Dimensionality of the Penn State Worry Questionnaire and its brief versions

Valencia, Pablo D.*, a ; Paredes-Angeles, Rubíb

Original Article

Abstract	Resumen	Tabla de Contenido

Worry, a form of repetitive negative thinking, has been studied using the Penn State Worry Questionnaire (PSWQ). The aim of the present study was to assess the dimensionality of the PSWQ, as well as three brief versions, in college students from Lima, Peru. The full PSWQ was administered to 290 participants. In order to achieve that goal, we examined the factor structure and reliability of several models: unidimensional, oblique, bifactor, as well as the brief versions. The results suggest that the PSWQ is an essentially unidimensional measure and it functions better without reversed items. Likewise, brief versions were found to have similarly adequate psychometric properties.

Palabras clave:
Worry, PSWQ, validity, university students | **Dimensionalidad del Penn State Worry Questionnaire y sus versiones breves.** La preocupación, una forma de pensamiento negativo repetitivo, ha sido estudiada mediante el Penn State Worry Questionnaire (PSWQ). El objetivo de este estudio fue evaluar la dimensionalidad del PSWQ y de tres versiones breves en estudiantes universitarios de Lima, Perú. El PWSQ completo fue administrado a 290 participantes. A fin de lograr el objetivo, se examinó la estructura factorial y confiabilidad de diversos modelos: unidimensional, oblicuo, bifactor, así como las versiones breves. Los resultados sugieren que el PSWQ es una medida esencialmente unidimensional y que funciona mejor sin ítems inversos. Igualmente, las versiones breves tienen propiedades psicométricas similarmente adecuadas.

Keywords:
preocupación; PSWQ; validez; estudiantes universitarios | |

Introduction

Contemporary approaches to psychopathology emphasize transdiagnostic psychological dimensions, that is, processes that are present to some degree in various diagnoses (Nolen-Hoeksema & Watkins, 2011; Watkins, 2015). One such process is repetitive negative thinking, which has strong associations with emotional disorders such as anxiety and depression (Ehring & Watkins, 2008). A specific form of repetitive negative thinking, worry, has been extensively studied in relation to generalized anxiety disorder (GAD), of which it is a central feature (Borkovec & Inz, 1990). The classical definition of worry describes it as "a chain of thoughts and images,

negatively affect-laden and relatively uncontrollable... [It] represents an attempt to engage in mental problem-solving on an issue whose outcome is uncertain but contains the possibility of one or more negative outcomes" (Borkovec, Robinson, Pruzinsky, & DePree, 1983, p. 10).

The evidence suggests that worry is predominantly verbal and abstract in content (Borkovec, Ray, & Stöber, 1998; Ehring & Watkins, 2008; Watkins, 2008), thus making it difficult for worriers to actually be in contact with unpleasant emotional experiences. That is, worry can be conceptualized as a form of experiential

a Universidad Nacional Mayor de San Marcos, Facultad de Psicología, Perú
b Universidad Nacional Mayor de San Marcos, Grupo de Estudios Avances en Medición Psicológica, Perú
*Enviar correspondencia a: Valencia, P. D. E-mail: pabvalenciam@gmail.com

avoidance (Hayes-Skelton & Eustis, 2020; Hayes, Wilson, Gifford, Follette, & Strosahl, 1996). Its detrimental consequences come from the fact that (1) avoided experiences cannot be processed and, therefore, the emotional reactions cannot be extinguished (Sibrava & Borkovec, 2006); and (2) worry is negatively reinforced since most things about which people worry never happen in reality (Borkovec et al., 1998).

Even though, as already mentioned, the study of worry originated in GAD research, it has extended its scope to other diagnoses, such as sleep disturbances, obsessive-compulsive disorder, alcohol dependence, post-traumatic stress disorder, and psychotic disorders (for a review, see Purdon & Harrington, 2006). In a recent study, for example, difficulties to control worry were associated with depression, anxiety, and suicide attempts (Gorday, Rogers, & Joiner, 2018). Indeed, worry has been found to act both as a mediator and a moderator of the relationship between intolerance of uncertainty and emotional distress (i.e., depression and anxiety; Dar, Iqbal, & Mushtaq, 2017). Moreover, worry has also been identified as a risk factor for persistent and worsening anxiety symptoms in depressed older adults (Spinhoven, van der Veen, Voshaar, & Comijs, 2017).

The mental health of college students is an area of increasing research interest (Auerbach et al., 2018; Duffy et al., 2019), which is significantly associated with relevant outcomes such as academic achievement (Bruffaerts et al., 2018) and dropout (Wang et al., 2015). It is, therefore, fundamental to identify relevant psychological predictors of mental health in this population. Worry has emerged as one of such predictors. For example, it has been shown to predict anxiety, depression, and insomnia symptoms in student samples (Zvolensky et al., 2019). Furthermore, worry mediates the relation between academic stressors and both anxiety and depression (Bauer, Braitman, Judah, & Cigularov, 2020). Also, regarding a contemporary problem such as problematic smartphone use, worry seems to be significantly related to it in college populations (Elhai, Rozgonjuk, Yildirim, Alghraibeh, & Alafnan, 2019). In sum, worry seems to be a relevant variable for the study of college students' mental health.

Due to its relevance for the study of psychopathology, it is necessary to have standardized measures of worry. The Penn State Worry Questionnaire (PSWQ) is a popular measure of trait worry (Meyer, Miller, Metzger, & Borkovec, 1990). It has been extensively used in research, and there exist Spanish versions of it (Rodríguez-Biglieri & Vetere, 2011; Sandín, Chorot, Valiente, & Lostao, 2009). However, there are still some open questions regarding this measure's dimensionality, which is a fundamental aspect for both reliability and validity evidence (Furr & Bacharach, 2014). A great deal of this research has been conducted with student samples, achieving similar results to those of community and clinical samples (Castillo, Macrini, Cheniaux, & Landeira-Fernandez, 2010; Crittendon & Hopko, 2006; DeLapp, Chapman, & Williams, 2016; Fresco, Heimberg, Mennin, & Turk, 2002; Hazlett-Stevens, Ullman, & Craske, 2004; Lim, Kim, Lee, & Kwon, 2008; Pallesen, Nordhus, Carlstedt, Thayer, & Johnsen, 2006; Zhong, Wang, Li, & Liu, 2009). Therefore, the use of student samples seems justified for early examination of the PSWQ's factor structure in a new population.

