Multi-National Evaluation of the Measurement Invariance of the Level of Personality

Functioning Scale – Brief Form 2.0: Comparison of Student and Community Samples Across

Seven Countries

Adam P. Natoli, PhD
Department of Psychology & Philosophy, Sam Houston State University
https://orcid.org/0000-0002-2324-9242
Email: apn017@shsu.edu

Bo Bach, PhD*
Center for Personality Disorder Research, Psychiatric Research Unit, Region Zealand, Slagelse, Denmark
Email: bobachsayad@gmail.com

Alex Behn, PhD*
School of Psychology, Pontificia Universidad Católica de Chile
Millennium Institute for Research in Depression and Personality (MIDAP)
Email: albehn@uc.cl

Marianne Cottin, PhD*
School of Psychology, Universidad Finis Terrae
Millennium Institute for Research in Depression and Personality (MIDAP)
https://orcid.org/0000-0002-9460-7190
Email: mcottin@uft.cl

Emanuela S. Gritti, PhD*
Department of Developmental Psychology and Socialization, University of Padova
Department of Psychology, University of Milano-Bicocca
https://orcid.org/0000-0001-6109-2699
Email: emanuela.gritti@unumib.it

Joost Hutsebaut, PhD*
Viersprong Institute for Studies on Personality Disorders, Halsteren
Tilburg University
https://orcid.org/0000-0003-0189-377
Email: joost.hutsebaut@deviersprong.nl

Nishtha Lamba, PhD*
Department of Psychology, Middlesex University Dubai
Email: n.lamba@mdx.ac.ae

Yann Le Corff, PhD*
Groupe de recherche et d'intervention sur les adaptations sociales de l'enfance
Département d'orientation professionnelle, Université de Sherbrooke
https://orcid.org/0000-0002-0483-2969
Email: yann.le.corff@usherbrooke.ca

Johannes Zimmermann, PhD*
Department of Psychology
University of Kassel, Kassel, Germany
https://orcid.org/0000-0001-6975-2356
Email: jz@uni-kassel.de

Mélanie Lapalme, PhD
Groupe de recherche et d'intervention sur les adaptations sociales de l'enfance
Département de psychoéducation, Université de Sherbrooke
https://orcid.org/0000-0001-9862-6033
Email: melanie.lapalme@usherbrooke.ca

*These authors, listed alphabetically, contributed equally to the manuscript, and should be considered as second authors.

Abstract

DSM-5's Level of Personality Functioning Scale (LPFS) was introduced as a dimensional rating of impairments in self- and interpersonal functioning, and the LPFS – Brief Form (LPFS-BF) was the first published corresponding self-report. The updated LPFS-BF 2.0 has been translated into several languages and international research supports many of the instrument's psychometric properties; however, its measurement invariance has only been evaluated across a few countries. This study expands previous studies as an introductory step in a global evaluation of the LPFS-BF 2.0's measurement invariance. Archival data (N = 5,618, 57% female) from seven countries (Canada, Chile, Denmark, Germany, Italy, United Arab Emirates, United States of America) were used for this study. Participants were recruited from both community (n = 4,677) and student (n = 941) populations. After confirming adequate model fit separately in the community and student samples, we evaluated a series of increasingly stringent model comparisons to test three aspects of measurement invariance (configural, metric, scalar) and then examined latent mean differences across countries. Full scalar invariance was supported in the community sample and partial scalar invariance was supported in the student sample. Evaluation of latent mean differences revealed multiple significant differences. Overall, the LPFS-BF 2.0 appears to assess self- and interpersonal functioning impairment similarly across the included countries. Findings are discussed through the lenses of the cultures from which participants were recruited, as well as in the context of alternative explanations. Limitations, plans for future research, and implications for both research and clinical practice are offered.

Keywords: Personality Functioning; Alternative Model for Personality Disorders; Culture; International; Measurement Invariance.

Public Significance Statement: The LPFS-BF 2.0 is a brief self-report measure of personality functioning that appears to assess self- and interpersonal functioning impairment similarly across many different countries, which offers encouraging evidence supporting its international use.

Multi-National Evaluation of the Measurement Invariance of the Level of Personality Functioning Scale – Brief Form 2.0: Comparison of Student and Community Samples Across Seven Countries

Dimensional models of personality disorders (PDs) have gained increased attention and demonstrate superiority over categorical models in both validity (Zimmermann et al., 2019) and clinical utility (Bornstein & Natoli, 2019). To this end, current classification systems of PDs have refined their diagnostic frameworks to introduce a dimensional perspective, such as the shifts seen in the chapter on PD and related traits in ICD-11 (WHO, 2018) and in the DSM-5's Alternative Model for Personality Disorders (AMPD; APA, 2013). The dimensional models in both of these classification systems emphasize the importance of assessing the severity of personality pathology based upon a concept of impairments in self- and interpersonal functioning. Thus, the DSM-5 presented the Level of Personality Functioning Scale (LPFS) to help assess impairment severity in a dimensional manner. The LPFS consists of 60 descriptors of severity spanning five levels, for 12 facets of self- and interpersonal impairment, and was originally intended to perform as a single, expert rated scale (Morey et al., 2013). Many instruments designed to measure DSM-5 personality functioning impairment have since been introduced (Birkhölzer et al., 2021). As summarized in recent reviews (e.g., Zimmermann et al., 2019; 2022), research with these instruments indicates that the structure of the 12 facets may be best described by two highly correlated factors of self- and interpersonal functioning, consistent with a strong general factor of severity. Research also demonstrates that DSM-5 personality functioning can be assessed with acceptable reliability in interviews and is strongly related to other clinical indicators of severity. Other aspects of the LPFS's validity are more controversial: For example, empirical investigations have produced mixed findings on whether the information provided by these measures is distinct or redundant with that available from measures of

maladaptive personality traits, fueling debates on the value of the LPFS (e.g., Widiger et al., 2019). In this study, we focus on an aspect of the LPFS that is still under-researched, namely measurement invariance across different cultural groups.

The Level of Personality Functioning Scale – Brief Form

The Level of Personality Functioning Scale – Brief Form (LPFS-BF; Hutsebaut et al., 2015) was the first published self-report instrument that provided an assessment of the LPFS. The original LPFS-BF was designed as an easy tool for patients to self-assess the severity of their personality dysfunction and their potential eligibility for PD treatment. This goal was reflected by a simple and convenient yes/no response scale. The items themselves were based on the DSM-5's LPFS and aimed to capture the core psychological concept behind each of the 12 facets. Although the list was rather simple and easy, its initial psychometric properties were acceptable to good (Hutsebaut et al., 2015). The LPFS-BF revealed a clear two-factor structure, interpretable as self- and interpersonal functioning. Internal consistency was fair to good, and associations with other measures of personality pathology were as expected. Nevertheless, basic adjustments were made to the LPFS-BF to upgrade the instrument and increase its utility for outcome research (Weekers et al., 2019). A Likert scale replaced the binary response scale, ranging from 1 (completely untrue) to 4 (completely true), which follows the response categories of the Personality Inventory for DSM-5 (PID-5; APA, 2013) and improved the instrument's sensitivity to change. Three items that did not perform well in the initial LPFS-BF were changed (items 4, 6, 11). The resulting LPFS-BF 2.0 has been translated into several languages (www.deviersprong.nl/LPFS-BF) and used in numerous studies around the world (see Table 1).

Current Use

The LPFS-BF 2.0 has been administered in community (e.g., Spitzer et al., 2021; Stone et al., 2021; Zimmermann et al., 2020), clinical (e.g., Bach & Hutsebaut, 2018; Weekers et al., 2019) and forensic (e.g., Bach & Hutsebaut, 2018) samples. Findings thus far continually support the two-factor structure of the LPFS-BF 2.0, referring to impairments in self- and interpersonal functioning. Studies often find these factors to be strongly correlated, consistent with the AMPD's assumption of a general dimension of severity. Internal consistencies of the total and subscale scores have been good to excellent, all above $\alpha = .70$ across studies. Construct validity has been supported by moderate to high associations with related measures of PD severity (e.g., Spitzer et al., 2021; Weekers et al., 2019), and Bach and Hutsebaut (2018) demonstrated the LPFS-BF 2.0's incremental validity over a well-validated measure of pathological traits (and vice-versa). Finally, Weekers and colleagues re-assessed a small sample of PD patients after a 3-month intensive psychodynamically-oriented inpatient treatment. They found large effect sizes ($d = 1.05$) for the LPFS-BF 2.0 total score, demonstrating its potential use as a routine outcome monitoring instrument. These findings support the LPFS-BF 2.0 as a psychometrically sound brief measure of personality impairment severity. Combined with its increasingly global use, open-access, brevity and ease of use, inclusion of only negatively worded items[1], parallel with popular diagnostic systems, and numerous current and forthcoming translations, the LPFS-BF 2.0 stands out as a suitable candidate for becoming a worldwide measure of personality functioning. Encouragingly, the International Consortium for Health Outcomes Measurement (ICHOM) has included the LPFS-BF 2.0 in the proposed standardized minimum set of outcomes

---

[1] Use of both positively and negatively worded items are known to contribute to cross-cultural measurement noninvariance due to such items being interpreted differently across countries (e.g., Schmitt & Allik, 2005).

(Prevolnik Rupel et al., 2021), thereby endorsing its potential for clinical use and outcome research across cultural and geographical settings.

International and Cross-Cultural Considerations

Recent meta-analytic investigations report fairly high global prevalence rates for PDs (ranging from 7.8% to 12.16%; Winsper et al., 2020), and have revealed high levels of heterogeneity in prevalence rates across countries. Environmental and sociocultural factors are believed to contribute to variations in the development, expression, and maintenance of personality pathology (see Turner et al., 2020), but few large-scale global studies of these effects have been conducted (e.g., Huang et al., 2009). However, sociocultural factors have been shown to account for country-level differences in variables related to self- and interpersonal functioning, including self-construal (Markus & Kitayama, 2010), personality traits (McCrae et al., 2005), and attachment styles (van IJzendoorn & Kroonenberg, 1988). Thus, levels of personality functioning may also differ across countries, and specific sociocultural or environmental factors might underlie that variability. Research seeking to explore this question requires access to a psychometrically sound measure of personality functioning, necessitating examination of that instrument's measurement invariance across countries before cross-cultural comparisons would be appropriate. Explicitly, any instrument used for cross-national or cross-cultural comparisons must first be shown to measure the same latent variable (e.g., personality functioning) in the same way for each group being compared (e.g., countries) because a lack of measurement invariance can inhibit meaningful comparisons or lead to erroneous conclusions.

