Original Research Article

Assessment
1–13

DOI: 10.1177/10731911251327255
journals.sagepub.com/home/asm
Sage

Normative Values and Psychometric Properties of the Penn State Worry Questionnaire in Substance Use Disorder Treatment Population

Martin Hochheimer[1], Justin C. Strickland[1], Jennifer D. Ellis[1], Jill A. Rabinowitz[2], J. Gregory Hobelmann[3], and Andrew S. Huhn[1]

Abstract

This study evaluated the Penn State Worry Questionnaire (PSWQ) as a tool for measuring worry and anxiety levels among individuals entering treatment for substance use disorders (SUDs). The sample included 75,047 individuals admitted to SUD treatment centers, with assessments conducted weekly. Individuals entering SUD treatment exhibited higher baseline levels of worry; however, worry levels declined over the course of treatment. The PSWQ demonstrated good internal consistency, high test-retest reliability, and good discriminant validity when correlated with measures of depression and stress. The factor structure analysis confirmed that the PSWQ measures the same underlying construct of worry in the SUD treatment population, with a single-factor model showing satisfactory fit. This extends the reach of the PSWQ to the SUD treatment population by reaffirming its reliability, validity, and factor structure, with the expectation of higher levels of worry compared to a non-SUD population at the beginning of treatment, which decline over time.

Keywords

substance use disorder, anxiety, Penn State Worry Questionnaire, psychometric properties, treatment population

Introduction

For individuals with substance use disorders (SUDs), continued drug and/or alcohol use is driven, in part, by a desire to relieve negative affective states associated with abstinence (Koob, 2015). Co-occurring mental health conditions, such as anxiety disorders, are common among persons with SUDs. Treating these conditions during SUD treatment can be challenging, as it is often difficult to distinguish between pre-existing or co-occurring mental health conditions and symptoms that arise as a direct consequence of long-term substance use. The presence of mental health symptoms at SUD treatment intake and their trajectory during treatment have been associated with treatment attrition and relapse (Boschloo et al., 2012; Ellis et al., 2022; Rabinowitz et al., 2023); yet, there is limited information regarding normative values for scales that measure mental health symptoms, such as anxiety, among individuals with SUD, which could help clinicians benchmark individuals as they enter and participate in SUD treatment.

Worry, or apprehensive expectation about potentially negative events, is defined as a clinically relevant symptom when "the intensity, duration, or frequency is out of proportion to the actual likelihood or impact of the anticipated event" (American Psychiatric Association, 2013, p. 222) and is a common feature of anxiety disorders, specifically generalized anxiety disorder (GAD). Among individuals with anxiety disorders, it is not the presence of worry in and of itself, which all humans experience, but rather the experience of high levels of worry that is often disproportionate to the events experienced. Worry in individuals with SUDs may differ from that observed in anxiety disorders due to the effects of substance use, withdrawal, and psychosocial stressors unique to SUD such as uncertainty regarding recovery, or external stressors such as legal or financial concerns (J. F. Kelly & Hoeppner, 2013; W. E. Kelly, 2008; Koob & Volkow, 2016; Schuckit, 2006). While worry is a core feature of GAD, the elevated worry experienced in early

[1]Johns Hopkins University School of Medicine, Baltimore, MD, USA
[2]Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ, USA
[3]Ashley Addiction Treatment, Havre de Grace, MD, USA

Corresponding Author:
Martin Hochheimer, Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, 5510 Nathan Shock Drive, Baltimore, MD 21224, USA.
Email: mhochhe1@jh.edu

SUD treatment may be influenced by withdrawal-related distress, uncertainty regarding recovery, or external stressors such as legal or financial concerns.

The Penn State Worry Questionnaire (PSWQ) was developed to measure the frequency and intensity of clinically relevant worry in the general population (Meyer et al., 1990) and to ideally aid clinicians in measuring the pervasiveness of worry as one of several symptoms that identify an anxiety disorder. The PSWQ is a 16-item scale that measures the generality, intensity, and uncontrollability of worry whereby individuals are asked to rate how typical certain patterns of worry are for them. Items are scored on a 1 to 5 rating scale (scores ranging from 16 to 80) with lower scores indicating lower levels of worry. There are 11 positively worded items and 5 negatively worded items, which are reverse coded.

The PSWQ has been widely used to measure clinically significant worry in the general population. Previous studies have suggested cutoff points at which worry moves from being a ubiquitous human experience to a clinically relevant symptom, for example scores above 45 (Behar et al., 2003) or 50 (Wuthrich et al., 2014) have been proposed as thresholds for identifying clinically relevant worry in the general population. Non-clinical, community samples typically have mean scores ranging from 35 to 50 (Brenes et al., 2022; Fortune et al., 2005; Meyer et al., 1990; Oliveira et al., 2023; Pallesen et al., 2006; Rodríguez-Biglieri & Vetere, 2011; van Rijsoort et al., 1999). These cutoffs and means provide context for interpreting worry levels measured by PSWQ in the general population. This study focuses on evaluating their potential application in identifying elevated worry within the unique context of individuals entering SUD treatment, where the meaning of PSWQ scores could be interpreted differently.

