Original Article

# Longitudinal Measurement Invariance of the DSM-5 Anxiety and Depression Severity Measures

Verónica Vidal-Arenas[1], Adrian J. Bravo[2], Jordi Ortet-Walker[1], Generós Ortet[1,3], Manuel I. Ibáñez[1,3], and Laura Mezquita[1,3]

[1]Department of Basic and Clinical Psychology and Psychobiology, Universitat Jaume I, Castelló de la Plana, Castellón, Spain
[2]Department of Psychological Sciences, William & Mary, Williamsburg, VA, USA
[3]Centre for Biomedical Research Network on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Castellón de la Plana, Spain

**Abstract.** There is scarce evidence on the psychometric properties of the Severity Measures (SMs) of the DSM-5 longitudinally, especially among Spanish young adults. Thus, the main objective of the present study was to analyze the Longitudinal Measurement Invariance (LMI) of the SMs for Depression, Generalized Anxiety, Social Anxiety, Panic, Separation Anxiety, Agoraphobia, and Specific Phobia scales in four assessment waves at 6-month intervals. A total of 564 Spanish young adults ($M_{age}$ = 21.5 years; 67.9% females) participated in the study. Confirmatory Factor Analyses (CFAs) were carried out at each wave. Once the adequacy of the factor structure of the models was confirmed, three levels of LMI were examined (i.e., configural, metric, and scalar). Finally, we examined internal consistency (i.e., Cronbach's α and ω) across assessment waves and gender groups. CFAs analyses supported a two-factor solution for the specific phobia scale and a one-factor solution for the rest of the SMs in baseline models. Evidence of LMI at the configural, metric, and scalar levels was observed for all seven SMs. Reliability indices ranged from .73 to .96. These results suggest that the DSM-5 SMs are useful assessment tools for longitudinal and follow-up studies in Spanish young adults.

**Keywords:** DSM-5 severity measures, depression, anxiety, longitudinal measurement invariance


The categorical diagnostic classification system for mental health disorders has been questioned for more than 10 years due to a lack of empirical data supporting the fundamental assumption that psychopathology refers to discrete phenomena (Kotov et al., 2021). For example, studies have consistently found evidence of continuity between psychopathological symptoms and normality (Haslam et al., 2020). Further, clinical features, and not etiological assumptions, tend to define the criteria evaluation system (Brown & Barlow, 2005; Krueger et al., 2018). To this end, even though the categories make clinical decisions easier, they only do so in presence versus absence terms (Krueger et al., 2018), which is problematic given that numerous threshold problems have been identified. For example, high rates of *Not Otherwise Specified* diagnoses have been applied (Kotov et al., 2021). Moreover, the temporal stability of taxometric diagnoses is very low (Bromet et al., 2011), thereby leading to a loss of clinical information and reliability (MacCallum et al., 2002; Markon et al., 2011). Consequently, the DSM-5 Task Force outlined the need to consider the dimensional approach of psychopathology while revising the new edition of the DSM (Kraemer, 2007). Although the categorical system was retained within the latest edition (i.e., the DSM-5), a new section provides several dimensional assessment tools (American Psychiatric Association [APA], 2013).

DSM-5 Section III comprises two distinct groups of measures: (1) Self-Rated Cross-Cutting Symptom Measures, which assess symptoms across diagnostic categories (see Narrow et al., 2013 for further details); and (2) Severity Measures (SMs), which assess disorder-specific symptoms, such anxiety-specific problems (social anxiety, agoraphobia, specific phobia, separation anxiety, panic, generalized anxiety disorder), depression, dissociative symptoms, and two

measures for problems related to stress (posttraumatic and acute stress symptoms). The SMs were designed as short self-reported measures to be administrated both at an initial intake interview and over time to track the severity of the individual's disorder and response to treatment (APA, 2013). Considering their short design (i.e., up to 10 items), the DSM-5 SMs could be useful to cover different psychopathological conditions related to internalizing symptoms in a short evaluation period. This point is especially relevant in some at-risk groups, such as college students, not only due to their high prevalence of distress problems but also their associations with, for example, suicidal thoughts and behaviors (Auerbach et al., 2019; Bravo et al., 2018). Thus, providing short self-reported measures, based on the DSM-5, to cover different psychopathological conditions related to distress (i.e., anxiety-related problems and depression SMs) could be of great utility. For example, a short assessment could help to reduce time spent on assessing individuals (National Guidelines Alliance, 2016), thereby reducing the waiting lists for psychological assistance services, a common barrier identified in undergraduate students to participate in psychological treatments (Vidourek et al., 2014).

Previous studies conducted across various countries have provided reliability and validity evidence of the anxiety SMs scores among samples of non-clinical young adults (US, LeBeau et al., 2012; Germany, Knappe et al., 2013; Spain, Vidal-Arenas et al., 2021) non-clinical adults (Brazil, DeSousa et al., 2017; Germany, Knappe et al., 2013; The Netherlands, Möller & Bögels, 2016), and clinical populations (US, LeBeau et al., 2012, 2016); Germany, Beesdo-Baum et al., 2012, Knappe et al., 2013). These studies have found evidence to support the unidimensionality for most of the anxiety SMs (DeSousa et al., 2017; Vidal-Arenas et al., 2021; Yalın Sapmaz et al., 2017). One exception is the Specific Phobia domain, which incorporates an anxiety and avoidance subscale (Vidal-Arenas et al., 2021). All these prior studies provided reliable evidence of its scores (Cronbach's α from .82 to .98) (LeBeau et al., 2012; Vidal-Arenas et al., 2021).

