psychol.* 12, 1–14. doi: 10.3389/fpsyg.2021.723366

Tesio, L., Caronni, A., Kumbhare, D., and Scarano, S. (2024). Interpreting results from Rasch analysis 1. The "most likely" measures coming from the model. *Disabil. Rehabil.* 46, 591–603. doi: 10.1080/09638288.2023.2169771

van Buuren, S., and Eekhout, I. (2023). Child development with the D-score: turning milestones into measurement. *Child Develop. D-score.* 5, 1–86. doi: 10.12688/gatesopenres.13222.2

Vitrikas, K., Savard, D., and Bucaj, M. (2017). Developmental delay: when and how to screen. *Am. Fam. Physician* 96:1.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York:Hadley Wickham Springer-Verlag.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Software* 4:1686. doi: 10.21105/joss.01686

Wickham, H., Hester, J., and Bryan, J. (2024). _readr: read rectangular text Data_. Rpackage version 2.1.5. Available online at: https://CRAN.R-project.org/package=readr (Accessed April 6, 2023).

Wilke, C. (2024). _cowplot: streamlined plot theme and plot annotations for 'ggplot2'_. R package version 1.1.3. Available online at: https://CRAN.R-project.org/package=cowplot (Accessed January 8, 2025).

World Health Organization (2007). International classification of functioning, disability, and health: Children & Youth Version: ICF-CY. Geneva: World Health Organization.

Yunilda, E., Gunardi, H., Medise, B. E., and Oswari, H. (2023). The Indonesian version of ages and stages questionnaire III accuracy compared to Bayley scales of infant development III. *Infant Child Dev.* 32:1. doi: 10.1002/icd.2387

Zuccarini, M., Guarini, A., Gibertoni, D., Suttora, C., Aceti, A., Corvaglia, L., et al. (2023). Describing communication profiles of low-risk preterm and full-term late talkers. *J. Commun. Disord.* 104:106336. doi: 10.1016/j.jcomdis.2023.106336

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 21 May 2025
DOI 10.3389/fpsyg.2025.1497827

Check for updates

OPEN ACCESS

EDITED BY
Victor Zaia,
Faculdade de Medicina do ABC, Brazil

REVIEWED BY
Maria Rita Sergi,
University of G.'d'Annunzio, Italy
Hugo W. F. Mak,
The University of Hong Kong, Hong Kong SAR, China
Marta Brás,
University of Algarve, Portugal

*CORRESPONDENCE
Jie Zhang
✉ ZHANGJ@BuffaloState.edu
Zhenyu Ma
✉ ma_zhenyu@gxmu.edu.cn
Guoxiang Chen
✉ chenguoxiang@gxmu.edu.cn

†These authors share first authorship

RECEIVED 17 September 2024
ACCEPTED 21 March 2025
PUBLISHED 21 May 2025

CITATION
Huang Q, Mo Q, He H, Bai X, Zhang J, Ma Z and Chen G (2025) Reliability and validity of the short quality of life scale among bank employees in Guangxi, China. *Front. Psychol.* 16:1497827. doi: 10.3389/fpsyg.2025.1497827

Reliability and validity of the short quality of life scale among bank employees in Guangxi, China

Qianwei Huang[1†], Qiqing Mo[2†], Huiming He[3], Xinyu Bai[4], Jie Zhang[5,6]*, Zhenyu Ma[1]* and Guoxiang Chen[7]*

[1]School of Public Health, Guangxi Medical University, Nanning, Guangxi, China, [2]Guilin People's Hospital, Guilin, Guangxi, China, [3]Department of Child, Adolescent Health and Maternal Care, School of Public Health, Capital Medical University, Beijing, China, [4]Guangxi Academy of Medical Sciences, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China, [5]School of Public Health, Shandong University, Jinan, Shandong, China, [6]Department of Sociology, State University of New York (SUNY) Buffalo State University, Buffalo, NY, United States, [7]Life Sciences Institute, Guangxi Medical University, Nanning, Guangxi, China

Background: This study aims to assess the reliability and validity of the Quality of Life Scale (QOLS-6) among bank employees in Guangxi, grounding the investigation in the theoretical framework of quality of life measurement and psychological well-being. Given the increasing importance of mental health in the workplace, understanding how psychological conditions impact life quality is critical. The QOLS-6, a widely used tool for measuring quality of life, has been shown to have potential for application in diverse populations.

Methods: A cluster sampling method was used in the study. A questionnaire survey was conducted among 3,974 employees of a bank in Guangxi Province of China. To evaluate its performance across different mental health conditions, 298 participants in the suicidal ideation group and 3,676 in the non-suicidal ideation group. The non-suicidal group was randomly divided into two subsamples, with subsample A (n = 1,838) for exploratory factor analysis (EFA) and subsample B (n = 1,838) for confirmatory factor analysis (CFA). Reliability and validity were assessed for each group, enhancing the understanding of QOLS-6's sensitivity and effectiveness in capturing quality of life variations across different psychological profiles. Cronbach's α coefficient was used to analyze the internal consistency reliability. EFA and CFA were used to examine the construct validity. Spearman correlations was used to evaluate the concurrent validity.

Results: The QOLS-6 demonstrated good internal consistency, with Cronbach's α values of 0.908 and 0.857 for the two groups, respectively, and 0.865 for the overall sample. Its construct validity was supported by high KMO values of 0.841, 0.805 and 0.818 for the suicidal ideation group, subsample A and total sample, respectively. Exploratory factor analysis of the QOLS-6 revealed a single-factor structure. However, confirmatory factor analysis demonstrated that a three-factor model provided a better fit. The QOLS-6 showed negative correlations with depression, loneliness, hopelessness, and job burnout while positively correlating with job satisfaction and family function.

Conclusion: The QOLS-6 has demonstrated strong reliability and validity among bank employees in Guangxi, China. While exploratory factor analysis suggested a one-factor structure, confirmatory factor analysis supported a three-factor model, indicating a multidimensional nature. Based on the theoretical framework of quality of life and the design of the scale's content, the three-factor model demonstrates statistical and theoretical validity. Additionally, the scale exhibited significant correlations with key psychological factors, further supporting its applicability. These findings suggest that the QOLS-6 is an effective tool for assessing quality of life in diverse psychological contexts.

KEYWORDS

quality of life, reliability, validity, suicidal ideation, China

Introduction

The World Health Organization defines quality of life as an individual's subjective evaluation of their position in the cultural and value system, as well as their relationships with their goals, expectations, standards, and concerns (WHOQOL, 1993). Within various demographic groups and across diverse cultural milieus, the quality of life plays a significant role in shaping suicidal ideation and behaviors have been reported by researches (He et al., 2020). Health-related quality of life is affected by psychological issues such as depression and suicidal ideation which has been shown significantly in stroke patients in the United States (Chen et al., 2024). In Nigeria, suicidal behavior among HIV-infected individuals is strongly associated with a diminished quality of life (Oladeji et al., 2017). Similarly, a clear link between decreased quality of life and suicidal ideation have been identified by the study in a Australian general community (Goldney et al., 2001). These findings underscore the close relationship between quality of life and suicidal ideation. Therefore, this study groups participants based on the presence or absence of suicidal ideation to more accurately assess quality of life across different psychological states. Assessing quality of life provides a comprehensive understanding of an individual's physical, psychological, and social well-being, which are closely linked to suicidal ideation (Ki et al., 2024). The quality of life scale serves as an effective tool for identifying individuals experiencing a diminished quality of life, who may be at risk for mental health issues and potentially suicidal ideation or behaviors (Chen et al., 2024). This allows for early detection and targeted intervention to address underlying issues such as depression, loneliness, and hopelessness, thereby reducing the likelihood of suicidal behavior.

The Guangxi region in China, situated in the multi-ethnic and less developed western part of the country. In recent years, banking professionals have been experiencing significant psychological stress because of the pressures associated with digital transformation and other factors (Li, 2022). A study reported that the prevalence of suicidal ideation among bank employees in Guangxi was 7.5%, with major depression, hopelessness, and negative coping styles identified as significant risk factors (He et al., 2020).

The quality of life of Chinese bank employees is influenced by multiple factors and exhibits significant regional disparities. Work-life imbalance is a critical factor contributing to the decline in their quality of life. Long hours in confined environments and strict performance evaluations, which are typical of the banking profession, severely disrupt the work-life balance of bank employees. Research has demonstrated a significant positive correlation between work-life balance and job satisfaction in the Shanghai region (Xue et al., 2024). Due to factors such as fewer bank branches, employees in county and rural areas are more susceptible to work-life imbalance, leading to a decline in their quality of life (Li and Liu, 2021). Environmental factors, particularly the work environment, have a substantial impact on the quality of life of bank employees (Gür et al., 2016). Furthermore, grassroots bank employees in different altitude regions show varying probabilities of abnormal results in physical examinations, which further affect their quality of life (Wang and Zhao, 2008). These disparities suggest that improving the quality of life of Chinese bank employees requires tailored strategies and a comprehensive consideration of various factors. Nonetheless, research on the quality of life of bank employees remains limited, especially regarding the validation of relevant assessment tools.

Various measurement tools provide quantitative bases for assessing the quality of life of bank employees. WRQoL scale focuses on work-related aspects of life quality, accurately capturing employees' quality of life within their work environment (Geçkil, 2021). The Quality of Work Life Scale measures multiple dimensions, including job satisfaction, overall well-being, and work stress (Tentama and Suwandi, 2020). The scale was evaluated using structural equation modeling, which demonstrated that its dimensions and indicators effectively captured the construction of work life quality with high validity. In the empirical analysis of bank employees' quality of life, Cronbach's α was employed to assess the reliability of the model, and the results indicated a satisfactory level of reliability. Although various quality of life scales, including QOLS (Burckhardt and Anderson, 2003), WHOQOL-BREF (Makovski et al., 2019), CASP-11-SG (Tan et al., 2023), CQ-11D (Zhou et al., 2024) and ProQOL (Galiana et al., 2020), each has certain limitations. The WHOQOL-BREF demonstrates strong cross-cultural validity but requires lengthy administration time, while the QOLS, CASP-11-SG and CQ-11D face challenges in cultural adaptation within Chinese contexts. ProQOL, while effective in assessing professional burnout, has a narrower focus. Given these limitations, we selected the QOLS-6 scale and conducted targeted validation to better suit the specific population of this study.

This study used the 6-item short form Quality of Life Scale (QOLS-6) developed by Phillips (Phillips et al., 2002). The QOLS-6 was showed good reliability and validity in a psychological analysis of the general Chinese population in 2002. The scale focuses on six dimensions of quality of life over the preceding 30 days, including physical well-being, psychological health, economic status, occupational satisfaction, family relationships, and social interactions. The QOLS-6 has been validated in a study of psychological autopsy of elderly individuals who died by suicide in rural China, demonstrating its reliability and effectiveness within this population (He et al., 2021). Currently, the QOLS-6 has demonstrated strong psychometric properties primarily among rural elderly populations in China, highlighting its potential for broader applicability. The scale features a concise structure with only six items, enhancing its practicality for large-scale assessments.

To the best of our knowledge, QOLS-6 was used for the first time in research among the bank employees. This study evaluates the

Abbreviations: QOLS-6, Quality of life scale; MBI-GS, Maslach burnout inventory-general survey; PHQ-9, Patient health questionnaire; ULS-6, University of California Los Angeles Loneliness Scale-6; BHS-4, Beck's hopelessness scale; APGAR, Family adaptive, partnership, growth, affection and resolve scale.

applicability of the QOLS-6 within bank employees in Guangxi, providing a foundation for future research to validate its generalizability across diverse occupational and regional groups. By demonstrating its reliability and validity, the study proposes the QOLS-6 as a practical tool for assessing quality of life in mental health research contexts. The following research hypotheses are proposed:

> *H1*: Based on the theoretical framework of quality of life and the design of the scale's content, three latent factors can be extracted from the quality of life scale in the bank employee population.
>
> *H2*: The quality of life scale demonstrates good reliability in the bank employee population.
>
> *H3*: The quality of life scale demonstrates good validity in the bank employee population.

Methods

Design and participants

In this study, a cross-sectional survey and cluster sampling method were adopted to conduct an online questionnaire survey among 5,000 employees of a bank in Guangxi from November to December 2019. Guangxi is an underdeveloped region in western China. The research sample covers several prefecture-level cities such as Nanning, Liuzhou, Guilin and Guigang. The work institutions involved include first and second-level branches and first and second-level sub-branches of banks, and the work institution levels involved cover both the front desk and middle and back offices of banks. The researchers sent each respondent a survey link via an electronic questionnaire and briefly introduced the study. The participants signed informed consent, and their identities remained anonymous prior to the commencement of the study. To prevent data loss and redundant responses, it was necessary for each respondent to complete all questions and submitted their response only once. Meanwhile, the survey included an automated logic verification feature to guarantee the integrity of responses. After the survey was completed, the data was double-checked, cleaned, and reviewed, resulting in the collection of 3,974 valid questionnaires, including 298 employees in the suicidal ideation group and 3,676 employees in the non-suicidal ideation group. This study was approved by the Medical Ethics Committee of Guangxi Medical University (Ethical Application Ref: KY20240151).

Measurements

Demographics

The collected demographic data includes sex, age, educational level (Associate Degree or below, Bachelor's Degree, Master's Degree or above), marital status (Single, Married, or Other), and years of working experience (≤3 years, 4–12 years, over 12 years). "Other" marital statuses include separation, divorce, and widowhood. The age grouping is based on the sample's median age of 35 years. Work experience groups are defined according to the requirement in China for new employees to sign a labor contract of at least three years at the beginning of their employment, combined with the division of career stages into early, mid, and late phases.

Quality of life scale (QOLS-6)

The QOLS-6 developed by Phillips (Phillips et al., 2002) assesses the respondent's quality of life in six dimensions over the preceding month: physical, psychological, economic, occupational, familial relationships, and social interactions. Participants rated each item on a 5-point Likert scale (1 = very poor, 5 = excellent), with total scores ranging from 6 to 30. A higher total score indicates a better quality of life.

Suicidal ideation and self-harming actions

The presence of suicidal ideation within the past year is assessed by asking the question "Have you seriously considered ending your own life within the past year?" (Wang, 2019). In addition, self-harming actions were evaluated with the question: "At any point in the past, have you taken any self-harming actions?" Participants responded with either "Yes" or "No."

Maslach burnout inventory-general survey (MBI-GS)

The scale measures three dimensions of emotional exhaustion, depersonalization, and diminished professional accomplishment. This is a self-report questionnaire with 15 items, rating on a 7-point scale (0 = never to 6 = each day). The higher the total score on the scale, the higher the level of job burnout (Li and Shi, 2003). The scale demonstrated good internal consistency, with a Cronbach's α coefficient of 0.884.