Importance of Measurement Invariance Examination Across Countries

The cross-national and cross-cultural comparability of a measure, such as the LPFS-BF 2.0, can be jeopardized in myriad ways, and three hierarchical levels of measurement invariance

are typically investigated (see Steenkamp & Baumgartner, 1998). To demonstrate measurement invariance, it is first necessary to prove that an identical factor structure can be imposed across groups (configural invariance), evidence for which suggests the instrument's basic organization of the measured latent variables are similar across countries. Regarding the LPFS-BF 2.0, invariance at the configural level would mean a similar pattern of item loadings on the instrument's two latent factors across countries: items 1-6 loading onto a self-functioning factor and items 7-12 loading onto an interpersonal functioning factor. If configural invariance is supported, it is then necessary to examine whether item loadings are equal across groups (metric invariance). Support for metric invariance indicates items and response scales maintain comparable meanings across countries, as well as equivalence in the relationships between the latent variable(s) and items observed in each group. Metric noninvariance suggests that one or more items are differentially affected by the latent variable across groups. For example, metric noninvariance for the LPFS-BF 2.0 could reveal that one or more items is more closely related to personality functioning in some countries than in others. Finally, scalar invariance is tested to evaluate the assumption that individuals from different groups with equal levels of the latent variable will similarly respond to a given item meant to tap that latent variable. Said another way, scalar invariance is tested to determine whether observed mean differences can be understood as reflecting differences in the latent variable and not differences in how the measurement instrument performs in different groups. Scalar invariance permits a valid comparison of latent means across groups; however, meaningful and valid cross-group comparisons are still possible when full scalar invariance is not supported. In such cases, there may be support for partial scalar invariance wherein at least a subset of test items (typically at least two items per latent variable; Byrne et al., 1989; Steenkamp & Baumgartner, 1998) are

found to function equivalently across groups. When this occurs, useful information on group differences may be extractable and multiple strategies are available to handle partial measurement noninvariance (see Davidov et al., 2014; Marsh et al., 2018).

At this time, no brief self-report instrument for assessing self- and interpersonal functioning (i.e., personality functioning) has demonstrated broad, multicultural measurement invariance. However, the LPFS-BF 2.0's measurement invariance across multiple languages has been evidenced. Specifically, a recent study (Le Corff et al., in press) examining the Dutch, English, French, and Spanish versions of the LPFS-BF 2.0 found support for configural and metric invariance across three samples (Belgian, Canadian, and Spanish). Interestingly, scalar invariance was fully supported when using within-sample comparisons (a subsample of the Canadian respondents filled both the English and French versions), while only partial scalar invariance was observed across samples (i.e., across countries), suggesting translations to be equivalent but that cultural or sample-dependent differences may compromise the comparability of scores across countries. Accordingly, evaluation of the LPFS-BF 2.0's measurement invariance across diverse countries is essential before data stemming from these naturally diverse populations can be adequately interpreted.

The Current Study

With the above as context, it is clear that formal tests of the LPFS-BF 2.0's international measurement invariance are necessary. Thus, the current study evaluated measurement invariance of the LPFS-BF 2.0 across seven countries from four distinct regions of the world (Europe, Middle East, North America, and South America). The included countries were Canada, Chile, Denmark, Italy, Germany, the United Arab Emirates (UAE), and the United States of America (USA). Measurement invariance was tested in both community and student

samples, as these are two of the most popularly studied populations. Specifically, this study investigated whether the LPFS-BF 2.0's factor structure is equivalent across countries (configural invariance); whether test items measure the latent factors of self- and interpersonal functioning equally across countries (metric invariance); and whether individuals from different countries who obtain identical test scores actually possess equal levels of personality functioning (scalar invariance). Measurement invariance across the countries included in the current study would offer initial evidence signifying the meaning of the LPFS-BF 2.0 scores are consistent across these groups, which would support the validity of observed latent mean differences (if full or partial scalar invariance is supported) and of the interpretation of scores in these populations. Replication and expansion – by including additional countries – of the current study would be necessary before more global generalizations can be made, but initial support for measurement invariance of the LPFS-BF 2.0 obtained from the current study would advocate for such future research. Conversely, should measurement invariance not be supported, findings can guide future revisions and/or culturally specific modifications of the LPFS-BF 2.0.

Method

Participants and Procedures by Country

The current study made use of archival data to evaluate measurement invariance of the LPFS-BF 2.0. All data were originally obtained via online studies wherein participants were required to complete the LPFS-BF 2.0 and various demographic items, often accompanied by additional study-specific self-report or implicit measures. Given our interest in testing the LPFS-BF 2.0's measurement invariance with respect to general populations, the frequency student samples are used in psychological research, and due to the fact that not all of our samples recruited student participants, we elected to conduct analyses separately for community samples

and student samples. Descriptive statistics of participants for each country, by sample type, are presented in Table 2. More detailed information on sample characteristics, including statistical analyses of group differences on demographic variables and skewness and kurtosis values of individual LPFS-BF 2.0 items within each subsample, can be found in our online supplementary documents (see OSF LINK). As to be expected in multi-sample studies using archival data, differences in participant recruitment strategies and data collection methods were present across countries. The following summarizes participant recruitment and data collection in each country. This study was not preregistered.

Canadian Sample. The Canadian sample included 2,505 adults from the Province of Québec who completed the French LPFS-BF 2.0 (Le Corff et al., in press) as part of a larger study on the assessment of the AMPD (see Le Corff et al., in press, for a full description of the sample and procedure). Potential participants were recruited via Léger 360, the largest Survey firm in Canada, by sending electronic invitations to a randomized subsample of Léger 360's panel of over 200,000 residents from the Province of Québec, Canada. Data collection began on 18 June and ended on 15 July 2019 when the questionnaire was closed due to reaching the target sample size of 2,500 completed questionnaires. Participants under 18 years of age or who reported a "very poor" understanding of the French language were excluded from the study. Participants who completed the questionnaires received an incentive in the form of points added to their Léger 360 account, which can be exchanged for cash, gifts, or participation in prize drawings. The study received approval from the ethics board from the Canadian author's research institution, and all participants signed an electronic informed consent form.

Chilean Samples. The Chilean sample included 494 adults from the Santiago region of Chile who completed the Spanish LPFS-BF 2.0 (Cottin et al., in preparation) as part of a larger

study on the relationship between adverse childhood experiences and the AMPD. Potential

participants were recruited between November 2020 and February 2021 from three populations:

a university student population, a community population, and a clinical population, the latter of

which was not included in the current study. Both student and community samples were

recruited through online social media announcements and advertisements. The total sample ($N =$

494) consisted of 281 community participants and 213 student participants. All participants

signed online informed consent forms and completed questionnaires online, and all study

participants were entered into a raffle for a chance to win a gift card as an incentive for their

participation. The study received approval from the ethics board from the Chilean author's

research institution.

Danish Sample. The Danish sample included 713 adults from across Denmark who

completed the Danish LPFS-BF 2.0 (originally, Bach et al., 2016) as part of a larger data

collection project in 2020 aimed at producing normative data for outcome measures used in

mental health care. A representative age- and gender-stratified sample ($N = 2000$) was randomly

drawn from the Civil Registration System in Denmark, comprising an equal number of men and

women born in each year between 1938–2001 (i.e., aged 18-79 years). Data were collected using

individual access to an online self-report system via an emailed link. Potential participants above

the age of 70 were recruited through standard mail due to their tendency to not regularly use

email. Non-responders were contacted a second time and, if they did not participate, their

information was destroyed. Slightly more females returned completed responses. The study

received approval from the local ethics board and data protection authority. No participation

incentives were offered, and no participant received compensation for completing the survey.

German Sample. The German sample included 924 adults from across Germany who completed the German LPFS-BF 2.0 as part of a larger data collection project aimed at establishing a common metric of PD severity (see Zimmermann et al., 2020 for a full description of the sample and procedure). Potential participants were recruited via the commercial survey provider clickworker between 15-26 May 2019, with participants selected so that the distribution of gender and age roughly corresponded to the general population in Germany. All participants received an allowance of €5 for completing the survey. Of the 924 individuals who completed the survey, 75 individuals were excluded because they completed the survey too quickly, carelessly, or based on the same IP address. The final sample consisted of 849 people. The study received approval from the ethics board from the German author's research institution.

Italian Samples. The Italian sample included 411 adults who completed the Italian LPFS-BF 2.0 (Gritti et al., in preparation) as part of a larger study designed to validate the Italian version of the LPFS-BF 2.0. Potential participants were recruited using snowball (chain) sampling in two populations between December 2020 and June 2021. To recruit a community sample, the link to the online study was posted on social media pages and disseminated through word of mouth. The student sample was recruited by presenting the opportunity to psychology students in class and by having students share the survey link with other students. In both cases, participants were encouraged to complete the study and then distribute the link to others over the age of 18. The final sample ($N = 411$) consisted of 187 community participants and 224 student participants. The study received approval from the ethics board from the Italian author's research institution. No participant received credit or compensation for completing the survey.

United Arab Emirates (UAE) Samples. The UAE sample included 525 adults from across the UAE who completed the English LPFS-BF 2.0 (originally, Hutsebaut et al., 2016) as

part of a larger study investigating links between personality and perceptions of different environments. Participants were recruited using convenience and snowball (chain) sampling in two populations during October and November 2020. Both the student and community samples were recruited by posting participation requests, which included a link to an online survey, to multiple social media platforms, such as WhatsApp, Facebook, and Instagram. Of the 525 participants who successfully completed the survey, 63 participants were removed due to failed attention checks. The final sample ($N = 462$) consisted of 142 community participants and 320 student participants. Consistent with the UAE's residents, the majority of participants in both samples were born in a country other than the UAE (community = 83.1%; student = 62.2%), with India being the most common country of birth (community = 57.0%; student = 37.5%). The study received approval from an ethics board at the UAE author's research institution. No participant received credit or compensation for completing the survey.

United States of America (USA) Sample. The USA sample included 201 adult university students who completed the English LPFS-BF 2.0 (originally, Hutsebaut et al., 2016) as part of a larger study investigating links between personality, mentalization, and perfectionism. Potential participants were recruited through a study participation program at a medium-sized public university in the southern USA between April and June 2021. Due to inclusion of a performance-based measure, participants were required to complete the study using a laptop or desktop computer with a physical keyboard. Participants were notified if they attempted the study on an incompatible device (e.g., touchscreen phone) and were asked to switch to a compatible device. The study received approval from the ethics board from the USA author's research institution and all participants received course credit for completing the survey.

Data Analysis

Confirmatory factor analysis (CFA) in Mplus 8.4 (Muthén & Muthén, 1998-2019) was used to examine a single multigroup model in each of our two aggregate samples by first specifying the LPFS-BF 2.0's two-factor model in an aggregate student sample and in an aggregate community sample. We used full information maximum likelihood estimation (FIML) to deal with missing data, and robust standard errors (MLR) to account for non-normality. We then evaluated a series of increasingly stringent model comparisons to test three aspects of measurement invariance: configural, metric, and scalar. Each model was also evaluated for model fit. The configural model tests whether an identical factor structure can be adequately applied across countries within a given sample type (student or community). In this case, model fit was evaluated by examining the chi-square test statistic, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) and interpreted in accordance with recommendations proposed by Hu and Bentler (1999): acceptable model fit is indicated by CFI values > .95; RMSEA values < .06; and SRMR values < .08. These cutoff values constitute only rough guidelines and are not considered strict rules (Marsh et al., 2004). The metric invariance model constrains (unstandardized) factor loadings to be equal across groups, and invariance is supported when model fit is not notably poorer than the configural model. The scalar invariance model tests for group differences in item intercepts and is examined by comparing model fit to the metric model. When comparing models, the chi-square difference test-statistic may suggest noninvariance as a result of small changes in model fit due to being overly sensitive when sample size is large (Chen, 2007). Thus, in addition to chi-square difference tests, we examined the following fit statistics using cutoffs from Chen (2007) to evaluate noninvariance: a change in CFI ≤ -.005 in addition to a change of ≥ .010 in RMSEA or a change of ≥ .025 in SRMR indicates metric noninvariance, and a CFI

change ≤ - .005 in addition to a change of ≥ .010 in RMSEA or a change of ≥ .005 in SRMR

indicates scalar noninvariance. Partial invariance was tested in instances when the change in fit

statistics failed to support full invariance by iteratively freeing parameters that accounted for the

largest source of misfit in the given model until partial measurement invariance was obtained.