The SM for assessing depressive symptoms incorporated in the new section of the DSM-5 (APA, 2013) is an adaptation from the 9-item Patient Health Questionnaire (PHQ-9; Löwe et al., 2004). Some studies have found a two-factor solution in which a somatic factor and a cognitive-affective factor are differentiated (Beard et al., 2016; Chilcot et al., 2013; Guo et al., 2017; Keum et al., 2018; Petersen et al., 2015). This two-factor solution is invariant across genders (Petersen et al., 2015) and time (Guo et al., 2017). However, specific items that load onto each factor differ across these prior studies (Beard et al., 2016; Chilcot et al., 2013; Guo et al., 2017; Keum et al., 2018). In combination with the fact that both factors are highly correlated (from .85 to .97; Beard et al., 2016; González-Blanch et al., 2018; Keum et al., 2018), some researchers suggest that a one-factor solution could better depict the structure of the PHQ-9 (Boothroyd et al., 2019). To this end, there is evidence of a one-factor solution among clinical (González-Blanch et al., 2018) and general populations (Kocalevent et al., 2013). This unidimensional structure has also been shown to be invariant across gender and differing age groups among Chinese adolescents (Leung et al., 2020).

Although previous studies have provided some preliminary validity and reliability evidence of the DSM-5 SMs scores, limited research has explored their psychometric properties over time. This is an important gap, as the dimensional measures of the DSM-5 were mainly created to circumvent the problem of the categorical diagnosis structure to detect changes in psychopathology across time when patients are attending therapy (APA, 2013), among other purposes.

For these reasons, the present study aimed to examine the psychometric properties of Depression, Generalized Anxiety, Social Anxiety, Panic, Separation Anxiety, Agoraphobia, and Specific Phobia scales from the DSM-5 over time in a sample of undergraduate students from Spain. Specifically, we examined the Longitudinal Measurement Invariance (LMI) of each SM across four assessment waves, and we provide reliable evidence of its scores at each wave. We expected that the structure of the SMs would be invariant over time, showing evidence of reliability at each assessment wave.

# Methods

## Participants and Procedure

College students from a university located in the eastern region of Spain participated in an 18-month longitudinal-funded project (code. UJI-A2017-18, Universitat Jaume I) regarding personal mental health (i.e., internalizing and externalizing behaviors), personality traits, and contextual risk factors. Through online surveys, four waves of data were collected at 6-month intervals from early February 2018 to early December 2019, using the Qualtrics platform. Participants completed informed consents and received financial compensation at each assessment wave (i.e., €5, €10, €15, and €15 at Time 1 [T1], 2 [T2], 3 [T3], and 4 [T4], respectively). Given the aim of the present study, only data from participants that completed the DSM-5 SMs were analyzed (T1, $n$ = 564; T2, $n$ = 362; T3, $n$ = 301; T4, $n$ = 279). Negligible differences in magnitude were observed in terms of age between the original and final sample ($d$ < .07). Participants were primarily females (T1 = 68.3%; T2 = 71.2%; T3 = 71.8%; T4 = 73.9%), with a mean

age of 21.6 (*SD* = 3.65). Significant differences, but small in magnitude (φ = .12), between those who completed all waves compared with those who only completed the first wave were observed in terms of gender groups.

## Instruments

### Depression Severity Measure

Adapted from the Patient Health Questionnaire short version (Löwe et al., 2004), this measure assesses depressive symptoms on a 4-point response scale (0 = *never*, 3 = *every day*) based on 9 items (APA, 2013). We made some modifications to the Spanish version used (APA, 2014), such that: (1) we changed the time frame of assessment to report the symptoms experienced "in the last 6 months" instead of "the last 7 days", and we also (2) adapted the statement for each scale to an online assessment format.

### Anxiety Severity Measures

The Spanish version of generalized anxiety, social anxiety, panic, separation anxiety, agoraphobia, and specific phobia SMs were included in the present study (Vidal-Arenas et al., 2021). Each scale is composed of 10 items, which are rated on a 5-point scale from 0 = *never* to 4 = *always*.

## Data Analysis

Before running Longitudinal Measurement Invariance (LMI) analyses, we examined the unidimensional structure for each measure at each wave using Confirmatory Factor Analyses (CFA), except for specific phobia, where we tested a two-factor structure. To evaluate overall model fit, we used the following criteria: Comparative Fit Index (CFI) > .90 (acceptable) > .95 (optimal), Tucker-Lewis Index (TLI) > .90 (acceptable) > .95 (optimal), Root Mean Square Error of Approximation (RMSEA) < .06 (Marsh et al., 2009). In addition, to examine the internal consistency of the SMs, we estimated Cronbach's αs and ordinal ωs with 95% CIs (Dunn et al., 2014) at each assessment wave.