Job satisfaction scale

The scale consists of six items, including satisfaction with promotion opportunities, colleagues, superiors, the job itself, income, and overall satisfaction. The answers can be given using a 5-point Likert Scale (1 = very dissatisfied to 5 = very satisfied). The higher the total score of the scale, the more satisfied the job (Schriesheim and Tsui, 1980). The Cronbach's α for the scale was 0.865, indicating satisfactory reliability.

Patient health questionnaire (PHQ-9)

PHQ-9 consists of nine depressive symptoms items that evaluate the state of the past two weeks, including both emotional (Items 1, 2, 4, 6, 8, 9) and physical (Items 3, 5, 7) dimensions. The specific items include "Over the past two weeks, how often have you felt little interest or pleasure in doing things?," etc. The answers can be given using a 4-point Likert Scale ranging from none to every day. The total score of the scale is 27 points, The higher total score indicating more severe depressive symptoms (Spitzer et al., 1999). In the study, the Cronbach's α for the scale was 0.918.

Chinese revision of short-form of the UCLA loneliness scale (ULSA-6)

Consisting of 6 items, it uses a 4-point Likert Scale (1 = never, 5 = always). The specific items include "Lack of company from others," "No one to turn to for help," "I feel neglected," etc. The total score of the scale ranges from 6 to 24 points, with higher scores indicating a stronger sense of loneliness (Zhou et al., 2012). The Cronbach's α for the scale was 0.913.

Beck hopeless scale (BHS)

The BHS evaluates an individual's mental state over the previous week, comprising four items rated on a 5-point Likert Scale from completely agree to completely disagreement. The specific items include, "I hope that in the future I can do the most important things well.," "My future is dark.," "I'm just unlucky and I'll always be." and "I am full of confidence about the future." The total score ranges from 4 to 20, with higher scores indicating a greater sense of hopelessness (Beck et al., 1974). The Cronbach's α for the scale was 0.797.

Affection and resolve scale (family APGAR)

The study utilized the Family APGAR Questionnaire developed by Smilkstein (1978), which evaluates family functionality across five dimensions: Adaptation, Partnership, Growth, Affection, and Resolve. The questionnaire employs a 3-point reverse scoring system, with responses of "often," "sometimes," and "rarely" scored as 2, 1, and 0 points, respectively. The total score ranges from 0 to 10, with higher scores indicating better family functionality. The Cronbach's α for the scale was 0.874.

Statistical analysis

Categorical variables were expressed as count and percentage, while continuous variables were expressed as mean ± standard deviation (Means ± SD). Descriptive analysis, chi-square test, independent sample t-test and Mann–Whitney test were employed to compare demographic characteristics between the groups.

Reliability was assessed using Cronbach's α coefficient, where values between 0.70 and 0.79 were deemed acceptable, 0.80 to 0.89 as good, and ≥ 0.90 as excellent (van Krugten et al., 2022). Corrected item-total score correlations were also calculated, with coefficients≥0.40 considered indicative of acceptable item consistency within the overall scale.

To ensure the robustness of the factor structure, the non-suicidal ideation group (N = 3,676) was randomly split into two subsamples: Subsample A (n = 1,838) for EFA and Subsample B (n = 1,838) for CFA. The structural validity of QOLS-6 was assessed through exploratory factor analysis (EFA). The suitability of EFA was confirmed by the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity, where KMO values >0.600 and p < 0.05 indicated adequacy (Ma et al., 2020). EFA was conducted using principal components extraction and varimax rotation, with factors identified based on eigenvalues >1 and the scree plot. A factor loading cut-off of ≥0.40 was applied to retain items contributing significantly to the factor structure. To ensure the accuracy of the results, a parallel analysis using Monte Carlo simulation was conducted, generating a series of data sets that simulate the experimental data. If the eigenvalues of the factors from the experimental data set are larger than those for the simulated data set, then it can be concluded that the respective factors are present in the data set (Sanatkar and Rubin, 2023). Confirmatory factor analysis (CFA) was used to evaluate the construct validity. Using chi-square test, root mean square error of approximation (RMSEA; values<0.08 are acceptable, ≤0.05 are ideal), standardized root mean square residual (SRMR; values ≤0.05 suggest a good fit), Bentler Comparative Fit Index (CFI), and Bentler Bonett normed fit index (NFI) (CFI and NFI values >0.90 suggest adequate fit) (Sun et al., 2019). For concurrent validity, Spearman correlation was used, with correlation coefficients below 0.2 considered absent, 0.2 to <0.35 as weak, 0.35 to <0.50 as moderate, and ≥ 0.50 as very strong (Haspels et al., 2023). Statistical analyses were performed using SPSS 27.0, Amos 26.0 and Mplus 8.3 software, and a two-tailed p-value of <0.05 was considered statistically significant.

Results

Descriptive analysis

Demographics

Among the 3,974 survey participants, there were 1,195 males (30.1%) and 2,779 females (69.9%). Among them, 298 (7.5%) reported suicidal ideation, as a suicidal ideation group, 3,676 (92.5%) of non-suicidal ideation workers as a non-suicidal ideation group. As shown in Table 1, it is evident that the prevalence of suicidal ideation is significantly higher among individuals aged ≤35 years (p = 0.001), and with 4–12 years of work experience (p = 0.038). The prevalence of self-harming actions was found to be higher among participants with suicidal ideation (p < 0.001). The scores of depression, hopelessness, loneliness, and job burnout in the group with suicidal ideation were significantly higher than those in the non-suicidal ideation group (p < 0.001), while their scores of job satisfaction and family function were significantly lower than those in the non-suicidal ideation group (p < 0.001).

Score of each item in quality of life scale

As shown in Table 2, total scores of quality of life in the suicide ideation group and the non-suicide ideation group were 18.59 ± 5.23 and 20.44 ± 3.80, respectively. Compared with the non-suicidal ideation group, the suicide ideation group had lower quality of life scores, indicating that the quality of life of the suicide ideation group was worse than that of the non-suicidal ideation group (p < 0.001).

Reliability

Internal consistency reliability

The Cronbach's α coefficient of the QOLS-6 scale was 0.865. For the suicidal ideation group and the non-suicidal ideation group, the Cronbach's α coefficients were 0.908 and 0.857, respectively. This indicates good internal consistency reliability.

Corrected item-total score correlation

The correlation coefficients between each item of QOLS-6 and the total scores of the suicidal ideation group and the non-suicidal ideation group were 0.673–0.829 and 0.512–0.769, respectively. Cronbach's α coefficients if item deleted ranged from 0.879–0.902 among suicidal ideation group and varied from 0.807–0.856 among non-suicidal ideation group. The details were shown in Table 3.

Validity

Construct validity

Exploratory factor analysis

For the overall sample, the KMO value was 0.818, and Bartlett's test of sphericity: χ^2 = 12160.210, p < 0.001. The exploratory factor

TABLE 1 Comparison of different characteristics between the suicidal ideation group and non-suicidal ideation group.

Characteristics	Suicidal Ideation Group (N = 298)	Non-suicidal Ideation Group (N = 3,676)	$\chi^2/Z/t$	p
Sex			0.704	0.401
Male	96 (32.2)	1,099 (29.9)		
Female	202 (67.8)	2,577 (70.1)		
Age groups			−3.783	**0.001**
≤35	194 (65.1)	1976 (53.8)		
>35	104 (34.9)	1700 (46.2)		
Years of working			−2.079	**0.038**
≤3 years	22 (7.4)	357 (9.7)		
4–12 years	201 (67.4)	2072 (56.4)		
>12 years	75 (25.2)	1,247 (33.9)		
Educational level			0.218	0.897
Associate degree or below	70 (23.5)	845 (23.0)		
Bachelor’s degree	223 (7.4)	2,780 (92.6)		
Master’s degree or above	5 (8.9)	51 (91.1)		
Marital status			5.855	0.054
Single	56 (18.8)	529 (14.4)		
Married	222 (74.5)	2,953 (80.3)		
Other	20 (6.7)	194 (5.3)		
Self-harming actions			−11.577	**<0.001**
Yes	99 (33.2)	55 (1.5)		
No	199 (66.8)	3,621 (98.5)		
Depression (Means ± SD)	11.51 ± 7.88	7.25 ± 5.19	9.180	**<0.001**
Hopeless (Means ± SD)	10.57 ± 3.40	8.26 ± 2.93	12.955	**<0.001**
Loneliness (Means ± SD)	14.86 ± 5.13	12.27 ± 4.33	8.464	**<0.001**
Job Burnout (Means ± SD)	42.99 ± 13.92	35.53 ± 12.22	8.982	**<0.001**
Job satisfaction (Means ± SD)	19.15 ± 3.95	20.66 ± 3.31	−6.391	**<0.001**
Family APGAR (Means ± SD)	5.68 ± 3.11	7.07 ± 2.65	−7.496	**<0.001**

APGAR, family adaptive, partnership, growth, affection and resolve scale. Bold values indicate $p < 0.05$.

TABLE 2 The score of each item for suicidal ideation group and non-suicidal ideation group in QOLS-6.

Item	Suicidal Ideation Group (Means ± SD)	Non-suicidal Ideation Group (Means ± SD)	t	p
1. How was your physical health in the last month?	2.92 ± 1.13	3.22 ± 0.95	−4.453	<0.001
2. How was your psychological health in the last month?	2.80 ± 1.19	3.30 ± 0.93	−7.135	<0.001
3. How was your economic status in the last month?	2.81 ± 1.08	2.98 ± 0.81	−2.659	0.008
4. How was your work in the last month?	3.07 ± 1.00	3.32 ± 0.78	−4.149	<0.001
5. How were your relationships with your family in the last month?	3.50 ± 1.00	3.87 ± 0.80	−6.194	<0.001
6. How were your relationships with others in the last month?	3.49 ± 0.91	3.76 ± 0.70	−4.992	<0.001
Total scores	18.59 ± 5.23	20.44 ± 3.80	−5.986	<0.001

analysis was determined to be suitable for both the suicidal ideation group (KMO = 0.841, Bartlett’s test of sphericity: χ^2 = 1264.004, $p < 0.001$) and the subsample A (KMO = 0.805, Bartlett’s test of sphericity: χ^2 = 5208.881, $p < 0.001$). These results indicate that the data met the necessary criteria for factor analysis, with adequate sampling adequacy and significant correlations among variables.

One common factor was extracted using exploratory factor analysis. The cumulative variance contribution rate was 60.37% for the total sample, 69.00% for the suicide ideation group, and 58.32%

TABLE 3 Corrected part-whole correlations and Cronbach's α if item deleted of QOLS-6 among suicidal ideation group and non-suicidal ideation group.

Items	Total Sample		Suicidal Ideation Group		Non-suicidal Ideation Group	
	R	α	*R*	α	*R*	α
1. How was your physical health in the last month?	0.678	0.840	0.751	0.891	0.666	0.830
2. How was your psychological health in the last month?	0.780	0.819	0.829	0.879	0.769	0.807
3. How was your economic status in the last month?	0.536	0.864	0.689	0.9	0.512	0.856
4. How was your work in the last month?	0.705	0.835	0.800	0.884	0.689	0.825
5. How were your relationships with your family in the last month?	0.612	0.851	0.673	0.902	0.598	0.841
6. How were your relationships with others in the last month?	0.678	0.842	0.751	0.892	0.663	0.832

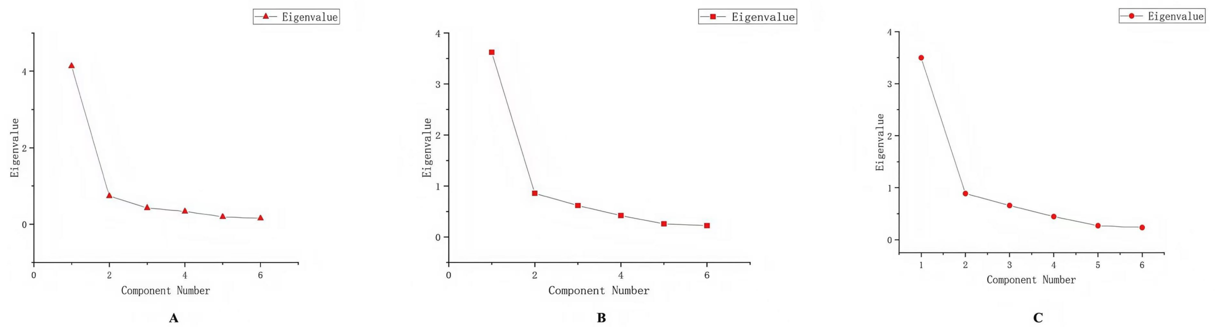

FIGURE 1
Scree plots from the exploratory factor analysis for **(A)** the total sample, **(B)** the suicidal ideation group, and **(C)** subsample A.

for the subsample A, respectively. Exploratory factor analysis extracted one common factor, with Eigenvalue >1.0 and a clear inflection point in the scree plot. Scree plots of exploratory factor analysis of total sample, suicidal ideation group and subsample A are shown in Figures 1A–C. The results of the parallel analysis showed that the eigenvalue of the first factor exceeded the eigenvalue of the first factor in the simulated data set. However, the eigenvalue of the second factor did not exceed the second factor in the simulated data. Consequently, we specified the extraction of only one factor using the promax method of oblique rotation. As shown in Figures 2A–C. Component matrix for each item of the QOLS-6 is shown in Table 4.

Confirmatory factor analysis

CFA was conducted on subsample B ($n = 1838$). As shown in the left panel of Figure 3, the fit indices for the single-factor model did not meet satisfactory standards ($\chi^2 = 891.837$, $p < 0.001$, CFI = 0.838 NFI = 0.837, SRMR = 0.048, RMSEA = 0.231). Considering that quality of life generally encompasses physical and mental health, economic status, and interpersonal relationships (The WHOQOL Group, 1998), and based on the specific content of the scale, correlations were added between items 1 and 2, items 3 and 4, and items 5 and 6. The fit indices improved ($\chi^2 = 33.757$, $p < 0.001$, CFI = 0.995, NFI = 0.994, SRMR = 0.012, RMSEA = 0.050). As shown in the right panel of Figure 3. Consequently, confirmatory factor analysis suggested that a three-factor solution was superior to a one-factor structure.

