



# The factor structure and measurement invariance of the PHQ-4 and the prevalence of depression and anxiety in a Southeast Asian context amid the COVID-19 pandemic

Norman B. Mendoza<sup>1</sup> · Cherry E. Frondoza<sup>1</sup> · John Ian Wilzon T. Dizon<sup>3</sup> · Jet U. Buenconsejo<sup>2</sup>

Accepted: 29 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

This study examined the psychometric properties of the Patient Health Questionnaire-4 (PHQ-4) as an ultra-brief screener of depression and anxiety in the Philippines during the COVID-19 outbreak. Data from 4,524 non-clinical community respondents aged 18-73 years old was collected online between March and July 2020. We evaluated the screener's factor structure, measurement invariance, and criterion-related validity using confirmatory factor analysis (CFA), multigroup CFA, and structural equation modeling (SEM), respectively. We also evaluated the accuracy of the PHQ-4 cut-off scores by comparing the them with the screeners' full scales (i.e., PHQ-9 and GAD-7). Using the cutoff scores of the screeners, we also estimated the prevalence rates of depression and anxiety. The PHQ-4 has good internal reliability (Cronbach's  $\alpha=0.82$ ). The CFA results show that the two-factor model has an excellent model fit that is superior to the one-factor model. The two-factor model held through increasingly constrained multi-group CFA models across gender, age, and geographical location groups, demonstrating measurement invariance. The SEM model supported the PHQ-4's theoretical association to stress, negative affect, and positive affect, supporting the screener's criterion-related validity. In estimating prevalence rates, among those screened by the PHQ-4 cut-off scores for depression ( $n = 1,905$ , 42.11%) and anxiety ( $n = 1,853$ , 40.96%), 81.78% and 94.06% were consequently screened by the PHQ-9 and GAD-7, respectively. This study supports the reliability, validity, and measurement invariance of the PHQ-4 as an ultra-brief screener of depression and anxiety in a large community sample in Southeast Asia. The inclusion of ultra-brief screeners in COVID-19-related studies and other human disasters, especially among non-clinical samples in low- and middle-income countries, is relevant for the sustainable evaluation and monitoring of the severity mental health symptoms leading to timely and effective mental health service provision.

**Keywords** Depression · Anxiety · PHQ-4 · COVID-19 · Measurement invariance · Philippines

## Introduction

Coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus-2 has brought tremendous disruptions in the lives of individuals worldwide. The most recent global estimates state that more than 127

million cases of COVID-19 have been confirmed, with more than 2.7 million recorded deaths due to the virus (World Health Organization, 2021). Although the COVID-19 has primarily affected individuals' physical health, the threat of contracting the virus, the unprecedented challenges brought about by socio-economic changes, and the public health measures implemented have triggered negative mental health outcomes such as anxiety and depression (Bendau et al., 2020; Choi et al., 2020; Lai et al., 2020; Li et al., 2020; Mendoza & Dizon, 2022; Petzold et al., 2020b; Salari et al., 2020; Schnell & Krampe, 2020; Shah et al., 2021). Despite the growing research on the mental health impact of the COVID-19 pandemic, the need for reliable and valid mental health screeners in the general population has been emphasized as one of the topmost concerns for mental health professionals to address during this pandemic (Chandu et al., 2020; Cortez et al., 2020). This is especially

✉ Norman B. Mendoza  
normanmendoza0421@gmail.com

<sup>1</sup> Department of Curriculum and Instruction, The Education University of Hong Kong, New Territories, Hong Kong, SAR, People's Republic of China

<sup>2</sup> Department of Special Education and Counselling, The Education University of Hong Kong, New Territories, Hong Kong, SAR, People's Republic of China

<sup>3</sup> Department of Psychology, College of Arts and Sciences, Angeles University Foundation, Angeles City, Philippines

true for low- and middle-income countries (e.g., Philippines; Ali et al., 2016; The World Bank, 2021).

The Patient Health Questionnaire-4 (PHQ-4) is one of the most commonly used mental health screeners for anxiety and depression (Kroenke et al., 2009). The PHQ-4 is an ultra-brief screener, consisting of two items from the PHQ-9 depression screener (i.e., PHQ-2; see Kroenke et al., 2003) and two items from the GAD-7 anxiety screener (i.e., GAD-2; see Kroenke et al., 2007). The PHQ-4 is associated with greater negative clinical outcomes (Ghaheri et al., 2020; Kroenke et al., 2009; Mills et al., 2015; Renovanz et al., 2019) and lower well-being and quality of life (Ghaheri et al., 2020; Kocalevent et al., 2014; Löwe et al., 2010; Renovanz et al., 2019).

Although the PHQ-4 has been validated in several samples globally (Cano-Vindel et al., 2018; Ghaheri et al., 2020; Khubchandani et al., 2016; Kim et al., 2021; Kocalevent et al., 2014; Kroenke et al., 2009; Löwe et al., 2010; Mills et al., 2015; Renovanz et al., 2019; Tibubos & Kröger, 2020), to date, no published study has examined the psychometric properties of PHQ-4 among community samples from the Philippines. Given that several studies have used the PHQ-4 to evaluate mental health outcomes during the COVID-19 pandemic (Bendau et al., 2020; Choi et al., 2020; Lai et al., 2020; Li et al., 2020; Petzold et al., 2020b; Schnell & Krampe, 2020), validating this instrument in the Philippines can extend its applicability as a tool for non-clinical samples and can shed light to the current state of mental health in a Southeast Asian context. The Philippines has recorded the highest number of confirmed cases (2,632,881) and recorded deaths (38,937) due to COVID-19 in the Western Pacific region (World Health Organization, 2021; as of October 10, 2021). The Philippines is also one of the countries with the longest lockdowns with stern community restrictions from March 2020 until the present (Biana & Joaquin, 2020; TIME, 2021). Recent observational studies have also pointed out the mental health burden in the Philippines during the pandemic (Bernardo et al., 2020).

Given the impact of the COVID-19 pandemic on people's mental health and owing to the importance of validating brief screeners for mental health, the present study seeks to examine the reliability, validity (i.e., construct and criterion-related validity), and measurement invariance (i.e., configural, metric, and scalar) of the PHQ-4 among a large non-clinical sample in the Philippines using data collected during the outbreak of the pandemic. Further, using the PHQ-4, this study aims to estimate the prevalence of dysfunctional levels of anxiety and depression during the outbreak of the COVID-19.