Once the adequacy of the factor structure of the models was confirmed, we next tested the LMI for each measure. In particular, three levels of measurement invariance were tested: (1) *configural* (test whether all items load on the proposed factor), (2) *metric* (test whether item-factor loadings are similar across time), and (3) *scalar* (test whether the unstandardized item thresholds are similar across time). To indicate a significant decrement in fit when testing for measurement invariance, we used model comparison criteria of ΔCFI/ΔTLI $\geq$ .01 (i.e., decrease indicates worse fit; Cheung & Rensvold, 2002) and ΔRMSEA $\geq$ .015 (i.e., increase indicates worse fit; Chen, 2007). For CFA and LMI, due to the non-normality observed in the data and the sample size (Li, 2016), a diagonally weighted least squares (WLSMV) model estimator was used.

Analyses were performed using Mplus 8.4, and SPSS v. 25. All data, analysis code, and research materials are available at DOI: 10.17605/OSF.IO/JZ4GE.

# Results

## Structure Validity Evidence and Reliability

Results from the CFAs of all SMs across the different waves of assessment are presented in Table 1. Overall, CFA analyses supported a two-factor solution for the specific phobia scale and a one-factor solution for the rest of the SMs at baseline models with acceptable to optimal fit indices (CFIs $\geq$ .927; TLIs $\geq$ .904; RMSEAs $\leq$ .177). The mean total score and internal consistency indexes for each DSM-5 severity measure at each time point (and across gender groups) are presented in Table 2. Cronbach's αs and ordinal ωs ranged from .73 to .96. Following the cut-offs provided (APA, 2013), higher distress-related symptoms were observed (i.e., 64,4% and 52% reported medium to moderate depressive and generalized anxiety symptoms) compared with fears related problems (e.g., mild-severe 25.7% panic symptoms) at the intake (see Supplemental Table 5 in the OSF deposit). Normative data across each DSM scale are available in Supplemental Tables 1–4 at DOI: 10.17605/OSF.IO/JZ4GE.

## Longitudinal Measurement Invariance

Once the structure of the SMs was established at each wave, further examination of LMI was performed (Table 1). Results showed good fit indexes for all the configurable models of the DSM-5 SMs (CFIs $\geq$ .931; TLIs $\geq$ .926; RMSEAs $\leq$ .088). When the constraints of the factor loading across waves were added, good fit indexes (CFIs $\geq$ .940; TLIs $\geq$ .938; RMSEAs $\leq$ .082) and an improvement of CFIs, TLIs, and RMSEAs compared with the previous model (i.e., configured) were found, which suggested metric invariance. The addition of constraints between the thresholds across the different assessment points of each scale also provided good fit indexes (CFIs $\geq$ .937; TLIs $\geq$ .938; RMSEAs $\leq$ .081) and negligible differences among CFI/TLI/RMSEA, suggesting scalar invariance.

# Discussion

The Severity Measures (SMs) from the DSM-5 to assess symptoms of depressive, generalized anxiety, social anxiety,

**Table 1.** Confirmatory Factor Analyses (CFAs) fit indexes and Longitudinal Measurement Invariance (LMI) testing of DSM-5 Severity Measures

| | Overall fit indices | | | | | Comparative fit indices | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\chi^2$ | *df* | CFI | TLI | RMSEA [90% CI] | Model comparison | ΔCFI | ΔTLI | ΔRMSEA |
| Depression | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 195.194* | 27 | .948 | .931 | .105 [.092, .119] | | | | |
| T2 | 130.884* | 27 | .959 | .946 | .104 [.086, .122] | | | | |
| T3 | 107.814* | 27 | .977 | .970 | .102 [.082, .122] | | | | |
| T4 | 128.923* | 27 | .976 | .968 | .117 [.097, .138] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,297.240* | 588 | .931 | .927 | .046 [.043, .050] | | | | |
| 2. Metric | 1,229.576* | 612 | .940 | .939 | .042 [.039, .046] | 1 vs. 2 | .009 | .012 | .004 |
| 3. Scalar | 1,286.366* | 639 | .937 | .938 | .042 [.039, .046] | 2 vs. 3 | −.003 | −.001 | .000 |
| Generalized anxiety | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 359.106* | 35 | .949 | .934 | .128 [.116, .140] | | | | |
| T2 | 224.414* | 35 | .963 | .953 | .123 [.108, .138] | | | | |
| T3 | 186.761* | 35 | .962 | .949 | .122 [.106, .140] | | | | |
| T4 | 137.655* | 35 | .975 | .967 | .103 [.086, .122] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,367.630* | 734 | .958 | .955 | .039 [.036, .042] | | | | |
| 2. Metric | 1,319.186* | 761 | .963 | .962 | .036 [.033, .039] | 1 vs. 2 | .005 | .007 | −.004 |
| 3. Scalar | 1,454.584* | 791 | .956 | .956 | .039 [.035, .041] | 2 vs. 3 | −.007 | −.005 | .003 |
| Social anxiety | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 315.375* | 35 | .971 | .962 | .118 [.106, .130] | | | | |
| T2 | 164.571* | 35 | .977 | .970 | .102 [.086, .117] | | | | |
| T3 | 122.967* | 35 | .981 | .975 | .093 [.076, .111] | | | | |
| T4 | 110.970* | 35 | .988 | .985 | .089 [.072, .108] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,193.164* | 734 | .976 | .975 | .033 [.030, .037] | | | | |
| 2. Metric | 1,170.140* | 761 | .979 | .978 | .031 [.027, .034] | 1 vs. 2 | .002 | .004 | −.003 |
| 3. Scalar | 1,323.302* | 791 | .972 | .973 | .035 [.031, .038] | 2 vs. 3 | −.006 | −.006 | .004 |
| Panic | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 338.337* | 35 | .977 | .971 | .124 [.112, .136] | | | | |
| T2 | 254.134* | 35 | .979 | .973 | .132 [.117, .148] | | | | |
| T3 | 115.925* | 35 | .991 | .989 | .089 [.072, .108] | | | | |
| T4 | 49.970* | 35 | .999 | .998 | .040 [.004, .063] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 982.202* | 734 | .992 | .991 | .024 [.020, .028] | | | | |
| 2. Metric | 995.111* | 761 | .992 | .992 | .023 [.019, .027] | 1 vs. 2 | .000 | .001 | −.001 |
| 3. Scalar | 1,060.554* | 791 | .991 | .991 | .025 [.021, .028] | 2 vs. 3 | −.001 | .000 | .001 |