Numerous studies have demonstrated the reliability and validity of the PSWQ scores in various populations, showing high internal consistency (Cronbach's α above .85) and strong test-retest reliability ($r > .74$ to $r > .84$) (T. A. Brown, 2003; T. A. Brown et al., 1992; Davey, 1993; Meyer et al., 1990; Pallesen et al., 2006; van Rijsoort et al., 1999) and high test-retest reliability ($r > .74$ in one study and $r > .84$ in others) (Meyer et al., 1990; Molina & Borkovec, 1994; Pallesen et al., 2006; Stöber & Bittencourt, 1998). The PSWQ previously demonstrated to have high convergent validity with the Worry Domains Questionnaire, (r = .61–.67) (Davey, 1993; Tallis et al., 1992; van Rijsoort et al., 1999) and the State Trait Anxiety Inventory (trait r = .64–.79 state r = .49) (Meyer et al., 1990; Pallesen et al., 2006; van Rijsoort et al., 1999), as well as shown discriminant validity with the Beck Depression Inventory (r = .36–.62) (Meyer et al., 1990; Pallesen et al., 2006; van Rijsoort et al., 1999) and the rumination (r = .67) and impulses (r = .60) subscales of the Padua Inventory Revised (van Rijsoort et al., 1999), which is a scale for obsessive compulsive disorder.

However, the psychometric properties of the PSWQ have not been extensively examined in SUD populations, where worry symptoms may arise from different underlying mechanisms than in the general population or those with primary anxiety disorders. Research suggests that substance use may mimic or exacerbate symptoms of anxiety (Schuckit, 2006) and that chronic substance use dysregulates the brain's stress and reward systems in ways that may alter how worry is experienced (Koob & Volkow, 2016). Additionally, psychosocial factors, such as gender differences in emotional regulation and stress coping strategies, may further shape the manifestation of worry in SUD populations (J. F. Kelly & Hoeppner, 2013). Given these neurobiological and psychosocial considerations, it is important to assess whether standard worry measures function equivalently in SUD populations compared to the general population.

The PSWQ was initially conceived as measuring one underlying construct that of "worry" but some early studies identified two latent factors: one made up of those items worded positively and the second of those items worded negatively which were described as "worry" (or "worry engagement") and "absence of worry" (Beck et al., 1995; T. A. Brown et al., 1992; Fresco et al., 2002; Meyer et al., 1990; Stoeber, 1995; van Rijsoort et al., 1999). However, a convincing case has been made that the one factor model should be retained. First, the construct of "absence of worry" is difficult to interpret, and second, the difference in response patterns should be thought of as an anomaly due to the positive and negative valence of the items rather than a difference in underlying constructs (T. A. Brown, 2003).

However, the psychometric properties of the PSWQ may not generalize uniformly across all populations. Worry is common among those in treatment for SUDs (Lai et al., 2015). This may be due to the differing physiological and psychosocial effects of the substance use and withdrawal. Anxiety and worry are commonly reported as symptoms of withdrawal (Vorspan et al., 2015). Thus, levels of worry and anxiety symptoms among those who are beginning treatment for SUDs need to be interpreted differently than in either a normative sample or one with anxiety disorders, such as GAD. These stressors faced by individuals in SUD treatment may lead to worry that is proportionate to their challenges, rather than excessive or irrational. Furthermore, during the prolonged period of withdrawal, the level of heightened anxiety and worry fluctuates due to the physiological effects of withdrawal (Bluthenthal et al., 2020; Chartoff & Carlezon, 2014; Jesse et al., 2017; Schuckit, 2014). For these reasons, it is possible that worry driven

by substance withdrawal may be qualitatively different from worry unrelated to withdrawal and may not be measured accurately by the PSWQ.

The PSWQ has become one of the most widely used self-report measure of worry over the last three decades (Johnco et al., 2022; Oliveira et al., 2023; Xie et al., 2023) due to its high internal consistency and strong psychometric support across various populations, including those with anxiety disorders, community samples, and older adults (Johnco et al., 2022; Oliveira et al., 2023). However, its psychometric properties have not been thoroughly examined within the context of a SUD population. Given the unique psychological and physiological stressors that individuals in SUD treatment face, as well as the potential overlap between substance-induced anxiety and pre-existing anxiety disorders (Fatséas et al., 2010; Garey et al., 2020; McHugh, 2015; Schuckit, 2006), it is important to explore whether the psychometric properties of the PSWQ might differ in this population. Specifically, we hypothesized that the elevated worry seen at the start of SUD treatment might affect the factor structure and lower the internal reliability of the PSWQ. Additionally, the evaluation of concurrent and discriminant validity was exploratory, as prior evidence to support specific associations in the SUD population is limited. Understanding these psychometric properties in this context is crucial, as clinicians rely on accurate and reliable measures to assess anxiety and worry levels, which can impact treatment planning and outcomes.