## The PHQ-4

The four items of the PHQ-4 refer to two symptoms of depression (i.e., depressed mood and loss of interest or anhedonia) and two symptoms of a generalized anxiety disorder (i.e., nervousness and uncontrollable worry). All

items are evaluated in terms of frequency in the past two weeks. The PHQ-4 has been validated in a variety of samples, including patients from primary-care clinics in the United States (Kroenke et al., 2009), patients with suspected psychological disorders from primary-care centres in Spain (Cano-Vindel et al., 2018), patients with an intracranial or brain tumour in Germany (Renovanz et al., 2019), and more recently, among infertile parents in Iran (Ghaheri et al., 2020). The PHQ-4 has also been examined among non-clinical samples from the general population to generate normative data. Examples of which include a nationally representative sample in Germany (Löwe et al., 2010), a sample of the general population in Columbia (Kocalevent et al., 2014), English- and Spanish-speaking Hispanic Americans in the U.S. (Mills et al., 2015), as well as undergraduate students in the U.S. (Khubchandani et al., 2016).

The PHQ-4's relationship with a wide range of positive and negative well-being variables has also been adequately observed and documented. Scores from the PHQ-4 are correlated with outcomes which includes low levels of self-esteem (Löwe et al., 2010), self-efficacy (Kocalevent et al., 2014), life satisfaction (Kocalevent et al., 2014; Löwe et al., 2010), resilience (Löwe et al., 2010), quality of life (Renovanz et al., 2019), well-being (Ghaheri et al., 2020), and high levels of hopelessness and distress (Kocalevent et al., 2014), and as well as perceived distress (Mills et al., 2015). The PHQ-4 has also been found to be associated with several medical-related outcomes such as functional impairment, disability days, and healthcare use (Kroenke et al., 2009), need for psycho-oncological support (Renovanz et al., 2019), and infertility duration and failure in previous infertility treatment (Ghaheri et al., 2020). Notably, anxiety and depression as measured by PHQ-4 were found to be significantly associated with an obsession with COVID-19 and Coronavirus anxiety (Choi et al., 2020).

Results of factor analytic strategies on PHQ-4 with various samples have generally supported its two-factor structure consisting of anxiety and depression and model invariance by gender and age groups (Ghaheri et al., 2020; Kocalevent et al., 2014; Kroenke et al., 2009; Löwe et al., 2010; Mills et al., 2015; Renovanz et al., 2019). A cut-off score of  $\geq 3$  in GAD-2 is reasonably sensitive in detecting generalized anxiety disorder (88%), panic disorder (76%), social anxiety disorder (70%), and posttraumatic stress disorder (59%; Kroenke et al., 2007). Further, GAD-2 was found to have 81–83% specificity among the said disorders (Kroenke et al., 2007). A cut-off score of  $\geq 3$  in PHQ-2 is 83% sensitive and 90% specific in detecting major depressive disorder (Kroenke et al., 2003). Similarly, using a computerized version of PHQ-4 among Spanish primary-care patients, Cano-Vindel and colleagues (2018) also recommended a cut-off score of 3 to obtain greater sensitivity in detecting anxiety and depression.

Overall, evidence supports the reliability and validity of the PHQ-4 as a screener for depression and anxiety in different samples. The PHQ-4 has been used for non-clinical samples both in COVID-19-related studies (e.g., Petzold et al., 2020a; Taylor et al., 2020) and otherwise (e.g., Cavanagh et al., 2018; Schmalbach et al., 2021), however, its psychometric properties including its reliability, factor structure, and measurement invariance, and criterion-related validity are yet to be examined. Moreover, previous validation studies of the PHQ-4 were focused on primarily Western and Middle Eastern countries. To the best of our knowledge, there remain no published studies that validated the PHQ-4 using a sample from the general population in Southeast Asia, particularly the Philippines, and not amid an ongoing global health crisis such as the COVID-19 pandemic.

## The Present Study

This study aims to address these research gaps by examining the psychometric properties of the PHQ-4 as an ultra-brief screener of anxiety and depression using data from a sample of Filipino adults collected in the first few months of the COVID-19 outbreak (i.e., March to July 2020). Specifically, we examined the scale's reliability and factor structure (i.e., one-factor vs. two-factor model), measurement invariance for different age, gender, and locale groups, and criterion-related validity by testing its association to other self-report measures known to be linked with depression and anxiety (i.e., stress, negative affect, and positive affect). Lastly, peripheral yet critical to the accurate prevalence estimates of depression and anxiety in the Philippines, those who met the cut-off score for the PHQ-4 subscales (i.e.,  $\geq 3$  for both PHQ-2 and GAD-2) were presented with the full PHQ-9 (depression) and GAD-7 (anxiety).

## Methods

### Participants

A total of 4,524 Filipino adults from ages 18-73 years old ( $M = 27.16$ ;  $SD = 7.61$ ) participated in the study. There were 3,382 (74.76%) female and 1,142 (25.24%) male participants. Majority of the participants were single ( $n = 3,351$ , 74.07%) while 22.72% ( $n = 1,028$ ) were married, and the rest ( $n = 145$ ) responded with "Others (e.g., widow/er, separated, annulled/divorced, etc.)". Almost half of the participants come from the capital Manila or the National Capital Region ( $n = 2,046$ ; 45.23%), followed by those coming from the capital's neighboring regions Southern Luzon ( $n = 806$ ; 17.82%) and Central Luzon ( $n = 614$ ; 13.57%).

## Procedures

The survey was administered online through a social media post in partnership with a local non-profit organization. The online survey was in English and was designed in Qualtrics, which allowed the respondents to obtain a copy of their responses to the survey, receive information on available mental health and emergency services, and learn tips on healthy coping strategies during the COVID-19 pandemic. The average completion time for the survey was 12 min. Data collection was conducted from March 2020 to July 2020, where over 86% of the participants responded in the first two months. Informed consent was sought for all participants, and those who did not consent were routed to an exit page that provides information on mental health support and linkages to care. No personal information from the respondents was collected. After participating in the survey, they were provided with links to COVID-19 information and mental health services. Participants who met the cut-off score of  $\geq 3$  for each of the PHQ-4 subscales (i.e., PHQ-2 and GAD-2) were evaluated further with the longer PHQ-9 and GAD-7. This procedure is recommended by previous PHQ-4 studies (e.g., Kroenke et al., 2003, 2007; Löwe et al., 2005, 2010) to further evaluate depression and anxiety symptom severity.

## Measures

PHQ-4. Depression and anxiety symptoms were measured using the PHQ-4 (Kroenke et al., 2009). The items for depression of the PHQ-4 are "Little interest or pleasure in doing things" and "Feeling down, depressed, or hopeless", and the items for anxiety are "Feeling nervous, anxious or on edge" and "Not being able to stop or control worrying". These items are responded to on a 4-point Likert scale, from 0 (*not at all*) to 3 (*nearly every day*). The cut-off score for the PHQ-4's subscales is greater than or equal to 3. In the present study, the internal reliability of the PHQ-4 is adequate ( $\alpha = 0.82$ ), and the reliability of its depression and anxiety subscales are  $\alpha = 0.71$  and  $\alpha = 0.83$ , respectively.