(Continued on next page)

panic, separation anxiety, agoraphobia, and specific phobia symptoms were created, among other purposes, to help clinicians in detecting changes in psychopathology across time. However, limited research has studied the longitudinal measurement invariance of the scales, which is a necessary step before using the measures in follow-up assessments. Thus, the aims of the present study were: (1) to test the structure of the SMs across different assessment waves and provide reliable evidence of its scores, (2) to examine the Longitudinal Measurement Invariance (LMI) of the SMs by examining three invariance levels (i.e., configural, metric, and scalar).

When the structure of the SMs was tested independently at each wave, acceptable-to-optimal fit indexes were

**Table 1.** (Continued)

| | Overall fit indices | | | | | Comparative fit indices | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\chi^2$ | *df* | CFI | TLI | RMSEA [90% CI] | Model comparison | ΔCFI | ΔTLI | ΔRMSEA |
| Separation anxiety | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 350.109* | 35 | .954 | .941 | .126 [.115, .139] | | | | |
| T2 | 200.652* | 35 | .967 | .958 | .115 [.100, .131] | | | | |
| T3 | 152.970* | 35 | .977 | .971 | .098 [.081, .116] | | | | |
| T4 | 82.945* | 35 | .992 | .989 | .071 [.052, .091] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,193.917* | 734 | .968 | .966 | .033 [.030, .037] | | | | |
| 2. Metric | 1,172.824* | 761 | .971 | .971 | .031 [.027, .034] | 1 vs. 2 | .004 | .005 | −.002 |
| 3. Scalar | 1,255.876* | 791 | .968 | .969 | .032 [.029, .035] | 2 vs. 3 | −.003 | −.001 | .001 |
| Agoraphobia | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 359.938* | 35 | .966 | .956 | .130 [.118, .142] | | | | |
| T2 | 135.529* | 35 | .985 | .981 | .089 [.074, .106] | | | | |
| T3 | 133.198* | 35 | .982 | .977 | .099 [.081, .117] | | | | |
| T4 | 107.721* | 35 | .989 | .986 | .087 [.069, .106] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,115.955* | 734 | .981 | .979 | .030 [.027, .034] | | | | |
| 2. Metric | 1,095.392* | 761 | .983 | .983 | .028 [.024, .032] | 1 vs. 2 | .002 | .003 | −.002 |
| 3. Scalar | 1,181.447* | 791 | .980 | .980 | .030 [.026, .033] | 2 vs. 3 | −.002 | −.002 | .001 |
| Specific phobia | | | | | | | | | |
| CFA | | | | | | | | | |
| T1 | 406.400* | 34 | .945 | .927 | .139 [.127, .152] | | | | |
| T2 | 414.827* | 34 | .928 | .904 | .177 [.162, .192] | | | | |
| T3 | 274.048* | 34 | .948 | .931 | .156 [.139, .174] | | | | |
| T4 | 174.123* | 34 | .971 | .961 | .123 [.105, .141] | | | | |
| LMI | | | | | | | | | |
| 1. Configural | 1,580.164* | 712 | .937 | .931 | .046 [.043, .050] | | | | |
| 2. Metric | 1,524.136* | 736 | .943 | .940 | .044 [.040, .047] | 1 vs. 2 | .006 | .008 | −.003 |
| 3. Scalar | 1,590.831* | 766 | .940 | .939 | .044 [.041, .047] | 2 vs. 3 | −.002 | .001 | −.001 |

*Note*. CFAs = Confirmatory Factor Analyses; LMI = Longitudinal Measurement Invariance. *$p < .001$.

observed, supporting the hypothesized two-factor solution for the specific phobia SM and the one-factor solution for the rest of the anxiety SMs (DeSousa et al., 2017; Vidal-Arenas et al., 2021) and the depression SM (González-Blanch et al., 2018; Kocalevent et al., 2013; Leung et al., 2020). The only fit index that was over the standard cut-offs was the RMSEA (Marsh et al., 2009). However, this result could be expected due to the non-normality distribution observed of the item scores from each scale (skewness $\geq 1.5$; kurtosis $\geq 3.0$) and sample size ($n \geq 500$) of the present study (Li, 2016) and it also similar to those found in previous studies about the structure of the DSM-5 SMs (DeSousa et al., 2017; Vidal-Arenas et al., 2021). For each SMs at each wave, adequate reliability coefficients were observed (> .73). Altogether, these results provide new evidence on the structure and reliability of the scores of the Spanish version of the DSM-5 SMs among Spanish undergraduates.