Concurrent validity

Table 5 shows significant correlations between QOLS-6 and other scales. The quality of life among bank employees is positively associated with family function and job satisfaction ($p < 0.001$), while showing negative associations with job burnout, depression, loneliness, and hopelessness ($p < 0.001$). In both the total sample and the suicidal ideation group, QOLS-6 exhibits a strong correlation with all other scales ($p < 0.01$). In the non-suicidal ideation group, QOLS-6 is strongly positively correlated with MBI-GS, PHQ-9, ULAS-6, and BHS-4 ($p < 0.01$), and moderately negatively correlated with APGAR and Job Satisfaction ($p < 0.01$).

Discussion

The results of this study showed that QOLS-6 had good reliability and validity. QOLS-6 is a practical tool to assess the quality of life among the bank occupational population.

The scale was also utilized in a psychological autopsy study conducted among elderly individuals at risk of suicide in rural China (He et al., 2021), demonstrating commendable reliability and validity. However, the mean quality of life scores observed within both the suicide group and non-suicidal ideation group were comparatively lower than those reported in this study, potentially attributable to varying conditions influencing quality of life across diverse groups. Consequently, it is imperative to conduct specific analyses when formulating intervention strategies. The study shows good internal

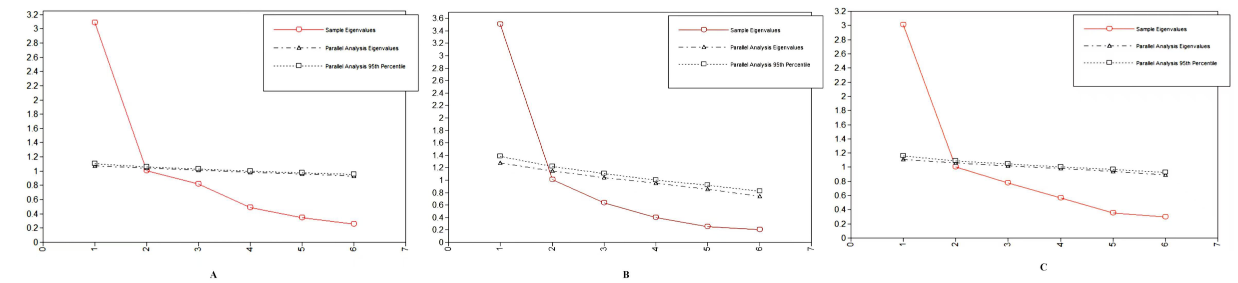

FIGURE 2
Parallel analysis plots from the exploratory factor analysis for **(A)** the total sample, **(B)** the suicidal ideation group, and **(C)** subsample A.

TABLE 4 Component matrix for each item of the QOLS-6.

Items	Total sample	Suicidal ideation group	Non-suicidal ideation group
	1	1	1
1. How was your physical health in the last month?	0.783	0.828	0.777
2. How was your psychological health in the last month?	0.86	0.888	0.854
3. How was your economic status in the last month?	0.659	0.778	0.639
4. How was your work in the last month?	0.807	0.869	0.797
5. How were your relationships with your family in the last month?	0.747	0.780	0.739
6. How were your relationships with others in the last month?	0.791	0.835	0.783

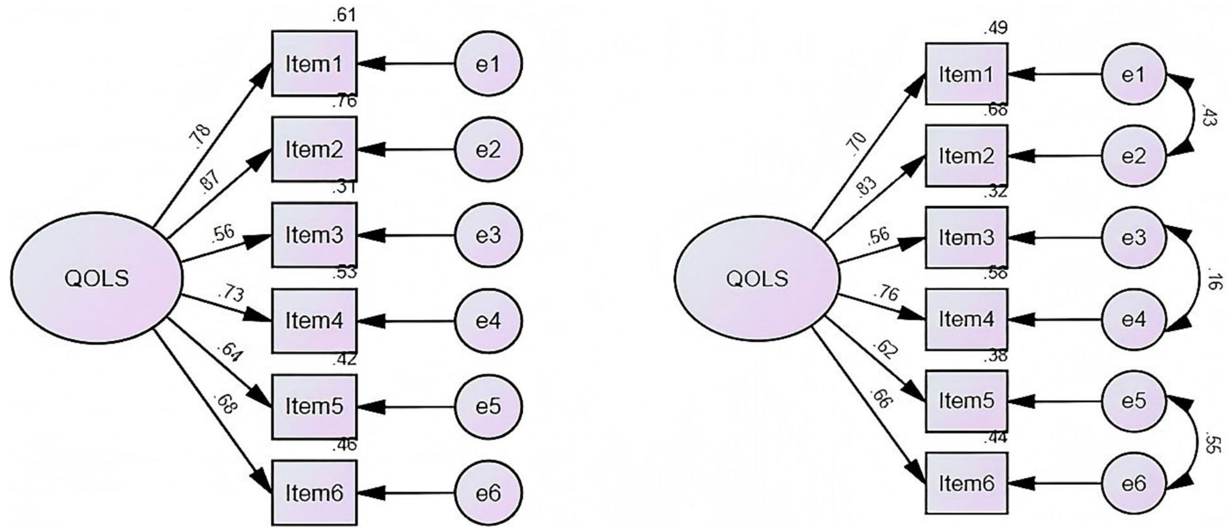

FIGURE 3
Pathway diagram of the confirmatory factor analysis of the one-factor (left panel) and three-factor (right panel) model for the QOLS-6.

consistency of QOLS-6 in overall sample, and both in the suicidal ideation group and the non-suicidal ideation group. One of the best methods for evaluating internal consistency is Cronbach's α coefficient, which were used to assess the reliability of QOLS-6. The Cronbach's α coefficient of the QOLS-6 scale was 0.865. For the suicidal ideation group and the non-suicidal ideation group, the Cronbach's α coefficients were 0.908 and 0.857, respectively. Also, QOLS-6 has good uniformity and stability in this study. The similar result was reported in the research among suicide elderly (He et al., 2021).

The split-sample EFA-CFA design within the non-suicidal group provides rigorous evidence for the unidimensionality of QOLS-6. The exploratory factor analysis supports a single-factor structure for the QOLS-6, with the analysis revealing only one common factor. This result differs from the finding of two factors

TABLE 5 Correlation coefficient between the QOLS-6 and other relevant scales.

Scales	QOLS-6		
	Total sample	Suicidal ideation group	Non-suicidal ideation group
	r	*r*	*r*
MBI-GS	−0.533***	−0.600***	−0.513***
PHQ-9	−0.633***	−0.732***	−0.601***
ULAS-6	−0.571***	−0.689***	−0.538***
BHS-4	−0.513***	−0.503***	−0.506***
APGAR	0.513***	0.516***	0.496***
Job Satisfaction	0.503***	0.586***	0.466***

QOLS-6, Quality of life scale; MBI-GS, Maslach burnout inventory-general survey; PHQ-9, Patient health questionnaire; ULAS-6, University of California Los Angeles Loneliness Scale-6; BHS-4, Beck's hopelessness scale; APGAR, family adaptive, partnership, growth, affection and resolve scale. ***$p < 0.001$.

identified in the research on elderly people who committed suicide in rural areas (He et al., 2021). Such a discrepancy may be attributed to differences in the characteristics of the populations studied. Confirmatory factor analysis revealed a three-factor structure (physical-mental health, economic-occupational status, and family-social relationships), which significantly improved model fit compared to the original single-factor model (CFI = 0.995 vs. 0.838, RMSEA = 0.050 vs. 0.231). This multidimensional structure aligns with the theoretical framework of quality of life proposed by the World Health Organization. Furthermore, the improved fit indices of the adjusted model suggest that incorporating correlations between items allows for a better capture of the underlying relationships between dimensions. A possible reason why EFA failed to extract three factors is the high correlation among measurement items. When different dimensions of the scale are closely related, EFA may underestimate the number of factors, leading to discrepancies between the factor structures identified by EFA and CFA (Sanatkar and Rubin, 2023). Moreover, EFA is a data-driven approach that assumes factors are uncorrelated by default (Mindrila, 2017). However, different dimensions of quality of life are often interrelated. As a result, EFA may fail to identify an underlying multifactor structure. In contrast, CFA is theory-driven and tests a predefined factor structure based on prior theoretical support. By accounting for correlations among factors, CFA provides a more accurate representation of the theoretical framework. Therefore, the findings support the classification of this scale into three distinct factor dimensions. This finding further validates the interrelationships among the different dimensions of the scale and provides valuable insights for future improvements to the scale.

In this study, quality of life was positive with family function and job satisfaction, negative with job burnout, depression, loneliness, and hopelessness. Depression can seriously lower the quality of life (Malhi and Mann, 2018). The sense of loneliness among elderly in China is negatively correlated with their quality of life (Zhu et al., 2018). Despair plays a mediating role between quality of life and suicide (Chen et al., 2022). A high quality of life reduces depression, hopelessness and suicidal ideation (Büsselmann et al., 2019). Good family function can improve the quality of life for older adults (He et al., 2021). Previous studies have shown similar results, so based on the correlation between these factors, the QOLS-6 also demonstrates good concurrent validity.

Quality of life is closely related to suicidal ideation, and a deep understanding of the relationship between them is helpful to prevent suicide. A large number of studies at home and abroad have proved that suicidal ideation is associated with low quality of life, including rural elderly people (He et al., 2021), community residents (van Spijker et al., 2020), Parkinson's patients (Santos-García et al., 2023) and so on. In this study, it shows that employees with suicidal ideation have a lower quality of life.

This study, by examining the reliability and validity of the QOLS-6 among bank employees in Guangxi, particularly distinguishing between groups with and without suicidal ideation, demonstrates the scale's applicability and reliability across different psychological states, holding significant practical implications. The findings not only validate the broad applicability of the scale in mental health assessments but also provide an effective tool for monitoring and intervention in high-stress occupational groups, such as bank employees. Through scientific evaluation of employees' quality of life, management can promptly identify and address potential mental health issues, enabling targeted support and intervention measures to enhance overall quality of life and job satisfaction.

This study also has limitations. Due to the use of a self-administered questionnaire, the study participants and subjectivity may cause certain bias in data quality and accuracy. Secondly, this study focuses on the bank workers in Guangxi, and further investigation is needed for other groups or countries of people, and more other factors should be taken into consideration. The study aimed to verify the reliability and validity of the QOLS-6 among bank employees, and thus did not conduct an in-depth analysis of the population with suicidal ideation. Future research could explore this topic more extensively, which would help to more comprehensively reveal the characteristics and psychological mechanisms of individuals with suicidal ideation, thereby providing a more scientific basis for formulating targeted intervention strategies.

Conclusion

This study, based on the theoretical framework of quality of life and the design of the scale's content, measured and analyzed the quality of life of bank employees. Confirmatory factor analysis (CFA) indicated that three latent factors could be extracted from the quality of life scale in this population, and the three-factor model exhibited a better fit than the single-factor model, aligning with theoretical expectations. Additionally, validity testing demonstrated that the scale performed well in terms of concurrent validity. Finally, reliability analysis showed that the internal consistency reliability of the scale reached a high level.

In conclusion, this study confirmed the applicability of the quality of life scale among bank employees in Guangxi, demonstrating its good reliability and validity. The scale can serve as an effective tool for assessing the quality of life of bank employees. However, further assessments of its reliability and validity across different populations and regions are recommended to better evaluate its factor structure.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Medical Ethics Committee of Guangxi Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

QH: Writing – original draft, Formal analysis. QM: Writing – review & editing. HH: Data curation, Investigation. XB: Data curation, Investigation. JZ: Writing – review & editing. ZM: Writing – review & editing, Project administration. GC: Writing – review & editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was provided by Guangxi Medical University 2020 Undergraduate Innovation and Entrepreneurship Training Program Project (Grant No.: 202010598173).

Acknowledgments

We are grateful to all the respondents for participating in this study. We also have all thank the investigators for the data collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Beck, A. T., Weissman, A., Lester, D., and Trexler, L. (1974). The measurement of pessimism: the hopelessness scale. *J. Consult. Clin. Psychol.* 42, 861–865. doi: 10.1037/h0037562

Burckhardt, C. S., and Anderson, K. L. (2003). The quality of life scale (QOLS): reliability, validity, and utilization. *Health Qual. Life Outcomes* 1:60. doi: 10.1186/1477-7525-1-60

Büsselmann, M., Nigel, S., Otte, S., Lutz, M., Franke, I., Dudeck, M., et al. (2019). High quality of life reduces depression, hopelessness, and suicide ideations in patients in forensic psychiatry. *Front. Psych.* 10:1014. doi: 10.3389/fpsyt.2019.01014

Chen, G., Mo, Q., Chen, X., Yu, B., He, H., Wang, G., et al. (2022). Quality of life, hopelessness, impulsivity, and suicide in the rural elderly in China: a moderated mediation analysis of psychological autopsy data. *Front. Public Health* 10:900156. doi: 10.3389/fpubh.2022.900156

Chen, H., Sun, Y., Song, Z., Lin, Z., Sang, Y., Fu, X. A., et al. (2024). Association of depressive symptom severity and suicidal ideation with health-related quality of life among stroke survivors, NHANES 2005-2018. *Psychiatry Res.* 336:115891. doi: 10.1016/j.psychres.2024.115891

Galiana, L., Oliver, A., Arena, F., De Simone, G., Tomás, J. M., Vidal-Blanco, G., et al. (2020). Development and validation of the short professional quality of life scale based on versions IV and V of the professional quality of life scale. *Health Qual. Life Outcomes* 18:364. doi: 10.1186/s12955-020-01618-3

Geçkil, T., and Şendoğdu, A. A. (2021). The impact of perceived organizational democracy on the quality of work life: an investigation in the banking sector. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 9, 497–507.

Goldney, R. D., Fisher, L. J., Wilson, D. H., and Cheok, F. (2001). Suicidal ideation and health-related quality of life in the community. *Med. J. Aust.* 175, 546–549. doi: 10.5694/j.1326-5377.2001.tb143717.x

Gür, K., Pinar, R., and Erol, S. (2016). Health related quality of life and related factors among bank call center employees. *Nobel Med*, 12.

Haspels, H. N., de Vries, M., and van den Akker-van Marle, M. E. (2023). The assessment of psychometric properties for the subjective wellbeing-5 dimensions (SWB-5D) questionnaire in the general Dutch population. *Qual. Life Res.* 32, 237–245. doi: 10.1007/s11136-022-03234-8

He, H., Bai, X., Chen, G., Zhang, Y., and Ma, Z. (2020). Incidence of suicidal intention and its related factors among employees in a bank of Guangxi. *J. Guangxi Med. Univ.* 37, 2270–2275. doi: 10.16190/j.cnki.45-1211/r.2020.12.031

He, H., Mo, Q., Bai, X., Chen, X., Jia, C., Zhou, L., et al. (2021). Studying suicide using proxy-based data: reliability and validity of a short version scale for measuring quality of life in rural China. *PeerJ* 9:e12396. doi: 10.7717/peerj.12396

Ki, M., Lapierre, S., Gim, B., Hwang, M., Kang, M., Dargis, L., et al. (2024). A systematic review of psychosocial protective factors against suicide and suicidality among older adults. *Int. Psychogeriatr.* 36, 346–370. doi: 10.1017/s104161022300443x

Li, Y. (2022). Carrying out humanistic Care for the Banking Industry under the current situation and psychological counseling. *Work J. Jilin Finan. Res.* 4, 53–56.