Stress subscale of the Depression, Anxiety, and Stress Scale (DASS-S). We used the 7-item stress subscale of DASS-21 (Lovibond & Lovibond, 1995). The DASS-S includes items measuring irritability, tension, and a tendency for heightened reactions to overly stressful events for a non-clinical sample (Antony et al., 1998). Participants rated the scale from 0 (*never*) to 3 (*almost always*), based on how much the items applied to them in the past week. DASS-S has adequate internal reliability in this study ( $\alpha = 0.88$ ). The construct validity of the unidimensional stress subscale ( $n = 4,036$ ) is supported by the following robust fit indices: SB  $\chi^2(14) = 330.12$ ,  $p < 0.001$ , CFI = 0.974, TLI = 0.962, RMSEA = 0.085 (C.I. = 0.077 to 0.093), SRMR = 0.027.

Positive and Negative Affect Schedule (PANAS). The PANAS scale (Watson et al., 1988) was used to measure positive and negative emotions. Participants rated ten emotion items based on the extent of how they felt each emotion in the past week, from a scale of 1 (*very slightly or not at all*) to 5 (*extremely*). Both the positive emotions subscale ( $\alpha=0.87$ ) and negative emotions subscale ( $\alpha=0.84$ ) were found to be internally consistent. The construct validity of the two-factor PANAS ( $n = 3,877$ ) is supported by the following robust fit indices: SB  $\chi^2(33) = 535.18$ ,  $p < 0.001$ , CFI = 0.969, TLI = 0.958, RMSEA = 0.068 (C.I. = 0.063 to 0.073), SRMR = 0.037.

PHQ-9 and GAD-7. The Patient Health Questionnaire-9 (Kroenke et al., 2001) assessed nine symptoms of depression. GAD-7. The Generalized Anxiety Disorder-7 (Spitzer et al., 2006) was used to measure seven anxiety symptoms. For the PHQ-9, respondents indicated the severity of their symptoms on a Likert-type scale ranging from 0 (*not at all*) to 3 (*all the time*). The total score ranged from 0 to 27, where a higher total score indicated greater depression symptom severity. Its reliability in the current study was high ( $\alpha=0.81$ ). The GAD-7 consists of seven items with the same response options to the PHQ-9 and provides an anxiety symptom severity score from 0 to 21. Its reliability in the present study was  $\alpha=0.79$ . The cut-off score for the PHQ-9 and GAD-7 is greater than or equal to 10 and 9, respectively.

## Data Analysis

Descriptive statistics, including testing normality assumptions, were conducted. Confirmatory factor analyses (CFA) in R lavaan package (Rosseel, 2012) were used to test the two-factor structure of the data (i.e., anxiety and depression). To account for non-normality, we used the maximum likelihood estimator with robust standard errors, and a Satorra-Bentler scaled test statistic to test the CFA. Satorra-Bentler chi-square tests ( $\chi^2$ ) were obtained. Since a non-significant  $\chi^2$  result is usually difficult to obtain with larger samples, other fit indices, including Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR) were also used to evaluate goodness-of-fit (Barrett, 2007). Models with CFI and TLI  $> 0.90$  and RMSEA  $< 0.08$  were deemed to have a good fit for the data (Hu & Bentler, 1995), while SRMR  $< 0.08$  was also deemed as a good fit to the data (Hu & Bentler, 1999).

To test measurement invariance, multigroup CFA was used to test models according to groups formed by key demographic characteristics (i.e., gender, age, and geographical location). Multigroup CFA was conducted using equaltestMI (Jiang & Mai, 2020). To determine measurement invariance, we followed Chen's (2007) recommendations for samples greater than 300: a change of CFI ( $\Delta$ CFI)

that is less than or equal to 0.01, supplemented by a change of RMSEA ( $\Delta$ RMSEA) that is less than or equal to 0.15 or a change in SRMR ( $\Delta$ SRMR) that is less than or equal to 0.03 will indicate invariance. The initial multigroup CFA model for each group allowed all factor loadings, uniqueness, and correlations to be freely estimated. Configural, metric, and scalar invariance were subsequently tested by constraining factor structure, factor loadings, and intercepts, respectively.

Structural equation modelling (SEM) was used to examine convergent validity. We tested the association between the PHQ-4 to stress, negative affect, and positive affect. Listwise deletion approach was used for participants with missing values (14.31%) which led to the analytic sample of  $n = 3,877$  for the SEM. Similar to the CFA models, CFI, TLI, RMSEA, and SRMR were also used to evaluate the SEM's model fit. Using a full SEM is essential to further examine the association between the latent constructs while controlling for item-level measurement errors (Mendoza & Yan, 2021; Yu & Hsu, 2013; Zumbo, 2014).

## Results

### The Two-factor Model of the PHQ-4 and its Measurement Invariance across Demographic Groups

Table 1 presents the summary statistics, bivariate correlations, and internal consistency rating of the scales. All constructs were normally distributed. Construct validity was tested by comparing the model fit of the two-factor to the unidimensional model of PHQ-4 using CFA. The two-factor model of PHQ-4 consists of depression (2 items) and anxiety (2 items), while the unidimensional model examined all items four items as a single factor (4 items). Table 2 shows the factor loadings of the 2-factor model and Table 3 shows that it has excellent model fit [ $\chi^2(1) = 0.590$ , CFI = 1.000, TLI = 1.000, RMSEA = 0.000, SRMR = 0.001, see Fig. 1], that is significantly better than the unidimensional model ( $\chi^2_{diff} = 434.09$ ,  $p < 0.001$ ).