Furthermore, measurement invariance was tested between each pair of countries within a given

sample type (i.e., student or community) when noninvariance was indicated, as well as between

student and community samples within countries when possible (i.e., Chile, Italy, UAE).

Following evaluation of measurement invariance, and once partial or full measurement

invariance was evidenced, latent mean differences were investigated by calculating Cohen's *d* as

an estimate of effect size. Copies of the materials and analysis code may be received by emailing

the corresponding author.

Results

Basic Model

We first tested the basic model of the LPFS-BF 2.0 separately in our student and

community samples. A two-factor structure of the LPFS-BF 2.0 was tested in each sample by

specifying items 1-6 to load onto one factor (self-functioning) and items 7-12 to load onto a

second factor (interpersonal functioning). Remaining consistent with previous findings, an *a*

priori covariance for the two factors was specified for both samples. Except for significant chi-

square tests, which might be the product of Type I errors due to our sample sizes, all examined

fit indexes confirmed a good fit of this basic model in both samples (see Table 3).

Measurement Invariance

Our models evaluated measurement invariance based on country identification in student

and community samples. Fit statistics for all invariance tests are displayed in Table 3 (see Table

4 for standardized factor loadings and internal consistency statistics). The configural model was shown to be a good fit to the data in each sample type (i.e., student or community). A comparison of the fit indices of the metric model to the configural model revealed no meaningful decrement in fit based on recommended criteria, supporting metric invariance in both sample types. This demonstration of full metric invariance suggests LPFS-BF 2.0 items are similarly related to latent self- and interpersonal functioning across the included countries for both student and community participants. Thus, the meaning of LPFS-BF 2.0 items can be said to be comparable. Full scalar invariance was supported in the community sample; however, there was substantial decrement in fit for the scalar model compared to the metric model in the student sample (ΔCFI = -.065, ΔRMSEA = +.015, ΔSRMR = +.013). After freely estimating indicator intercepts for two and three items for Italy and Chile, respectively, partial scalar invariance was supported in the student sample (ΔCFI = -.025, ΔRMSEA = +.004, ΔSRMR = +.004). The specific items for which intercepts were freed for each country in the student sample and the accompanying fit statistics are reported in Table 3.

As only partial scalar invariance was reached within the student sample, measurement invariance between each pair of countries within the student sample was tested (complete results are reported in supplementary Table S6). Full scalar invariance was supported when comparing UAE and USA student samples, whereas partial scalar invariance was reached for all other comparisons after freeing the intercept for one to four items. Specifically, the intercepts for items 4 and 5 were freed when comparing Chile and Italy; item 11 for Chile and UAE; items 2, 6, 8, and 11 for Chile and USA; items 8, 11, and 12 for Italy and UAE; and items 7, 8, and 12 for Italy and USA. Finally, full scalar invariance between student and community samples was supported

within each of the three countries that collected data from both sample types (i.e., Chile, Italy, UAE; see Table S7).

Evaluation of latent mean differences across countries revealed a broad range of effect sizes representing the standardized difference between countries (see Tables 5 and 6). Latent mean comparisons across countries can only be confidently interpreted for the community sample due to a lack of support for full scalar invariance in the student sample. Consequently, latent mean comparisons across countries in the student sample are interpreted with caution as only partial scalar invariance was supported. Finally, the latent factors of self-functioning and interpersonal functioning were found to be significantly positively correlated overall ($r = .807$), within the student ($r = .753$) and community ($r = .821$) samples, as well as within each country ($r = .738$ to $.899$; see Table S8).

General Discussion

The measurement invariance of the LPFS-BF 2.0 was evaluated in two sample types (student and community) across a total of seven countries: Canada, Chile, Denmark, Italy, Germany, UAE, and USA. Overall, the LPFS-BF 2.0 appears to assess self- and interpersonal functioning similarly across this diverse collection of countries. Full configural invariance was demonstrated, suggesting an identical factor structure can be adequately applied across countries. Metric invariance of the LPFS-BF 2.0 was also fully supported across countries in both samples, suggesting its items are similarly related to latent self- or interpersonal functioning in each country (i.e., the meaning of items seem to be similar). Lastly, full scalar invariance was supported in community samples and partial scalar invariance was achieved in student samples. This designates very high similarity across countries in how a given item's response scale relates to latent levels of self- or interpersonal functioning in the community samples and partial

similarity in the student samples. Full scalar invariance in the community samples suggests valid score comparisons can be made and indicates that equivalent cutoff scores across the included countries are likely valid. Conversely, the partial scalar invariance found in the student samples suggests a response to a given item might not correspond to the same latent score in each country (i.e., response scales may not be entirely equivalent). Hence, there appears to be some degree of systematic variability across countries in how the response scale for certain items relates to values of latent self- or interpersonal functioning in the student samples. Although methodological differences (discussed below) cannot be ruled out as the source of this noninvariance, individual comparisons between countries propose language and/or the results of instrument translation may be contributing features. Specifically, full scalar invariance was supported between UAE and USA student samples, which both completed the English version of the LPFS-BF 2.0, but only partial scalar invariance was reached for the remaining comparisons – each included comparing data collected using different LPFS-BF 2.0 translations (English, Italian, Spanish). Language can be considered an inherent cultural difference, but the observed variability might also be due to systematic response bias based on cultural differences (e.g., Hamid et al., 2001). Interestingly, similar noninvariance was not observed across community samples that also used these translations. Lastly, examination of latent means in both samples revealed several significant differences between countries. Altogether, the encouraging findings of this study, especially if replicated and expanded, have many important implications (also discussed below).

Discussion of Latent Mean Differences

Cross-cultural generalizability research, including cross-cultural comparisons, is uniquely complex and vulnerable to methodological and psychometric shortcomings that interfere with

determining which inferences can be appropriately drawn from cross-cultural comparisons (see

Deffner et al., 2021). Recruitment procedures and sampling methods varied across countries in

the current study, and sample-level differences in certain demographic variables were evident

(see online supplementary materials for statistical analysis of group differences). Accordingly,

interpretation of latent mean differences in terms of cultural or country-level differences in self-

ratings of personality functioning (i.e., in ratings of the self-descriptive sentences contained on

the LPFS-BF 2.0) should only be understood as preliminary, as any observed differences may

reflect differences in methodology and/or sample composition. With this in mind, examination of

latent mean differences revealed three countries that consistently deviated from all others,

allowing relatively more confidence in the interpretation of these differences. Within the student

sample, ratings of both self- and interpersonal functioning impairment were significantly lower

among the Chilean sample. Among the community samples, Denmark consistently produced

significantly lower ratings of both self- and interpersonal functioning impairment compared to

other countries. We also observed significantly higher ratings of interpersonal functioning

impairment in the UAE sample compared to all other community samples. There are multiple

plausible interpretations of these findings, some of which pertain to cultural differences, some of

which are based on methodological differences across data collections, and some reflecting

differences in sample composition. Given this core limitation of the current study, the following

possible explanations of these findings are offered as speculative and require further

investigation before they can be offered with confidence.

     The Chilean student sample's consistently lower self- and interpersonal functioning

impairment ratings contradict published self-report epidemiological data, which reports a higher

level of incidence of mental disorders among the Chilean adult population compared to the adult

populations of the remaining three countries (Italy, UAE, USA; Institute for Health Metrics and Evaluation, 2019). These findings also diverge from a recent study of mental health in university students that showed a marked increase in levels of common mental health disorders in Chilean students between 2015 and 2017 (Alamo, et. al., 2020). However, personality functioning was not assessed in these studies and our findings may reflect country-level differences in students' personality functioning specifically, above and beyond other mental health disorders. At the time of writing this paper, there was no known epidemiological data on personality pathology in Chile. The observed latent mean differences might, however, be explained by dissimilarities in study method across samples. Comparable recruitment strategies were used for the Chilean, Italian, and UAE samples (i.e., included postings on social media), whereas participants for the USA sample were recruited through an online study participation program at the researcher's university. Perhaps most distinctively, the Chilean data collection used the Spanish translation of the LPFS-BF 2.0, a methodological difference that could also be considered a cultural difference, and differed from the remaining student samples by its use of a raffle as a participation incentive. No incentives were used to collect the Italian or UAE samples, and all USA sample participants received course credit. Lastly, the significant latent mean differences may be explainable by sample differences. These findings are not likely explainable by differences in gender, as the gender compositions of the student samples were roughly similar and the prevalence of PDs in student samples have been shown to be comparable between men and women (e.g., Le Corff et al., 2021). Conversely, younger individuals have been shown to report greater levels of personality functioning impairment (e.g., Grant et al., 2008; Meaney et al., 2016; Spitzer et al., 2021), and it's plausible that the lower ratings of impairment observed in the significantly older Chilean student sample reflects this trend. However, it should be noted that we do not have

demographic data for approximately 23% of the Chilean student sample and these results are based on the subsample of Chilean students for whom this information was known. Any of the above-mentioned differences in method or sample composition could account for some, or all, of the Chilean student sample's significantly lower self- and interpersonal functioning impairment ratings.

Among the community samples, Denmark consistently produced significantly lower ratings of both self- and interpersonal functioning impairment compared to each of the other countries, whereas significantly higher ratings of interpersonal functioning impairment were consistently observed in the UAE sample. As with the student sample, latent mean differences across the community sample may reflect cultural differences, differences in study method, and/or dissimilar sample compositions. For instance, cultural factors may underlie the lower personality functioning impairment scores observed in the Danish sample, which seems consistent with Denmark being regularly ranked among the happiest countries in the world (Helliwell et al., 2020). Yet, Denmark's incidence of mental illness is comparable to other European countries (Institute for Health Metrics and Evaluation, 2019). Country-level factors, such as social environment and population composition, might also (partially) explain the significantly higher interpersonal functioning impairment scores observed for the UAE community sample. Approximately 88% of the UAE's residents are born outside of the country (i.e., expatriates; Edarabia, 2021), a characteristic also reflected in the UAE community sample (expatriates composed 83.1% of the community sample). Fewer close familial bonds and a transitory lifestyle may have contributed to UAE participants' endorsements of items meant to measure interpersonal functioning impairment (VanderWielen, 2001). Contradicting this, though, is meta-analytic evidence indicating a greater prevalence of PDs in Western countries than non-

Western countries (Winsper et al., 2020). Methodological differences and dissimilarities in

sample composition should be equally considered as possible explanations for the significant

latent mean differences that emerged across our community samples. For example, three of our

samples were collected in effort to produce normative data (Canada, Denmark, Germany), using

strategies to ensure representative (or approximately representative) samples while the remaining

countries recruited participants using convenience and/or snowball sampling methods. As such,

sample-level differences in certain demographic variables (e.g., age) may have contributed to the

current findings (Roberts et al., 2006).