While the psychometric properties of the PSWQ have been evaluated in many different populations (Chorpita et al., 1997; Crittendon & Hopko, 2006; Kertz et al., 2014; Puccinelli et al., 2023) and translated into many languages (Ediati & Utari, 2019; Motooka et al., 2009; Pallesen et al., 2006; Ruiz et al., 2018), it has not been evaluated within those entering and participating in SUD treatment sample whose experience of anxiety may be qualitatively different (S. A. Brown & Schuckit, 1988; J. F. Kelly & Hoeppner, 2013; Koob & Volkow, 2016). The aim of this study was to determine the normative values, internal consistency, test-retest reliability, discriminant validity, and factor structure of the PSWQ among individuals in treatment for SUD, for each of eight substances during the first 4 weeks of treatment.

Method

Sample and Data Collection

The sample consisted of 75,047 unique individuals who were admitted to 113 SUD treatment centers located across the United States with the majority in the South and West regions, between January 1, 2015 and November 9, 2020. Individuals were treated in one of three levels of care: supervised withdrawal, intensive outpatient, or residential treatment; data were only considered for the first admission to treatment and for the first level of care (if multiple were present). Assessments were delivered on a weekly basis to patients electronically (e.g., on computer or tablet), and data were collected by a third-party treatment outcomes provider (Trac9). De-identified data were provided for research through a data transfer agreement, which was acknowledged by the Johns Hopkins University School of Medicine Institutional Review Board as non-human subjects research.

Data was collected over the course of treatment (for up to 20 weeks). Patients were asked to complete an assessment approximately every week, but circumstances of treatment meant that not every individual was given every assessment, and some were given more than one survey in a week. Cross-sectional analysis was conducted at intake using individuals from all levels of care and intake was defined as the first assessment given during the first 3 days of treatment. Longitudinal analysis of the first 4 weeks of treatment was limited to those who were in either intensive outpatient or residential treatment because these treatment programs are typically structured to last at least 1 month. This analysis did not include those in supervised withdrawal (i.e., detoxification) programs due to the short-term nature of these programs as well as the notion that anxiety can be a symptom of withdrawal. Weeks in treatment were defined as the first assessment given 3 days before or after treatment days 7, 14, 21, and 28.

Statistical Analysis

Normative Values. Means, standard deviations, and percentiles of the PSWQ were calculated at intake, and means (with standard deviations) and medians were calculated after 1, 2, 3, and 4 weeks of treatment. It is important to note that the sample is large, and almost all null hypothesis statistical tests would be significant at the .05 level, indicating that differences between groups would exist within the SUD population, but these differences were not necessarily clinically relevant. Therefore, comparisons between groups were limited to omnibus tests with post hoc tests, only conducted when the effect size was medium according to Cohen's guidelines (Acock, 2014; Cohen, 2013; Fey et al., 2023; Nieminen, 2022) expressed as $\eta > .06$, or Cohen's $\omega > .30$.

Reliability. Reliability was tested at intake using Cronbach's α and McDonald's ρ. Cronbach's α is the commonly reported measure of internal consistency; however, it is not appropriate when the instrument being

evaluated has two or more factors since each factor may have its own underlying source of variation. As noted above, the factor structure of the PSWQ has been debated; therefore, we are additionally reporting the McDonald's ρ to account for the possibility of a two-factor structure. Test re-test reliability was assessed in light of the natural decline in anxiety that is associated with engaging in SUD treatment. Therefore, in addition to reporting the correlation between the scores across weeks, the relative change from week to week was gauged by determining the intraclass correlation coefficients (ICC) (3,k). The ICC(3,k) is the two-way mixed-effects model that measures absolute agreement between the tests given using the same instrument, with scores of .75 to .90 indicating good reliability and scores above .90 indicating excellent reliability (Koo & Li, 2016; McGraw & Wong, 1996; Shrout & Fleiss, 1979).