Even though initial assessment of the PSWQ's dimensionality suggested a one-factor structure (Figure 1a; Brown, Antony, & Barlow, 1992), some other exploratory studies proposed two dimensions: worry engagement (11 positively keyed items) and absence of worry (5 negatively keyed items; Beck, Stanley, & Zebb, 1995; van Rijsoort, Emmelkamp, & Vervaeke, 1999). Later, using confirmatory methods and a sample composed of American undergraduates, this two-factor structure (Figure 1b) was shown to have better fit than the strictly unidimensional one (CFI = .950 and RMSEA = .065 versus CFI = .912 and RMSEA = .086), therefore concluding that the PSWQ is not a unidimensional measure (Fresco et al., 2002). However, this approach has been criticized for being theoretically meaningless, because the two proposed factors seem to be an artifact caused by the different wording of items (Brown, 2003). Moreover, there seems to be a very strong correlation between factors (Fresco et al., 2002; Gana, Martin, Canouet, Trouillet, & Meloni, 2002; Pallesen et al., 2006; Verkuil & Brosschot, 2012), which again indicates that they overlap and are probably indistinguishable.

An alternative approach to the problem of the PSWQ's dimensionality considers the influence of method effects due to the presence of negatively keyed items. This was first analyzed by adding covariances between the error terms of the reversed items to the unidimensional model (Brown, 2003). In a similar manner, method effects have also been tested by modeling a residual "reverse wording" factor, as shown in Figure 1d. According to Chen, West, and Sousa's (2006) terminology, this is an *incomplete bifactor model*, since all of the items are influenced by the general factor (Worry), but only some of them are influenced by a residual factor (Reverse Wording).

Several studies have found the incomplete bifactor model to have good fit and to provide a parsimonious interpretation of the PSWQ's data (Hazlett-Stevens et al., 2004; Kertz, Lee, & Björgvinsson, 2014; Lim et al., 2008; Verkuil & Brosschot, 2012; Zhong et al., 2009). Still, other researchers have also found good fit for a *complete bifactor model* (i.e., all the items are influenced by the general factor and by one residual factor, as shown in Figure 1c; Chen et al., 2006), which includes an additional residual factor for directly worded items (Carter et al., 2005; Castillo et al., 2010; Pallesen et al., 2006).
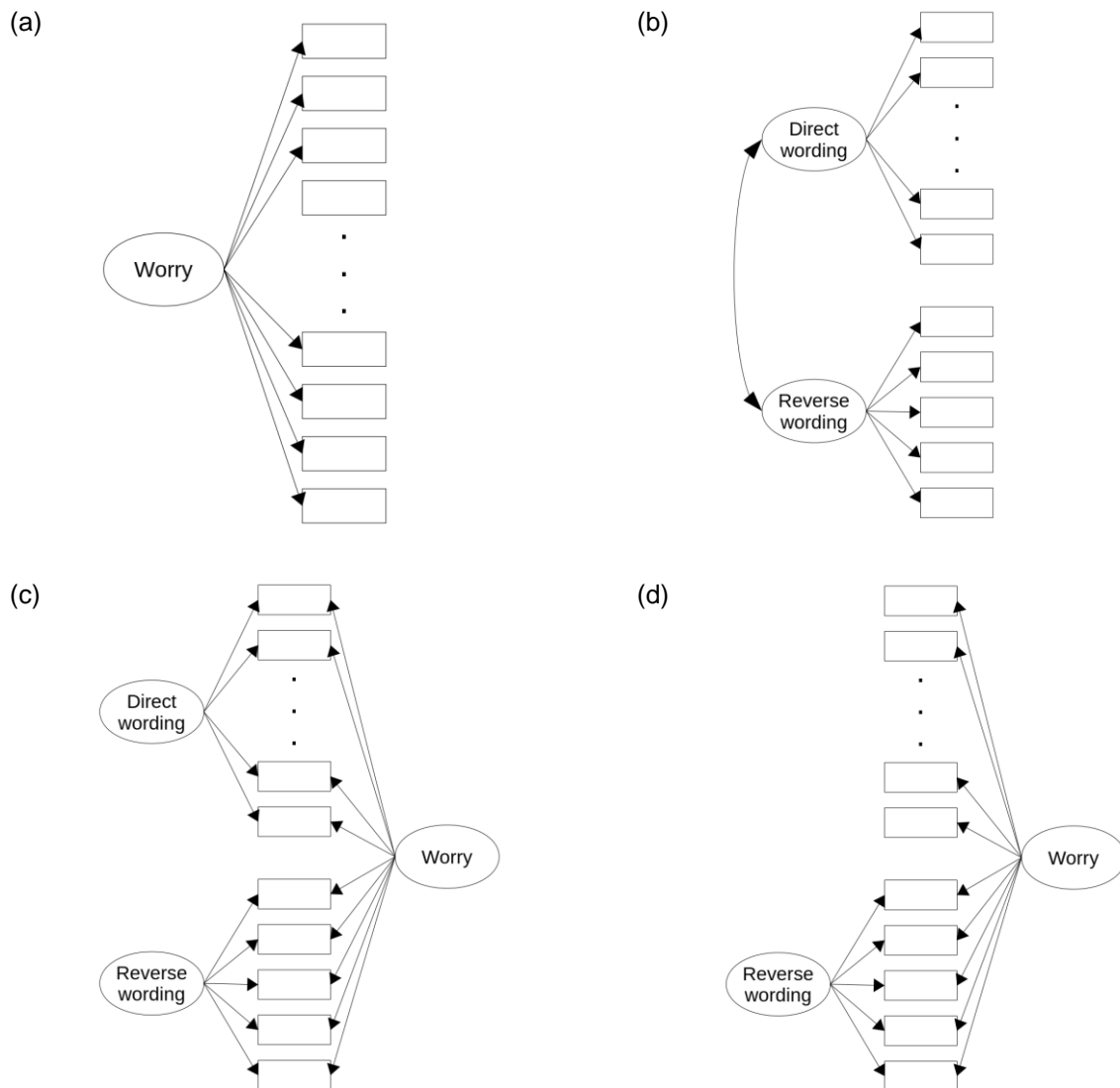

Figure 1. (a) Strictly unidimensional model; (b) two-factor model; (c) complete bifactor model; (d) incomplete bifactor model.