Limitations and Future Research

Despite several strengths of the current study, limitations are present. Most notably,

recruitment procedures and sampling methods varied across countries, preventing our ability to

determine whether the observed latent mean differences in personality functioning are reflective

of actual differences across countries or if these findings are the product of differences in

methodology or sample composition. Similarly, differences in item intercepts across the

noninvariant groups in the student sample also cannot be confidently interpreted as being the

result of cross-cultural differences (and/or translation artifacts) due to this limitation. This

limitation, in light of otherwise promising findings, necessitates follow-up investigation.

Fortunately, our findings offer encouragement for continued research and the groundwork for an

informative and systematic cross-cultural study of the LPFS-BF 2.0. Continuation of this line of

work must involve addressing the limitations above and would benefit from a greater sampling

of countries, measures of country-level explanatory variables, and outcome measures that would

permit evaluation of the LPFS-BF 2.0's convergent and discriminant validity. Many of the

current limitations can be addressed by standardizing data collection methods, specifically by

using uniform sampling methodology (e.g., participant recruitment strategies, selection processes, incentives), parallel test batteries administered under comparable conditions (e.g., tests administered online via the same survey platform), and simultaneous data collection. Still, cross-cultural generalizability research is uniquely complex and vulnerable to methodological and psychometric shortcomings that interfere with determining which inferences can be appropriately drawn from cross-cultural comparisons. As such, a causal framework for cross-cultural generalizability of findings may be advantageous in future studies (see Deffner et al., 2021).

Implications

Current findings, should they be replicated, and future investigations of the LPFS-BF 2.0 stand to have several important implications for both research and practice. Although most mental health professionals still use a categorical model of PDs, the body of evidence supporting the scientific (Zimmermann et al., 2019) and practical (Bornstein & Natoli, 2019) superiority of dimensional models will inevitably lead to their wider adoption and the increased need to directly assess self- and interpersonal functioning impairment severity. The feasibility of the LPFS-BF 2.0 is a major strength in comparison to many other measures of personality functioning, as it represents a very short and therefore economical self-report measure. The brevity of this instrument allows for assessing self- and interpersonal functioning without overburdening study participants or patients, a benefit to the already burdened public health facilities around the world. The current findings further support the value of the LPFS-BF 2.0.

The LPFS-BF 2.0 is especially useful in countries using (or those who will be using) the ICD-11 (WHO, 2018) and DSM-5 (APA, 2013) because of its direct parallel with these diagnostic systems. However, it may be necessary to combine the LPFS-BF 2.0 with other

measures to capture certain aspects of ICD-11's personality dysfunction construct in its entirety (e.g., aspects of harm to self/others, behavioral control, reality testing, and global psychosocial impairment). Encouragingly, the LPFS-BF 2.0 (in combination with WHODAS for psychosocial impairment) was recently selected by an international research group as part of a "standard set" of self-report outcome measures designed to match the ICD-11 personality dysfunction classification (Prevolnik Rupel et al., 2021). Against the background of the introduction of ICD-11, the AMPD of the DSM-5, and the psychometric validation of a general severity factor across models and measures of personality dysfunction, the current findings offer preliminary support for international adoption of the LPFS-BF 2.0 as a standard instrument for routine, economical screenings and diagnostics in various practice contexts. Replication and expansion of the current findings to a broader collection of countries would strengthen the evidence supporting such a widespread adoption. Future evaluations of measurement invariance and cross-cultural differences would also need to include clinical samples before findings can be generalized to this key population.

In some countries, such as Chile and the UAE, personality pathology is not routinely measured by mental health service providers. The lack of reliable and valid instruments for rapidly screening for personality dysfunction exacerbates this disparity and interferes with the research necessary to better understand and treat personality pathology in these countries. The availability of short, psychometrically sound measures, such as the LPFS-BF 2.0, are most likely to move the needle in the areas of detection and local research. The current findings support that effort and encourage the wider acceptance and use of the LPFS-BF 2.0, which could help address the deficit in personality pathology research and treatment in many countries. What is more, one can postulate that the current results suggest the LPFS-BF 2.0 might also perform similarly in the

immigrant populations across the world that originate from these countries; however, future research is necessary before this generalization can be made.

A final implication of the current findings is the added support for the translated versions of the LPFS-BF 2.0, at least within community samples. The LPFS-BF 2.0 has been translated into numerous languages (all versions are available at www.deviersprong.nl/LPFS-BF) and is the only, or one of a select few, measures of personality functioning aligning with popular diagnostic systems that has been translated into certain languages and subsequently validated. For instance, at the time of writing this article, only two such measures had been validated in the French language, the LPFS-BF 2.0 (see Le Corff et al., in press) and the Self and Interpersonal Functioning Scale (Gamache et al., 2019). In light of the current results, combined with previous research, the LPFS-BF 2.0 can be said to provide clinicians and researchers with a very short screening instrument for problems in personality functioning that seems to perform similarly in Danish, English, French, German, Italian, and Spanish. However, a portion of the variability observed in our student samples may be attributable to language differences. Future studies should continue to examine measurement invariance across the numerous translated versions of the LPFS-BF 2.0. The current study, as well as the recent research by Le Corff and colleagues (in press), can serve as starting points for continuing this line of inquiry.

References

Álamo, C., Antúnez, Z., Baader, T., Kendall, J., Barrientos, M., & Barra, D. D. L. (2020). The sustained increase of mental health symptoms in Chilean university students over three years. *Revista Latinoamericana de Psicología*, *52*, 71-80.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (5th ed). American Psychiatric Association.

Bach, B., & Anderson, J. (2020). Patient-reported ICD-11 personality disorder severity and *DSM-5* level of personality functioning. *Journal of Personality Disorders*, *34*(2), 231-249. https://doi.org/10.1521/pedi_2018_32_393

Bach, B., & Hutsebaut, J. (2018). Level of Personality Functioning Scale–Brief Form 2.0: Utility in capturing personality problems in psychiatric outpatients and incarcerated addicts. *Journal of Personality Assessment*, *100*(6), 660–670.

Bach, B., Brown, T. A., Mulder, R. T., Newton-Howes, G., Simonsen, E., & Sellbom, M. (2021). Development and initial evaluation of the ICD-11 personality disorder severity scale: PDS-ICD-11. *Personality and Mental Health*, *15*(3), 223-236.

Bach, B., Simonsen, S., & Simonsen, E. (2016). Danish translation of Level of Personality Functioning Scale–Brief Form 2.0 (LPFS–BF 2.0).

Biberdzic, M., Tang, J., & Tan, J. (2021). Beyond difficulties in self-regulation: The role of identity integration and personality functioning in young women with disordered eating behaviours. *Journal of Eating Disorders*, *9*(93), 1-10. https://doi.org/10.1186/s40337-021-00398-5

Bliton, C. F., Roche, M. J., Pincus, A. L., & Dueber, D. (2022). Examining the structure and validity of self-report measures of DSM-5 Alternative Model for Personality Disorders Criterion A. Journal of Personality Disorders, 36, 157-182. https://doi.org/10.1521/pedi_2021_35_531

Bornstein, R. F., & Natoli, A. P. (2019). Clinical utility of categorical and dimensional perspectives on personality pathology: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment, 10*, 479-490. https://doi.org/10.1037/per0000365

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504.

Cottin, M., Muzard, A., Hernández, C., Mundt, A., Behn, A. (in preparation). Validation of the Spanish translation of the Level of Personality Functioning Scale-Brief Form 2.0 (LPFS-BF 2.0) in a Chilean community and clinical sample. School of Psychology, Universidad Finis Terrae.

Davidov, E., et al. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55-75. https://doi.org/10.1146/annurev-soc-071913-043137

Deffner, D., Rohrer, J. M., & McElreath, R. (2021). *A causal framework for cross-cultural generalizability.* PsyArXiv. https://doi.org/10.31234/osf.io/fqukp

Edarabia. *Population of the UAE* (2021). Retrieved from https://www.edarabia.com/population uae/. Accessed on 20th May, 2021.

Eskildsen, A., Reinholt, N., van Bronswijk, S., Brund, R. B., Christensen, A. B., Hvenegaard, M., Ardendt, M., Alro, A., Poulsen, S., Rosenberg, N. K., Markus, M.J.H., & Arnfred, S. (2020). Personalized psychotherapy for outpatients with major depression and anxiety disorders: Transdiagnostic versus diagnosis-specific group cognitive behavioural therapy. *Cognitive Therapy and Research*, *44*, 988-1001. https://doi.org/10.1007/s10608-020-10116-1

Gamache, D., Savard, C., Leclerc, P, & Côté, A. (2019). Introducing a short self-report for the assessment of DSM-5 level of personality functioning for personality disorders: The Self and Interpersonal Functioning Scale. *Personality Disorders: Theory, Research, and Treatment, 10*(5), 438-447. https://doi.org/10.1037/per0000335

Grant, B. F., (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: Results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry 69*(4), 533-545. https://doi.org/10.4088/jcp.v69n0404

Gritti, E. S., De Carli, P., Hutsebaut, J., Simonelli, A., & Zimmermann, J. (in preparation). Cross-cultural adaptation of the Personality Functioning Scale- Brief Form 2.0 (LPFS-BF 2.0): Temporal stability and construct validity of the Italian version.

Hamid, P. N., Lai, J. C. L., & Cheng, S.-T. (2001). Response bias and public and private self-consciousness in Chinese. *Social Behavior and Personality: An International Journal, 29(8)*, 733–742. https://doi.org/10.2224/sbp.2001.29.8.733

Hopwood, C. J., Good, E. W., & Morey, L. C. (2018). Validity of the DSM–5 Levels of Personality Functioning Scale–Self Report. *Journal of Personality Assessment, 100*, 650-659. https://doi.org/10.1080/00223891.2017.1420660

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, Y., Kotov, R., de Girolamo, G., Preti, A., Angermeyer, M., Benjet, C., Demyttenaere, K., de Graaf, R., Gureje, O., Karam, A. N., Lee, S., Lépine, J. P., Matschinger, H., Posada-Villa, J., Suliman, S., Vilagut, G., & Kessler, R. C. (2009). DSM-IV personality disorders in the WHO

World Mental Health Surveys. *British Journal of Psychiatry, 195*(1), 46–53.

https://doi.org/10.1192/bjp.bp.108.058552

Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (2016). Development and preliminary

psychometric evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM–5

level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). *Personality

Disorders: Theory, Research, and Treatment*, *7*(2), 192-197. https://doi.org/10.1037/per0000159

Hutsebaut, J., Weekers, L. C., Tuin, N., Apeldoorn, J. S., & Bulten, E. (2021). Assessment of ICD-

11 Personality Disorder severity in forensic patients using the Semi-structured Interview for

Personality Functioning DSM-5 (STiP-5.1): Preliminary findings. *Frontiers in Psychiatry*, *12*.

https://doi.org/10.3389/fpsyt.2021.617702

Huxley, E., Seaton, D. C., & Grenyer, B. F. (2021). Remembered childhood invalidation as a

predictor of narcissism, personality functioning, and wellbeing. *Personality and Individual

Differences*, *175*. https://doi.org/10.1016/j.paid.2021.110686

Institute for Health Metrics and Evaluation (2019). Global Health Data Exchange [Data set].