Demonstrating the measurement invariance of the two-factor PHQ-4, multigroup CFA results showed that the 2-factor model is invariant across gender, age group, and geographic location (Table 4). Specifically, between males ( $n = 1,142$ ) and females ( $n = 3,382$ ), the configural model (constraining factor structure across groups) has good fit indices. Both the metric model (constraining all factor loadings across groups) and the scalar model (constrains all item intercepts) had a  $\Delta$ CFI less than 0.01 and a  $\Delta$ RMSEA less than 0.15 or  $\Delta$ SRMR less than 0.03. Consistent invariance results were found on the PHQ-4's two-factor model among those under the upper median age bracket ( $n = 2,180$ ) and lower median age bracket ( $n = 2,344$ ), as well as those from

**Table 1** Pairwise bivariate correlations, descriptive statistics, and internal consistencies

	1	2	3	4	5	6
1. PHQ-4 <sup>a</sup>	(0.82)					
2. PHQ-2 <sup>a</sup> (depression)	0.89	(0.71)				
3. GAD-2 <sup>a</sup> (anxiety)	0.90	0.60	(0.83)			
4. Positive Emotions <sup>b</sup>	-0.43	-0.48	-0.29	(0.87)		
5. Negative Emotions <sup>b</sup>	0.68	0.58	0.64	-0.33	(0.84)	
6. Stress <sup>c</sup>	0.69	0.60	0.64	-0.38	0.72	(0.88)
Mean	4.98	15.11	13.88	14.44	13.65	8.75
SD	3.22	5.66	4.10	4.72	4.93	5.14
Skewness	0.42	0.06	-0.09	0.14	0.12	0.29
Kurtosis	2.27	2.30	2.11	2.41	2.12	2.37

All correlations were significant at  $p < 0.001$ ; scores shown in parentheses on diagonal are internal consistency reliabilities of the scales; a = 4,524 responses; b = 3,877; c = 4,036

**Table 2** The items and factor loadings of the two-factor CFA of the PHQ-4

	Factor loading	SE	p-value
Depression (PHQ-2)			
PHQ1. Little interest or pleasure in doing things	0.64	---	---
PHQ2. Feeling down, depressed, or hopeless	0.85	0.04	0.000
Anxiety (GAD-2)			
PHQ3. Feeling nervous, anxious or on edge	0.82	---	---
PHQ4. Not being able to stop or control worrying	0.87	0.02	0.000

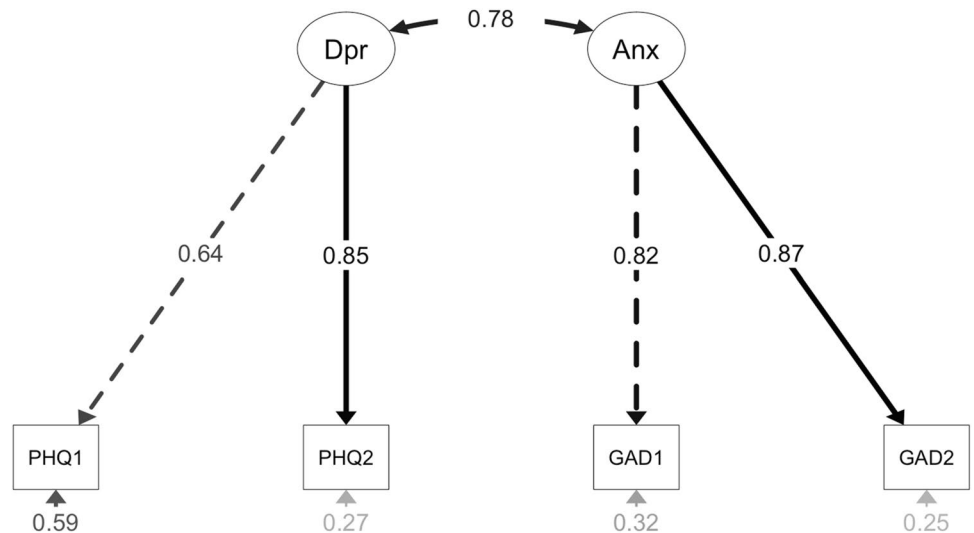
The covariance between depression and anxiety is  $B = 0.78$ ,  $p < 0.001$

**Table 3** Confirmatory factor analysis (CFA) comparing fit indices of the two-factor vs. one-factor model of PHQ-4

PHQ-4 Factor Structure (n=4,524)	SB $\chi^2$	df	CFI	TLI	RMSEA	SRMR	$\chi^2$ difference vs. hypothesized model
Two-factor model (hypothesized model)	0.590	1	1.000	1.000	0.000	0.001	
One-factor model	434.676	2	0.938	0.815	0.219	0.046	434.09***

\*\*\*  $p < 0.001$

**Fig. 1** The two-factor CFA model for PHQ-4. Dpr = depression, Anx = anxiety





**Table 4** Multigroup CFA results demonstrating the measurement invariance of the 2-factor model across age, gender, and geographical location

Model constrains	SB $\chi^2$	df	CFI	TLI	RMSEA	SRMR	$\Delta$ CFI	$\Delta$ RMSEA	$\Delta$ SRMR	Invariance
2-factor model	0.590	1	1.000	1.000	0.000	0.001	---	---	---	---
Across gender										
Males only (n=1,142)	0.741	1	1.000	1.000	0.000	0.003	---	---	---	---
Females only (n=3,382)	1.884	1	1.000	0.999	0.016	0.002	---	---	---	---
Configural invariance	2.625	2	1.000	0.999	0.012	0.003	---	---	---	Invariant
Metric invariance	3.293	4	1.000	1.000	0.000	0.005	0.000	0.000	0.002	Invariant
Scalar invariance	5.103	6	1.000	1.000	0.000	0.006	0.000	0.000	0.001	Invariant
Across age groups										
25 years old or younger (n=2,344)	0.007	1	1.000	1.000	0.000	0.000	---	---	---	---
26 years old or older (n=2,180)	0.512	1	1.000	1.000	0.000	0.002	---	---	---	---
Configural invariance	0.520	2	1.000	1.000	0.000	0.001	---	---	---	Invariant
Metric invariance	10.619	4	0.999	0.997	0.027	0.015	-0.001	0.027	0.014	Invariant
Scalar invariance	23.298	6	0.997	0.995	0.036	0.018	-0.002	0.009	0.003	Invariant
Across geographical location										
Manila (Capital; n=2,046)	0.972	1	1.000	1.000	0.000	0.002	---	---	---	---
Other regions (n=2,478)	0.010	1	1.000	1.002	0.000	0.000	---	---	---	---
Configural invariance	0.982	2	1.000	1.000	0.000	0.001	---	---	---	Invariant
Metric invariance	2.056	4	1.000	1.000	0.000	0.000	0.000	0.000	-0.001	Invariant
Scalar invariance	13.842	6	0.999	0.998	0.024	0.011	-0.001	0.024	0.011	Invariant

All SB $\chi^2$  tests were significant at  $p < 0.001$

the capital Manila ( $n = 2,046$ ) and other regions ( $n = 2,478$ ). Overall, the results suggest evidence for measurement invariance of the two-factor PHQ-4 given the consistent model fit indices despite the increasing levels of model constrain.