Once the structure of the scales was confirmed cross-sectionally, LMI was tested. The LMI analyses revealed that the factor solution tested for SMs (i.e., one-factor solution for depressive, generalized anxiety, social anxiety, panic, separation anxiety, and agoraphobia symptoms scales, and two-factor solution for the specific phobia symptoms scale) were invariant across time, such that all items loaded on the proposed factor (i.e., configural-invariance), item-factor loadings were similar across time (i.e., metric-invariance) and the unstandardized item thresholds were similar across time (i.e., scalar-invariance). Then, these findings support the utility of the DSM-5 anxiety and depression SMs for follow-up assessments.

The present study has some limitations that should be considered. The first limitation is that our participants were a small sample of undergraduates from Spain; therefore, future studies among other populations (e.g., clinical samples), and larger sample sizes are needed to investigate

**Table 2.** Descriptive statistics and reliability coefficients

| | Depression | Generalized Anxiety | Social anxiety | Separation anxiety | Panic | Agoraphobia | Specific phobia | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Anxiety | Avoidance |
| T1 | | | | | | | | |
| Cronbach's α (CI) | .83 (.809, .851) | .90 (.891, .915) | .92 (.911, .931) | .90 (.885, .910) | .93 (.924, .941) | .92 (.904, .925) | .88 (.861, .892) | .82 (.795, .844) |
| Ordinal ω (CI) | .84 (.808, .859) | .91 (.887, .921) | .93 (.910, .937) | .90 (.880, .919) | .93 (.914, .948) | .92 (.900, .930) | .88 (.861, .904) | .83 (.794, .856) |
| Mean (*SD*) | | | | | | | | |
| Total | 7.05 (4.30) | 7.22 (6.10) | 6.63 (6.56) | 4.16 (5.29) | 3.45 (5.79) | 4.16 (5.31) | 4.35 (4.56) | 3.93 (3.70) |
| Male | 7.17 (4.34) | 6.58 (5.80) | 6.62 (6.29) | 4.43 (5.69) | 3.64 (5.96) | 4.34 (5.35) | 3.62 (4.28) | 3.50 (3.58) |
| Female | 6.99 (4.28) | 7.52 (6.22) | 6.64 (6.68) | 4.03 (5.09) | 3.35 (5.72) | 4.07 (5.39) | 4.70 (4.65) | 4.13 (3.74) |
| T2 | | | | | | | | |
| Cronbach's α (CI) | .86 (.835, .879) | .91 (.896, .924) | .92 (.902, .928) | .92 (.909, .933) | .93 (.917, .939) | .93 (.916, .938) | .89 (.867, .904) | .82 (.787, .848) |
| Ordinal ω (CI) | .86 (.832, .888) | .91 (.893, .930) | .93 (.910, .937) | .92 (.896, .943) | .93 (.911, .948) | .93 (.910, .949) | .89 (.859, .915) | .84 (.780, .870) |
| Mean (*SD*) | | | | | | | | |
| Total | 6.20 (4.23) | 5.81 (5.87) | 5.48 (5.95) | 3.67 (5.39) | 2.91 (5.34) | 3.29 (5.14) | 3.97 (4.46) | 3.54 (3.65) |
| Male | 6.16 (4.32) | 4.78 (5.47) | 5.25 (5.93) | 4.13 (5.79) | 2.71 (5.13) | 3.28 (5.20) | 3.98 (5.08) | 3.29 (3.74) |
| Female | 6.22 (4.20) | 6.24 (5.99) | 5.57 (5.98) | 3.47 (5.22) | 2.99 (5.44) | 3.30 (5.12) | 3.96 (4.19) | 3.64 (3.62) |
| T3 | | | | | | | | |
| Cronbach's α (CI) | .89 (.868, .906) | .90 (.886, .918) | .92 (.903, .931) | .90 (.886, .919) | .93 (.923, .945) | .92 (.906, .933) | .84 (.810, .867) | .73 (.669, .773) |
| Ordinal ω (CI) | .89 (.863, .914) | .91 (.881, .926) | .92 (.895, 938) | .91 (.874, .928) | .94 (.914, .953) | .92 (.896, .941) | .87 (.829, .903) | .83 (.785, .869) |
| Mean (*SD*) | | | | | | | | |
| Total | 6.59 (4.80) | 5.50 (5.72) | 4.82 (5.62) | 2.98 (4.40) | 2.42 (4.61) | 3.14 (4.77) | 3.64 (3.95) | 3.16 (3.34) |
| Male | 6.56 (5.23) | 4.95 (5.68) | 5.10 (5.95) | 3.89 (5.23) | 3.21 (5.67) | 4.15 (5.91) | 3.23 (4.27) | 2.87 (3.09) |
| Female | 6.61 (4.64) | 5.71 (5.70) | 4.72 (5.51) | 2.63 (4.00) | 2.12 (4.12) | 4.77 (4.22) | 3.80 (3.83) | 3.27 (3.43) |
| T4 | | | | | | | | |
| Cronbach's α (CI) | .88 (.853, .896) | .92 (.899, .929) | .93 (.922, .945) | .92 (.901, .930) | .95 (.944, .960) | .93 (.913, .939) | .87 (.841, .890) | .88 (.852, .899) |
| Ordinal ω (CI) | .88 (.855, .904) | .92 (.888, .937) | .94 (.917, .954) | .92 (.894, .937) | .96 (.941, .968) | .93 (.907, .950) | .87 (.829, .905) | .88 (.831, .912) |
| Mean (*SD*) | | | | | | | | |
| Total | 5.83 (4.36) | 5.37 (5.77) | 4.17 (5.80) | 2.86 (4.51) | 2.21 (4.92) | 2.92 (4.96) | 3.55 (3.92) | 2.95 (3.53) |
| Male | 5.19 (4.51) | 5.16 (6.05) | 4.74 (6.44) | 2.99 (5.15) | 2.48 (5.47) | 3.42 (5.70) | 3.33 (4.52) | 2.58 (3.51) |
| Female | 6.06 (4.29) | 5.44 (5.68) | 3.97 (5.55) | 2.81 (4.27) | 2.11 (4.72) | 2.75 (4.68) | 3.63 (3.70) | 3.08 (3.53) |