University of Washington. https://www.ghdx.healthdata.org/gbd-results-tool

Konjusha, A., Hopwood, C. J., Price, A. L., Masuhr, O., & Zimmermann, J. (2021). Investigating

the transdiagnostic value of subjective emptiness. *Journal of Personality Disorders, 35*(5).

https://doi.org/10.1521/pedi_2021_35_510

Le Corff, Y., Aluja, A., Rossi, G., Lapalme, M., Forget, K., Rolland, J.-P., et Garcia, L. F. (in

press). Construct validity of the Dutch, English, French and Spanish LPFS-BF 2.0: Measurement

invariance across language and gender, and criterion validity. *Journal of Personality Disorders*.

Le Corff, Y., Martin-Storey, A., Touchette, L., Lapalme, M., & Forget, K. (2021). Validation of a

French translation of the McLean screening instrument for borderline personality disorder,

invariance across genders, and association with depression, trauma symptoms, and substance use among university students. *Journal of Personality Disorders, 35,* 605-617. https://doi.org/10.1521/pedi_2020_34_494

Liggett, J., & Sellbom, M. (2018). Examining the DSM-5 alternative model of personality disorders operationalization of obsessive–compulsive personality disorder in a mental health sample. *Personality Disorders: Theory, Research, and Treatment*, *9*(5), 397–407. https://doi.org/10.1037/per0000285

Markus, H., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives in Psychological Science, 5*, 420–430. https://doi.org/10.1177/1745691610375557

Marsh, H. W., Guo, J., Parker, P., Nagengast, B., Asparouhov, T., Muthén, B. O., & Dicke, T. (2018). What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychological Methods, 23*(3), 524–545. https://doi.org/10.1037/met0000113

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indices and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

McCabe, G. A., & Widiger, T.A. (2020). A comprehensive comparison of the ICD-11 and DSM–5 section III personality disorder models. *Psychological Assessment*, *32*(1), 72–84. https://doi.org/10.1037/pas0000772

McCrae, R., Terraciano, A., & 79 members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology, 89*(3), 407–425. https://doi.org/10.1037/0022-3514.89.3.407

Meaney, R., Hasking, P., & Reupert, A. (2016). Borderline personality disorder symptoms in college students: The complex interplay between alexithymia, emotional dysregulation and rumination. *PloS one*, *11*, e0157294. https://doi.org/10.1371/journal.pone.0157294

Morey, L. C., Bender, D. S., & Skodol, A. E. (2013). Validating the proposed Diagnostic and Statistical Manual of Mental Disorders, 5th edition, severity indicator for personality disorder. *Journal of Nervous and Mental Disease, 201(9)*, 729–735. https://doi.org/10.1097/NMD.0b013e3182a20ea8

Müller, S., Wendt, L. P., Spitzer, C., Masuhr, O., Back, S. N., & Zimmermann, J. (2021). A critical evaluation of the Reflective Functioning Questionnaire (RFQ). *Journal of Personality Assessment*. Advanced online publication. https://doi.org/10.1080/00223891.2021.1981346

Müller, S., Wendt, L. P., & Zimmermann, J. (2021). Development and validation of the Certainty about Mental States Questionnaire (CAMSQ): A self-report measure of mentalizing oneself and others. *Assessment*. Advanced online publication. https://doi.org/10.1177/10731911211061280

Muthén, L. K., & Muthén, B. O. (1998-2019). *Mplus User's Guide*. (8th ed.). Authors.

Prevolnik Rupel, V., Jagger, B., Fialho, L.S., Chadderton, L., Gintner, T., Arntz, A., Baltzersen, A., Blazdell, J., van Busschbach, J., Cencelli, M., Chanen, A., Delvaux, C., van Gorp, F., Langford, L., McKenna, B., Moran, P., Pacheco, K., Sharp, C. Wang, W.,… Crawford, M. J. (2021). Standard set of patient-reported outcomes for personality disorder. *Quality of Life Research, 30,* 3485-3500. https://doi.org/10.1007/s11136-021-02870-w

Rishede, M. Z., Juul, S., Bo, S., Gondan, M., Bjerrum Møeller, S., & Simonsen, S. (2021). Personality functioning and mentalizing in patients with subthreshold or diagnosed borderline personality disorder: Implications for ICD-11. *Frontiers in Psychiatry*, *12*, 408. https://doi.org/10.3389/fpsyt.2021.634332

Roche, M. J., & Jaweed, S. (2021). Comparing measures of Criterion A to better understand incremental validity in the Alternative Model of Personality Disorders. *Assessment.* Advanced online publication. https://doi.org/10.1177/10731911211059763

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Rossi, G., Weekers, L. C., & Hutsebaut, J. (2021). Resilient, undercontrolled, and overcontrolled personality types based upon DSM-5 maladaptive personality traits. *Heliyon*, *7*(5). https://doi.org/10.1016/j.heliyon.2021.e06938

Schetsche, C. (2021). Traducción y adaptación de la Level of Personal Functioning Scale-Brief Form 2.0. *Ciencias Psicológicas*, *15*(2), 15-33. http://dx.doi.org/10.22235/cp.v15i2.2387

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89*, 623–642. https://doi.org/10.1037/0022-3514.89.4.623

Somma, A., Borroni, S., Gialdi, G., Carlotta, D., Emanuela Giarolli, L., Barranca, M., ... & Fossati, A. (2020). The inter-rater reliability and validity of the Italian translation of the structured clinical interview for DSM-5 Alternative Model for Personality Disorders Module I and Module II: A preliminary report on consecutively admitted psychotherapy outpatients. *Journal of Personality Disorders*, *34* (Supplement C), 95-123. https://doi.org/10.1521/pedi_2020_34_511

Somma, A., Krueger, R. F., Markon, K. E., Gialdi, G., Frau, C., Boscaro, L., Liberatore, G., & Fossati, A. (2020). Body dysmorphic disorder in the perspective of the Alternative DSM-5 Model for Personality Disorder: A study on Italian community-dwelling women. *The Journal of Nervous and Mental Disease*, *208*(12), 974-981. DOI: 10.1097/NMD.0000000000001239

Spitzer, C., Müller, S., Kerber, A., Hutsebaut, J., Brähler, E., & Zimmermann, J. (2021). Die

deutsche Version der Level of Personality Functioning Scale-Brief Form 2.0 (LPFS-BF):

Faktorenstruktur, konvergente Validität und Normwerte in der Allgemeinbevölkerung.

Psychotherapie - Psychosomatik - Medizinische Psychologie, 71(7), 284-293.

https://doi.org/10.1055/a-1343-2396

Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-

national consumer research. *Journal of Consumer Research, 25*(1), 78-90.

https://doi.org/10.1086/209528

Stone, L. E., Segal, D. L., & Noel, O. R. (2021). Psychometric evaluation of the Levels of

Personality Functioning Scale-Brief Form 2.0 among older adults. *Personality Disorders:*

Theory, Research, and Treatment, *12*(6), 526-533 https://doi.org/10.1037/per0000413

Stricker, J., & Pietrowsky, R. (2021). Incremental validity of the ICD-11 personality disorder model

for explaining psychological distress. *Personality Disorders: Theory, Research, and Treatment*.

Advanced online publication. https://doi.org/10.1037/per0000489

Turner, B. J., Prud'homme, J., & Legg, N. (2020). Environmental and sociocultural influences on

personality disorders. In C. W. Lejuez & K. L. Gratz (Eds.), *The Cambridge handbook of*

personality disorders (pp. 50–64). Cambridge University Press.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World

Population Prospects, Online Edition. Rev. 1.

van IJzendoorn, M. H., & Kroonenberg, P. M. (1988). Cross-cultural patterns of attachment: A

meta-analysis of the strange situation. *Child Development, 59*(1), 147-156.

Vanderveren, E., Bogaerts, A., Claes, L., Luyckx, K., & Hermans, D. (2021). Narrative coherence of turning point memories: Associations with psychological well-being, identity functioning, and personality disorder symptoms. *Frontiers in Psychology*, *12*.

VanderWielen, J. J. (2001). Cognitive appraisal, coping, and the psychological and sociocultural adjustment of expatriates. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62*(3-B), 1630.

Waugh, M. H., McClain, C. M., Mariotti, E. C., Mulay, A. L., DeVore, E. N., Lenger, K. A., Russell, A. N., Florimbio, A. R., Lewis, K. C., Ridenour, J. M., & Beevers, L. G. (2021) Comparative content analysis of self-report scales for Level of Personality Functioning. *Journal of Personality Assessment, 103*(2), 161-173. https://doi.org/10.1080/00223891.2019.1705464

Weekers, L. C., Hutsebaut, J., & Kamphuis, J. H. (2019). The Level of Personality Functioning Scale-Brief Form 2.0: Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health*, *13*(1), 3–14. https://doi.org/10.1002/pmh.1434

Weekers, L. C., Verhoeff, S. C. E., Kamphuis, J. H., & Hutsebaut, J. (2020). Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM–5. *Personality Disorders: Theory, Research, and Treatment*, *12*(4), 312-319. https://doi.org/10.1037/per0000454

WHR (2020). *World Happiness Report 2020.* J. F. Helliwell, R. Layard, J. D. Sachs & J.-E. De Neve (Eds.). Sustainable Development Solutions Network.

Widiger, T. A., Bach, B., Chmielewski, M., Clark, L. A., DeYoung, C., Hopwood, C. J., Kotov, R., Krueger, R. F., Miller, J. D., Morey, L. C., Mullins-Sweatt, S. N., Patrick, C. J., Pincus, A. L., Samuel, D. B., Sellbom, M., South, S. C., Tackett, J. L., Watson, D., Waugh, M. H., Wright, A.,

… Thomas, K. M. (2019) Criterion A of the AMPD in HiTOP. *Journal of Personality Assessment, 101*, 345-355. https://doi.org/10.1080/00223891.2018.1465431

Winsper, C., Bilgin, A., Thompson, A., Marwaha, S., Chanen, A. M., Singh, S. P., Wang, A., & Furtado, V. (2020). The prevalence of personality disorders in the community: A global systematic review and meta-analysis. *The British Journal of Psychiatry*, *216*(2), 69–78. https://doi.org/10.1192/bjp.2019.166

Woehler, L. L. (2020). Internet addiction and the relationship to self and interpersonal functioning within the Alternative Model for Personality Disorders (AMPD): Implications for Psychosocial development. https://aura.antioch.edu/etds/619

World Health Organization (2018). *International Statistical Classification of Disease and Related Health Problems* (11th ed.). World Health Organization.

Zaeimzadeh, L., Lotfi, M., & Amini, M. (2020). Role of personality functioning in predicting interpersonal and emotional problems. *Journal of Mazandaran University of Medical Sciences*, *30*(190), 97-108.

Zimmermann, J., Hopwood, C. J., & Krueger, R. F. (2022). The DSM-5 Level of Personality Functioning Scale. In R. F. Krueger & P. H. Blaney (Eds.), *Oxford Textbook of Psychopathology* (4th ed.). Oxford University Press.