Li, X., and Liu, Y. (2021). Thoughts on the work-life balance of employees in county-level banks in second-tier cities and countermeasures. *China Collect. Econ.* 27, 90–91.

Li, C., and Shi, K. (2003). Influence of fairness of distribution and fairness of procedure on job burnout. *Acta Psychol. Sin.* 5, 677–684. doi: 10.1016/j.jpsychores.2012.12.017

Ma, Z., He, Q., Nie, G., Jia, C., and Zhou, L. (2020). Reliability and validity of short Beck hopelessness scale in psychological autopsy study among Chinese rural elderly. *Int. Psychogeriatr.* 32, 525–531. doi: 10.1017/s1041610219001315

Makovski, T. T., Schmitz, S., Zeegers, M. P., Stranges, S., and van den Akker, M. (2019). Multimorbidity and quality of life: systematic literature review and meta-analysis. *Ageing Res. Rev.* 53:100903. doi: 10.1016/j.arr.2019.04.005

Malhi, G. S., and Mann, J. J. (2018). Depression. *Lancet* 392, 2299–2312. doi: 10.1016/s0140-6736(18)31948-2

Mindrila, D. (2017). Exploratory factor analysis: An overview. *Exploratory factor analysis*, 1–25.

Oladeji, B. D., Taiwo, B., Mosuro, O., Fayemiwo, S. A., Abiona, T., Fought, A. J., et al. (2017). Suicidal behavior and associations with quality of life among HIV-infected patients in Ibadan, Nigeria. *J. Int. Assoc. Provid. AIDS Care* 16, 376–382. doi: 10.1177/2325957415617829

Phillips, M. R., Yang, G., Zhang, Y., Wang, L., Ji, H., and Zhou, M. (2002). Risk factors for suicide in China: a national case-control psychological autopsy study. *Lancet* 360, 1728–1736. doi: 10.1016/s0140-6736(02)11681-3

Sanatkar, S., and Rubin, M. (2023). An exploratory investigation of the reliability and validity of the independent-interdependent problem-solving style scale. *Int. J. Psychol.* 58, 30–41. doi: 10.1002/ijop.12878

Santos-García, D., de Deus Fonticoba, T., Cores Bartolomé, C., Feal Panceiras, M. J., García Díaz, I., Íñiguez Alvarado, M. C., et al. (2023). Suicidal ideation among people with Parkinson's disease and comparison with a control group. *Int. J. Geriatr. Psychiatry* 38:e5919. doi: 10.1002/gps.5919

Schriesheim, C., and Tsui, A. S. (1980). Development and validation of a short satisfaction instrument for use in survey feedback interventions. Western Academy of Management Meeting.

Smilkstein, G. (1978). The family APGAR: a proposal for a family function test and its use by physicians. *J. Fam. Pract.* 6, 1231–1239.

Spitzer, R. L., Kroenke, K., and Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA* 282, 1737–1744. doi: 10.1001/jama.282.18.1737

Sun, Y., Luk, T. T., Wang, M. P., Shen, C., Ho, S. Y., Viswanath, K., et al. (2019). The reliability and validity of the Chinese short Warwick-Edinburgh mental well-being scale in the general population of Hong Kong. *Qual. Life Res.* 28, 2813–2820. doi: 10.1007/s11136-019-02218-5

Tan, L. T., Østbye, T., Visaria, A., and Malhotra, R. (2023). Derivation, and establishment of the validity and reliability, of the CASP-11-SG quality of life scale among community-dwelling older adults. *Qual. Life Res.* 32, 295–306. doi: 10.1007/s11136-022-03238-4

Tentama, F., and Suwandi, G. F. (2020). The quality of work life scale: a psychometric study and its application for the employee. *Int. J. Sci.* 50, 68–81.

The WHOQOL Group (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. *Psychol. Med.* 28, 551–558. doi: 10.1017/s0033291798006667

van Krugten, F. C. W., Busschbach, J. J. V., Versteegh, M. M., Hakkaart-van Roijen, L., and Brouwer, W. B. F. (2022). The mental health quality of life questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Qual. Life Res.* 31, 633–643. doi: 10.1007/s11136-021-02935-w

van Spijker, B. A. J., Batterham, P. J., Calear, A. L., Wong, Q. J. J., Werner-Seidler, A., and Christensen, H. (2020). Self-reported disability and quality of life in an online Australian community sample with suicidal thoughts. *J. Affect. Disord.* 263, 707–714. doi: 10.1016/j.jad.2019.11.050

Wang, C. X. (2019). The affect of impulsivity and academic achievement on suicidal ideation in adolescents: a cross-lag analysis [Master's thesis], Tianjin Normal university]. Available online at: https://kns.cnki.net/kcms2/article/abstract?v=m4MC-GHErBuG0UDToV56IYNqmJ8Yc6_emSP9NAKeM3uwvSjjFpxCYmVnSTGivdnY9pw113VXGeItYPx-202V_ICa03212qApJPlI0zp7R7VWhiUvVIKJzfhH5rxTbE1-KqrQinNhRnHfPBWA0R1hs6st3LOKyrRWew7PvGf9f_PbSG0Zpvy5W9e2y4kA9DlNIn216cDD8rs=&uniplatform=NZKPT&language=CHS

Wang, H. L., and Zhao, X. M. (2008). Analysis of the health status of Bank employees in different altitude areas. *Qinghai Med. J.* 5, 57–58.

WHOQOL (1993). Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL). *Qual. Life Res.* 2, 153–159. doi: 10.1007/BF00435734

Xue, H., Omar Din, N. N., Balakrishna, S., and Zainal, N. (2024). Relationship between motivation determinants and job satisfaction among the employees of Bank of China, Shanghai. *Int. J. Acad. Res. Bus. Soc. Sci.* 14, 825–840. doi: 10.6007/IJARBSS/v14-i1/20475

Zhou, L., Li, Z., Hu, M., and Xiao, S. (2012). Reliability and validity of ULS-8 loneliness scale in elderly samples in a rural community. *J. Central South Univ.* 37, 1124–1128. doi: 10.3969/j.issn.1672-7347.2012.11.008

Zhou, J., Xu, L., Pan, J., Wang, M., Zhou, P., Wang, W., et al. (2024). A comparative study of Chinese medicine quality of life assessment scale (CQ-11D) and EQ-5D-5L and SF-6D scales based on Chinese population. *Qual. Life Res.* 33, 113–122. doi: 10.1007/s11136-023-03512-z

Zhu, Y., Liu, J., Qu, B., and Yi, Z. (2018). Quality of life, loneliness and health-related characteristics among older people in Liaoning province, China: a cross-sectional study. *BMJ Open* 8:e021822. doi: 10.1136/bmjopen-2018-021822

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 June 2025
DOI 10.3389/fpsyg.2025.1504554

Check for updates

OPEN ACCESS

EDITED BY
Karina M. Shreffler,
University of Oklahoma Health Sciences Center, United States

REVIEWED BY
Richard Xu,
Hong Kong Polytechnic University, Hong Kong SAR, China
Chiara Remondi,
Sapienza University of Rome, Italy

*CORRESPONDENCE
Victor Zaia
✉ victorzaia@gmail.com

RECEIVED 30 September 2024
ACCEPTED 10 June 2025
PUBLISHED 27 June 2025

CITATION
Zaia V, Gremigni P and Casu G (2025) Validation and dyadic invariance of the Infertility Concerns Questionnaire among Brazilian couples.
Front. Psychol. 16:1504554.
doi: 10.3389/fpsyg.2025.1504554

Validation and dyadic invariance of the Infertility Concerns Questionnaire among Brazilian couples

Victor Zaia[1,2]*, Paola Gremigni[3] and Giulia Casu[3]

[1]Postgraduate Program in Health Sciences, Centro Universitário FMABC, Santo André, Brazil, [2]Ideia Fértil Institute of Reproductive Health, Centro Universitário FMABC, Santo André, Brazil, [3]Department of Psychology, University of Bologna, Bologna, Italy

Introduction: Infertility often brings profound emotional challenges, intertwining the desire for parenthood with the complexities of communication within and beyond the couple. Existing scales fall short of fully capturing these nuances, prompting the development of the Infertility Concern Questionnaire (ICQ), which uniquely addresses both the parental desire and communication dimensions.

Methods: The ICQ was completed by 350 opposite-sex infertile couples from Brazil, aged 22–58 years, who were undergoing fertility treatment. Both partners of all couples completed the ICQ, and a subsample of 168 couples also completed a measure of infertility-related stress. The factor structure of the ICQ was evaluated using confirmatory factor analysis, and measurement invariance across couple members was tested within a dyadic framework.

Results: Configural, metric, partial scalar, and strict measurement invariance was supported for the ICQ's theoretical two-factor model of infertility concerns as made of parental (PAR - Examines the central role that the theme of procreation assumes in the construction of individual identity) and communication (COM - Investigates the significance of interpersonal sharing of reflections and experiences related to the decision to have—or not to have—children) dimensions. No between-partner differences were found in ICQ latent factor means. Reliability was found to be acceptable for both dimensions across genders. PAR correlated positively with intrapersonal infertility-related stress in both males and females, while COM correlated negatively with interpersonal infertility-related stress in females. Significant associations between ICQ scores and type of infertility, history of miscarriage, and treatment failures were observed among females.

Discussion: The findings provide initial evidence of the ICQ's validity and reliability, suggesting its potential use by professionals in fertility treatment settings to better assess and address the distinct emotional and communicative concerns of individuals and couples experiencing infertility.

KEYWORDS

infertility, desire for parenthood, communication, psychometric validation, dyadic invariance, assisted reproductive technology

1 Introduction

The World Health Organization (WHO) estimates that 15% of couples of reproductive age globally experience infertility (WHO WHO, 2006). Infertility negatively impacts individuals both intrapersonally and interpersonally (Casu and Gremigni, 2016). Intrapersonally, it affects quality of life, life satisfaction, mental health, self-esteem, and sexuality. Many individuals also grapple with guilt and a deep sense of failure, reflecting the heavy emotional toll infertility imposes on one personal sense of self (Ho et al., 2020; Taebi et al., 2021; Luca et al., 2021). Interpersonally, infertility often leads to social stigma and pressure to achieve parenthood, which can deteriorate social relationships, reduce access to social support, and lead to social isolation (Daniluk, 1988; Shreffler et al., 2020; Chamorro et al., 2024; Kiesswetter et al., 2020). Therefore, addressing both the internal and social challenges of infertility is essential when assessing psychological adjustment in both men and women (Braverman et al., 2024; Sharma and Shrivastava, 2022; Gameiro et al., 2015).

The diagnosis of infertility typically occurs after a year of unsuccessful attempts to conceive and marks the beginning of a treatment journey (Boivin et al., 2023), which can reshape life decisions and place the desire for parenthood at the forefront of personal priorities, profoundly influencing and altering various aspects of infertile individuals' lives (Massarotti et al., 2019; Zurlo et al., 2023). To address infertility, more than half of affected couples pursue medical interventions, such as assisted reproductive technology (ART) treatments (Boivin et al., 2007). ART encompasses a variety of medical and biological techniques to facilitate fertilization and support pregnancy, both *in vivo* and *in vitro* (Jain and Singh, 2024). However, ART does not guarantee success, with success rates ranging between 30 and 40%, depending on factors like infertility type, patient age, and the specific ART methods employed (Präg et al., 2017). The physical, financial, emotional, and time-consuming burdens of ART, alongside its uncertain outcomes, introduce significant additional stress (WHO WHO, 2006; Taebi et al., 2021; Ozturk et al., 2021) and contribute to a substantial percentage of couples discontinuing treatment prematurely (Pedro et al., 2017; Assaysh-Oberg et al., 2023; Shen et al., 2024). Women, in particular, tend to experience greater emotional distress throughout ART treatment (Bose et al., 2021; Tavousi et al., 2022).

The desire for parenthood is a complex, deeply personal aspiration that varies widely among individuals and cultures. Parenthood is frequently associated with personal fulfillment, the continuation of family traditions and values, and the desire to nurture and raise a child (Miscioscia et al., 2017). Unfulfilled parenthood aspirations can lead to feelings of sadness, frustration, and even depression (Assaysh-Oberg et al., 2023; Batz et al., 2023; Biggs et al., 2024; Canada and Schover, 2012). Conversely, the prospect of becoming a parent can bring joy and a sense of purpose. Cultural norms and societal expectations play a central role in shaping the desire for parenthood. In many cultures, having children is seen as a natural and essential part of life, while in others, greater emphasis may be placed on personal or professional achievements (Cetre et al., 2016; Järvinen, 1998; Harkness and Super, 2002). Furthermore, biological instincts and the natural drive to reproduce, along with life circumstances such as financial stability, relationship status, and career goals, also influence the desire for parenthood and decisions around it (Baiocco and Laghi, 2013; Vignoli et al., 2022; Beaujouan, 2020). For many individuals, the desire to become a parent is not only about personal fulfillment but also about the emotional and relational wish to share the parenting experience with a partner, other family members, or close friends (Miscioscia et al., 2017; Baiocco and Laghi, 2013; Miller, 1995).

However, many individuals experience difficulty discussing infertility with people outside their family, and some even find it hard to talk about it with their spouses (Zurlo et al., 2023; Sormunen et al., 2018). These communication barriers often increase emotional distress and contribute to social isolation, making it crucial to identify those at risk of developing emotional problems due to these difficulties (Xie et al., 2022). Effective communication in the context of infertility goes beyond simply exchanging information; it involves providing emotional support, fostering mutual understanding, and working collaboratively to navigate the infertile condition (Jansen et al., 2022). Since infertility is frequently accompanied by feelings of grief, frustration, and anxiety, open communication between partners allows for the sharing of these emotions, enabling mutual support and reducing feelings of isolation. Discussing fears, hopes, and expectations openly can strengthen the emotional bond between partners, enhancing their resilience in coping with the challenges of infertility (Kielek-Rataj et al., 2020; Schmidt, 2006). Additionally, effective communication can assist partners in managing expectations and facilitating joint decisions about treatment and future plans (Gameiro and Finnigan, 2017).