the replicability and generalizability of our findings. Moreover, further longitudinal studies should consider the inclusion of time-invariant and time-varying variables (e.g., SES), due to its effects in between and within-individual change in psychopathology (Wickrama et al., 2016), both psychological (e.g., personality), personal (e.g., clinical antecedents) and sociocontextual (e.g., ethnicity/race).

Despite these limitations, the present study provides evidence for the structure, reliability, and longitudinal measurement invariance of the Spanish DSM-5 anxiety and depression SMs. These results suggest that the DSM-5 SMs are useful tools to dimensionally assess the symptoms related to depression and anxiety DSM-5 disorder over time. The advantages of this are clear both at a clinical level, reducing the time spent on assessing individuals (National Guidelines Alliance, 2016) and facilitating improved monitoring of these symptoms across different phases of treatment; and at the research level, by decreasing participant fatigue in longitudinal studies, and helping the researcher cover a significant number of internalizing symptoms briefly.

## References


American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. APA.

American Psychiatric Association. (2014). *Diagnostic and statistical manual of mental disorders* (5th ed.). Editorial Médica Panamérica.

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Lee, S., Lochner, C., McLafferty, M., Nock, M. K., Petukhova, M. V., Pinder-Amaker, S., Rosellini, A. J., Sampson, N. A., Vilagut, G., … Kessler, R. C. (2019). Mental disorder comorbidity and suicidal thoughts and behaviors in the World Health Organization World Mental Health Surveys International College Student initiative. *International Journal of Methods in Psychiatric Research, 28*(2), Article e1752. https://doi.org/10.1002/mpr.1752

Beard, C., Hsu, K. J., Rifkin, L. S., Busch, A. B., & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders, 193*, 267–273. https://doi.org/10.1016/J.JAD.2015.12.075

Beesdo-Baum, K., Klotsche, J., Knappe, S., Craske, M. G., Lebeau, R. T., Hoyer, J., Strobel, A., Pieper, L., & Wittchen, H. U. (2012). Psychometric properties of the dimensional anxiety scales for DSM-V in an unselected sample of German treatment seeking patients. *Depression and Anxiety, 29*(12), 1014–1024. https://doi.org/10.1002/da.21994

Boothroyd, L., Dagnan, D., & Muncer, S. (2019). PHQ-9: One factor or two? *Psychiatry Research, 271*, 532–534. https://doi.org/10.1016/J.PSYCHRES.2018.12.048

Bravo, A. J., Villarosa-Hurlocker, M. C., Pearson, M. R., Prince, M. A., Madson, M. B., Henson, J. M., Looby, A., Gonzalez, V. M., Henslee, A. M., Cuttler, C., Wong, M. M., & McChargue, D. E. (2018). College student mental health: An evaluation of the DSM-5 self-rated Level 1 cross-cutting symptom measure. *Psychological Assessment, 30*(10), 1382–1389. https://doi.org/10.1037/PAS0000628

Bromet, E. J., Kotov, R., Fochtmann, L. J., Carlson, G. A., Tanenberg-Karant, M., Ruggero, C., & Chang, S. W. (2011). Diagnostic shifts during the decade following first admission for psychosis. *American Journal of Psychiatry, 168*(11), 1186–1194. https://doi.org/10.1176/APPI.AJP.2011.11010048/ASSET/IMAGES/LARGE/AJP_168_11_1186_F002.JPEG