Discriminant Validity. Discriminant validity indicates that a specific instrument, designed to measure a single construct, will not be as effective in measuring a similar but different construct (Rönkkö & Cho, 2022). Though the PSWQ is a measure of "worry" rather than of GAD, previous research tested the PSWQ's ability to discriminate between GAD (which is primarily a worry disorder [American Psychiatric Association, 2013]) and depression, since these constructs overlap but are not the same (Meyer et al., 1990; Nitschke et al., 2001; Pallesen et al., 2006). Following this reasoning, in this exploratory analysis, the PSWQ was tested at intake to see if it discriminated between worry and depression measured by the Centers for Epidemiological Studies Depression Scale (CES-D) (a 20-item self-report scale, with individuals reporting how often in the last week they experienced specific symptoms of depression on a scale from 0 "*not at all*" to 3 "*a lot*"). Similarly, worry and stress are overlapping constructs (Bergeria et al., 2021; W. E. Kelly, 2008), and the discriminant validity of the PSWQ was evaluated in its correlation with the Perceived Stress Scale, 10-item version (PSS-10) (a 10-item scale, with individuals grading how often they experienced specific emotions associated with stress on a scale from 0 "*never*" to 4 "*very often*").

While there has been discussion as to the best way to measure discriminant validity (Rönkkö & Cho, 2022), the most interpretable is by using a correlation coefficient. When two measurements instruments are highly correlated, they are likely measuring the same underlying concept which indicates convergent validity if both instruments are measuring the same construct (e.g., the PSWQ and the State Trait Anxiety Inventory) or that the instrument is not able to discriminate between two separate constructs (such as stress or depression and worry). A low correlation coefficient indicates that two instruments are measuring two separate constructs. When two constructs are different but linked (such as worry and stress), the measures must be correlated to some degree, yet not be so correlated as to indicate a single underlying construct. The exact benchmark depends on the theoretical expectation of the degree of overlap between the two concepts (Clark & Watson, 1995, 2019; Rönkkö & Cho, 2022). Given the theoretical overlap between these three constructs (i.e., stress, depression and worry) a moderate correlation ($<.7$) will indicate good discriminant validity. For context, in a review Cheung et al. (2024) mention that no cutoff is universally accepted but that correlations above .80 have been typically found to indicate poor discriminant validity.

A second method that was used to measure discriminant validity is the Heterotrait-Monotrait Ratio of Correlations (HTMT), which compares the correlations between items measuring different constructs (heterotrait) with the items measuring the same construct (monotrait). The HTMT is the ratio of the average heterotrait correlations to the average montrait correlations. With overlapping constructs it is necessary (as with the correlation coefficient) to understand the HTMT within the context of the expected convergence and divergence of the two constructs; again previous studies have assumed that HTMT below .85 would indicate good discriminant validity (Clark & Watson, 1995; Henseler et al., 2015; Kline, 2023; Voorhees et al., 2016).

Factor Structure. As discussed above, the PSWQ is assumed to be a single latent construct measuring worry, but since 11 items are positively worded and five are negatively worded, there is likely a response bias (T. A. Brown, 2003; Marsh, 1996). The model that we tested was based on responses at intake, assuming a single factor but with correlated residuals for the negatively worded questions, which was previously shown to be the best model in other populations (T. A. Brown, 2003). Additionally, a second two-factor model was tested for comparison.

Goodness of fit was evaluated using standard metrics delineated in T. A. Brown (2015): root mean square error of approximation (RMSEA) < 0.08, standardized root mean square residual (SRMR) < 0.05, comparative fit index (CFI) > 0.9, and Tucker-Lewis index (TLI) > 0.9. All analyses were done in the R programing language (R Core Team, 2023) using the "psych" (Revelle, 2023), and "lavaan" packages (Rosseel, 2012).

Sensitivity to Change. To evaluate the changes in the PSWQ scores over time, a linear mixed-effects model was employed. The model included week as a fixed effect and participant as a random effect to account for the repeated measures design and individual variability in baseline scores. Pairwise comparisons between weeks were conducted using Tukey's method to adjust for multiple comparisons, allowing us to determine which specific weeks showed statistically significant differences in worry levels. The effect size measured as Cohen's *d* was also used to describe the magnitude of the group differences.

In addition to the pairwise comparisons, the within-person sensitivity to change was evaluated using the Standardized Response Mean (SRM) to assess the magnitude of change in PSWQ scores over time. SRM values were calculated by dividing the mean change in PSWQ scores between consecutive weeks by the standard deviation of the change, providing insight into the clinical relevance of changes in worry levels during treatment. For the SRM and Cohen's *d* the benchmarks of 0.2 was considered small, 0.5 medium, and 0.8 or greater as large (Cohen, 2013).

Results

Sample

The largest part of the sample consisted of those whose primary substance was alcohol (48.6%), followed by those who primarily used opioids, 25.4% (12.6% used heroin/fentanyl and 12.8 used commercial opioids), then 17.8% who primarily used stimulants (7.1% cocaine and 10.7% methamphetamines, 1.3% commercial stimulants). Far fewer individuals reported that their primary substance was benzodiazepines (2.9%) or cannabis (3.8%). Most of the samples were collected from residential treatment (57.2%) or supervised withdrawal (35.5%) with intensive outpatient accounting for only 7.3%. Demographically, the sample was primarily male (67.4%), White (81.8%), not Hispanic/Latino (91.8%) and while there were statically significant ($p < .001$) differences between the groups based on primary substance, the effect sizes were either small or negligible following Cohen's guidelines (Acock, 2014; Cohen, 2013; Fey et al., 2023; Nieminen, 2022). See Table 1 for all demographic differences between groups.