Zimmermann, J., Kerber, A., Rek, K., Hopwood, C. J., & Krueger, R. F. (2019). A brief but comprehensive review of research on the Alternative DSM-5 Model for Personality Disorders. *Current Psychiatry Reports, 21:92*.

Zimmermann, J., Müller, S., Bach, B., Hutsebaut, J., Hummelen, B., & Fischer, F. (2020). A common metric for self-reported severity of personality disorder. *Psychopathology*, *53*(3), 161-171. https://doi.org/10.1159/000507377

Table 1
Studies wherein the LPFS-BF and/or the LPFS-BF 2.0 were included

Country	Country Population	Country Median Age	Country Sex Breakdown	Study Reference	Population Studied	N	Sample Age μ (SD)	Sample Sex Breakdown
Argentina	44,781,000	31.5	M = 48.8% F = 51.2%	Schetsche (2021)	Community	361	41.01 (15.58)	F = 56.2%
Australia	25,203,000	37.9	M = 49.8% F = 50.2%	Biberdzic et al. (2021)	Student	247	19.39 (1.70)	F = 88.7%
				Huxley et al. (2021)	Community	291	27.61 (13.20)	F = 81.1%
				Liggett & Sellbom (2018)	Clinical	214	22.47 (8.43)	F = 72.4%
Belgium	11,539,000	41.9	M = 49.6% F = 50.4%	Rossi et al. (2021)	Clinical	192	36.71 (11.10)	F = 66.1%
				Vanderveren et al. (2021)	Community	333	22.56 (3.13)	F = 72.1%
Denmark	5,772,000	42.3	M = 49.7% F = 50.3%	Bach & Anderson (2020)	Clinical	150	32.50 (10.6)	F = 33%
				Bach & Hutsebaut (2018)	Clinical/Forensic	228	31.50 (10.03)	F = 42%
				Eskildsen et al. (2020)	Clinical	291	32.22 (11.01)	F = 64.6%
				Rishede et al. (2021)	Clinical	116	32 (NA)	F = 94%
Germany	83,517,000	45.7	M = 49.4% F = 50.6%	Konjusha et al. (2021)	Clinical	43	34.60 (12.1)	F = 71%
					Community	114	27.60 (9.0)	F = 71%
				Müller, Wendt, & Zimmermann (2021)	Community	401	45.90 (16.6)	F = 49%
				Spitzer et al. (2021)	Community	2470	49.50 (17.5)	F = 50.1%
				Zimmermann et al. (2020)	Community	849	42.60 (16.1)	F = 49.2%
Islamic Republic of Iran	82,914,000	32.0	M = 50.5% F = 49.5%	Zaeimzadeh et al. (2020)	Community	245	range: 18-60	F = 56.3%
Italy	60,550,000	47.3	M = 48.7% F = 51.3%	Somma et al. (2020a)	Clinical	88	36.47 (14.04)	F = 54.5%
				Somma et al. (2020b)	Community	420	31.74 (12.15)	F = 100%
Netherlands	17,097,000	43.3	M = 49.8% F = 50.2%	Hutsebaut et al. (2016)	Clinical	240	33.97 (10.42)	F = 66.3%

Country	Population	Age	Sex	Study	Sample	N	M (SD)	Sex
				Hutsebaut et al. (2021)	Clinical	30	38.43 (11.70)	F = 89.7%
				Weekers et al. (2019)	Clinical	201	36.2 (11.0)	F = 65.2%
					Clinical	47	NA	NA
				Weekers et al. (2021)	Clinical	84	15.60 (1.39)	F = 89.3%
					Community	12	15.08 (1.16)	F = 75%
New Zealand	4,783,000	38.0	M = 49.2% F = 50.8%	Bach et al. (2021)	Clinical	87	27.96 (10.54)	F = 61.5%
United States of America	329,065,000	38.3	M = 49.5% F = 50.5%	Bach & Anderson (2020)	Clinical	150	32.50 (10.6)	F = 33%
				Bach et al. (2021)	Community	428	45.70 (17.3)	F = 50.9%
				Bliton et al. (2022)	Student	608	19.38 (2.37)	F = 51.2%
				McCabe & Widiger (2020)	Community	300	36.51 (10.36)	F = 54%
				Müller, Wendt, Spitzer, et al. (2021)	Community	862	34.90 (11.7)	F = 47%
				Müller, Wendt, & Zimmermann (2021)	Community	403	45.20 (16.6)	F = 52%
				Roche & Jaweed (2021)	Student	204	18.90 (1.47)	F = 69%
				Stone et al. (2021)	Community	130	64.61(3.73)	F = 65%
				Stricker & Pietrowsky (2021)	Community	226	36.95 (11.45)	F = 35%
				Waugh et al. (2021)	Experts	23	29.36 (8.23)	NA
				Woehler (2020)	Community	363	range: 25-39	F = 48.8%

Note. For sex, M = male and F = female. Descriptives of country population, age, and sex breakdown were extracted from United Nations (2019). The German sample used in Zimmermann et al. (2020) also appeared in Spitzer et al. (2021) and, thus, is only reported once in this table. Two subsamples of patients were used in Weekers et al. (2019), but demographics were only provided for the larger subsample. Stricker and Pietrowsky (2021) included the LPFS-BF 2.0 only in the second of their two separate community samples, and only this second sample is reported in the table. Waugh et al. (2021) evaluated the content validity of several self-report measures of personality functioning, including the LPFS-BF 2.0, by inviting 23 raters to assess these tests. This table is the product of a systematic search conducted in late 2021 and updated in mid 2022.

Table 2
Participant Descriptive Statistics

Variables	Community						Student			
	Canada (*N* = 2,505)	Chile (*N* = 281)	Denmark (*N* = 713)	Germany (*N* = 849)	Italy (*N* = 187)	UAE (*N* = 142)	Chile (*N* = 213)	Italy (*N* = 224)	UAE (*N* = 320)	USA (*N* = 201)
Gender, *n* (%)										
Male	1217 (48.6%)	71 (25.3%)	325 (45.6%)	424 (49.9%)	48 (25.7%)	64 (45.1%)	30 (14.1%)	38 (17.0%)	72 (22.5%)	47 (23.4%)
Female	1279 (51.1%)	209 (74.4%)	388 (54.4%)	418 (49.2%)	139 (74.3%)	78 (54.9%)	143 (67.1%)	185 (82.6%)	246 (76.9%)	149 (74.1%)
Non-binary	9 (0.4%)	1 (0.4%)	0 (0%)	7 (0.8%)	0 (0%)	0 (0%)	0 (0%)	1 (0.4%)	2 (0.6%)	5 (2.5%)
Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	40 (18.8%)	0 (0%)	0 (0%)	0 (0%)
Age, *M* (*SD*)	46.1 (16.4)	36.3 (9.9)	N/A	42.7 (16.1)	43.4 (15.9)	36.5 (12.1)	23.3 (5.6)	23.0 (5.7)	19.3 (1.6)	21.0 (4.1)
< 30, *n* (%)	476 (19.0%)	59 (21.0%)	156 (21.9%)	219 (25.8%)	41 (21.9%)	52 (36.6%)	154 (72.3%)	207 (92.4%)	320 (100%)	194 (96.5%)
30-39, *n* (%)	510 (20.4%)	151 (53.7%)	61 (8.6%)	155 (18.3%)	44 (23.5%)	22 (15.5%)	14 (6.6%)	10 (4.5%)	0 (0%)	4 (2.0%)
40-49, *n* (%)	472 (18.8%)	38 (13.5%)	111 (15.6%)	143 (16.8%)	36 (19.3%)	42 (29.6%)	4 (1.9%)	5 (2.2%)	0 (0%)	3 (1.5%)
50-59, *n* (%)	423 (16.9%)	21 (7.5%)	149 (20.9%)	185 (21.8%)	31 (16.6%)	22 (15.5%)	1 (0.5%)	1 (0.4%)	0 (0%)	0 (0%)
60-69, *n* (%)	387 (15.4%)	10 (3.6%)	138 (19.4%)	115 (13.5%)	22 (11.8%)	1 (0.7%)	0 (0%)	1 (0.4%)	0 (0%)	0 (0%)
70+, *n* (%)	236 (9.4%)	2 (0.7%)	98 (13.7%)	32 (3.8%)	13 (7.0%)	1 (0.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Missing, *n* (%)	1 (<0.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (1.4%)	40 (18.8%)	0 (0%)	0 (0%)	0 (0%)
LPFS-BF 2.0										
SF, *M* (*SD*)	1.85 (0.72)	1.94 (0.82)	1.62 (0.67)	1.93 (0.79)	1.91 (0.73)	2.07 (0.75)	2.14 (0.80)	2.39 (0.67)	2.56 (0.74)	2.37 (0.70)
IF, *M* (*SD*)	1.84 (0.62)	1.69 (0.65)	1.56 (0.55)	1.85 (0.60)	1.80 (0.55)	2.06 (0.57)	1.72 (0.68)	1.88 (0.44)	2.15 (0.56)	2.04 (0.54)

Note. LPFS-BF 2.0 = Level of Personality Functioning Scale – Brief Form 2.0; SF = average self-functioning impairment score on the LPFS-BF 2.0; IF = average interpersonal functioning impairment score on the LPFS-BF 2.0; N/A = not applicable; UAE = United Arab Emirates; USA = United States of America.

Table 3
Measurement Invariance Results

	Chi-Square (*df*)	CFI	RMSEA	SRMR	ΔChi-square (*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariant?
Basic Model									
Student	198.73*** (53)	.946	.054	.036					
Community	631.41*** (53)	.967	.048	.027					
Country									
Student (*ns* = 191 from Chile, 224 from Italy, 320 from UAE, 201 from USA)									
Configural	423.59*** (212)	.922	.065	.056					
Metric	492.65*** (242)	.908	.067	.070	70.19*** (30)	-.014	+.002	+.014	Yes
Scalar	697.95*** (272)	.843	.082	.083	214.30*** (30)	-.065	+.015	+.013	No
Partial Scalar[1]	663.52*** (271)	.855	.079	.080	170.87*** (29)	-.053	+.012	+.010	No
Partial Scalar[2]	641.74*** (270)	.863	.077	.077	149.09*** (28)	-.045	+.010	+.007	No
Partial Scalar[3]	621.59*** (269)	.870	.075	.077	128.94*** (27)	-.038	+.008	+.007	No
Partial Scalar[4]	602.15*** (268)	.877	.073	.075	109.50*** (26)	-.031	+.006	+.005	No
Partial Scalar[5]	584.01*** (267)	.883	.071	.074	91.36*** (25)	-.025	+.004	+.004	Yes (P)
Community (*ns* = 2505 from Canada, 267 from Chile, 713 from Denmark, 849 from Germany, 187 from Italy, 142 from UAE)									
Configural	1076.53*** (318)	.961	.055	.034					
Metric	1200.38*** (368)	.957	.054	.045	115.54*** (50)	-.004	-.001	+.011	Yes
Scalar	1533.89*** (418)	.942	.059	.049	373.49*** (50)	-.015	+.005	+.004	Yes