### Criterion-related Evidence for Construct Validity

Given the internal validity of the two-factor structure of the PHQ-4, we ran correlation analysis and SEM to test its convergent validity. Correlation analysis shows that the depression ( $r = 0.60$ ,  $p < 0.01$ ) and anxiety subscales ( $r = 0.64$ ,  $p < 0.01$ ) of the PHQ-4 were positively correlated to stress. Depression ( $r = 0.58$ ,  $p < 0.01$ ) and anxiety symptoms ( $r = 0.64$ ,  $p < 0.01$ ) were also positively correlated with negative affect. Depression ( $r = -0.48$ ,  $p < 0.01$ ) and anxiety symptoms ( $r = -0.29$ ,  $p < 0.01$ ) had significant negative relationship with positive affect. These suggests convergent validity across the criterion-related constructs.

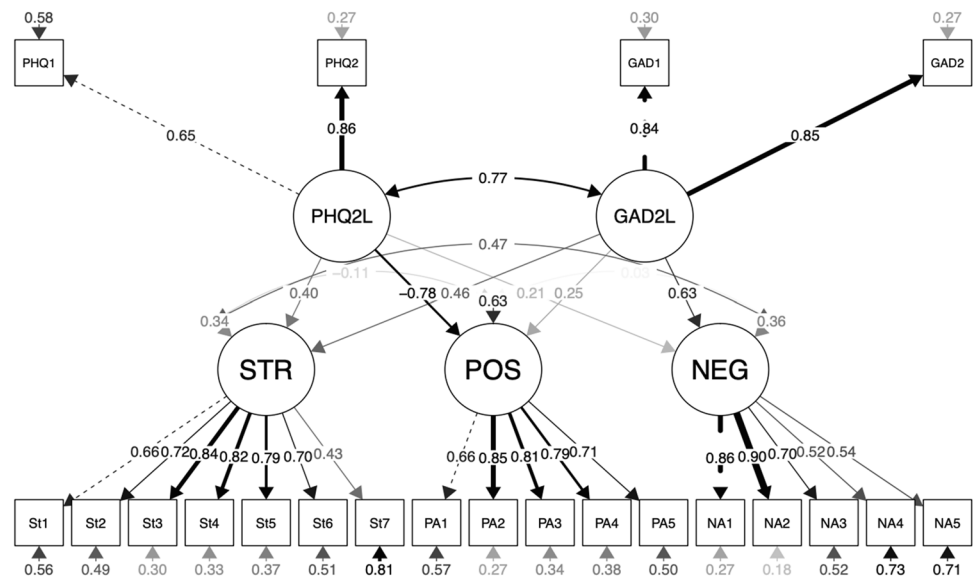
In the SEM, stress, positive affect, and negative affect were simultaneously regressed to depression and anxiety to explore latent correlations that account for item-level measurement errors (see Fig. 2). The SEM had good fit to the data [ $\chi^2(177) = 2764.17$ , CFI = 0.945, TLI = 0.935, RMSEA = 0.061, SRMR = 0.04]. The findings from the SEM show that the relationship between the latent constructs was consistent with the pairwise correlations. The depression subscale positively correlated with stress ( $B = 0.40$ ,  $p < 0.001$ ) and negative affect ( $B = 0.48$ ,  $p < 0.001$ ),

and negatively correlated with positive affect ( $B = -0.76$ ,  $p < 0.001$ ). The anxiety subscale had a positive association with stress ( $B = 0.41$ ,  $p < 0.001$ ) and negative affect ( $B = 0.40$ ,  $p < 0.001$ ). The latent correlation of the PHQ-4's anxiety subscale and positive affect was positive with relatively weak strength ( $B = 0.23$ ,  $p < 0.001$ ). This could be attributed to model complexity with having two highly correlated exogenous variables (i.e., depression and anxiety). Still, based on the bivariate correlations, the anxiety subscale and positive affect were negatively correlated. Overall, the SEM supports the criterion-related validity of the PHQ-4 and its subscales.

### Prevalence Estimates

Respectively, 42.11% and 40.96% of the respondents were screened for depression and anxiety by the PHQ-2 and GAD-2. Among those who were screened by PHQ-2 ( $n = 1,905$ ), 81.78% would also be screened by the PHQ-9, whereas among those screened by GAD-2 ( $n = 1,853$ ), 94.06% would also be screened by the GAD-7. This suggests the accuracy of the PHQ-4 as an ultra-brief screener for depression and anxiety. The specific prevalence estimates based on the longer screeners are 34.44% for depression ( $n = 1,558$  out of 4,524) and 38.53% for anxiety ( $n = 1,743$  out of 4,524).

**Fig. 2** Structural equation model demonstrating the criterion-related validity of PHQ-4. Darker/thicker paths illustrate higher/stronger latent covariances. PHQ2L = depression, GAD2L = anxiety, STR = stress, POS = positive affect, NEG = negative affect



## Discussion

This study evaluated the psychometric properties of the PHQ-4 as an ultra-brief instrument for evaluating mental health symptoms of a large non-clinical sample amid the COVID-19 pandemic. Overall, the findings support the instrument’s reliability and validity. Specifically, the two-factor structure of the PHQ-4 was found to be superior to the one-factor structure. The two-factor structure of the PHQ-4 held across configural, metric, and scalar invariance tests demonstrating the scales’ robust ability to assess depression and anxiety symptoms across age, gender, and locale. The depression and anxiety subscales of the PHQ-4 were both negatively correlated with positive affect and positively correlated with stress and negative affect, supporting its criterion-related validity. For the accuracy of the cut-off of the PHQ-4 subscales, 81.78% (n = 1,558) of those who were screened for depression (PHQ-2) were screened by the PHQ-9, whereas 94.06% (n = 1,743) who were screened for anxiety (GAD-2) were also screened by the GAD-7.

The results support the structural validity of the two-factor model (i.e., anxiety and depression) over the one-factor model of PHQ-4 as suggested by Kroenke et al. (2009) and validated by succeeding investigations in Germany (Löwe et al., 2010), Columbia (Kocalevent et al., 2014), U.S. (Mills et al., 2015), Spain (Cano-Vindel et al., 2018), and Iran (Ghaheri et al., 2020). Likewise, similar to the results of previous studies, the two-factor structure of PHQ-4 was also found to be consistent across gender and age groups (Cano-Vindel et al., 2018; Kocalevent et al., 2014; Löwe et al., 2010; Renovanz et al., 2019). Novel to this study is the examination of measurement invariance of the two-factor structure of PHQ-4 according to respondents’ geographical location (i.e., Manila or NCR and non-NCR residents).