Infertility often carries social stigma, which can shape how individuals discuss their condition with family, friends, and the broader community. Open and honest communication not only helps to reduce this stigma but also fosters a support network that is vital for emotional well-being (Taebi et al., 2021; Bose et al., 2021; Kielek-Rataj et al., 2020; Schmidt, 2006; Wang et al., 2024; Tang et al., 2022; Schmidt et al., 2005). When individuals can openly express the struggles of infertility and its treatments, they are more likely to receive social support, which plays a vital role in improving well-being (Massarotti et al., 2019; Sormunen et al., 2018; Schmidt, 2006; O'Connell et al., 2021; Chu et al., 2021). Therefore, cultivating open communication within relationships and social networks is essential for alleviating the emotional burden of infertility.

The desire for parenthood and communication difficulties are central concerns for many infertile individuals, influencing both their emotional well-being and their ability to cope with infertility and its treatment. Understanding these concerns in patients seeking ART can enable healthcare professionals, including psychologists, doctors, and nurses, to offer more tailored support strategies, addressing both the psychological and social needs of patients for more effective care.

In this context, various measures related to infertility experiences have been proposed, focusing on different psychological dimensions. The primary themes identified include the desire for parenthood, marital relationships, infertility-related stress, social relationships, anxiety, depression, and overall mental health (Tavousi et al., 2022; Nik Hazlina et al., 2022; Starc et al., 2019). General scales, such as the World Health Organization Well-Being Index (WHO-5) (Omani-Samani et al., 2019), the World Health Organization Quality of Life Assessment (WHOQOL) (WHO, 1995), and the Beck Depression Inventory (BDI) (Beck et al., 1961), were used also in the context of infertility. These are broad and comprehensive scales designed to assess mental health and quality of life, applicable to various conditions as well as the general population. Scales specifically tailored to the context of infertility were reviewed by Kitchen et al. (2017). These include, among others, the Fertility Problem Inventory (FPI) (Newton et al., 1999), the Fertility Problem Stress (FPS) (Abbey et al., 1992), the Fertility Quality of Life (FERTIQOL) (Boivin et al., 2011), the Infertility Questionnaire (IFQ) (Bernstein et al., 1985), and Parenthood Motivation List (PML) (van Balen and Trimbos-Kemper, 1995). Additionally, we considered the Infertility Related Stress Scale (IRSS) (Casu and Gremigni, 2016; Casu et al., 2021) a scale measuring stress specifically related to infertility, validated in the context of assisted reproductive technology.

While these existing scales provide valuable insights into various aspects of infertility, they often focus on specific themes such as stress, quality of life, or motivation. However, none comprehensively assess the unique interplay between the desire for parenthood and the communicative aspects of dealing with infertility. In this context, we propose the Infertility Concern Questionnaire (ICQ), a concise self-report scale designed to fill this gap. The ICQ stands apart from existing measures, such as the FPI, by addressing two dimensions that have not been fully explored together in previous instruments: the parental dimension, which focuses on the desire to become a parent, and the communicative dimension, which emphasizes the role of communication, both within and outside the couple, in enhancing well-being. What distinguishes the ICQ from other tools is its focus on the individual's personal experience of infertility, rather than the consequences of infertility.

The present study aimed to develop and psychometrically test the Infertility Concern Questionnaire (ICQ) within the Brazilian context. We assessed content and face validity, validity based on factor structure and dyadic measurement invariance, reliability, and validity based on relations to other variables. Dyadic measurement invariance refers to the consistency of a questionnaire's measurement properties across different members of dyads, such as romantic couples or parent–child pairs. This ensures that the questionnaire measures the same constructs in the same way for both members, enabling valid comparisons. Establishing dyadic invariance is crucial, as it ensures that any observed differences or similarities in responses reflect actual differences in the constructs being measured, rather than variations in how the questionnaire functions for each dyad member (Claxton et al., 2015). In this study, we tested dyadic invariance across males and females in infertile couples to ensure that the ICQ measures infertility concerns consistently for both partners.

2 Materials and methods

2.1 Participants and procedures

This cross-sectional study was conducted with couples undergoing ART at the Ideia Fértil Institute, Center for Human Reproduction and Genetics, located at Centro Universitário FMABC/Faculdade de Medicina do ABC in Santo André, Brazil. Participants were required to be at least 18 years old and have started ART treatment. Exclusion criteria included patients diagnosed with cancer, those pursuing ART for fertility preservation, single-parent families, and same-sex couples, as their reasons for treatment were not related to infertility.

All couples were personally invited to participate in the study by the first author. They were informed about the research details, and those agreeing to participate provided written informed consent. It was emphasized that participation was voluntary, and withdrawal or refusal to participate would have no impact on their treatment. Data collection was conducted using paper-and-pencil forms, with participants' anonymity maintained.

A total of 376 couples attending the infertility clinic who met the inclusion criteria were invited to participate in the study. Of these, 350 couples (93%) agreed to participate and completed the questionnaire, while the remaining couples declined due to lack of interest. Participants were between 22 and 58 years old (mean age: women = 34.78 ± 4.81; men = 37.28 ± 6.31), and 50% were women. Both members of all couples were administered the Infertility Concern Questionnaire (ICQ), and approximately half of the couples (n = 168), randomly selected, were also given the Infertility Related Stress Scale–Brazilian Portuguese (IRSS-BP). Sociodemographic and infertility-related characteristics were collected from all participants. Sociodemographic information included gender, age, ethnicity, education, employment status, household income, and relationship length. Infertility-related information included the type of infertility (i.e., primary infertility, defined as never having conceived despite at least 12 months of attempting conception, or secondary infertility, defined as having had at least one prior conception but subsequently being unable to conceive after at least 12 months of trying), history of miscarriage, infertility diagnosis (i.e., male factor, female factor, both male and female factor, or unexplained), duration of infertility, type of current ART treatment, and whether couples had experienced previous failed ART cycles.

The study was approved by the Ethics Committee of the Centro Universitario FMABC (approval number 88192218.3.0000.0082/2.675.077) and adhered to the principles of the Helsinki Declaration.

2.2 Measures

2.2.1 Infertility concern questionnaire

By reviewing existing infertility-specific scales, we identified several relevant instruments. The Fertility Problem Inventory (FPI, 46 items) assesses infertility-related stress across five domains: social, sexual, and relationship concerns, need for parenthood, and rejection of a childfree lifestyle (Newton et al., 1999); a Brazilian Portuguese version is available (Ribeiro, 2007). The Fertility Problem Stress Scale (FPS, 9 items) measures perceived infertility-related stress (Abbey et al., 1992), but no Brazilian Portuguese version exists. The Fertility Quality of Life tool (FERTIQOL, 24 items) evaluates the impact of infertility across four dimensions: emotional, relational, mind/body, and social (Boivin et al., 2011); a Brazilian Portuguese version is available from the original authors. The Infertility Questionnaire (IFQ, 21 items) focuses on self-esteem, blame/guilt, and sexuality (Bernstein et al., 1985), while the Parenthood Motivation List (PML, 18 items) assesses six motivations for having children (van Balen and Trimbos-Kemper, 1995); neither is available in Brazilian or Portuguese versions.

Based on this review, we identified a gap in the available instruments: the absence of a brief and accessible instrument in Brazilian Portuguese specifically designed to assess parenthood-related concerns and communication about the infertility experience in the context of medically assisted reproduction. To address this gap, we developed the Infertility Concern Questionnaire (ICQ) as a concise and easily applicable self-report scale, drawing inspiration from existing tools that assess experiences related to infertility (Newton et al., 1999; Boivin et al., 2011; Casu et al., 2021). We proposed 18 items aimed at exploring emotional and relational aspects, as well as individual and couple experiences associated with infertility, with particular attention to the desire for parenthood and communication about the infertility experience. Each item is rated on a 7-point Likert scale, ranging from 1 = "Completely false for me" to 7 = "Completely true for me."

The ICQ items were submitted to a focus group consisting of seven Brazilian psychologists, all PhD students in health psychology, to assess content validity. Their objective was to evaluate the conceptual construction of the items, ensuring that each item accurately captured and measured the intended aspects as fully described. A minimum concordance rate of 80% was required for an item and its domain to be retained. Each focus group member reviewed the items to determine whether they aligned with the proposed instrument and to identify the appropriate domain – Parenthood (PAR) or Communication (COM) (Meijering et al., 2013). This level of agreement was achieved for 7 of the 18 originally proposed items, 4 of which pertained to the PAR domain (items 1, 2, 3, and 4) and 3 to the COM domain (items 5, 6, and 7). For the retained items, the group confirmed their face validity.

Subsequently, a focus group was conducted with a sample from the target population, invited by the authors. The group consisted of nine participants (56% women), aged between 20 and 45 years, with varied ethnic and socioeconomic backgrounds and the session lasted 1 h. The objective was to assess whether the items were clear, relevant, and meaningfully captured the participants' experiences related to infertility. The participants did not suggest any modifications to the items and expressed full agreement with the ICQ format initially proposed by the expert focus group. Based on this feedback, the final version of the ICQ was approved for the next phase of validation and administered to the research target audience.

2.2.2 Infertility related stress scale

The Infertility Related Stress Scale–Brazilian Portuguese (IRSS-BP) (Casu et al., 2021) is a 12-item self-report measure designed to assess the level of stress that infertility imposes on different aspects of life. This instrument has been validated in Brazilian Portuguese within the context of assisted reproductive treatment and is freely accessible. It consists of two 6-item scales referring to the intrapersonal (e.g., mental well-being) and interpersonal (e.g., friends) domains of life. Each item is rated on a 7-point scale from 1 (not at all) to 7 (a great deal). Higher scores indicate a higher level of infertility-related stress. McDonald's ω in this study for the intrapersonal domain was 0.95 for males and 0.94 for females, and for the interpersonal domain was 0.96 for males and 0.96 for females.

2.3 Data analysis

Descriptive statistics were used to characterize the study sample. Using confirmatory factor analysis (CFA), the ICQ theoretical two-factor model and an alternative one-factor model were first tested in males and females separately. Dyadic measurement invariance of the selected model was then examined. To account for non-independence of observations at the factor and item levels, the factor model for males and females was connected through correlations between the latent factors, and error terms of parallel items were correlated between couple members. Increasingly restrictive models were compared that incrementally constrained model parameters to be equal across couple members, representing configural (equal factor structure), metric (equal factor loadings), scalar (equal intercepts), and strict (equal error variances) invariance. To test between-partner differences in latent factor means in case of achieved scalar invariance, we set the latent means as equal across couple members and compared this model against the scalar invariance model. Model fit was evaluated based root mean square error of approximation (RMSEA) and standardized residual mean root (SRMR), with values of ≤0.08 indicating acceptable fit, and comparative fit index (CFI), with values ≥ 0.90 as indicative of acceptable fit (Hu and Bentler, 1999). We evaluated measurement invariance based on both statistical and practical significance. Statistically, measurement invariance was considered achieved if nested models showed a non-significant decrease in model fit, as indicated by a non-significant χ^2 difference test ($\Delta\chi^2$). From a practical perspective, measurement invariance was considered achieved if the worsening of model fit was sufficiently small, as indicated by a decrement in CFI (ΔCFI) of ≤ 0.010, supplemented by increases in RMSEA (ΔRMSEA) and SRMR (ΔSRMR) of ≤0.015 and ≤0.010, respectively (Chen, 2007). If full measurement invariance was not achieved, partial measurement invariance was considered, which is viable when at least two items per latent construct have invariant parameters (Steenkamp and Baumgartner, 1998). To identify which items were responsible for the lack of invariance, equality constraints on the measurement parameters were sequentially relaxed. These nested models were compared using $\Delta\chi^2$ tests, with an adjusted α-level (Bonferroni correction). The effects coding identification method was used for measurement invariance testing, allowing latent parameters to be estimated in a non-arbitrary way by constraining the factor loadings to average 1 and the intercepts to sum to 0 (Little et al., 2006). The sample size was established *a priori* to ensure at least 5–10 observations for each estimated parameter in the baseline configural invariance model (Kline, 2016).

The estimation method used in the CFA was maximum likelihood (ML), which assumes multivariate normality. Since ML is more sensitive to deviations caused by kurtosis (Ryu, 2011), we evaluated multivariate normality using Mardia's coefficient of multivariate kurtosis. The standardized value obtained was 8.022, which falls below the commonly accepted threshold of 10 for the application of ML estimation (Kline, 2016).

Reliability was assessed using McDonald's ω (Gadermann et al., 2012), with acceptable values set at 0.70 or higher. Corrected item-total correlations were also calculated, with acceptable values greater than 0.30 (Nunnally and Bernstein, 1994).

To assess validity based on relations to other variables, Pearson's correlations were computed to examine the relationships between the ICQ and IRSS-BP, while ANOVAs were performed to test associations between ICQ scores AND infertility-related characteristics separately in males and females. Effect size values were interpreted as follows: correlations of 0.20 were considered small, 0.30 medium, and 0.50 large; Cohen's d values of 0.20, 0.50, and 0.80 indicated small, medium, and large effects, respectively (Cohen, 1988).

Significance level was set at $p < 0.05$. CFAs were conducted in Mplus 8.4. All other analyses were performed using IBM SPSS 29.

3 Results

3.1 Participants' characteristics

The sample consisted of 350 opposite-sex infertile couples (i.e., 50% women), with participants aged between 22 and 58 years (mean age: women = 34.78 ± 4.81; men = 37.28 ± 6.31). About 60% of both males and females identified as being of European descent, while 27% reported mixed ethnicity. Half of the males and two thirds of the females were highly educated, holding a degree or post-degree, and the majority of both males (93.7%) and females (82.9%) were employed. More than half of the couples had a medium household income (53.1%) and had been together for 5–9 years (55.1%). Seventy percent of the couples experienced primary infertility, and 14% had experienced previous miscarriages. One-third of the couples had unexplained infertility, and two-thirds had been trying to conceive for more than 2 years. Around half of the couples had experienced previous failed ART cycles. Participants' characteristics are detailed in Table 1.

3.2 Dyadic invariance of the ICQ

The proposed two-factor model for the ICQ showed an acceptable fit for both males [$\chi^2(13) = 30.966$, $p = 0.003$; RMSEA = 0.063, 90% CI 0.034–0.092, $p = 0.205$; SRMR = 0.035; CFI = 0.969] and females [$\chi^2(13) = 36.772$, $p = 0.0004$; RMSEA = 0.072, 90% CI 0.045–0.100, $p = 0.083$; SRMR = 0.051; CFI = 0.942]. The fit of the two-factor model was better than the alternative one-factor model for both males [$\Delta\chi^2(1) = 89.127$, $p < 0.001$] and females [$\Delta\chi^2(1) = 84.642$, $p < 0.001$]. The conceptual two-factor model was thus selected and subjected to testing of dyadic measurement invariance (Table 2).