Normative Values

The scores of the PSWQ ranged from 17 to 80 and the mean score for the whole sample was 54.85 ($SD = 13.84$). While the omnibus ANOVA test indicated significant differences between groups with different primary substance, the omnibus effect size was negligible ($\eta^2 < .001$). However, it was noticeable that those who identified benzodiazepine as their primary substance had comparatively elevated mean anxiety of 59.38 ($SD = 13.36$). Similarly, the median score was 55 for the whole sample and when examined by primary substance the medians were all between 52 and 57, except those who identified benzodiazepine as their primary substance whose median score was 61. We therefore evaluated the effect size between the benzodiazepine group and the rest of the sample, which demonstrated that there was a small effect (Cohen's $d = 0.3$).

Over time, the mean score of the PSWQ declined by more than 10 points, with the largest decline within the primary benzodiazepine group and smallest among the primary cannabis group. It should be noted that the decline was greatest during the first week of treatment but became less pronounced each successive week (see Table 2 and Figure 1). The patterns of these changes as well as the mean scores for all substances remained broadly consistent, even when the sample was restricted to those who remained in treatment for all five surveys. See Supplemental Material for results.

Reliability

Both Cronbach's α and McDonald's ω were high ($\alpha \geqslant .93$ and $\omega \geqslant .95$), for both the sample as a whole and when grouped by primary substance, indicating good internal consistency. Retesting over the course of the first 4 weeks of treatment as measured by the correlation coefficient from the intake to the end of the first week was acceptable ($r = .76$ for the whole sample ranging from 0.71 primary benzodiazepine to 0.79, primary cannabis) and improved weekly to $r = .87$ for the correlation between the third week and fourth week tests (see Table 3. for complete reliability statistics). The ICC(3,k) was also high, 0.93 for the whole sample and ranging from 0.92 (primary benzodiazepine and methamphetamine) to 0.94 (primary alcohol).

Discriminant Validity

Discriminant validity was assessed at intake, and the PSWQ was found to correlate with both the CES-D and the PSS-10 at $r = .60$ ranging between $r = .55$ and $r = .66$ for the CES-D and between $r = .55$ and $r = .65$ for the PSS-10 across substances. Similarly, the HTMT for the whole sample was 0.60 for a comparison to the CES-D (ranging from 0.52 to 0.66 for each of the substances) and 0.59 when compared to the PSS-10 (ranging from 0.56 to 0.65 for each of the individual substances)

Table 1. Demographic Characteristics of the Sample at Intake.

Characteristic mean (*SD*); *n* (%)	Alcohol, *N* = 36,505	Benzodiazepine, *N* = 2,211	Cocaine, *N* = 5,336	Heroin, *N* = 9,470	Marijuana, *N* = 2,870	Methamphetamine, *N* = 8,044	Opioids, *N* = 9,630	Stimulants, *N* = 975	*p*-Value	Effect size
Age	45 (12)	34 (13)	42 (13)	35 (9)	31 (12)	36 (10)	35 (11)	35 (10)	<.001[a]	$\eta^2 < .001$
Level of care									<.001[b]	$\omega^2 = .049$
Supervised withdrawal	13,981 (38%)	951 (43%)	1,155 (22%)	4,338 (46%)	494 (17%)	1,210 (15%)	4,316 (45%)	185 (19%)		
Intensive outpatient	2,842 (7.8%)	155 (7.0%)	483 (9.1%)	321 (3.4%)	435 (15%)	497 (6.2%)	615 (6.4%)	106 (11%)		
Residential	19,682 (54%)	1,105 (50%)	3,698 (69%)	4,811 (51%)	1,941 (68%)	6,337 (79%)	4,699 (49%)	684 (70%)		
Gender									<.001[b]	$\omega^2 = .004$
Female	11,869 (33%)	866 (39%)	1,524 (29%)	3,171 (33%)	692 (24%)	2,839 (35%)	2,944 (31%)	360 (37%)		
Male	24,576 (67%)	1,335 (60%)	3,803 (71%)	6,279 (66%)	2,153 (75%)	5,174 (64%)	6,661 (69%)	611 (63%)		
Other	60 (0.2%)	10 (0.5%)	9 (0.2%)	20 (0.2%)	25 (0.9%)	31 (0.4%)	25 (0.3%)	4 (0.4%)		
Ethnicity									<.001[b]	$\omega^2 = .002$
Hispanic/Latino	2,658 (7.3%)	202 (9.1%)	549 (10%)	743 (7.8%)	303 (11%)	650 (8.1%)	1,015 (11%)	52 (5.3%)		
Not Hispanic/Latino	33,847 (93%)	2,009 (91%)	4,787 (90%)	8,727 (92%)	2,567 (89%)	7,394 (92%)	8,615 (89%)	923 (95%)		
Race									<.001[c]	$\omega^2 = .051$
African American	2,773 (7.6%)	77 (3.5%)	1,665 (31%)	831 (8.8%)	459 (16%)	413 (5.1%)	712 (7.4%)	87 (8.9%)		
Asian	263 (0.7%)	27 (1.2%)	47 (0.9%)	41 (0.4%)	44 (1.5%)	41 (0.5%)	59 (0.6%)	14 (1.4%)		
Native American	427 (1.2%)	22 (1.0%)	69 (1.3%)	96 (1.0%)	59 (2.1%)	129 (1.6%)	112 (1.2%)	15 (1.5%)		
Native Hawaiian/Pacific Islander	92 (0.3%)	9 (0.4%)	8 (0.1%)	19 (0.2%)	15 (0.5%)	37 (0.5%)	22 (0.2%)	1 (0.1%)		
Other	2,091 (5.7%)	157 (7.1%)	450 (8.4%)	671 (7.1%)	250 (8.7%)	512 (6.4%)	822 (8.5%)	48 (4.9%)		
White	30,859 (85%)	1,919 (87%)	3,097 (58%)	7,812 (82%)	2,043 (71%)	6,912 (86%)	7,903 (82%)	810 (83%)		
Discharge against medical advice	3,017 (8.3%)	268 (12%)	666 (12%)	1,813 (19%)	367 (13%)	1,693 (21%)	1,391 (14%)	111 (11%)	<.001	$\omega^2 = .020$