In addition to analyses of the internal structure of the full PSWQ, researchers have also examined

the possibility of reducing the number of items. In this way, faster responses could be obtained, while also maintaining good psychometric properties and achieving a clearer factor structure. One first obvious possibility is to retain only the 11 directly worded items, and this has indeed been tested with acceptable results (CFIs ≥ .94, RMSEA ≤ .09; Padros-Blazquez, Gonzalez-Betanzos, Martinez-Medina, & Wagner, 2018; Ruiz, Monroy-Cifuentes, & Suárez-Falcón, 2018; Sandín et al., 2009). Others have developed even briefer versions of the PSWQ, which can be of use in time-limited situations when respondent burden is an issue (e.g., very long protocols). One of these ultra-brief versions is the one proposed by Hopko et al. (2003), who retained only 8 items of the original PSWQ (2, 4, 5, 6, 7, 9, 12, 13). This 8-item version (named PSWQ-A) has shown good psychometric properties in a number of studies (Crittendon & Hopko, 2006; DeLapp et al., 2016; Padros-Blazquez et al., 2018; Wuthrich, Johnco, & Knight, 2014). In a similar vein, Topper, Emmelkamp, Watkins, & Ehring (2014) proposed a 5-item version of the PSWQ, which is composed of items 4, 5, 6, 12, and 13. However, this version has received less attention than Hopko et al.'s PSWQ-A (2003).

It is worth mentioning that even briefer versions have been proposed. For instance, some authors selected items 4, 14, and 15 to form an ultra-brief version which showed similar associations to other variables as the original PSWQ (Berle et al., 2011; Kertz et al., 2014). Moreover, for situations in which only one item can be used, it has been suggested that item 15 best suits this need (Schroder, Clark, & Moser, 2017). These 3- and 1-item versions, however, were not analyzed in the present report, since measurement models would be just-identified and their fit could not be assessed without setting additional restrictions.

The aim of this study was to examine the dimensionality of the PSWQ by comparing competing factor structures (one dimension, two dimensions, and bifactor models). In addition, brief versions (PSWQ-11, PSWQ-A, and PSWQ-5) were also analyzed. The adequacy of each model was tested through fit indices and reliability estimates. To the best of our knowledge, this is the first study to examine the PSWQ's factor structure in Peruvian undergraduates, and thus we expect our results to contribute to emotion regulation research in this context. As research shows, worry is a significant predictor of students' mental health (Bauer et al., 2020; Zvolensky et al., 2019), which in turn predicts academic achievement (Bruffaerts et al., 2018) and school dropout (Wang et al., 2015). Thus, examination and refinement of a widely used scale like the PSWQ will allow researchers to use this measure in contexts where it has not been used before (e.g., Peruvian undergraduates). Furthermore, as psychometricians have long acknowledged, there is no such thing as a "validated test," since validity evidence should be continually collected from different populations (Furr & Bacharach, 2014).

Methods

Participants

Participants were 290 psychology students (184 women, 63%) from one large public university in Lima, Peru. Their ages ranged between 15 and 32 years (M = 20.39, SD = 2.45). Most of them (68%) were first- and second-year students, but there was also an important proportion of fifth-year students (18%). In Peru, psychology programs usually last five years.

Measure

Penn State Worry Questionnaire (PSWQ; Meyer et al., 1990). The PSWQ is a 16-item measure of trait worry. Five items are reverse scored; thus they must be recoded before computing a global composite score. All items are responded using a Likert-type scale of five categories (1 = *Not at all typical of me*, 5 = *Very typical of me*). The PSWQ was created as a unidimensional measure; however, as previously mentioned, other possibilities exist and should be tested. In this study, we used the Spanish version of the questionnaire (Sandín et al., 2009), which has been previously used with Latin American samples (Padros-Blazquez et al., 2018; Ruiz et al., 2018). Before collecting our data, we unsystematically asked some potential users about the clarity of the items. No changes were made to the original translation. The psychometric properties of the PSWQ in our sample are detailed in the Results section.

Procedure

The PSWQ was administered in paper-and-pencil format in participants' classrooms after they

read an information sheet and gave their informed consent. All the data were collected by the first author. First, he entered the classroom and explained the purpose of the study. Second, he described what was expected from people who accepted to participate. Later, he handed the questionnaires to everyone present. The first part of each booklet gave details about the study and encouraged participants to continue only if they were willing to do so. Moreover, they were urged not to write down their names, so that the data could remain anonymous. Finally, the researcher collected the questionnaires. Even though we did not record the percentage of people who declined to take part in the study, anecdotal observations were made that this number was very low. Participants received neither financial benefits nor course credit for their participation. All the data were collected anonymously. According to ethical standards, only participants who accepted to participate were included in the study.

Data analysis

All the analyses were performed using R software, version 3.5.3 (R Core Team, 2019). Specifically, the following packages were used: psych 1.8.12 (Revelle, 2019) and lavaan 0.6-5 (Rosseel, 2012).

First, the items were analyzed descriptively, including mean, standard deviation, skewness, and kurtosis. As to skewness and kurtosis coefficients, values between -1 and +1 suggest that there is not a large deviation from univariate normality (Ferrando & Anguiano-Carrasco, 2010).