The configural invariance model showed acceptable fit, indicating that the ICQ had a similar factor structure between couple members.

TABLE 1 Sociodemographic and infertility-related characteristics of the participants.

Characteristics	Males		Females		Couple	
	n	%	*n*	%	*n*	%
Age, *M* (SD, range)	37.28	(6.31, 23–58)	34.78	(4.81, 22–54)		
Ethnic group						
European	213	60.9	211	60.3		
Mixed-ethnicity	97	27.7	96	27.4		
African	22	6.3	30	8.6		
Middle Eastern/Asian	11	3.1	5	1.4		
Native South American	2	0.6	0	0		
Education, high	185	52.9	239	68.3		
Job status, employed	328	93.7	290	82.9		
Income						
Low					73	20.9
Medium					186	53.1
High					91	26.0
Relationship length						
1–4 years					58	16.6
5–9 years					193	55.1
>9 years					99	28.3
Infertility type, primary					247	70.6
History of miscarriage					50	14.3
Infertility diagnosis						
Male					89	25.4
Female					86	24.6
Both male and female					59	16.9
Unexplained					116	33.1
Infertility duration						
1–2 years					122	34.9
3–4 years					113	32.3
5–6 years					74	21.1
>6 years					41	11.7
Current ART treatment, IVF					244	69.7
Previous ART failures					183	52.3

ART, assisted reproductive technology; IVF, *in vitro* fertilization.

The correlation between parenthood (PAR) and communication (COM) latent factors was 0.546 among males and 0.337 among females ($p < 0.001$), indicating that the two factors were moderately correlated but still independent. Correlations between males' and females' latent factors ranged from 0.176 ($p < 0.05$) to 0.425 ($p < 0.001$). Correlations between parallel item residuals ranged from 0.023 ($p = 0.694$) to 0.297 ($p < 0.001$). Full metric invariance was achieved, as constraining all factor loadings to be equal across couple members did not worsen model fit. This indicates that ICQ items have equivalent meaning across couple members in terms of the strength of the association between each item and its latent factor, and that the items are measuring the latent factors using the same metric scale across partners. However, full scalar invariance was not supported, as constraining all item intercepts to be equal across couple members resulted in a significant loss of fit compared to the metric invariance model, with a significant $\Delta\chi^2$ and a ΔCFI greater than 0.010. Subsequent analyses ($\Delta\chi^2$ tests with Bonferroni-corrected $\alpha = 0.007$) revealed that the intercepts of item 4 ("The infertility problem has changed my life projects") were non-invariant, being higher for females (0.715) than for males (0.259) [$\Delta\chi^2(1) = 9.229$, $p = 0.002$]. When the equality constraint on this item's intercept was released, the $\Delta\chi^2$ test remained significant ($p = 0.029$), but the changes in CFI, RMSEA, and SRMR were sufficiently small, supporting partial scalar invariance. This suggests that mean item differences across couple members are due to differences in the latent factor being measured. The strict invariance model, in which the error variances of items showing full scalar invariance were constrained to be equal across couple members, was also achieved. Therefore, the ICQ latent factors are measured with the same measurement error across both partners. Estimates for the measurement parameters from the strict invariance model are displayed in Table 3.

Since partial scalar invariance was supported, we proceeded to test for differences in latent factor means between couple members. The addition of equality constraints on the latent factor means resulted in non-significant differences compared to a model without these constraints. Thus, males and females had similar factor means in both PAR (males: $M = 4.713$, $SE = 0.117$; females: $M = 4.632$, $SE = 0.111$) and COM (males: $M = 4.387$, $SE = 0.087$; females: $M = 4.429$, $SE = 0.085$).

3.3 Reliability

McDonald's ω was 0.76 for males and 0.83 for females for PAR, and 0.85 for males and 0.73 for females for COM. Item-total correlations for PAR ranged from 0.38 to 0.63 among males and from 0.44 to 0.51 among females. For COM, item-total correlations ranged from 0.43 to 0.53 among males and from 0.37 to 0.48 among females.

3.4 Validity based on relations to other variables

For males, a significant positive correlation was observed between PAR and intrapersonal stress, with a small effect size. For females, PAR correlated significantly and positively with both intrapersonal and infertility-related stress, with small-to-medium and medium effect sizes, respectively. Among females only, there was a significant

TABLE 2 Dyadic measurement invariance of the ICQ.

Level of invariance	df	χ^2	Δdf	$\Delta\chi^2$	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
Configural	64	119.944	–	–	0.949	–	0.050	–	0.046	–
Metric	69	126.045	5	6.101^{ns}	0.948	0.001	0.049	0.001	0.049	0.003
Scalar	74	146.031	5	19.986**	0.935	0.013	0.053	0.004	0.051	0.002
Scalar partial[a]	73	136.802	4	10.757*	0.942	0.006	0.050	0.001	0.051	0.002
Strict[b]	79	140.051	6	3.249^{ns}	0.945	0.003	0.047	0.003	0.053	0.002
Equal factor means	75	138.061	2	0.992^{ns}	0.943	0.001	0.049	0.001	0.051	0.000

$^{ns}p > 0.05$, $*p < 0.05$, $**p \leq 0.001$.

[a]Intercept of item 4 freely estimated.

[b]Intercept and error variance of item 4 freely estimated.

TABLE 3 Strict invariance model parameter estimates.

ICQ item	Factor loading	Intercept	Error variance
ICQ1. There is nothing left in life without children.	1.087 (0.051)	−0.495 (0.227)	2.748 (0.182)
ICQ2. Having a child is a central issue in my life	1.194 (0.050)	−0.808 (0.208)	1.225 (0.133)
ICQ3. Having children gives you something to live for.	0.923 (0.044)	1.303 (0.200)	1.475 (0.108)
ICQ4. The infertility problem has changed my life projects[a].	0.796 (0.059)	0.247 (0.363)/0.703 (0.354)	4.197 (0.336)/3.273 (0.270)
ICQ5. It helps me to talk about the fact that we have trouble having kids.	1.115 (0.064)	−0.548 (0.286)	2.236 (0.233)
ICQ6. I enjoy spending time with friends who have children.	0.803 (0.058)	1.107 (0.263)	2.959 (0.200)
ICQ7. It helps me to share with other couples who have the same problem.	1.081 (0.067)	−0.558 (0.303)	2.959 (0.251)

Unstandardized estimates (standard errors) are presented; all factor loadings and error variances were significant at $p < 0.001$.

[a]Estimates before and after the slash refer to males and females, respectively. This item showed non-invariant intercepts in the full scalar invariance models.

negative correlation, of small effect size, was found between COM and interpersonal infertility stress. The correlation coefficients are presented in Table 4.

As shown in Table 5, significant associations were found between females' ICQ scores and the infertility-related characteristics of infertility type, history of miscarriage, and previous failed cycles. Specifically, female participants with secondary infertility reported significantly, slightly higher COM scores than those with primary infertility ($d = 0.23$). Females with a history of miscarriage and previous failed ART cycles reported significantly, slightly lower PAR scores than those without such histories ($d = 0.33$ and 0.32, respectively). No significant effects were observed for other infertility-related characteristics in females, and no significant associations were found among males.

4 Discussion

The ICQ demonstrated content and face validity, supported by a multi-stage qualitative evaluation involving expert psychologists and confirmed through focus groups with the target population of infertile individuals. Results from CFAs and measurement invariance tests supported the proposed two-correlated factor model of Parenthood (PAR) and Communication (COM) and confirmed that the ICQ functions consistently across couple members. Specifically, the ICQ demonstrated full configural and metric invariance, along with partial scalar and strict invariance. The two-factor structure was invariant across couple members, with invariant factor loadings for all items, and invariant intercepts and residual variances for 86% (6/7) of its items. Notably, this proportion aligns well with the established standards for partial measurement invariance (Steenkamp and Baumgartner, 1998). Overall, these findings indicate that the ICQ works equivalently for both male and female partners, who conceptualize PAR and COM in similar ways. For both partners, the desire for parenthood represents a central life goal that shapes their sense of purpose and future plans. Parenthood is seen not only as a source of joy and personal fulfillment but also as a key part of their identity and life projects. Infertility, in turn, has a profound impact on these aspirations, altering life trajectories and priorities. Communication, on the other hand, is viewed as an essential means of emotional support

TABLE 4 Correlations with IRSS (n = 168) and descriptive statistics (n = 350) of ICQ.

IRSS domain	Males		Females	
	ICQ PAR	ICQ COM	ICQ PAR	ICQ COM
IRSS Intrapersonal	0.19*	0.04	0.25**	−0.12
IRSS Interpersonal	0.08	−0.03	0.10	−0.16*
M (SD)	18.67 (6.33)	12.86 (4.90)	18.36 (5.91)	15.59 (4.95)

IRSS, Infertility Related Stress Scale; ICQ, Infertility Concern Questionnaire; PAR, parenthood; COM, communication; correlations were computed on a subsample of couples (n = 168) that also completed the IRSS; score range was 4–28 for ICQ PAR and 4–21 for ICQ COM. M, median; SD, standard deviation.
*$p < 0.05$, **$p \leq 0.001$.

and coping, helping couples navigate the emotional and psychological challenges associated with infertility. For both partners, communication serves as a vital tool for building mutual understanding, sharing experiences, and finding relief through connections with others who face similar struggles.

Because the most stringent level of measurement invariance (i.e., strict invariance) was achieved, there is evidence that the ICQ provides equivalent precision in measurement across couple members. This allows for meaningful interpretations of between-partner comparisons on observed means and covariance structures (Millsap and Meredith, 2007). Additionally, since partial scalar invariance was reached, comparisons of ICQ latent factor means between members of infertile couples are valid. Results indicated that male and female partners had similar underlying levels of PAR and COM. This is in line with previous evidence showing that men desire parenthood as much as women (Lewis, 1988; Nitsche and Hayford, 2020; Sylvest et al., 2018; Hammarberg et al., 2017). Additionally, other studies have found no significant difference in how partners of infertile couples perceived their communication styles when dealing with stressors (Wang et al., 2024).

However, despite these broad similarities, it is important to acknowledge that men and women may experience infertility in distinct ways. The ICQ did not achieve full measurement invariance, indicating that some differences persist in how infertility-related concerns affect each partner's life projects (item 4). This may be partially explained by the social burden disproportionately placed on women, who often face greater expectations and pressures regarding childbearing (Caddy et al., 2023; Mills, 2010).

Reliability coefficient, McDonald's ω, and corrected item-total correlations, indicated acceptable levels of internal consistency for both ICQ dimensions across males and females, meeting established criteria (Gadermann et al., 2012; McNeish, 2018).

As evidence of validity based on relations to other variables, correlation analysis revealed significant, though weak, relationships between the ICQ and IRSS-BP domains. In both males and females, desire for parenthood correlated positively with the intrapersonal domain of infertility-related stress, but not with the interpersonal domain. This finding is expected, as the PAR dimension is inherently more intrapersonal in nature. The unfulfilled desire for parenthood among infertile individuals disrupts personal development, often leading to emotional turmoil, a sense of loss of control over life, and feelings of failure in identity building (Nik Hazlina et al., 2022; Alamin et al., 2020). Furthermore, the inability to have a child has been linked to reduced life satisfaction (Kiesswetter et al., 2020; Alamin et al., 2020; McQuillan et al., 2022) – one of the key components measured by the intrapersonal domain of the IRSS-BP. The absence of significant associations between PAR and interpersonal infertility stress is coherent with the notion that the stressful implications of an unachieved desire for parenthood primarily affect one's sense of self, rather than social life (Casu and Gremigni, 2016; Casu et al., 2021; Swanson and Braverman, 2021). The significant negative correlation between COM and infertility-related stress in the interpersonal domain, observed only in women, suggests that women who are better able to communicate about their infertility challenges experience less infertility-specific

TABLE 5 Associations between infertility-related characteristics and ICQ scores.

Characteristics	Males		Females	
	PAR	COM	PAR	COM
Infertility type	$F(1,348) = 1.10$	$F(1,348) = 2.18$	$F(1,348) = 0.11$	$F(1,348) = 3.91$*
Primary	18.88 (6.12)	13.01 (4.70)	19.31 (5.39)	12.97 (4.72)
Secondary	19.63 (5.98)	13.91 (6.30)	19.10 (5.76)	14.06 (4.66)
History of miscarriage	$F(1,348) = 0.77$	$F(1,348) = 0.61$	$F(1,348) = 4.69$*	$F(1,348) = 0.01$
No	18.99 (6.16)	13.36 (5.29)	19.51 (5.46)	13.30 (4.77)
Yes	19.80 (5.64)	12.74 (4.86)	17.70 (5.46)	13.22 (4.49)
Infertility diagnosis	$F(3,346) = 0.39$	$F(3,346) = 0.55$	$F(3,346) = 1.44$	$F(3,346) = 2.50$
Male	18.82 (6.63)	13.33 (4.73)	19.00 (4.93)	12.25 (4.85)
Female	19.66 (6.18)	13.84 (6.21)	19.73 (5.98)	13.16 (4.50)
Both male and female	19.24 (5.91)	12.85 (4.79)	18.07 (5.17)	14.14 (5.02)
Unexplained	18.84 (5.70)	13.03 (5.04)	19.68 (5.65)	13.75 (4.53)
Infertility duration	$F(3,346) = 1.59$	$F(3,346) = 0.67$	$F(3,346) = 0.34$	$F(3,346) = 0.90$
1–2 years	18.95 (6.19)	12.84 (5.56)	19.42 (5.29)	12.88 (4.55)
3–4 years	18.31 (6.18)	13.22 (5.18)	18.3 (5.71)	13.61 (4.77)
5–6 years	20.08 (6.17)	13.85 (5.10)	19.39 (5.71)	13.05 (4.87)
>6 years	19.98 (5.10)	13.68 (4.58)	19.63 (5.20)	14.05 (4.85)
Current ART treatment	$F(1,348) = 3.04$	$F(1,348) = 3.28$	$F(1,348) = 1.53$	$F(1,348) = 0.10$
IUI	18.25 (6.27)	12.51 (5.04)	18.70 (5.52)	13.17 (4.72)
IVF	19.48 (5.97)	13.61 (5.29)	19.49 (5.47)	13.34 (4.73)
Previous ART failures	$F(1,348) = 0.57$	$F(1,348) = 1.17$	$F(1,348) = 8.91$**	$F(1,348) = 0.06$
No	19.36 (5.81)	12.96 (5.81)	20.16 (5.21)	13.22 (4.73)
Yes	18.87 (4.64)	13.56 (4.64)	18.42 (5.62)	13.35 (4.73)

ICQ, Infertility Concern Questionnaire; PAR, parenthood; COM, communication; IUI, intrauterine insemination; IVF, in vitro fertilization; score range was 4–28 for PAR and 4–21 for COM.
*$p < 0.05$, **$p < 0.01$.

stress in their relationships with family, friends, and others. For women, effective communication may help alleviate feelings of isolation, leading to stronger social support and less strain in their interpersonal interactions. In contrast, women who struggle to communicate about infertility may experience higher stress due to unresolved tensions or lack of support in these relationships. This highlights the key role communication plays in buffering the social stress women face when dealing with infertility, as also reported in previous studies (Daniluk, 1988; Chamorro et al., 2024; Kielek-Rataj et al., 2020; Wang et al., 2024). The lack of a significant association between COM and interpersonal infertility-related stress in men can be understood in light of cultural stereotypes and gender role expectations. These norms tend to place greater emphasis on motherhood and childbearing as central to women's social identity, while men are culturally less encouraged to discuss infertility in emotional or relational terms (Taebi et al., 2021; Ying et al., 2015; Greil et al., 2010). As a result, men may not rely on communication for emotional support in the same way women do, and their communication about infertility may not significantly impact their interpersonal relationships or reduce social stress. Therefore, the COM domain, which focuses on emotional connection and support, may not serve as a strong buffer against interpersonal stress for men. These findings underscore the ICQ's potential to guide the development of gender-sensitive interventions, tailoring psychological support strategies to the distinct communication and emotional needs of men and women.