[a]ANOVA.
[b]Pearson's Chi-squared test.
[c]Fisher's exact test for count data with simulated *p*-value (based on 2,000 replicates).

Table 2. Norms at Intake and During the First 4 Weeks of Treatment.

Norms	Primary substance									
	Whole sample	Alcohol	Benzodiazepine	Cannabis	Cocaine	Heroin	Methamphetamine	Opioids	Stimulants	*p*-Value effect size
Intake										
n	54,436	27,120	1,587	1,993	3,645	6,884	5,803	6,745	659	
Mean (SD)	54.9 (13.8)	55.0 (13.7)	59.4 (13.4)	53.3 (14.2)	52.9 (14.0)	54.3 (14.5)	56.0 (14.1)	54.3 (13.3)	56.3 (13.9)	$p < .001$ $\eta^2 < .001$
Minimum	17	17	17	17	17	17	17	17	18	
Median	55	55	61	53	52	54	57	54	57	
75th percentile	65	65	70	65	64	65	67	64	67	
90th percentile	74	73	76	72	72	74	75	73	74	
95th percentile	77	77	78	76	76	78	78	76	76	
99th percentile	80	80	80	80	80	80	80	80	80	
Maximum	80	80	80	80	80	80	80	80	80	
Means of PSWQ at the end of the first 4 weeks										
Week 1	51.1 (13.7)	51.3 (13.5)	55.1 (13.3)	50.1 (14.3)	49.9 (13.5)	50.1 (13.7)	51.8 (14.0)	50.8 (13.7)	52.5 (13.8)	$p < .001$ $\eta^2 < .001$
Week 2	47.9 (13.7)	48.0 (13.7)	51.8 (14.0)	47.1 (13.5)	47.3 (13.5)	47.2 (13.7)	48.1 (13.9)	47.6 (13.7)	48.5 (13.7)	$p < .001$ $\eta^2 < .001$
Week 3	45.6 (13.7)	45.8 (13.7)	48.9 (14.0)	45.1 (13.6)	45.0 (13.7)	44.8 (13.6)	45.3 (13.9)	45.2 (13.7)	46.6 (13.8)	$p < .001$ $\eta^2 < .001$
Week 4	44.2 (13.8)	44.5 (13.8)	46.8 (13.6)	44.1 (13.5)	43.8 (13.7)	43.4 (13.5)	43.7 (13.8)	43.9 (14.1)	45.1 (14.2)	$p < .001$ $\eta^2 < .001$

Note. PSWQ = Penn State Worry Questionnaire.