Although Mills and colleagues (2015) found evidence supporting the measurement invariance of the two-factor structure of PHQ-4 among English- and Spanish-speaking Hispanic Americans in the U.S., these respondents were not grouped according to locale. This is particularly important in the context of the Philippines since the outbreak of COVID-19 started in its capital (i.e., Manila). This study is important in that mental health symptoms might vary with respect to a respondent’s geographical proximity or distance from the outbreak. The Philippines is an archipelago composed of three major geographical areas: Luzon, Visayas, and Mindanao. Each geographical area is further subdivided into regions and provinces. Manila is in Luzon and is composed of 16 cities. Compared to other regions, Manila is more urbanized and overpopulated. The largest international airport in the Philippines (i.e., Ninoy Aquino International Airport) is also located in Manila. Since the first case of COVID-19 in the Philippines was detected in Manila last January 2020 (Edrada et al., 2020), it is possible that mental health concerns are higher in the country’s capital. Despite the potential differences between Manila and other regions, the excellent model fit of the two-factor structure of the PHQ-4 held. This demonstrates the PHQ-4’s utility and validity in both “at-risk” and “low risk” locations.

The SEM model shows the theoretical association between the PHQ-4 subscales and criterion-related constructs. This is aligned with previous studies that link PHQ-4 to positive and negative psychological outcomes (Choi et al., 2020; Ghaheri et al., 2020; Kocalevent et al., 2014; Löwe et al., 2010; Mills et al., 2015; Schnell & Krampe, 2020; Renovanz et al., 2019) which further supports the criterion validity of PHQ-4.

Related to the prevalence estimates, the 34.44% prevalence for depression is similar to recent meta-analytic studies

(Bueno-Notivol et al., 2021; Salari et al., 2020) but higher than other pooled prevalence (Krishnamoorthy et al., 2020). The prevalence of anxiety in the Philippines (38.53%) might be higher than what's observed in other recent prevalence (Krishnamoorthy et al., 2020; Salari et al., 2020) or observational studies (Bernardo et al., 2020; Bernardo & Mendoza, 2021). These prevalence estimates mean that about one-in-three could be screened for depression or anxiety during the outbreak of the COVID-19 pandemic in the Philippines. More recent pooled prevalence in the South Asian context shows similar depression and anxiety prevalence with the current study (Hossain et al., 2021). This finding is instrumental for future prevalence studies that investigate the mental health effects of the COVID-19 pandemic to aid in the prevention, potential intervention, and sustainable monitoring of mental health symptoms.

### Limitations and Future Research Direction

The present investigation provided specific evidence on the psychometric robustness of the PHQ-4 in the Philippine context using a relatively large sample collected during the COVID-19 pandemic. We note our study limitations below to guide researchers and practitioners in interpreting the study results. First, the use of cross-sectional data does not allow to test of the predictive validity and test-retest reliability of PHQ-4. Second, although this investigation utilized a large community sample of Filipino respondents, access to the survey was limited to those who have an internet connection and may not accurately represent the entire Filipino population (i.e., sampling bias). Third, although excellent fit indices for the two-factor PHQ-4 and its measurement invariance across demographic groups highlights the statistical rigor of the screener, this could be attributed to having only two manifest items loaded onto each of the two latent factors. However, studies showing that one or two observed indicators should suffice as long as they are theoretically meaningful indicators (Hayduk & Littvay, 2012). In the case of the PHQ-4, the Diagnostic Statistical Manual-5 (DSM-5; American Psychiatric Association, 2013) identified (1) depressed mood and (2) loss of interest and pleasure as the two quintessential diagnostic criteria for clinical depression. For generalized anxiety disorder, feeling (1) excessive and (2) uncontrollable worry were the two key symptoms. These diagnostic symptoms inform the PHQ-4 and its longer versions. Hence, the statistical rigor of the PHQ-4 is also clinically and theoretically supported and informed. Still, the current study can benefit from replication studies to revisit the comparison between the two- and the one-factor structure of the PHQ-4. Finally, the assessment of anxiety and depression using PHQ-4 was not supplemented by standard clinical interviews to validate the prevalence of anxiety and depressive disorders among the participants who were

screened by the PHQ-4 or their full-scale counterparts (i.e., GAD-7 and PHQ-9). The reliance on self-reported data has inherent limitations. Still, the prevalence estimates of this study lend knowledge to the higher mental health symptom severity among Filipinos during the COVID-19 outbreak.

Despite the abovementioned limitations, the present study has noteworthy strengths. To date, this is the first investigation to examine the psychometric properties of PHQ-4 during the COVID-19 pandemic, and results show that it can be a valid mental health screener. Also, this is the first study to validate PHQ-4 in the Philippines using a considerably large sample. Aside from using criterion-related instruments to establish external validity, the present study also used the longer versions of the scales (i.e., GAD-7 and PHQ-9) to evaluate the accuracy of the PHQ-4 subscales and to accurately determine the prevalence rates of anxiety and depression in the Philippines. We recommend future researchers to include standard clinical interviews to further establish the specificity and sensitivity ratio of the PHQ-4. Also, employing a longitudinal research design to aid in examining temporal psychometric properties of PHQ-4 (i.e., test-retest reliability and predictive validity) is recommended. Finally, given the current distribution of the respondents by gender and locale, succeeding studies may investigate recruiting a more nationally representative sample by exploring the prevalence of mental health symptoms in rural and other hard-to-reach locations through non-Web-based methods (e.g., paper-pen data collection).

### Conclusions

The PHQ-4 is a reliable, valid, and cost-effective measure of depression and anxiety symptoms. Especially during health crises, the use of brief screeners for mental health symptoms is invaluable for accurate estimation of mental health prevalence rates to initiate adequate and timely psychological support. The validity of the PHQ-4 applies to age, gender, and local groups suggesting its overall applicability and utility, specifically in a low- and middle-income country situated in Southeast Asia. It is of practical importance to use valid screeners of depression and anxiety symptoms among non-clinical samples or general population. The use of valid mental health screeners reduces participant burden in large-scale data collection, enables researchers to rapidly estimate the prevalence and severity of mental health symptoms, aids in timely intervention and psychological support, and offers a sustainable means to monitor and evaluate mental health symptoms during global health crises and other human disasters.

**Acknowledgements** The authors express their sincerest gratitude to the volunteers and professionals at LifeRisksPH—a non-profit youth organization registered under the Philippine's National Youth Commission. They are focused on stigma reduction on mental health and the prevention and psychoeducation for substance use and suicide. The authors also sincerely thank Dr KWAN Lok Yin Joyce, for training the



doctoral students among the authors in the use of R and for providing support in the preliminary stages of the analysis of this study.