Analyses of associations with infertility-related characteristics revealed significant associations between ICQ scores and infertility type, history of miscarriage, and previous failed cycles among women. Women with secondary infertility reported slightly higher COM scores than those with primary infertility, suggesting that they may be more inclined to use communication as a coping mechanism. Having experienced pregnancy or childbirth before, women with secondary infertility are likely more familiar with the emotional toll of reproductive challenges and may have developed stronger communication skills with their partners and social networks. This allow them to handle the emotional difficulties of infertility through open dialog, seeking emotional support and connecting with others who face similar struggles. In contrast, women with primary infertility, who have never conceived, may feel a deeper sense of stigma and societal pressure regarding motherhood, making it harder for them to seek emotional support through communication, which can lead to greater social isolation and alienation, as previously reported (Taebi et al., 2021; Raque-Bogdan and Hoffman, 2015). Female participants with previous miscarriages or failed ART cycles reported slightly lower

PAR scores than those without such experiences. This finding is consistent with existing literature on psychological adaptation to infertility. As individuals repeatedly face reproductive losses, such as treatment failures or miscarriages, they may begin to revise their aspirations for parenthood, adjusting their desires in response to the growing perception that achieving their goal may be unlikely (Gray et al., 2013; Sousa-Leite et al., 2022; Boivin et al., 1995).

The ICQ has potential for both research and clinical applications. In research settings, it can serve as a valuable tool for studying the psychological and relational aspects of infertility, particularly in exploring how parenthood aspirations and communication around infertility interact. It can also contribute to evaluating interventions aimed at improving the personal and social adjustment of infertile couples by tracking changes in these dimensions over time. Moreover, researchers can use the ICQ to compare different populations and identify key factors influencing emotional resilience in infertility contexts.

In clinical practice, the ICQ can be integrated into routine assessments to provide a comprehensive understanding of infertile couples' emotional and communicative concerns. Its practical integration into routine assessments could occur during the initial stages of a couple's engagement with assisted reproduction services, allowing clinicians to quickly gather valuable insights and tailor support accordingly. This allows healthcare providers to tailor interventions more effectively, addressing specific needs and improving patient outcomes. The use of the ICQ does not require specialized training; any healthcare provider familiar with diagnostic and psychometric scales would be sufficiently equipped to administer it. Furthermore, its scoring system is intuitive and facilitates the interpretation of the results obtained.

By identifying distinct dimensions of parenthood desire and infertility-related communication skills, the ICQ can help counselors and therapists develop personalized counseling strategies, enhancing the emotional support offered to couples as they navigate the psychological challenges of infertility. The ICQ also highlights the critical role of communication in managing infertility. Healthcare providers can use insights from the questionnaire to design targeted communication training programs that foster better mutual support and understanding between partners. Furthermore, it can be used in support groups, identifying common concerns and communication challenges among participants. Group leaders can then structure discussions and activities that address these issues in meaningful ways.

Finally, the ICQ can inform the development of policies and programs that address the psychosocial aspects of infertility. This can lead to more holistic reproductive health services that ensure emotional and communicative needs are adequately met.

In summary, the ICQ can be a versatile tool in research, clinical practice, group counseling, and policy development, helping healthcare providers support couples by addressing both their desire for parenthood and the vital role of communication in their infertility journey.

4.1 Limitations and future directions

While this study provides valuable preliminary insights into the psychometric functioning and usability of the ICQ, several limitations should be acknowledged. First, Brazil is characterized by substantial cultural and socioeconomic diversity across its regions, which may limit the generalizability of the findings. Although the use of a single measure of infertility-related stress provided useful evidence for construct validity, it may not fully capture the multifaceted nature of infertility-related experiences. Future research should incorporate a broader range of validated instruments—such as those assessing mental health, self-efficacy, and perceived social support from both partners and the wider social network—to strengthen the robustness and comprehensiveness of the validation process. In particular, construct validity for the ICQ could be further examined in future studies through the use of tools such as the Decision Regret Scale (Xu et al., 2020), given the significant impact that assisted reproductive treatment decisions can have on patients' lives.

Additionally, more evidence is needed in relation to variables that are especially relevant to the experience of male partners, such as perceived social stigma, self-image, and beliefs about masculinity (Hanna and Gough, 2020; Hughes et al., 2024). Test–retest reliability was not assessed in this study, representing a limitation in evaluating the instrument's stability over time. Longitudinal studies could address this gap by administering the ICQ at multiple stages of the ART process, thus assessing both its temporal stability and its sensitivity to clinically meaningful changes (e.g., treatment failures, pregnancy, or treatment pauses). Future research should also explore how the ICQ relates to other psychological measures and sociodemographic or clinical characteristics, potentially enhancing its psychometric robustness and clinical applicability. Moreover, future investigations could examine the specific role of gender in communication as a coping strategy for infertility—potentially through the integration of qualitative interviews or evaluations of couple communication quality.

Finally, the sample was recruited from a single reproductive health center, which may restrict the external validity of the results. Expanding future studies to include participants from multiple clinics and geographically diverse regions would enhance the generalizability and applicability of the findings to broader populations.

5 Conclusion

The findings of this study provide initial evidence of the validity and reliability of the ICQ as a tool for rapidly assessing the desire for parenthood and communication about infertility, both within and outside the couple. Notably, the dimensionality and measurement properties of the ICQ were invariant across male and female partners of infertile couples. With support for (partial) scalar and strict invariance, Brazilian researchers and clinicians can confidently use the ICQ to explore potential differences in desire for parenthood and infertility-related communication issues between male and female members of infertile couples, assured that any observed differences reflect true, conceptually meaningful variations rather than measurement biases. Future research should further investigate whether dyadic invariance holds across different types of infertility (i.e., primary vs. secondary) and ART treatments (e.g., intrauterine insemination or *in vitro* fertilization), to enhance our understanding of how infertility concerns manifest in different clinical contexts. Additionally, more studies are required to gather further evidence of the ICQ's validity based on relations to relevant variables and to assess its temporal stability and sensitivity to change.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found: https://doi.org/10.17605/OSF.IO/2QYPJ.

Ethics statement

The studies involving humans were approved by Ethics Committee of the Centro Universitario FMABC (approval number 88192218.3.0 000.0082/2.675.077). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

VZ: Conceptualization, Data curation, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing. PG: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. GC: Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) under Grant 2019/17853-2.

Acknowledgments

We thank all participants and Caio Parente Barbosa for their availability and for making this study possible.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Abbey, A., Halman, L. J., and Andrews, F. M. (1992). Psychosocial, treatment, and demographic predictors of the stress associated with infertility. *Fertil. Steril.* 57, 122–128. doi: 10.1016/S0015-0282(16)54787-6

Alamin, S., Allahyari, T., Ghorbani, B., Sadeghitabar, A., and Karami, M. T. (2020). Failure in identity building as the main challenge of infertility: a qualitative study. *J. Reprod. Infertil.* 21, 49–58.

Assaysh-Oberg, S., Borneskog, C., and Ternstrom, E. (2023). Women's experience of infertility and treatment – a silent grief and failed care and support. *Sex. Reprod. Healthc.* 37:100879. doi: 10.1016/j.srhc.2023.100879

Baiocco, R., and Laghi, F. (2013). Sexual orientation and the desires and intentions to become parents. *J. Fam. Stud.* 19:9. doi: 10.5172/jfs.2013.19.1.90

Batz, F., Lermer, E., Lech, S., O'Malley, G., Zati Zehni, A., Zenz-Spitzweg, D., et al. (2023). Publisher correction: the psychological burden of COVID-19 on the desire for parenthood in minoritized sexual identities: a study on depressive symptoms and family planning in Germany. *BMC Public Health* 23:609. doi: 10.1186/s12889-023-15385-5

Beaujouan, E. (2020). Latest-late fertility? Decline and resurgence of late parenthood across the low-fertility countries. *Popul. Dev. Rev.* 46, 219–247. doi: 10.1111/padr.12334

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571. doi: 10.1001/archpsyc.1961.01710120031004

Bernstein, J., Potts, N., and Mattox, J. H. (1985). Assessment of psychological dysfunction associated with infertility. *J. Obstetric Gynecol. Neonatal. Nurs.* 14, S63–S66. doi: 10.1111/j.1552-6909.1985.tb02803.x

Biggs, S. N., Halliday, J., and Hammarberg, K. (2024). Psychological consequences of a diagnosis of infertility in men: a systematic analysis. *Asian J. Androl.* 26, 10–19. doi: 10.4103/aja202334

Boivin, J., Bunting, L., Collins, J. A., and Nygren, K. G. (2007). International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum. Reprod.* 22, 1506–1512. doi: 10.1093/humrep/dem046

Boivin, J., Oguz, M., Duong, M., Cooper, O., Filipenko, D., Markert, M., et al. (2023). Emotional reactions to infertility diagnosis: thematic and natural language processing analyses of the 1000 dreams survey. *Reprod. Biomed. Online* 46, 399–409. doi: 10.1016/j.rbmo.2022.08.107

Boivin, J., Takefman, J., and Braverman, A. (2011). The fertility quality of life (Ferti QoL) tool: development and general psychometric properties. *Hum. Reprod.* 26, 2084–2091. doi: 10.1093/humrep/der171

Boivin, J., Takefman, J. E., Tulandi, T., and Brender, W. (1995). Reactions to infertility based on extent of treatment failure. *Fertil. Steril.* 63, 801–807. doi: 10.1016/S0015-0282(16)57485-8

Bose, S., Roy, B., and Umesh, S. (2021). Marital duration, and fertility-related stress as predictors of quality of life: gender differences among primary infertile couples. *J. Hum. Reproduct. Sci.* 14, 184–190. doi: 10.4103/jhrs.jhrs_233_20

Braverman, A. M., Davoudian, T., Levin, I. K., Bocage, A., and Wodoslawsky, S. (2024). Depression, anxiety, quality of life, and infertility: a global lens on the last decade of research. *Fertil. Steril.* 121, 379–383. doi: 10.1016/j.fertnstert.2024.01.013

Caddy, C., Temple-Smith, M., and Coombe, J. (2023). Who does what? Reproductive responsibilities between heterosexual partners. *Cult. Health Sex.* 25, 1640–1658. doi: 10.1080/13691058.2023.2173800

Canada, A. L., and Schover, L. R. (2012). The psychosocial impact of interrupted childbearing in long-term female cancer survivors. *Psycho-Oncology* 21, 134–143. doi: 10.1002/pon.1875

Casu, G., and Gremigni, P. (2016). Screening for infertility-related stress at the time of initial infertility consultation: psychometric properties of a brief measure. *J. Adv. Nurs.* 72, 693–706. doi: 10.1111/jan.12830

Casu, G., Zaia, V., Montagna, E., de Padua Serafim, A., Bianco, B., Barbosa, C. P., et al. (2021). The infertility-related stress scale: validation of a Brazilian-Portuguese version and measurement invariance across Brazil and Italy. *Front. Psychol.* 12:784222. doi: 10.3389/fpsyg.2021.784222

Cetre, S., Clark, A. E., and Senik, C. (2016). Happy people have children: choice and self-selection into parenthood. *Eur. J. Popul.* 32, 445–473. doi: 10.1007/s10680-016-9389-x

Chamorro, P. P., Herruzo, J., Pino, M. J., and Casas-Rosal, J. C. (2024). Coping, social support and medical factors on psychosocial impact in couples experiencing infertility. *J. Sex Marital Ther.* 50, 197–215. doi: 10.1080/0092623X.2023.2269983

Chu, X., Geng, Y., Zhang, R., and Guo, W. (2021). Perceived social support and life satisfaction in infertile women undergoing treatment: a moderated mediation model. *Front. Psychol.* 12:651612. doi: 10.3389/fpsyg.2021.651612

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*. 14, 464–504. doi: 10.1080/10705510701301834

Claxton, S. E., DeLuca, H. K., and van Dulmen, M. H. M. (2015). Testing psychometric properties in dyadic data using confirmatory factor analysis: current practices and recommendations. *Test. Psychom. Methodol. Appl. Psychol.* 22:18. doi: 10.4473/TPM22.2.2

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *2nd* Edn. New York: Lawrence Erlbaum Associates.