Brown, T. A., & Barlow, D. H. (2005). Dimensional versus categorical classification of mental disorders in the fifth edition of the diagnostic and statistical manual of mental disorders and beyond: Comment on the special section. *Journal of Abnormal Psychology, 114*(4), 551–556.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chilcot, J., Rayner, L., Lee, W., Price, A., Goodwin, L., Monroe, B., Sykes, N., Hansford, P., & Hotopf, M. (2013). The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research, 75*(1), 60–64. https://doi.org/10.1016/J.JPSYCHORES.2012.12.012

DeSousa, D. A., Moreno, A. L., Osório, F. L., Crippa, J. A. S., LeBeau, R., Manfro, G. G., Salum, G. A., & Koller, S. H. (2017). Psychometric properties of the dimensional anxiety scales for DSM-5 in a Brazilian community sample. *International Journal of Methods in Psychiatric Research, 26*(3), Article e1531. https://doi.org/10.1002/mpr.1531

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399–412. https://doi.org/10.1111/bjop.12046

González-Blanch, C., Medrano, L. A., Muñoz-Navarro, R., Ruíz-Rodríguez, P., Moriana, J. A., Limonero, J. T., Schmitz, F., & Cano-Vindel, A. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLoS One, 13*(2), Article e0193356. https://doi.org/10.1371/JOURNAL.PONE.0193356

Guo, B., Kaylor-Hughes, C., Garland, A., Nixon, N., Sweeney, T., Simpson, S., Dalgleish, T., Ramana, R., Yang, M., & Morriss, R. (2017). Factor structure and longitudinal measurement invariance of PHQ-9 for specialist mental health care patients with persistent major depressive disorder: Exploratory Structural Equation Modelling. *Journal of Affective Disorders, 219*, 1–8. https://doi.org/10.1016/j.jad.2017.05.020

Haslam, N., McGrath, M. J., Viechtbauer, W., & Kuppens, P. (2020). Dimensions over categories: A meta-analysis of taxometric research. *Psychological Medicine, 50*(9), 1418–1432. https://doi.org/10.1017/S003329172000183X

Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse US college Students. *Psychological Assessment, 30*(8), 1096–1106. https://doi.org/10.1037/PAS0000550

Knappe, S., Klotsche, J., Strobel, A., Lebeau, R. T., Craske, M. G., Wittchen, H. U., & Beesdo-Baum, K. (2013). Dimensional anxiety scales for DSM-5: Sensitivity to clinical severity. *European Psychiatry: The Journal of the Association of European Psychiatrists, 28*(7), 448–456. https://doi.org/10.1016/j.eurpsy.2013.02.001

Kocalevent, R. D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in

the general population. *General Hospital Psychiatry, 35*(5), 551–555. https://doi.org/10.1016/J.GENHOSPPSYCH.2013.04.006

Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., Eaton, N. R., Forbes, M. K., Hallquist, M. N., Latzman, R. D., Mullins-Sweatt, S. N., Ruggero, C. J., Simms, L. J., Waldman, I. D., Waszczuk, M. A., & Wright, A. G. C. (2021). The Hierarchical Taxonomy of Psychopathology (HiTOP): A quantitative nosology based on consensus of evidence. *Annual Review of Clinical Psychology, 17*(1), 83–108. https://doi.org/10.1146/annurev-clinpsy-081219-093304

Kraemer, H. C. (2007). DSM categories and dimensions in clinical and research contexts. *International Journal of Methods in Psychiatric Research, 16*(S1), S8–S15. https://doi.org/10.1002/mpr.211

Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., Simms, L. J., Widiger, T. A., Achenbach, T. M., Bach, B., Bagby, R. M., Bornovalova, M. A., Carpenter, W. T., Chmielewski, M., Cicero, D. C., Clark, L. A., Conway, C., DeClercq, B., DeYoung, C. G., … Zimmermann, J. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry, 17*(3), 282–293. https://doi.org/10.1002/wps.20566

LeBeau, R. T., Glenn, D. E., Hanover, L. N., Beesdo-Baum, K., Wittchen, H.-U., & Craske, M. G. (2012). A dimensional approach to measuring anxiety for DSM-5. *International Journal of Methods in Psychiatric Research, 21*(4), 258–272. https://doi.org/10.1002/mpr.1369

LeBeau, R. T., Mesri, B., & Craske, M. G. (2016). The DSM-5 Social Anxiety Disorder Severity Scale: Evidence of validity and reliability in a clinical sample. *Psychiatry Research, 244*, 94–96. https://doi.org/10.1016/J.PSYCHRES.2016.07.024

Leung, D. Y. P., Mak, Y. W., Leung, S. F., Chiang, V. C. L., & Loke, A. Y. (2020). Measurement invariances of the PHQ-9 across gender and age groups in Chinese adolescents. *Asia-Pacific Psychiatry, 12*(3), Article e12381. https://doi.org/10.1111/APPY.12381

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. https://doi.org/10.3758/S13428-015-0619-7

Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders, 81*(1), 61–66. https://doi.org/10.1016/S0165-0327(03)00198-8

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin, 137*(5), 856–879. https://doi.org/10.1037/A0023678

Marsh, H. W., Hau, K. T., & Wen, Z. (2009). In search of golden rules: Comment on hypothesistesting approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341. https://doi.org/10.1207/S15328007SEM1103_2