## References

- Ali, G. C., Ryan, G., & De Silva, M. J. (2016). Validated screening tools for common mental disorders in low and middle income countries: A systematic review. *PLoS ONE*, *11*(6), 1–14. <https://doi.org/10.1371/journal.pone.0156939>
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric Properties of the 42-Item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychological Assessment*, *10*(2), 176–181. <https://doi.org/10.1037/1040-3590.10.2.176>
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815–824. <https://doi.org/10.1016/j.paid.2006.09.018>
- Bendau, A., Petzold, M. B., Pyrkosch, L., Mascarell Maricic, L., Betzler, F., Rogoll, J. ... Plag, J. (2020). Associations between COVID-19 related media consumption and symptoms of anxiety, depression and COVID-19 related fear in the general population in Germany. *European Archives of Psychiatry and Clinical Neuroscience*, 0123456789. <https://doi.org/10.1007/s00406-020-01171-6>
- Bernardo, A. B. I., & Mendoza, N. B. (2021). Measuring hope during the COVID-19 outbreak in the Philippines: Development and validation of the state locus-of-Hope scale short form in Filipino. *Current Psychology*, *40*(11), 5698–5707. <https://doi.org/10.1007/s12144-020-00887-x>
- Bernardo, A. B. I., Mendoza, N. B., Simon, P. D., Cunanan, A. L. P., Dizon, J. I. W. T., Tarroja, M. C. H. ... Saplala, J. E. G. (2020). Coronavirus Pandemic Anxiety Scale (CPAS-11): development and initial validation. *Current Psychology*, *1*. <https://doi.org/10.1007/s12144-020-01193-2>
- Biana, H. T., & Joaquin, J. J. B. (2020). The mental health of medical workers in Wuhan, China dealing with the 2019 novel coronavirus. *Journal of Public Health*. Oxford University Press. [https://doi.org/10.1016/S2215-0366\(20\)30047-X](https://doi.org/10.1016/S2215-0366(20)30047-X)
- Bueno-Notivol, J., Gracia-García, P., Olaya, B., Lasheras, I., López-Antón, R., & Santabábara, J. (2021). Prevalence of depression during the COVID-19 outbreak: A meta-analysis of community-based studies. *International Journal of Clinical and Health Psychology*, *21*(1). <https://doi.org/10.1016/j.ijchp.2020.07.007>
- Cano-Vindel, A., Muñoz-Navarro, R., Medrano, L. A., Ruiz-Rodríguez, P., González-Blanch, C., Gómez-Castillo, M. D., ... Santolaya, F. (2018). A computerized version of the Patient Health Questionnaire-4 as an ultra-brief screening tool to detect emotional disorders in primary care. *Journal of Affective Disorders*, *234*(January), 247–255. <https://doi.org/10.1016/j.jad.2018.01.030>
- Cavanagh, K., Churchar, A., O'Hanlon, P., Mundy, T., Votolato, P., Jones, F., ... Strauss, C. (2018). A randomised controlled trial of a brief online mindfulness-based intervention in a non-clinical population: Replication and extension. *Mindfulness*, *9*(4), 1191–1205. <https://doi.org/10.1007/s12671-017-0856-1>
- Chandu, V. C., Marella, Y., Panga, G. S., Pachava, S., & Vadapalli, V. (2020). Measuring the Impact of COVID-19 on mental health: A scoping review of the existing scales. *Indian Journal of Psychological Medicine*, *42*(5), 421–427. <https://doi.org/10.1177/0253717620946439>
- 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>
- Choi, E., Lee, J., & Lee, S. A. (2020). Validation of the Korean version of the obsession with COVID-19 scale and the Coronavirus anxiety scale. *Death Studies*. <https://doi.org/10.1080/07481187.2020.1833383>
- Cortez, P. A., Joseph, S. J., Das, N., Bhandari, S. S., & Shoib, S. (2020). Tools to measure the psychological impact of the COVID-19 pandemic: What do we have in the platter? *Asian Journal of Psychiatry*, *53*. Elsevier B.V. <https://doi.org/10.1016/j.ajp.2020.102371>
- Edrada, E. M., Lopez, E. B., Villarama, J. B., Villarama, S., Dagoc, E. P., Smith, B. F. ... Solante, J. R., R. M (2020). First COVID-19 infections in the Philippines: A case report. *Tropical Medicine and Health*, *48*(1). <https://doi.org/10.1186/s41182-020-00203-0>
- Ghaehri, A., Omani-Samani, R., Sepidarkish, M., Hosseini, M., & Maroufizadeh, S. (2020). The four-item patient health questionnaire for anxiety and depression: A validation study in infertile patients. *International Journal of Fertility and Sterility*, *14*(3), 234–239. <https://doi.org/10.22074/ijfs.2020.44412>
- Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*, *12*. <https://doi.org/10.1186/1471-2288-12-159>
- Hossain, M. M., Rahman, M., Trisha, N. F., Tasnim, S., Nuzhath, T., Hasan, N. T., et al. (2021). Prevalence of anxiety and depression in South Asia during COVID-19: A systematic review and meta analysis. *Heliyon*, *7*(4). e06677. <https://doi.org/10.1016/j.heliyon.2021.e06677>
- Hu, L. T., & Bentler, P. M. (1995). *Evaluating model fit* (R. H. Hoyle (ed.)). Structural equation modeling: Concepts, issues, and applications. Sage Publications, Inc.
- 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>
- Jiang, G., & Mai, Y. (2020). *equaltestMI: Examine measurement invariance via equivalence testing and projection method*. R Package Version 0.6.0. <https://cran.r-project.org/web/packages/equaltestMI/index.html>. Accessed 13 Sept 2021.
- Khubchandani, J., Brey, R., Kotecki, J., Kleinfelder, J. A., & Anderson, J. (2016). The psychometric properties of PHQ-4 depression and anxiety screening scale among college students. *Archives of Psychiatric Nursing*, *30*(4), 457–462. <https://doi.org/10.1016/j.apnu.2016.01.014>
- Kim, H. W., Shin, C., Lee, S. H., & Han, C. (2021). Standardization of the Korean version of the patient health Questionnaire-4 (PHQ-4). *Clinical Psychopharmacology and Neuroscience*, *19*(1), 104–111. <https://doi.org/10.9758/CPN.2021.19.1.104>
- Kocalevent, R. D., Finck, C., Jimenez-Leal, W., Sautier, L., & Hinz, A. (2014). Standardization of the Colombian version of the PHQ-4 in the general population. *BMC Psychiatry*, *14*(1), 1–8. <https://doi.org/10.1186/1471-244X-14-205>
- Krishnamoorthy, Y., Nagarajan, R., Saya, G. K., & Menon, V. (2020). Prevalence of psychological morbidities among general population, healthcare workers and COVID-19 patients amidst the COVID-19 pandemic: A systematic review and meta-analysis. *Psychiatry Research*, *293*. <https://doi.org/10.1016/j.psychres.2020.113382>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9 validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The patient health questionnaire-2: Validity of a two-item depression screener. *Medical Care*, *41*(11), 1284–1292. <https://doi.org/10.1097/01.MLR.0000093487.78664.3C>
- Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The