In order to examine the underlying latent structure of the PSWQ, confirmatory factor analyses (CFA) were conducted. Seven models were tested: Model 1, a one-factor model with all 16 items (Figure 1a; Meyer et al., 1990); Model 2, a two-factor model (Direct Wording and Reverse Wording) as reported by Fresco et al. (2002), which included the covariance between these two factors (Figure 1b); Model 3, a complete bifactor model, with one general worry factor and two other factors linked to the 11 positively worded and the 5 negatively worded items (Figure 1c); Model 4, an incomplete bifactor model, with a general worry factor and a residual factor for the 5 negatively worded items (Figure 1d). Likewise, the three

abbreviated versions were also examined: Model 5, the PSWQ-11 (Sandín et al., 2009); Model 6, the 8-item PSWQ-A (Hopko et al., 2003); and Model 7, the PSWQ-5 (Topper et al., 2014).

These CFAs used the weighted least squares mean and variance adjusted (WLSMV) estimator, which is considered appropriate for categorical variables (Brown, 2015). Furthermore, the chi-square statistic (χ^2) is reported to assess the fit of the models. Nevertheless, this index is sensitive to sample size so other goodness-of-fit measures were used (Browne & Cudeck, 1992; Hu & Bentler, 1999): the comparative fit index (CFI > .95), the Tucker-Lewis index (TLI > .95), the weighted root mean square residual (WRMR < 1; DiStefano, Liu, Jiang, & Shi, 2018), as well as the root mean square error of approximation (RMSEA < .06) and its 90% confidence interval. For the complete bifactor model, strength indices were calculated: explained common variance (ECV; Reise, Scheines, Widaman, & Haviland, 2013), percentage of uncontaminated correlations (PUC; Reise et al., 2013), omega hierarchical (ωh) for the general factor, and omega hierarchical for specific factors (ωhs; Zinbarg, Yovel, Revelle, & McDonald, 2006).

Traditionally, reliability has been estimated using coefficient alpha. However, the adequacy of this estimate depends on several assumptions, two of them being tau-equivalence and the absence of residual correlations (Raykov, 2012). Since these two assumptions seldom hold, other coefficients have been proposed, the most popular of which is, perhaps, coefficient omega (Dunn, Baguley, & Brunsden, 2014). In this study, the reliability of the scores was computed using categorical omega, a variation of coefficient omega that was proposed specifically for nonlinear latent variable models (Green & Yang, 2009).

Results

Table 1 presents the descriptive analyses of the items. It can be seen that item 15 (*M* = 2.48; *SD* = 1.03) and item 10 (*M* = 3.84; *SD* = 1.07) had the lowest and the highest mean values, respectively. Moreover, all skewness and kurtosis values were between -1 and +1.

Table 1.
Descriptive statistics of the PSWQ

Items	Mean	Standard deviation	Skewness	Kurtosis
1. If I do not have enough time to do everything, I do not worry about it. *Cuando no dispongo de tiempo suficiente para hacer todo lo que tengo que hacer, no me preocupo por ello.*	3.58	1.25	-0.56	-0.81
2. My worries overwhelm me. *Me agobian mis preocupaciones.*	3.11	1.04	0.09	-0.65
3. I do not tend to worry about things. *No suelo preocuparme por las cosas.*	3.71	1.00	-0.52	-0.31
4. Many situations make me worry. *Son muchas las circunstancias que hacen que me sienta preocupado(a).*	2.92	1.01	0.13	-0.76
5. I know I should not worry about things, but I just cannot help it. *Sé que no debería estar tan preocupado(a) por las cosas, pero no puedo hacer nada por evitarlo.*	2.73	1.08	0.23	-0.81
6. When I am under pressure I worry a lot. *Cuando estoy bajo estados de tensión tiendo a preocuparme muchísimo.*	3.33	1.13	-0.32	-0.82
7. I am always worrying about something. *Siempre estoy preocupado(a) por algo.*	2.52	1.04	0.43	-0.42
8. I find it easy to dismiss worrisome thoughts. *Me resulta fácil eliminar mis pensamientos de preocupación.*	3.18	1.06	-0.23	-0.68
9. As soon as I finish one task, I start to worry about everything else I have to do. *Tan pronto como termino una tarea, enseguida empiezo a preocuparme sobre alguna otra cosa que debo hacer.*	2.76	1.11	0.20	-0.79
10. I never worry about anything. *Nunca suelo estar preocupado(a).*	3.84	1.07	-0.55	-0.57
11. When there is nothing more I can do about a concern, I do not worry about it anymore. *Cuando no puedo hacer nada más sobre algún asunto, no vuelvo a preocuparme más de él.*	3.30	1.07	-0.29	-0.71
12. I have been a worrier all my life. *Toda mi vida he sido una persona muy preocupada.*	2.55	1.06	0.30	-0.71
13. I notice that I have been worrying about things. *Soy consciente de que me he preocupado excesivamente por las cosas.*	3.08	1.17	-0.20	-0.91
14. Once I start worrying, I cannot stop. *Una vez que comienzan mis preocupaciones no puedo detenerlas.*	2.51	1.03	0.35	-0.73
15. I worry all the time. *Estoy preocupado(a) constantemente.*	2.48	1.03	0.42	-0.47
16. I worry about projects until they are all done. *Cuando tengo algún proyecto no dejo de preocuparme hasta haberlo efectuado.*	3.34	1.02	-0.22	-0.53

As presented in Table 2, the strictly unidimensional model (Model 1) had acceptable fit when tested with comparative indices (CFI, TLI), but slightly mediocre fit when examined with absolute indices (WRMR, RMSEA). On the other hand, the oblique two-factor model (Model 2) showed better fit on all the indices, even though the correlation between both factors was large (φ = .67). Regarding Model 3 (complete bifactor model), it had acceptable goodness of fit values (Table 2). However, strength indices (ECV = .76, PUC = .46, ωh = .82) showed that the

measurement model was essentially unidimensional. Moreover, omega hierarchical coefficients for the specific factors were very low (Direct Wording: ωhs = .11; Reverse Wording: ωhs = .37). Table 2 also presents indices for the incomplete bifactor model (Model 4), which had worse fit than the complete bifactor model and was therefore not considered further. As to the three brief versions (Models 5–7). all of them provided good fit of the data (Table 2). Details about all the models tested are presented in Table 3.