Note. df = degrees of freedom; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; (P) = partial invariance; otherwise, "Yes" indicates that full invariance was established; UAE = United Arab Emirates; USA = United States of America. [1]Intercept freed for Item 11 in Chile; [2]Intercept freed for Item 8 in Italy; [3]Intercept freed for Item 2 in Chile; [4]Intercept freed for Item 5 in Chile; [5]Intercept freed for Item 12 in Italy.
***p < .001

Table 4

Standardized loadings of LPFS-BF 2.0 items onto their respective parent factor

Item[2]	Community						Student			
	Canada	Chile	Denmark	Germany	Italy	UAE	Chile	Italy	UAE	USA
1) I often do not know who I really am.	0.759	0.728	0.716	0.784	0.653	0.646	0.801	0.744	0.767	0.754
2) I often think very negatively about myself.	0.785	0.779	0.727	0.797	0.732	0.728	0.736	0.665	0.689	0.707
3) My emotions change without me having a grip on them.	0.783	0.775	0.755	0.803	0.685	0.677	0.697	0.590	0.612	0.655
4) I have no sense of where I want to go in my life.	0.736	0.756	0.716	0.742	0.687	0.628	0.624	0.529	0.566	0.568
5) I often do not understand my own thoughts and feelings.	0.836	0.824	0.820	0.852	0.729	0.685	0.799	0.685	0.689	0.761
6) I often make unrealistic demands on myself.	0.719	0.693	0.653	0.722	0.641	0.637	0.625	0.530	0.594	0.597
7) I often have difficulty understanding the thoughts and feelings of others.	0.677	0.652	0.627	0.646	0.541	0.452	0.662	0.405	0.465	0.516
8) I often find it hard to stand it when others have a different opinion.	0.641	0.555	0.547	0.571	0.440	0.450	0.500	0.256	0.352	0.396
9) I often do not fully understand why my behavior has a certain effect on others.	0.699	0.689	0.660	0.682	0.542	0.505	0.728	0.473	0.572	0.629
10) My relationships and friendships never last long.	0.668	0.628	0.630	0.612	0.516	0.436	0.637	0.376	0.421	0.511
11) I often feel very vulnerable when relations become more personal.	0.701	0.683	0.601	0.629	0.528	0.475	0.677	0.346	0.564	0.556
12) I often do not succeed in cooperating with others in a mutually satisfactory way	0.737	0.693	0.684	0.704	0.579	0.489	0.676	0.350	0.504	0.529
Internal Consistency Statistics (McDonald's ω)										
Self-functioning Impairment	.898	.892	.872	.904	.852	.833	.870	.801	.817	.836
Interpersonal Functioning Impairment	.844	.816	.799	.808	.714	.649	.822	.492	.661	.705

Note. This table only reports standardized loadings of items onto their respective parent factor. Items 1-6 were specified to load onto one factor (level of self-functioning) and items 7-12 were specified to load onto a second factor (level of interpersonal functioning); UAE = United Arab Emirates; USA = United States of America.

Table 5
Latent mean differences across countries with student samples

	Chile	Italy	UAE	USA
Chile (*n* = 191)	-	-0.256* (-0.449, -0.062)	-0.361* (-0.542, -0.181)	-0.248* (-0.447, -0.050)
Italy (*n* = 224)	0.220* (0.026, 0.413)	-	-0.110 (-0.281, 0.061)	0.024 (-0.166, 0.215)
UAE (*n* = 320)	0.431* (0.250, 0.612)	0.178* (0.006, 0.349)	-	0.136 (-0.040, 0.313)
USA (*n* = 201)	0.447* (0.246, 0.647)	0.093 (-0.098, 0.283)	-0.103 (-0.279, 0.074)	-

Note. Latent mean differences are reported using Cohen's *d*, with positive values indicating the latent mean of the country identified by the row label is greater than that of the country identified by the column label and negative values indicating the latent mean of the country identified by the column label is greater than the country identified by the row label. Latent mean differences for self-functioning are reported in the upper diagonal and latent mean differences for interpersonal functioning are reported in the lower diagonal. UAE = United Arab Emirates; USA = United States of America.
*Significant at *p* < .05.

Table 6
Latent mean differences across countries with community samples

	Canada	Chile	Denmark	Germany	Italy	UAE
Canada (*n* = 2,505)	-	-0.063 (-0.189, 0.064)	0.172* (0.088, 0.255)	-0.064 (-0.142, 0.014)	-0.048 (-0.196, 0.101)	-0.149 (-0.318, 0.021)
Chile (*n* = 267)	-0.137* (-0.263, -0.011)	-	0.300* (0.159, 0.441)	0.019 (-0.119, 0.156)	0.024 (-0.162, 0.212)	-0.160 (-0.364, 0.044)
Denmark (*n* = 713)	-0.289* (-0.372, -0,205)	-0.174* (-0.315, -0.034)	-	-0.400* (-0.500, -0.299)	-0.274* (-0.436, -0.113)	-0.400* (-0.581, -0.219)
Germany (*n* = 849)	0.004 (-0.074, 0.082)	0.419* (0.281, 0.558)	0.458* (0.357, 0.559)	-	0.037 (-0.121, 0.196)	-0.453* (-0.632, -0.274)
Italy (*n* = 187)	-0.067 (-0.215, 0.082)	0.099 (-0.088, 0.287)	0.247* (0.085, 0.408)	-0.231* (-0.389, -0.072)	-	-0.181 (-0.399, 0.038)
UAE (*n* = 142)	0.223* (0.054, 0.392)	0.516* (0.310, 0.723)	0.562* (0.380, 0.744)	0.801* (0.620, 0.982)	0.403* (0.183, 0.623)	-

Note. Latent mean differences are reported using Cohen's *d*, with positive values indicating the latent mean of the country identified by the row label is greater than that of the country identified by the column label and negative values indicating the latent mean of the country identified by the column label is greater than the country identified by the row label. Latent mean differences for self-functioning are reported in the upper diagonal and latent mean differences for interpersonal functioning are reported in the lower diagonal. UAE = United Arab Emirates.
*Significant at *p* < .05.

<u>**Online Supplementary Materials**</u>

Multi-National Evaluation of the Measurement Invariance of the Level of Personality Functioning Scale – Brief

Form 2.0: Comparison of Student and Community Samples Across Seven Countries

Copies of the Level of Personality Functioning Scale - Brief Form 2.0, available translations, and essential

articles are available at the following link: www.deviersprong.nl/LPFS-BF

Cross-Country Comparison of Gender Proportions

A series of tests of equality of proportions were carried out using R's 'prop.test' function (R Core Team, 2020)

to formally test whether the gender proportions were equal across countries, performed separately for our

student and community samples. The test of equality of proportions failed to reject the null that the gender

proportions (females:males) across countries in our student sample were the same, χ^2 (3) = 5.04, p = .169.

Conversely, gender proportions were found to significantly differ across countries in our community sample, χ^2

(5) = 93.27, p < .001. Specifically, significant differences were observed between Canada and Chile, χ^2 (1) =

54.50, Canada and Italy, χ^2 (1) = 36.30, Chile and Denmark, χ^2 (1) = 33.46, Chile and Germany, χ^2 (1) = 52.26,

Chile and the UAE, χ^2 (1) = 15.94, Denmark and Italy, χ^2 (1) = 23.39, Germany and Italy, χ^2 (1) = 36.57, and

between Italy and the UAE, χ^2 (1) = 12.68 (all at p < .001).

Cross-Country Comparison of Age

Mean age was compared across countries by conducting a one-way independent analysis of variance (ANOVA)

separately in our student and community samples. Due to unequal sample sizes and differences in population

variances, the Games-Howell procedure was used for post-hoc analyses. Mean age could not be calculated for

the Danish sample because only participants' age group memberships were recorded. Thus, the Danish sample

was not included in the initial ANOVA; however, a third ANOVA was conducted for our community sample

wherein estimated ages were entered for the Danish sample by assigning each participant a mid-range age based on their age group (e.g., participants in the 30-39 age range group were assigned an age of 34.5). Results of these analyses are as follows: There was a significant effect of country on student sample age, $F(3, 892) = 47.43$, $p < .001$, $\omega^2 = .14$; post-hoc tests using Tukey's procedure revealed several significant group differences (see Table S1). There was also a significant effect of country on community sample age, $F(4, 3956) = 37.12$, $p < .001$, $\omega^2 = .04$ when the Danish sample was excluded; post-hoc tests using Tukey's procedure revealed several significant group differences (see Table S2). There was also a significant effect of country on community sample age when the estimated ages for the Danish sample were included, $F(5, 4668) = 40.90$, $p < .001$, $\omega^2 = .04$; post-hoc tests using Tukey's procedure revealed several significant group differences (see Table S3).

References

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from: https://www.R-project.org/

Table S1

| Comparison | Mean Difference | 95% CI | | t (df) |
		Lower	Upper	
Chile				
Italy	-0.05	-1.45	1.35	-0.09 (358.96)
UAE	3.67***	2.65	4.69	9.34 (167.07)
USA	1.95***	0.71	3.19	4.05 (298.85)
Italy				
UAE	3.72***	2.70	4.74	9.45 (248.11)
USA	2.00***	0.76	3.24	4.15 (404.45)
UAE				
USA	-1.72***	-2.51	-0.94	-5.65 (239.25)

Note. A positive mean difference value indicates a relatively greater mean age was observed for the leading country in the comparison; UAE = United Arab Emirates; USA = United States of America.
***p < .001

Table S2

Comparison		Mean Difference	95% CI		t (df)
			Lower	Upper	
Canada					
	Chile	9.81***	7.96	11.66	14.52 (474.12)
	Germany	3.46***	1.71	5.21	5.39 (1490.31)
	Italy	2.74	-0.58	6.06	2.27 (216.72)
	UAE	9.59***	6.63	12.54	8.94 (169.10)
Chile					
	Germany	-6.35***	-8.56	-4.14	-7.86 (784.24)
	Italy	-7.07***	-10.65	-3.49	-5.42 (292.10)
	UAE	-0.22	-3.47	3.02	-0.19 (234.71)
Germany					
	Italy	-0.72	-4.25	2.82	-0.56 (276.24)
	UAE	6.13***	2.94	9.32	5.28 (228.90)
Italy					
	UAE	6.84***	2.60	11.09	4.43 (324.92)

Note. A positive mean difference value indicates a relatively greater mean age was observed for the leading country in the comparison.
***p < .001

Table S3

Comparison		Mean Difference	95% CI		t (df)
			Lower	Upper	
Canada					
	Chile	9.81***	7.88	11.74	14.52 (474.12)
	Denmark	-3.19***	-5.60	-1.10	-4.37 (1096.87)
	Germany	3.46***	1.63	5.29	5.39 (1490.31)
	Italy	2.74	-0.73	6.22	2.27 (216.72)
	UAE	9.59***	6.50	12.68	8.94 (169.10)
Chile					
	Denmark	-13.00***	-15.51	-10.48	-14.76 (870.71)
	Germany	-6.35***	-8.66	-4.04	-7.86 (784.24)
	Italy	-7.07***	-10.81	-3.33	-5.42 (292.10)
	UAE	-0.22	-3.61	3.17	-0.19 (234.71)
Denmark					
	Germany	6.65***	4.21	9.09	7.78 (1464.77)
	Italy	5.93***	2.11	9.75	4.45 (313.82)
	UAE	12.78***	9.30	16.25	10.55 (267.31)
Germany					
	Italy	-0.72	-4.41	2.98	-0.56 (276.24)
	UAE	6.13***	2.79	9.46	5.28 (228.90)
Italy					
	UAE	6.84***	2.41	11.28	4.43 (324.92)

Note. A positive mean difference value indicates a relatively greater mean age was observed for the leading country in the comparison; UAE = United Arab Emirates; USA = United States of America.