- PHQ-4. *Psychosomatics*, 50(6), 613–621. [https://doi.org/10.1016/s0033-3182\(09\)70864-3](https://doi.org/10.1016/s0033-3182(09)70864-3)
- Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of internal medicine*, 146(5), 317–325. <https://doi.org/10.7326/0003-4819-146-5-2007-3060-00004>.
- Lai, A. Y., kwan, Lee, L., Wang, M. P., Feng, Y., Lai, T. T., kwan, Ho, L. M. ... Lam, T. H. (2020). Mental Health Impacts of the COVID-19 Pandemic on International University Students, Related Stressors, and Coping Strategies. *Frontiers in Psychiatry*, 11(November). <https://doi.org/10.3389/fpsy.2020.584240>
- Li, H. Y., Cao, H., Leung, D. Y. P., & Mak, Y. W. (2020). The Psychological Impacts of a COVID-19 outbreak on college students in China: A longitudinal study. *International Journal of Environmental Research and Public Health*, 17(3933). <https://doi.org/10.1157/13090384>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression, Anxiety, Stress Scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Löwe, B., Kroenke, K., & Gräfe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, 58(2), 163–171. <https://doi.org/10.1016/j.jpsychores.2004.09.006>
- Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K. ... Brähler, E. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders*, 122(1–2), 86–95. <https://doi.org/10.1016/j.jad.2009.06.019>
- Mendoza, N. B., & Yan, Z. (2021). Validation of a subject-specific student Self-Assessment Practice Scale (SaPS) Among Secondary School Students in the Philippines. *Journal of Psychoeducational Assessment*, 073428292199437. <https://doi.org/10.1177/0734282921994374>
- Mendoza, N. B., & Dizon, J. I. W. T. (2022). Prevalence of severe anxiety in the Philippines Amid the COVID-19 outbreak. *Journal of Loss and Trauma*, 0(0), 1–3. <https://doi.org/10.1080/15325024.2021.2023298>
- Mills, S. D., Fox, R. S., Pan, T. M., Malcarne, V. L., Roesch, S. C., & Sadler, G. R. (2015). Psychometric evaluation of the patient health questionnaire-4 in Hispanic Americans. *Hispanic Journal of Behavioral Sciences*, 37(4), 560–571. <https://doi.org/10.1177/0739986315608126>
- Petzold, M. B., Bendau, A., Plag, J., Pyrkosch, L., Maricic, L. M., Rogoll, J. ... Ströhle, A. (2020a). Development of the COVID-19-Anxiety Questionnaire and first psychometric testing. *BJPsycho Open*, 6(5), 1–4. <https://doi.org/10.1192/bjo.2020.82>
- Petzold, M. B., Bendau, A., Plag, J., Pyrkosch, L., Mascarell Maricic, L., Betzler, F., ... Ströhle, A. (2020b). Risk, resilience, psychological distress, and anxiety at the beginning of the COVID-19 pandemic in Germany. *Brain and Behavior*, 10(9), 1–10. <https://doi.org/10.1002/brb3.1745>
- Renovanz, M., Soebianto, S., Tsakmaklis, H., Keric, N., Nadji-Ohl, M., Beutel, M. ... Hickmann, A. K. (2019). Evaluation of the psychological burden during the early disease trajectory in patients with intracranial tumors by the ultra-brief Patient Health Questionnaire for Depression and Anxiety (PHQ-4). *Supportive Care in Cancer*, 27(12), 4469–4477. <https://doi.org/10.1007/s00520-019-04718-z>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Salari, N., Hosseini-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M. ... Khaledi-Paveh, B. (2020). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Globalization and Health*, 16(1), 1–11.
- Schmalbach, B., Altmann, U., Peitsch-Vuorilehto, H., Brähler, E., & Strauß, B. (2021). A short version of the body experience questionnaire (FBek): Results of clinical and non-clinical samples. *Psychotherapie, Psychosomatik, Medizinische Psychologie*. <https://doi.org/10.1055/A-1503-5036>
- Schnell, T., & Krampe, H. (2020). Meaning in life and self-control buffer stress in times of COVID-19: Moderating and mediating effects with regard to mental distress. *Frontiers in Psychiatry*, 11(September), 1–16. <https://doi.org/10.3389/fpsy.2020.582352>
- Shah, S. M. A., Mohammad, D., Qureshi, M. F. H., Abbas, M. Z., & Aleem, S. (2021). Prevalence, psychological responses and associated correlates of depression, anxiety and stress in a global population, during the coronavirus disease (COVID-19) pandemic. *Community Mental Health Journal*, 57(1), 101–110. <https://doi.org/10.1007/s10597-020-00728-y>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Taylor, S., Landry, C. A., Paluszek, M. M., Fergus, T. A., McKay, D., & Asmundson, G. J. G. (2020). Development and initial validation of the COVID stress scales. *Journal of Anxiety Disorders*, 72(April), 102232. <https://doi.org/10.1016/j.janxdis.2020.102232>
- The World Bank (2021). *Philippines | Data*. <https://data.worldbank.org/country/PH>. Accessed 26 Sept 2021.
- Tibubos, A. N., & Kröger, H. (2020). A cross-cultural comparison of the ultrabrief mental health screeners PHQ-4 and SF-12 in Germany. *Psychological Assessment*, 32(7), 690–697. <https://doi.org/10.1037/pas0000814>
- TIME (2021). *Inside One of the World's Longest COVID-19 Lockdowns | Time*. <https://time.com/5945616/covid-philippines-pandemic-lockdown/>. Accessed 30 March 2021.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- World Health Organization (2021). *WHO Coronavirus (COVID-19) Dashboard*. <https://covid19.who.int/table>. Accessed 23 March 2021.
- Yu, S. C., & Hsu, W. H. (2013). Applying structural equation modeling methodology to test validation: An example of cyberspace positive psychology scale. *Quality and Quantity*, 47(6), 3423–3434. <https://doi.org/10.1007/s11135-012-9730-3>
- Zumbo, B.D. (2014). Structural equation modeling and test validation. In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri and J.L. Teugels (Eds.), *Statsref: Statistics reference online*. Wiley. <https://doi.org/10.1002/9781118445112.stat06521>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.