Table 2.
Fit indices of the PSWQ's models

Model	χ^2 (df)	χ^2/df	CFI	TLI	WRMR	RMSEA [90% CI]
Model 1: One factor	284.272 (104)	2.73	.977	.973	1.056	.077 [.067, .088]
Model 2: Two factors	188.437 (103)	1.829	.989	.987	0.832	.054 [.041, .066]
Model 3: Complete bifactor	124.462 (88)	1.414	.995	.994	0.612	.038 [.021, .052]
Model 4: Incomplete bifactor	181.080 (99)	1.829	.989	.987	0.797	.054 [.041, .066]
Model 5: PSWQ-11	100.468 (44)	2.283	.993	.991	0.805	.067 [.049, .084]
Model 6: PSWQ-A	32.140 (20)	1.607	.997	.996	0.539	.046 [.009, .074]
Model 7: PSWQ-5	8.928 (5)	1.786	.998	.996	0.382	.052 [.000, .107]

Note. PSWQ = Penn State Worry Questionnaire. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, SRMR = Standardized Root Mean Square Residual, RMSEA = Root Mean Square Error of Approximation, CI = Confidence Intervals.

Coefficient omega was used to assess reliability of the unidimensional and oblique models (Models 1, 2, 5–7; Table 3). Model 1 had the lowest reliability (ω = .65). Moreover, in Model 2, reliability was strong for engagement of worry

(ω = .93) but weak for absence of worry (ω = .29). Results suggested that internal consistency was good for all three brief versions, but the highest value was obtained from the PSWQ-11 (ω = .92; Table 3).

Table 3.
Factor loadings and reliability of the PSWQ's models

Items	Model 1	Model 2*		Model 3			Model 4		Model 5	Model 6	Model 7
		DW	RW	G	DW	RW	G	RW			
1	.26		.36	.24		.34	.22	.35			
2	.72	.72		.72	.15		.72		.72	.74	
3	.38		.51	.35		.51	.32	.53			
4	.76	.76		.73	.23		.76		.77	.78	.75
5	.80	.80		.82	.13		.80		.80	.82	.85
6	.70	.70		.75	.00		.70		.68	.70	.69
7	.83	.83		.78	.29		.83		.83	.84	
8	.54		.74	.54		.38	.50	.42			
9	.52	.52		.50	.16		.52		.52	.51	
10	.52		.71	.68		.41	.48	.44			
11	.24		.34	.22		.37	.20	.38			
12	.78	.79		.68	.41		.79		.79	.77	.75
13	.75	.75		.71	.24		.75		.75	.74	.77
14	.85	.85		.72	.44		.85		.85		
15	.86	.86		.66	.68		.86		.86		
16	.64	.64		.64	.13		.64		.64		
λ_{mean}	.63	.75	.53	.61	.26	.40	.62	.42	.75	.74	.76
ω	.65	.93	.29						.92	.89	.85

Note. G = general worry factor; DW = direct wording factor; RW = reverse wording factor; λ_{mean} = mean of factor loadings; ω = categorical omega coefficient. *The inter-factor correlation was .67

Discussion

The present study examined the dimensionality of the PSWQ, as well as existing brief versions. Consistent with previous reports, the PSWQ was essentially unidimensional in our data (Brown, 2003; Castillo et al., 2010; Hazlett-Stevens et al., 2004; Pallesen et al., 2006). Moreover, all the brief versions fit a strictly unidimensional structure with good levels of reliability. That is, only one global score should be computed for the PSWQ, and brief versions may be better measures of worry (i.e., with less measurement error).

Traditionally, in order to control for acquiescence, psychometricians have included reverse-scored items in their tests (Furr & Bacharach, 2014). However, this approach has been questioned for creating spurious multidimensionality, and some methodologists have even proposed that reversed items be avoided in most situations (Suárez-Alvarez et al., 2018). Alternatively, others have developed methods for determining whether data are essentially unidimensional, that is, whether they can be treated as unidimensional even though, strictly speaking, they are not (Reise et al., 2013). When our PSWQ data were examined with these methods, we found clear support for essential

unidimensionality. Therefore, one practical implication of our results is that researchers and practitioners should consider only one global worry score, and not two dimensions as sometimes has been proposed (e.g., Fresco et al., 2002).

Even though the two-factor model did show good fit in our data, statistical fit should not be the primary reason why we select a model. Indeed, as Brown (2003) pointed out, the supposed bidimensionality of the PSWQ seems to be a methodological artifact due to reverse-wording of some items. This reasoning is further supported by the large correlation between both factors (φ = .67 in our data). The complete bifactor model (Figure 1c), by including both the two specific factors and the global one, allows determination of which of these factors are actually relevant. As stated, our results indicate that, regarding the PSWQ, only a general Worry dimension should be considered. This is consistent with theoretical accounts of worry, which consider it to be a unitary cognitive phenomenon (Borkovec et al., 1998).

The complete bifactor model (Model 3; Figure 1c), besides showing the robustness of the general dimension, also provides additional information at the item level. Specifically, some items seemed to be almost perfect indicators of the global factor (e.g., item 6: *When I am under pressure I worry a lot*), whereas others seemed less connected to the global factor and more to their specific factors (e.g. item 11: *When there is nothing more I can do about a concern, I do not worry about it anymore*). The two items with the lowest loadings on the global factor (items 1 and 11) have both complex phrasings, which include double negatives. Therefore, it is possible that these items are very influenced by other, unmeasured variables, such as reading skills. Also, it is worth noting that the Reverse Wording factor showed more consistent loadings (all λs > .30) than the Direct Wording one, again suggesting that all the reversed items are similarly influenced by other variables besides worry. On the other hand, most of the directly worded items had stronger loadings on the global factor. Therefore, another practical implication of our results is that only these items should be retained if worry is to be measured with less measurement error.