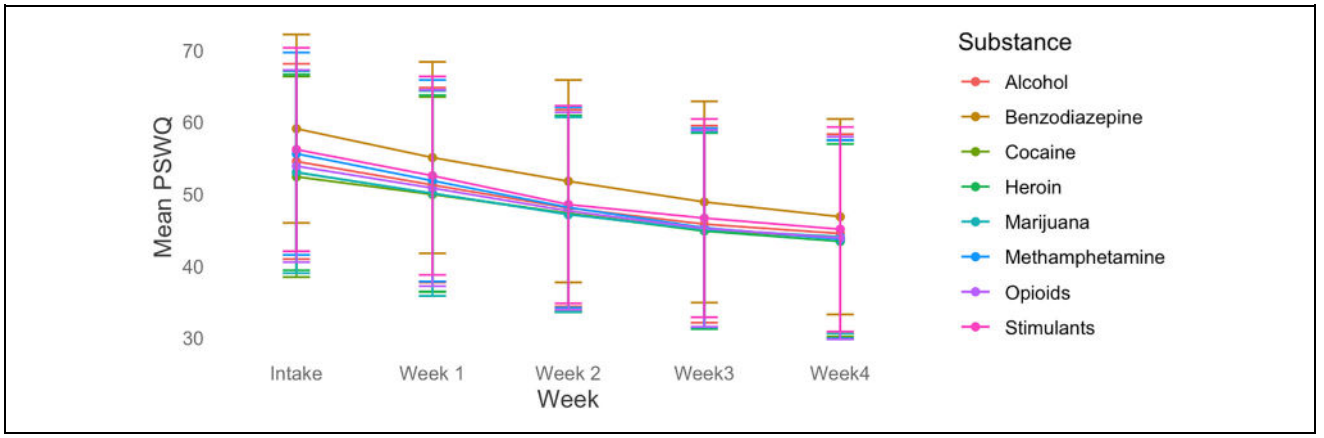

Figure 1. Change in Mean PSWQ During Treatment.

Factor Structure

A confirmatory factor analysis (CFA) model with a single latent factor of worry but with correlated residuals of the negatively worded questions was found to have good fit for both the RMSEA = 0.059, SRMR = 0.02, CFI = 0.970, and TLI = 0.962. The results were consistent when each primary substance group was analyzed alone (see Table 3). We also tested a two-factor model with similar results (RMSEA = 0.061, SRMSR = 0.03, CFI = 0.966, TLI = 0.960) (see Table 3 and Supplemental Material for results by primary substance). These findings support the one and two factor structure validity of the PSWQ in this sample.

Sensitivity to Change

A linear mixed-effects model revealed a significant effect of week on PSWQ scores ($p < .001$). Pairwise comparisons using Tukey's method showed significant reductions in worry levels between each consecutive week (all p-values $< .001$). The largest decrease in worry occurred between intake and week four, with a mean difference of 12.06 ($p < .001$). Cohen's d indicated a large effect size between intake and week four ($d = 0.78$), reflecting a substantial reduction in worry over time.

In addition, within-person the sensitivity to change was evaluated using the SRM, which confirmed a meaningful reduction in worry levels. The SRM values ranged from −0.254 between intake and week one to −0.780 between intake and week four, indicating moderate to large effects across the treatment period. These findings suggest that the PSWQ is sensitive to detecting clinically meaningful changes in worry levels over time, with the largest reductions occurring later in treatment (see Table 4).

Discussion

In contrast to previous research finding that mean scores on the PSWQ in a normative population of range between 35 and 50 (Brenes et al., 2022; Fortune et al., 2005; Meyer et al., 1990; Oliveira et al., 2023; Pallesen et al., 2006; Rodríguez-Biglieri & Vetere, 2011; van Rijsoort et al., 1999), the population entering SUD treatment (i.e., at intake) had a higher mean PSWQ score of 54.85 ($SD = 13.84$). One possible interpretation of this finding is that individuals entering SUD treatment represent a distinct population, who on average have elevated levels of worry. Our sample displayed higher baseline levels of worry than the population at large, and these levels declined over the course of SUD treatment, which may indicate that elevated worry at intake could possibly be a feature of substance use in which case the proposed cutoffs of 45 (Behar et al., 2003) or 50 (Wuthrich et al., 2014) to indicate clinically relevant worry may not be suitable for an SUD treatment sample. However, by the end of the second week of treatment all primary substance groups experienced worry levels below the 50 cutoff (except the primary benzodiazepine group which had a mean of 51.77) and by the end of the fourth week of treatment, all groups had a mean below the 45 point cut off (except the benzodiazepine group, $M = 46.83$, and the prescription stimulants group $M = 45.09$) implying the utility of the PSWQ in identifying clinically relevant worry after some time in SUD treatment.

This study's findings align with previous research and extend the reach of PSWQ by reaffirming its reliability, validity, and factor structure to the SUD treatment population. The study's internal consistency results, indicated by McDonalad's ω and Cronbach's α, are in agreement with prior investigations, which consistently report values of $\alpha > .85$ (T. A. Brown, 2003; Davey,

Table 3. Descriptive Statistics Reflecting Reliability, Validity, and Model Fit Indices of Latent Factor Structure.