***$p < .001$

Table S4
Skewness values at the item level for each subsample

	Community						Student			
	Canada	Chile	Denmark	Germany	Italy	UAE	Chile	Italy	UAE	USA
Skewness (SE)										
Item 1	0.991	1.060	2.061	1.073	1.152	0.530	0.659	0.143	-0.086	-0.031
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 2	0.580	0.267	0.765	0.517	0.299	0.445	-0.141	-0.418	-0.250	-0.120
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 3	0.619	0.736	1.142	0.654	0.197	0.141	0.383	-0.089	-0.211	0.071
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 4	0.777	0.933	1.284	0.690	0.852	0.803	1.100	0.257	0.246	0.493
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 5	0.733	0.817	1.393	0.708	0.686	0.335	0.232	0.286	-0.164	-0.088
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 6	0.438	0.291	0.784	0.430	0.769	0.232	0.373	0.094	-0.376	-0.223
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 7	0.411	1.100	1.254	0.601	0.643	0.551	1.335	0.750	0.575	0.382
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 8	0.553	0.741	0.955	0.676	0.284	-0.116	0.729	0.183	0.403	0.423
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 9	0.368	1.073	1.004	0.532	0.453	-0.040	0.945	0.459	0.059	0.244
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.137)	(0.172)
Item 10	1.086	1.641	2.160	1.017	1.720	1.060	1.581	1.083	0.642	0.935
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)
Item 11	0.516	1.147	0.970	0.467	0.630	-0.175	0.765	0.199	-0.489	-0.337
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.204)	(0.176)	(0.163)	(0.136)	(0.172)
Item 12	0.741	1.372	1.687	0.905	1.276	1.049	1.237	0.734	0.643	0.728
	(0.049)	(0.149)	(0.092)	(0.084)	(0.178)	(0.203)	(0.176)	(0.163)	(0.136)	(0.172)

Note. UAE = United Arab Emirates; USA = United States of America.

Table S5
Kurtosis values at the item level for each subsample

	Community						Student			
	Canada	Chile	Denmark	Germany	Italy	UAE	Chile	Italy	UAE	USA
Kurtosis (SE)										
Item 1	-0.011 (0.098)	-0.273 (0.297)	3.403 (0.183)	0.054 (0.168)	0.014 (0.354)	-1.127 (0.404)	-0.850 (0.350)	-1.065 (0.324)	-1.249 (0.272)	-1.060 (0.341)
Item 2	-0.769 (0.098)	-1.233 (0.297)	-0.615 (0.183)	-0.912 (0.168)	-1.155 (0.354)	-1.084 (0.404)	-1.284 (0.350)	-0.642 (0.324)	-1.061 (0.272)	-1.010 (0.341)
Item 3	-0.609 (0.098)	-0.524 (0.297)	0.173 (0.183)	-0.557 (0.168)	-1.173 (0.354)	-1.055 (0.404)	-1.123 (0.350)	-0.754 (0.324)	-0.939 (0.272)	-0.897 (0.341)
Item 4	-0.435 (0.098)	-0.436 (0.297)	0.544 (0.183)	-0.713 (0.168)	-0.475 (0.354)	-0.647 (0.404)	0.098 (0.350)	-1.101 (0.324)	-1.230 (0.272)	-0.721 (0.341)
Item 5	-0.382 (0.098)	-0.689 (0.297)	0.973 (0.184)	-0.566 (0.168)	-0.800 (0.354)	-1.232 (0.404)	-1.075 (0.350)	-0.809 (0.324)	-1.190 (0.272)	-0.990 (0.341)
Item 6	-0.869 (0.098)	-1.156 (0.297)	-0.353 (0.184)	-0.844 (0.168)	-0.468 (0.354)	-1.071 (0.404)	-1.122 (0.350)	-0.936 (0.324)	-0.825 (0.272)	-0.890 (0.341)
Item 7	-0.691 (0.098)	0.296 (0.297)	0.875 (0.184)	-0.194 (0.168)	-0.344 (0.354)	-1.049 (0.404)	0.870 (0.350)	0.185 (0.324)	-0.729 (0.272)	-0.462 (0.341)
Item 8	-0.446 (0.098)	-0.217 (0.297)	0.121 (0.184)	-0.319 (0.168)	-0.810 (0.354)	-1.392 (0.404)	-0.331 (0.350)	-0.729 (0.324)	-0.765 (0.272)	-1.068 (0.341)
Item 9	-0.812 (0.098)	0.121 (0.297)	-0.092 (0.185)	-0.309 (0.168)	-0.859 (0.354)	-1.017 (0.404)	-0.349 (0.350)	-0.453 (0.324)	-1.018 (0.272)	-0.845 (0.341)
Item 10	0.543 (0.098)	1.671 (0.297)	4.176 (0.185)	0.126 (0.168)	2.196 (0.354)	0.096 (0.404)	1.490 (0.350)	0.273 (0.324)	-0.753 (0.272)	0.197 (0.341)
Item 11	-0.811 (0.098)	0.232 (0.297)	-0.180 (0.185)	-0.942 (0.168)	-0.894 (0.354)	-1.102 (0.406)	-0.857 (0.350)	-1.261 (0.324)	-0.613 (0.272)	-0.741 (0.341)
Item 12	-0.258 (0.098)	0.856 (0.297)	2.126 (0.185)	0.222 (0.168)	1.006 (0.354)	0.124 (0.404)	0.581 (0.350)	0.017 (0.324)	-0.337 (0.272)	-0.132 (0.341)

Note. UAE = United Arab Emirates; USA = United States of America.

Table S6
Measurement Invariance Results for Student Sample

	Chi-Square (*df*)	CFI	RMSEA	SRMR	ΔChi-square (*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Intercept Freed
Chile vs Italy									
Configural	186.43*** (106)	.936	.060	.053					
Metric	215.81*** (116)	.921	.064	.069	30.57*** (10)	-.015	+.004	+.016	
Scalar	268.09*** (126)	.887	.074	.074	54.70*** (10)	-.034	+.010	+.005	
Partial Scalar	233.07*** (124)	.914	.065	.072	17.26* (8)	-.014	+.001	+.003	Items 4, 5
Chile vs UAE									
Configural	209.75*** (106)	.935	.062	.053					
Metric	244.33*** (116)	.920	.066	.066	36.96*** (10)	-.015	+.004	+.013	
Scalar	333.36*** (126)	.870	.080	.078	91.89*** (10)	-.050	+.014	+.012	
Partial Scalar	298.77*** (125)	.891	.074	.070	54.44*** (9)	-.029	+.008	+.004	Item 11
Chile vs USA									
Configural	165.02*** (106)	.958	.053	.049					
Metric	175.91*** (116)	.958	.051	.055	10.13 (10)	-.000	+.002	+.006	
Scalar	251.44*** (126)	.911	.071	.076	77.92*** (10)	-.047	+.020	+.021	
Partial Scalar	192.01*** (122)	.951	.054	.059	16.10* (6)	-.007	+.003	+.004	Items 2, 6, 8, 11
Italy vs UAE									
Configural	258.77*** (106)	.882	.073	.060					
Metric	266.81*** (116)	.884	.069	.063	6.41 (10)	+.002	-.004	+.003	
Scalar	337.57*** (126)	.837	.079	.073	76.98*** (10)	-.047	+.010	+.010	
Partial Scalar	295.96*** (123)	.867	.072	.065	29.15*** (7)	-.017	+.003	+.002	Items 8, 11, 12
Italy vs USA									
Configural	214.00*** (106)	.901	.069	.059					
Metric	235.80*** (116)	.890	.070	.069	21.88* (10)	-.011	+.001	+.010	
Scalar	321.75*** (126)	.820	.086	.083	90.86*** (10)	-.070	+.016	+.014	
Partial Scalar	274.12*** (123)	.861	.076	.073	38.32*** (7)	-.029	+.006	+.004	Items 7, 8, 12
UAE vs USA									
Configural	237.80*** (106)	.909	.069	.058					
Metric	264.98*** (116)	.898	.070	.068	27.95** (10)	-.011	+.001	+.010	
Scalar	292.93*** (126)	.885	.071	.070	28.33** (10)	-.013	+.001	+.002	N/A

Note. *df* = degrees of freedom; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; (P) = partial invariance; otherwise, "Yes" indicates that full invariance was established; UAE = United Arab Emirates; USA = United States of America. *$p < .05$ **$p < .01$ ***$p < .001$

Table S7
Measurement Invariance Results

	Chi-Square (*df*)	CFI	RMSEA	SRMR	ΔChi-square (*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariant?
Basic Model									
Chile (*n* = 458)	93.71*** (53)	.977	.041	.035					
Italy (*n* = 411)	159.71*** (53)	.902	.070	.052					
UAE (*n* = 462)	139.94*** (53)	.929	.060	.048					
Sample									
Chile (*ns* = 191 from Student; 267 from Community)									
Configural	149.51** (106)	.976	.042	.042					
Metric	171.53*** (116)	.970	.046	.057	22.91* (10)	-.006	+.004	+.015	Yes
Scalar	203.84*** (126)	.958	.052	.056	35.77*** (10)	-.012	+.006	-.001	Yes
Italy (*ns* = 224 from Student; 187 from Community)									
Configural	226.88*** (106)	.885	.074	.059					
Metric	251.71*** (116)	.871	.075	.072	25.28** (10)	-.014	+.001	+.013	Yes
Scalar	286.37*** (126)	.848	.079	.076	36.07*** (10)	-.023	+.004	+.004	Yes
UAE (*ns* = 320 from Student; 142 from Community)									
Configural	189.66*** (106)	.927	.058	.055					
Metric	199.61*** (116)	.928	.056	.058	8.28 (10)	+.001	-.002	+.003	Yes
Scalar	220.74*** (126)	.918	.057	.062	21.53* (10)	-.010	+.001	+.004	Yes

Note. df = degrees of freedom; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; UAE = United Arab Emirates.
*$p < .05$ **$p < .01$ ***$p < .001$

Table S8
*Correlation coefficients representing the relationship between
latent self- and interpersonal functioning impairment factors*

Sample	Correlation Coefficient
Student	
Chile (*n* = 191)	.833
Italy (*n* = 224)	.688
UAE (*n* = 320)	.763
USA (*n* = 201)	.738
Community	
Canada (*n* = 2,505)	.837
Chile (*n* = 267)	.827
Denmark (*n* = 713)	.820
Germany (*n* = 849)	.772
Italy (*n* = 187)	.763
UAE (*n* = 142)	.899
Combined	
Student (*n* = 936)	.753
Community (*n* = 4,663)	.821
Full Sample (*n* = 5,599)	.807

Note. All correlations significant at $p < .001$; UAE = United
Arab Emirates; USA = United States of America.