A detailed examination of the strictly unidimensional model (Model 1) shall enrich our discussion on essential unidimensionality of the PSWQ. First, it can be noted that the factor loadings of this model are very heterogeneous (ranging from .24 to .86). Following our previous reasoning, this result seems to imply that some items are good indicators of worry, while others are more influenced by external variables. Second, it can also be seen that the reliability of the 16-item PSWQ's composite score is rather low (ω = .65). Since longer scales are usually associated with increased reliability, this result may seem paradoxical at first sight.

In order to understand the preceding result, it should be noted that, in this study, reliability was estimated using categorical omega, a special method proposed for non-linear models (Green & Yang, 2009). Even though other methods originally proposed for linear factor analytical models are routinely applied to non-linear models (Gadermann, Guhn, & Zumbo, 2012), this approach has been criticized (Chalmers, 2017). Therefore, we selected categorical omega as a better suited method for estimating reliability. An important characteristic of this method is that the total variance of the composite score (i.e., the denominator of the formula) is calculated from item polychoric correlations, which is equivalent to calculating it from observed data. Therefore, categorical omega can be seen as similar in logic to hierarchical omega in that the denominator includes all the observed variance, while the numerator only covers variance related to the factor of interest (i.e., worry; Kelley & Pornprasertmanit, 2016). This explains why the reliability estimate of Model 1 is so low: The "signal" (i.e., true score Worry variance) is small compared to the "noise" (i.e., unmodeled variance).

Alongside the bifactor models, the previous finding suggests that the reversed items are contributing much of this "noise." Accordingly, retaining only variables that are worded in a direct way should reduce measurement error and, therefore, increase reliability. This is exactly what was intended with the brief versions of the PSWQ. Regarding these, we found comparable results for the 11, 8, and 5-item versions. The highest reliability coefficient was obtained by the PSWQ-11 (ω = .92), which is expected given that reliability estimates such as omega are influenced

by scale length. The other two brief versions also had good reliability, and hence could be useful in situations where even the PSWQ-11 is not short enough.

The current results should be interpreted considering several limitations. First, a non-probability sampling method was employed, and the sample consisted of psychology students only; thus, its homogeneous nature restricts the generalizability of these findings. Second, sample size was determined by practical reasons (i.e., how many participants could be evaluated). Although our data (n = 290) would be enough according to popular rules-of-thumb which state that minimum sample size is 200, it is important to bear in mind that such guidance can often be misleading (Kline, 2016). Therefore, future studies should include larger sample sizes. A third limitation, also related to sample size, is that measurement invariance between genders could not be tested. This is relevant because, as an anonymous reviewer correctly pointed out, affect-related variables such as worry tend to be strongly influenced by culture, of which gender differences are an important part. Fourth, although the dimensionality was evaluated, other sources of validity evidence are necessary. Fifth, the three brief versions were not administered separately from the full version. It is possible, then, that item position could have influenced the results (Podsakoff, MacKenzie, & Podsakoff, 2012). On the other hand, having only one application avoids possible complications associated with repeated measures. Finally, the fact that no language adaptation process was formally conducted could pose a threat to our results. However, it is important to bear in mind that other Latin American studies have also used the same Spanish translation (e.g., Padros-Blazquez et al., 2018).

To sum up, this study provides initial evidence of the factor structure of the PSWQ in a Peruvian student sample. Our results show that only one global dimension should be considered when using the PSWQ. Furthermore, shorter versions that only retain the directly worded items seem preferable. However, it is recommended to conduct research on this measure in different contexts, such as clinical samples, before our conclusions can be applied widely. For the prevention of psychological disorders, the shorter versions of the PSWQ may also be effective as time-efficient screeners.

Acknowledgments

References

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., … Kessler, R. C. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology*, *127*(7), 623–638. doi: 10.1037/abn0000362

Bauer, E. A., Braitman, A. L., Judah, M. R., & Cigularov, K. P. (2020). Worry as a mediator between psychosocial stressors and emotional sequelae: Moderation by contrast avoidance. *Journal of Affective Disorders*, *266*, 456–464. doi: 10.1016/j.jad.2020.01.092

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1995). Psychometric properties of the Penn State Worry Questionnaire in older adults. *Journal of Clinical Geropsychology*, *1*(1), 33–42.

Berle, D., Starcevic, V., Moses, K., Hannan, A., Milicevic, D., & Sammut, P. (2011). Preliminary validation of an ultra-brief version of the Penn State Worry Questionnaire. *Clinical Psychology & Psychotherapy*, *18*(4), 339–346. doi: 10.1002/cpp.724

Borkovec, T. D., & Inz, J. (1990). The nature of worry in generalized anxiety disorder: a predominance of thought activity. *Behaviour Research and Therapy*, *28*(2), 153–158. doi: 10.1016/0005-7967(90)90027-g

Borkovec, T. D., Ray, W. J., & Stöber, J. (1998). Worry: A cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioral processes. *Cognitive Therapy and Research*, *22*(6), 561–576. doi: 10.1023/A:1018790003416

Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy*, *21*(1), 9–16. doi: 10.1016/0005-7967(83)90121-3

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire:

Multiple factors or method effects? *Behaviour Research and Therapy, 41*(12), 1411–1426. doi: 10.1016/S0005-7967(03)00059-7

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, USA: Guilford Press.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1992). Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. *Behaviour Research and Therapy, 30*(1), 33–37. doi: 10.1016/0005-7967(92)90093-V

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258. doi: 10.1177/0049124192021002005

Bruffaerts, R., Mortier, P., Kiekens, G., Auerbach, R. P., Cuijpers, P., Demyttenaere, K., … Kessler, R. C. (2018). Mental health problems in college freshmen: Prevalence and academic functioning. *Journal of Affective Disorders, 225*, 97–103. doi: 10.1016/J.JAD.2017.07.044

Carter, M. M., Sbrocco, T., Miller, O., Suchday, S., Lewis, E. L., & Freedman, R. E. K. (2005). Factor structure, reliability, and validity of the Penn State Worry Questionnaire: differences between African-American and White-American college students. *Journal of Anxiety Disorders, 19*(8), 827–843. doi: 10.1016/J.JANXDIS.2004.11.001