Daniluk, J. C. (1988). Infertility: intrapersonal and interpersonal impact. *Fertil. Steril.* 49, 982–990. doi: 10.1016/S0015-0282(16)59948-8

Gadermann, A. M., Guhn, M., and Zumbo, B. D. (2012). Estimating or estimating ordinal r dinal reliability for like liability for liker ert-type and or t-type and ordinal item dinal itemrresponse data: a conceptual, empirical, and response data: a conceptual, empirical, and practical guide. *Pract. Assess. Res. Evaluatio.* 17:13. doi: 10.7275/n560-j767

Gameiro, S., Boivin, J., Dancet, E., de Klerk, C., Emery, M., Lewis-Jones, C., et al. (2015). ESHRE guideline: routine psychosocial care in infertility and medically assisted reproduction-a guide for fertility staff. *Hum. Reprod.* 30, 2476–2485. doi: 10.1093/humrep/dev177

Gameiro, S., and Finnigan, A. (2017). Long-term adjustment to unmet parenthood goals following ART: a systematic review and meta-analysis. *Hum. Reprod. Update* 23, 322–337. doi: 10.1093/humupd/dmx001

Gray, E., Evans, A., and Reimondos, A. (2013). Childbearing desires of childless men and women: when are goals adjusted? *Adv. Life Course Res.* 18, 141–149. doi: 10.1016/j.alcr.2012.09.003

Greil, A. L., Slauson-Blevins, K., and McQuillan, J. (2010). The experience of infertility: a review of recent literature. *Sociol. Health Illn.* 32, 140–162. doi: 10.1111/j.1467-9566.2009.01213.x

Hammarberg, K., Collins, V., Holden, C., Young, K., and McLachlan, R. (2017). Men's knowledge, attitudes and behaviours relating to fertility. *Hum. Reprod. Update* 23, 458–480. doi: 10.1093/humupd/dmx005

Hanna, E., and Gough, B. (2020). The social construction of male infertility: a qualitative questionnaire study of men with a male factor infertility diagnosis. *Sociol. Health Illn.* 42, 465–480. doi: 10.1111/1467-9566.13038

Harkness, S., and Super, C. M. (2002). Culture and parenting. Handbook of parenting: Biology and ecology of parenting, vol. *2. 2nd* Edn. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 253–280.

Ho, T. T. T., Le, M. T., Truong, Q. V., Nguyen, V. Q. H., and Cao, N. T. (2020). Psychological burden in couples with infertility and its association with sexual dysfunction. *Sex. Disabil.* 38:10. doi: 10.1007/s11195-019-09612-4

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Hughes, S., Oswald, F., and Pedersen, C. L. (2024). The virility-fertility tradeoff: effects of fatherhood on (precarious) masculinity, sexual esteem, and sexual depression. *Psychol. Sex.* 16:6863. doi: 10.1080/19419899.2024.2366863

Jain, M., and Singh, M. (2024). Assisted reproductive technology (ART) techniques. Treasure Island (FL): Stat Pearls.

Jansen, C., Kuhlmann, E., Scharli, P., Schick, M., Ditzen, B., Langer, L., et al. (2022). 'A sorrow shared …': a qualitative content analysis of what couples with recurrent miscarriages expect from one another and their families and friends. *Hum Reprod Open.* 2022:hoac032. doi: 10.1093/hropen/hoac032

Järvinen, T. (1998). The family, culture and educational choices. *Young* 6, 2–14. doi: 10.1177/110330889800600201

Kielek-Rataj, E., Wendolowska, A., Kalus, A., and Czyzowska, D. (2020). Openness and communication effects on relationship satisfaction in women experiencing infertility or miscarriage: a dyadic approach. *Int. J. Environ. Res. Public Health* 17:5721. doi: 10.3390/ijerph17165721

Kiesswetter, M., Marsoner, H., Luehwink, A., Fistarol, M., Mahlknecht, A., and Duschek, S. (2020). Impairments in life satisfaction in infertility: associations with perceived stress, affectivity, partnership quality, social support and the desire to have a child. *Behav. Med.* 46, 130–141. doi: 10.1080/08964289.2018.1564897

Kitchen, H., Aldhouse, N., Trigg, A., Palencia, R., and Mitchell, S. (2017). A review of patient-reported outcome measures to assess female infertility-related quality of life. *Health Qual. Life Outcomes* 15:86. doi: 10.1186/s12955-017-0666-0

Kline, R. (2016). Principles and practice of structural equation modeling. *4th* Edn. New York, NY: The Guilford Press.

Lewis, J. M. (1988). The transition to parenthood: II. Stability and change in marital structure. *Fam. Process* 27, 273–283. doi: 10.1111/j.1545-5300.1988.00273.x

Little, T. D., Slegers, D. W., and Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Struct. Equ. Modeling* 13, 59–72. doi: 10.1207/s15328007sem1301_3

Luca, G., Parrettini, S., Sansone, A., Calafiore, R., and Jannini, E. A. (2021). The Inferto-sex syndrome (ISS): sexual dysfunction in fertility care setting and assisted reproduction. *J. Endocrinol. Investig.* 44, 2071–2102. doi: 10.1007/s40618-021-01581-w

Massarotti, C., Gentile, G., Ferreccio, C., Scaruffi, P., Remorgida, V., and Anserini, P. (2019). Impact of infertility and infertility treatments on quality of life and levels of anxiety and depression in women undergoing in vitro fertilization. *Gynecol. Endocrinol.* 35, 485–489. doi: 10.1080/09513590.2018.1540575

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychol. Methods* 23, 412–433. doi: 10.1037/met0000144

McQuillan, J., Passet-Wittig, J., Greil, A. L., and Bujard, M. (2022). Is perceived inability to procreate associated with life satisfaction? Evidence from a German panel study. *Reprod Biomed Soc Online.* 14, 87–100. doi: 10.1016/j.rbms.2021.09.004

Meijering, J. V., Kampen, J. K., and Tobi, H. (2013). Quantifying the development of agreement among experts in Delphi studies. *Technol. Forecast. Soc. Change* 80, 1607–1614. doi: 10.1016/j.techfore.2013.01.003

Miller, W. B. (1995). Childbearing motivation and its measurement. *J. Biosoc. Sci.* 27, 473–487. doi: 10.1017/S0021932000023087

Mills, M. (2010). Gender roles, gender (in)equality and fertility: an empirical test of five gender equity indices. *Can. Stud. Popul.* 37:30. doi: 10.25336/P6131Q

Millsap, R. E., and Meredith, W. (2007). "Factorial invariance: historical perspectives and new problems" in Factor analysis at 100: Historical developments and future directions. eds. R. Cudeck and R. C. Mac Callum. *1st* ed (London: Lawrence Erlbaum Associates Publishers), 131–152.

Miscioscia, M., Blavier, A., Pagone, P. R., and Simonelli, A. (2017). The desire of parenthood: intuitive co-parental behaviors and quality of couple relationship among Italian and Belgian same-sex and opposite-sex couples. *Front. Psychol.* 8:110. doi: 10.3389/fpsyg.2017.00110

Newton, C. R., Sherrard, W., and Glavac, I. (1999). The fertility problem inventory: measuring perceived infertility-related stress. *Fertil. Steril.* 72, 54–62. doi: 10.1016/S0015-0282(99)00164-8

Nik Hazlina, N. H., Norhayati, M. N., Shaiful Bahari, I., and Nik Muhammad Arif, N. A. (2022). Worldwide prevalence, risk factors and psychological impact of infertility among women: a systematic review and meta-analysis. *BMJ Open* 12:e057132. doi: 10.1136/bmjopen-2021-057132

Nitsche, N., and Hayford, S. R. (2020). Preferences, partners, and parenthood: linking early fertility desires, marriage timing, and achieved fertility. *Demography* 57, 1975–2001. doi: 10.1007/s13524-020-00927-y

Nunnally, J. C., and Bernstein, I. H. (1994). Psychometric theory. *3rd* Edn. New York: McGraw-Hill, Inc.

O'Connell, S. B. L., Gelgoot, E. N., Grunberg, P. H., Schinazi, J., Da Costa, D., Dennis, C. L., et al. (2021). 'I felt less alone knowing i could contribute to the forum': psychological distress and use of an online infertility peer support forum. *Health Psychol. Behav. Med.* 9, 128–148. doi: 10.1080/21642850.2021.1884556

Omani-Samani, R., Maroufizadeh, S., Almasi-Hashiani, A., Sepidarkish, M., and Amini, P. (2019). The WHO-5 well-being index: a validation study in people with infertility. *Iran. J. Public Health* 48, 2058–2064.

Ozturk, R., Herbell, K., Morton, J., and Bloom, T. (2021). "the worst time of my life": treatment-related stress and unmet needs of women living with infertility. *J. Community Psychol.* 49, 1121–1133. doi: 10.1002/jcop.22527

Pedro, J., Sobral, M. P., Mesquita-Guimaraes, J., Leal, C., Costa, M. E., and Martins, M. V. (2017). Couples' discontinuation of fertility treatments: a longitudinal study on demographic, biomedical, and psychosocial risk factors. *J. Assist. Reprod. Genet.* 34, 217–224. doi: 10.1007/s10815-016-0844-8

Präg, P., Mills, M., Tanturri, M., Monden, C., and Pison, G. (2017). The demographic consequences of assisted reproductive technologies. Oxford: Hal Open Science.

Raque-Bogdan, T. L., and Hoffman, M. A. (2015). The relationship among infertility, self-compassion, and well-being for women with primary or secondary infertility. *Psychol. Women Q.* 39, 484–496. doi: 10.1177/0361684315576208

Ribeiro, A. C. (2007) Adaptação do inventário de problemas de fertilidade para homens e mulheres inférteis [Doctoral]: Universidade de São Paulo;

Ryu, E. (2011). Effects of skewness and kurtosis on normal-theory based maximum likelihood test statistic in multilevel structural equation modeling. *Behav. Res. Methods* 43, 1066–1074. doi: 10.3758/s13428-011-0115-7

Schmidt, L. (2006). Infertility and assisted reproduction in Denmark. Epidemiology and psychosocial consequences. *Dan. Med. Bull.* 53, 390–417.

Schmidt, L., Holstein, B. E., Christensen, U., and Boivin, J. (2005). Communication and coping as predictors of fertility problem stress: cohort study of 816 participants who did not achieve a delivery after 12 months of fertility treatment. *Hum. Reprod.* 20, 3248–3256. doi: 10.1093/humrep/dei193

Sharma, A., and Shrivastava, D. (2022). Psychological problems related to infertility. *Cureus* 14:e30320. doi: 10.7759/cureus.30320

Shen, Q., Wang, B., He, T., Li, S., Peng, E., and Lei, J. (2024). Factors associated with discontinuation in fertility treatment: a systematic scoping review. *J. Assist. Reprod. Genet.* 41, 409–421. doi: 10.1007/s10815-023-02982-x

Shreffler, K. M., Gallus, K. L., Peterson, B., and Greil, A. L. (2020). "Couples and infertility" in *The Handbook of Systemic Family Therapy*. eds. K. Wampler, R. Miller, R. Seedall, L. McWey, A. Blow, M. Rastogi et al. (Hoboken: John Wiley and Sons) 385–406.

Sormunen, T., Aanesen, A., Fossum, B., Karlgren, K., and Westerbotn, M. (2018). Infertility-related communication and coping strategies among women affected by primary or secondary infertility. *J. Clin. Nurs.* 27, e335–e344. doi: 10.1111/jocn.13953

Sousa-Leite, M., Fernandes, M., Reis, S., Costa, R., Figueiredo, B., and Gameiro, S. (2022). Feasibility and acceptability of psychosocial care for unsuccessful fertility treatment. *Health Expect.* 25, 2902–2913. doi: 10.1111/hex.13598

Starc, A., Trampus, M., Pavan Jukic, D., Rotim, C., Jukic, T., and Polona, M. A. (2019). Infertility and sexual dysfunctions: a systematic literature review. *Acta Clin. Croat.* 58, 508–515. doi: 10.20471/acc.2019.58.03.15

Steenkamp, J. B. E. M., and Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *J. Consum. Res.* 25, 78–90.

Swanson, A., and Braverman, A. M. (2021). Psychological components of infertility. *Fam. Court. Rev.* 59, 67–82. doi: 10.1111/fcre.12552

Sylvest, R., Koert, E., Birch Petersen, K., Malling, G. M. H., Hald, F., Nyboe Andersen, A., et al. (2018). Attitudes towards family formation among men attending fertility counselling. *Reprod Biomed Soc Online*. 6, 1–9. doi: 10.1016/j.rbms.2018.06.001

Taebi, M., Kariman, N., Montazeri, A., and Alavi Majd, H. (2021). Infertility stigma: a qualitative study on feelings and experiences of infertile women. *Int. J. Fertil. Steril.* 15, 189–196. doi: 10.22074/IJFS.2021.139093.1039

Tang, N., Jia, Y., Zhao, Q., Liu, H., Li, J., Zhang, H., et al. (2022). Influencing factors of dyadic coping among infertile women: a path analysis. *Front. Psych.* 13:830039. doi: 10.3389/fpsyt.2022.830039

Tavousi, S. A., Behjati, M., Milajerdi, A., and Mohammadi, A. H. (2022). Psychological assessment in infertility: a systematic review and meta-analysis. *Front. Psychol.* 13:961722. doi: 10.3389/fpsyg.2022.961722

van Balen, F., and Trimbos-Kemper, T. C. (1995). Involuntarily childless couples: their desire to have children and their motives. *J. Psychosom. Obstet. Gynaecol.* 16, 137–144. doi: 10.3109/01674829509024462

Vignoli, D., Minello, A., Bazzani, G., Matera, C., and Rapallini, C. (2022). Narratives of the future affect fertility: evidence from a laboratory experiment. *Eur. J. Popul.* 38, 93–124. doi: 10.1007/s10680-021-09602-3

Wang, Q., Jia, D., Gao, Y., Zhou, M., Zhao, X., Qin, R., et al. (2024). Relationship between stigma and infertility-related stress among couples undergoing AID: the mediating role of communication patterns. *Stress. Health* 40:e3412. doi: 10.1002/smi.3412

WHO (1995). The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. *Soc. Sci. Med.* 41, 1403–1409.

WHO, W. H. O. (2006). Reproductive health indicators: guidelines for their generation, interpretation and analysis for global monitoring (WHO, Ed.). Geneve: WHO Press.

Xie, Y., Ren, Y., Niu, C., Zheng, Y., Yu, P., and Li, L. (2022). The impact of stigma on mental health and quality of life of infertile women: a systematic review. *Front. Psychol.* 13:1093459. doi: 10.3389/fpsyg.2022.1093459

Xu, R. H., Zhou, L. M., Wong, E. L., Wang, D., and Chang, J. H. (2020). Psychometric evaluation of the Chinese version of the decision regret scale. *Front. Psychol.* 11:583574. doi: 10.3389/fpsyg.2020.583574

Ying, L. Y., Wu, L. H., and Loke, A. Y. (2015). Gender differences in experiences with and adjustments to infertility: a literature review. *Int. J. Nurs. Stud.* 52, 1640–1652. doi: 10.1016/j.ijnurstu.2015.05.004

Zurlo, M. C., Cattaneo Della Volta, M. F., and Vallone, F. (2023). Paths towards parenthood after repeated treatment failures: a comparative study on predictors of psychological health outcomes in infertile couples persisting in treatments or opting for adoption. *Front. Psychol.* 14:1147926. doi: 10.3389/fpsyg.2023.1147926