Möller, E. L., & Bögels, S. M. (2016). The DSM-5 Dimensional Anxiety Scales in a Dutch non-clinical sample: Psychometric properties including the Adult Separation Anxiety Disorder Scale. *International Journal of Methods in Psychiatric Research, 25*(3), 232–239. https://doi.org/10.1002/mpr.1515

Narrow, W. E., Clarke, D. E., Kuramoto, J., Kraemer, H. C., Kupfer, D. J., Greiner, L., & Regier, D. A. (2013). DSM-5 field trials in the United States and Canada, Part III: Development and reliability testing of a cross-cutting symptom assessment for DSM-5. *The American Journal of Psychiatry, 170*(1), 71–82. https://doi.org/10.1176/APPI.AJP.2012.12071000

National Guidelines Alliance. (2016). *Mental Health Problems in People with Learning Disabilities: Prevention, Assessment and Management*. National Institute for Health and Care Excellence (NICE). http://www.ncbi.nlm.nih.gov/pubmed/27683921

Petersen, J. J., Paulitsch, M. A., Hartig, J., Mergenthal, K., Gerlach, F. M., & Gensichen, J. (2015). Factor structure and measurement invariance of the Patient Health Questionnaire – 9 for female and male primary care patients with major depression in Germany. *Journal of Affective Disorders, 170*, 138–142. https://doi.org/10.1016/J.JAD.2014.08.053

Vidal-Arenas, V. (2022). *Five-factor model of personality & internalising psychopathology: New contributions for evaluation and aetiological models*. (Doctoral dissertation). Universitat Jaume I.

Vidal-Arenas, V., Bravo, A. J., Ortet-Walker, J., Ortet, G., Ribes, M. I., & Guillamón, L. M. (2023). *Longitudinal examination of DSM-5 anxiety and depression scales* [Data, Materials]. https://OSF.IO/JZ4GE

Vidal-Arenas, V., Ortet-Walker, J., Ibáñez, M. I., Ortet, G., & Mezquita, L. (2021). Selfreported DSM-5 anxiety severity measures: Evidence of validity and reliability in Spanish youths. *Psicothema, 33*(2), 312–319. https://doi.org/10.7334/PSICOTHEMA2020.398

Vidourek, R. A., King, K. A., Nabors, L. A., & Merianos, A. L. (2014). Students' benefits and barriers to mental health help-seeking. *Health Psychology and Behavioral Medicine, 2*(1), 1009–1022. https://doi.org/10.1080/21642850.2014.963586

Wickrama, K. K. A. S., Lee, T. K., O'Neal, C. W., & Lorenz, F. O. (2016). *Higher-order growth curves and mixture modeling with Mplus: A practical guide*. Routledge.

Yalın Sapmaz, Ş., Ergın, D., Şen Celasın, N., Karaarslan, D., Öztürk, M., Özek Erkuran, H., Köroğlu, E., & Aydemır, Ö. (2017). Validity and reliability of the Turkish version of DSM-5 Social Anxiety Disorder Severity Scale – Child Form. *Noropsikiyatri Arsivi, 54*(4), 354–357. https://doi.org/10.5152/npa.2017.17027

**History**
Received October 3, 2022
Revision received June 26, 2023
Accepted July 6, 2023
Published online December 31, 2023
EJPA Section / Category Clinical Psychology

**Conflict of Interest**
This study contains part of the content published in the doctoral thesis that is not under embargo (Vidal-Arenas, 2022).

**Authorship**
Verónica Vidal-Arenas: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft. Adrian J. Bravo: Supervision, Writing – Review & Editing. Jordi Ortet-Walker: Writing – Review & Editing. Generós Ortet: Resources, Funding acquisition, Writing – Review & Editing. Manuel I. Ibáñez: Supervision, Funding acquisition, Writing – Review & Editing. Laura Mezquita: Conceptualization Methodology, Investigation, Funding acquisition, Writing – Review & Editing.

**Open Science**
We report how we determined our sample size, all data exclusions (if any), all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses including all tested models. If we use inferential tests, we report exact *p* values, effect sizes, and 95% confidence or credible intervals.

Open Data: We confirm that there is sufficient information for independent research to reproduce all of the reported results (Vidal-Arenas et al., 2023).
The information needed to reproduce all of the reported results are openly accessible at https://osf.io/jz4ge/?view_only=ea57acad6e6a4a6e86fa5a57caf493c9
Open Materials: We confirm that there is sufficient information for independent research to reproduce all of the reported methodology (Vidal-Arenas et al., 2023).
Preregistration of Studies and Analysis Plans: This study was not preregistered.

**Funding**
Verónica Vidal-Arenas has been supported by a grant from Universitat Jaume I (PREDOC/18/12). This study was also supported by grants UJI-B2022-29 from the Universitat Jaume I, AICO/2021/052 from the Valencian Autonomous Government, and PID2022-141808OB-I00 from the Ministry of Science and Innovation of Spain (MICINN).

**Laura Mezquita**
Department of Basic and Clinical Psychology and Psychobiology
Universitat Jaume I
Vicent Sos Baynat
12071 Castelló de la Plana
Spain
lmezquit@uji.es