Castillo, C., Macrini, L., Cheniaux, E., & Landeira-Fernandez, J. (2010). Psychometric properties and latent structure of the Portuguese version of the Penn State Worry Questionnaire. *The Spanish Journal of Psychology, 13*(1), 431–443. doi: 10.1017/S113874160000398X

Chalmers, R. P. (2017). On misconceptions and the limited usefulness of ordinal alpha. *Educational and Psychological Measurement, 78*(6), 1056–1071. doi: 10.1177/0013164417727036

Chen, F. F., West, S., & Sousa, K. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41*(2), 189–225. doi: 10.1207/s15327906mbr4102_5

Crittendon, J., & Hopko, D. R. (2006). Assessing worry in older and younger adults: Psychometric properties of an abbreviated Penn State Worry Questionnaire (PSWQ-A). *Journal of Anxiety Disorders, 20*(8), 1036–1054. doi: 10.1016/j.janxdis.2005.11.006

Dar, K. A., Iqbal, N., & Mushtaq, A. (2017). Intolerance of uncertainty, depression, and anxiety: Examining the indirect and moderating effects of worry. *Asian Journal of Psychiatry, 29*, 129–133. doi: 10.1016/j.ajp.2017.04.017

DeLapp, R. C. T., Chapman, L. K., & Williams, M. T. (2016). Psychometric properties of a brief version of the Penn State Worry Questionnaire in African Americans and European Americans. *Psychological Assessment, 28*(5), 499–508. doi: 10.1037/pas0000208

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the Weighted Root Mean Square Residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 453–466. doi: 10.1080/10705511.2017.1390394

Duffy, A., Saunders, K. E. A., Malhi, G. S., Patten, S., Cipriani, A., McNevin, S. H., … Geddes, J. (2019). Mental health care for university students: a way forward? *The Lancet Psychiatry, 6*(11), 885–887. doi: 10.1016/S2215-0366(19)30275-5

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. doi: 10.1111/bjop.12046

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy, 1*(3), 192–205. doi: 10.1521/ijct.2008.1.3.192

Elhai, J. D., Rozgonjuk, D., Yildirim, C., Alghraibeh, A. M., & Alafnan, A. A. (2019). Worry and anger are associated with latent classes of problematic smartphone use severity among college students. *Journal of Affective Disorders, 246*, 209–216. doi: 10.1016/j.jad.2018.12.047

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo, 31*(1), 18–33.

Fresco, D. M., Heimberg, R. G., Mennin, D. S., & Turk, C. L. (2002). Confirmatory factor analysis of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 40*(3), 313–323. doi: 10.1016/S0005-7967(00)00113-3

Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An introduction*. New York,

USA: SAGE.

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, *17*, 1-13. doi: 10.7275/n560-j767

Gana, K., Martin, B., Canouet, M.-D., Trouillet, R., & Meloni, F. (2002). Factorial structure of a French version of the Penn State Worry Questionnaire. *European Journal of Psychological Assessment*, *18*(2), 158–164. doi: 10.1027/1015-5759.18.2.158

Gorday, J. Y., Rogers, M. L., & Joiner, T. E. (2018). Examining characteristics of worry in relation to depression, anxiety, and suicidal ideation and attempts. *Journal of Psychiatric Research*, *107*, 97–103. doi: 10.1016/j.jpsychires.2018.10.004

Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155–167. doi: 10.1007/s11336-008-9099-3

Hayes, S. C., Wilson, K. G., Gifford, E. V, Follette, V. M., & Strosahl, K. D. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, *64*(6), 1152–1168. doi: 10.1037/0022-006X.64.6.1152

Hayes-Skelton, S. A., & Eustis, E. H. (2020). Experiential avoidance. En J. S. Abramowitz & S. M. Blakey (Eds.), *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 115–131). USA: American Psychological Association.

Hazlett-Stevens, H., Ullman, J. B., & Craske, M. G. (2004). Factor structure of the Penn State Worry Questionnaire: Examination of a method factor. *Assessment*, *11*(4), 361–370. doi: 10.1177/1073191104269872

Hopko, D. R., Reas, D. L., Beck, J. G., Stanley, M. A., Wetherell, J. L., Novy, D. M., & Averill, P. M. (2003). Assessing worry in older adults: Confirmatory factor analysis of the Penn State Worry Questionnaire and psychometric properties of an abbreviated model. *Psychological Assessment*, *15*(2), 173–183. doi: 10.1037/1040-3590.15.2.173

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modelling*, *6*(1), 1–55. doi: 10.1080/10705519909540118

Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods*, *21*(1), 69–92. doi: 10.1037/a0040086

Kertz, S. J., Lee, J., & Björgvinsson, T. (2014). Psychometric properties of abbreviated and ultra-brief versions of the Penn State Worry Questionnaire. *Psychological Assessment*, *26*(4), 1146–1154. doi: 10.1037/a0037251

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, USA: The Guilford Press.

Lim, Y. J., Kim, Y. H., Lee, E. H., & Kwon, S. M. (2008). The Penn State Worry Questionnaire: Psychometric properties of the Korean version. *Depression and Anxiety*, *25*(10), E97–E103. doi: 10.1002/da.20356

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*(6), 487–495. doi: 10.1016/0005-7967(90)90135-6

Nolen-Hoeksema, S., & Watkins, E. R. (2011). A heuristic for developing transdiagnostic models of psychopathology. *Perspectives on Psychological Science*, *6*(6), 589–609. doi: 10.1177/1745691611419672

Padros-Blazquez, F., Gonzalez-Betanzos, F., Martinez-Medina, M. P., & Wagner, F. (2018). Psychometric characteristics of the original and brief version of the Penn State Worry Questionnaire (PSWQ) in Mexican samples. *Actas Españolas de Psiquiatría*, *46*(4), 117–124.

Pallesen, S., Nordhus, I. H., Carlstedt, B., Thayer, J. F., & Johnsen, T. B. (2006). A Norwegian adaptation of the Penn State Worry Questionnaire: Factor structure, reliability, validity and norms. *Scandinavian Journal of Psychology*, *47*(4), 281–291. doi: 10.1111/j.1467-9450.2006.00518.x

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*(1), 539–569. doi:

10.1146/annurev-psych-120710-100452

Purdon, C., & Harrington, J. (2006). Worry in psychopathology. En G. C. L. Davey & A. Wells (Eds.), *Worry and its psychological disorders: Theory, assessment and treatment* (pp. 41–50). New Jersey, USA: Wiley Publishing.

Raykov, T. (2012). Scale construction and development using structural equation modeling. En R. H. Hoyle (Ed.), *Handbook of structural equation modeling*, (pp. 472–492). New York, USA: Guilford Press.

R Core Team. (2019). R: A language and environment for statistical computing (version 3.5.3) [software]. Vienna: R Foundation of Statistical Computing. Recuperado de: https://www.r-project.org

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement, 73*(1), 5–26. doi: 10.1177/0013164412449831

Revelle, W. (2019). Psych: Procedures for Personality and Psychological Research (version 1.8.12) [software]. Recuperado de: https://cran.r-project.org/package=psych

Rodríguez-Biglieri, R., & Vetere, G. L. (2011). Psychometric characteristics of the Penn State Worry Questionnaire in an Argentinean sample: A cross-cultural contribution. *The Spanish Journal of Psychology*, *14*(1), 452–463. doi: 10.5209/rev_SJOP.2011.v14.n1.41

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. doi: 10.18637/jss.v048.i02

Ruiz, F. J., Monroy-Cifuentes, A., & Suárez-Falcón, J. C. (2018). Penn State Worry Questionnaire-11 validity in Colombia and factorial equivalence across gender and nonclinical and clinical samples. *Anales de Psicología / Annals of Psychology*, *34*(3), 451–457. doi: 10.6018/analesps.34.3.300281

Sandín, B., Chorot, P., Valiente, R. M., & Lostao, L. (2009). Validación española del cuestionario de preocupación PSWQ: estructura factorial y propiedades psicométricas. *Revista de Psicopatología y Psicología Clínica*, *14*(2), 107–122. doi: 10.5944/rppc.vol.14.num.2.2009.4070

Schroder, H. S., Clark, D. A., & Moser, J. S.

(2017). Screening for problematic worry in adults with a single item from the Penn State Worry Questionnaire. *Assessment*, *26*(2), 336–346. doi: 10.1177/1073191117694453

Sibrava, N. J., & Borkovec, T. D. (2006). The cognitive avoidance theory of worry. En G. C. L. Davey & A. Wells (Eds.), *Worry and its psychological disorders: Theory, assessment and treatment* (pp. 239–256). New Jersey, USA: Wiley Publishing.

Spinhoven, P., van der Veen, D. C., Voshaar, R. C. O., & Comijs, H. C. (2017). Worry and cognitive control predict course trajectories of anxiety in older adults with late-life depression. *European Psychiatry*, *44*, 134–140. doi: 10.1016/j.eurpsy.2017.05.002

Suárez-Alvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., & Muñiz, J. (2018). Using reversed items in Likert scales: A questionable practice. *Psicothema*, *30*(2), 149–158. doi: 10.7334/psicothema2018.33

Topper, M., Emmelkamp, P. M. G., Watkins, E., & Ehring, T. (2014). Development and assessment of brief versions of the Penn State Worry Questionnaire and the Ruminative Response Scale. *British Journal of Clinical Psychology*, *53*(4), 402–421. doi: 10.1111/bjc.12052

van Rijsoort, S., Emmelkamp, P., & Vervaeke, G. (1999). The Penn State Worry Questionnaire and the Worry Domains Questionnaire: structure, reliability and validity. *Clinical Psychology & Psychotherapy*, *6*(4), 297–307. doi: 10.1002/(SICI)1099-0879(199910)6:4<297::AID-CPP206>3.0.CO;2-E

Verkuil, B., & Brosschot, J. F. (2012). The online version of the Dutch Penn State Worry Questionnaire: Factor structure, predictive validity and reliability. *Journal of Anxiety Disorders*, *26*(8), 844–848. doi: 10.1016/j.janxdis.2012.08.002

Wang, H., Yang, C., He, F., Shi, Y., Qu, Q., Rozelle, S., & Chu, J. (2015). Mental health and dropout behavior: A cross-sectional study of junior high students in northwest rural China. *International Journal of Educational Development*, *41*, 1–12. doi: 10.1016/j.ijedudev.2014.12.005

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206. doi:

10.1037/0033-2909.134.2.163

Watkins, E. R. (2015). An alternative transdiagnostic mechanistic approach to affective disorders illustrated with research from clinical psychology. *Emotion Review*, *7*(3), 250–255. doi: 10.1177/1754073915575400

Wuthrich, V. M., Johnco, C., & Knight, A. (2014). Comparison of the Penn State Worry Questionnaire (PSWQ) and abbreviated version (PSWQ-A) in a clinical and non-clinical population of older adults. *Journal of Anxiety Disorders*, *28*(7), 657–663. doi: 10.1016/j.janxdis.2014.07.005

Zhong, J., Wang, C., Li, J., & Liu, J. (2009). Penn State Worry Questionnaire: Structure and psychometric properties of the Chinese version. *Journal of Zhejiang University SCIENCE B*, *10*(3), 211-218. doi: 10.1631/jzus.B0820189

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating Generalizability to a Latent Variable Common to All of a Scale's Indicators: A comparison of estimators for ωh. *Applied Psychological Measurement, 30*(2), 121-144. doi: 10.1177/0146621605278814

Zvolensky, M. J., Kauffman, B. Y., Bogiaizian, D., Viana, A. G., Bakhshaie, J., & Peraza, N. (2019). Worry among Latinx college students: relations to anxious arousal, social anxiety, general depression, and insomnia. *Journal of American College Health*, 1–8. Publicación anticipada en línea. doi: 10.1080/07448481.2019.1686004