Analyses	Primary substance								
	Whole sample	Alcohol	Benzodiazepine	Cannabis	Cocaine	Heroin	Methamphetamine	Opioids	Stimulants
Internal consistency									
Cronbach's α	.94	.94	.93	.94	.93	.94	.94	.93	.94
McDonald's ω	.95	.95	.95	.95	.95	.96	.95	.95	.95
Stability over time									
Correlation									
Intake, Week 1	0.76	0.77	0.71	0.79	0.74	0.73	0.74	0.74	0.77
Week 1, Week 2	0.81	0.83	0.77	0.79	0.79	0.78	0.78	0.78	0.79
Week 2, Week 3	0.84	0.86	0.80	0.83	0.8	0.81	0.82	0.82	0.84
Week 3, Week 4	0.87	0.89	0.85	0.83	0.85	0.85	0.83	0.85	0.83
ICC(3,k)[a]	0.93	0.94	0.92	0.93	0.93	0.92	0.92	0.93	0.93
Discriminant validity									
CES-D[b]									
Correlation	.60	.61	.55	.66	.59	.52	.60	.55	.59
HTMT[c]	0.60	0.62	0.55	0.66	0.59	0.52	0.60	0.55	0.59
PSS-10[d]									
Correlation	.60	.61	.55	.65	.59	.55	.59	.57	.58
HTMT[c]	0.59	0.61	0.56	0.65	0.59	0.55	0.59	0.57	0.58
Confirmatory factor analysis—one factor solution									
χ^2	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Comparative fit index	0.970	0.970	0.967	0.966	0.966	0.969	0.973	0.968	0.957
Tucker-Lewis index	0.962	0.962	0.958	0.956	0.957	0.960	0.966	0.960	0.945
Root mean square error of approximation	0.059	0.059	0.059	0.061	0.060	0.065	0.055	0.059	0.069
Standardized root mean square residual	0.020	0.020	0.025	0.024	0.021	0.020	0.019	0.021	0.029

[a]Inter class correlation (3,k).
[b]Centers for epidemiological studies depression scale.
[c]Heterotrait-Monotrait ration of correlations.
[d]Perceived stress scale.

Table 4. Pairwise Comparisons of PSWQ Scores Between Weeks.

Contrast	Mean difference	SE	z.ratio	*p*-Value	Cohen's *d*
Intake—week 1	3.88	0.049	79.124	<.0001	0.26
Intake—week 2	7.40	0.054	137.773	<.0001	0.53
Intake—week 3	10.12	0.057	177.194	<.0001	0.68
Intake—week 4	12.06	0.065	184.009	<.0001	0.78
Week 1—week 2	3.53	0.052	68.351	<.0001	0.26
Week 1—week 3	6.24	0.055	113.698	<.0001	0.47
Week 1—week 4	8.18	0.064	128.714	<.0001	0.60
Week 2—week 3	2.71	0.056	48.165	<.0001	0.20
Week 2—week 4	4.65	0.065	72.066	<.0001	0.34
Week 3—week 4	1.94	0.066	29.311	<.0001	0.14

Note. PSWQ = Penn State Worry Questionnaire; SE = standard error.

1993; Meyer et al., 1990; Pallesen et al., 2006; van Rijsoort et al., 1999). Similarly, the study demonstrates the PSWQ is stable during treatment, as evidenced by the correlation coefficients over different weeks, aligning with the high values reported in earlier research (Meyer et al., 1990; Molina & Borkovec, 1994; Pallesen et al., 2006; Stöber & Bittencourt, 1998). The exploration of discriminant validity further concurs with previous studies, revealing correlations with the CES-D and PSS-10, which corroborates previous findings that the PSWQ discriminates between GAD and Major Depressive Disorder when compared to measures like the Beck Depression Inventory and the Padua Inventory Revised (Davey, 1993; Meyer et al., 1990; Pallesen et al., 2006; van Rijsoort et al., 1999). Similarly, the factor structure in the population beginning SUD treatment was comparable to that of other populations, whereby a single factor model when the residuals of the negatively worded items were correlated, had satisfactory fit, as did the two-factor model. However, it remains an open question whether the PSWQ captures the same underlying construct in this population as in the general population.

Sensitivity to change analysis, as well as pairwise comparisons (which were significant) demonstrated a clinically meaningful difference between intake and week four, with both SRM and Cohen's *d* of 0.78, confirming that the PSWQ is sensitive to detecting meaningful clinical changes in worry over the first 4 weeks of treatment.

This study demonstrates that in SUD treatment initiates (regardless of primary substance) the PSWQ measures the same underlying construct of worry as within other populations, with similar reliability found in previous studies (Meyer et al., 1990; Pallesen et al., 2006), though on average, SUD treatment initiates have higher scores corresponding to more worry. These findings may be generalizable to other individuals entering SUD treatment since they are based on a large sample from several areas of the United States, though the generalizability is limited by the use of a convenience sample rather than one that was random and nationally representative, and location, socioeconomic and demographic factors were not available for the study. Another limitation was that this study did not attempt to assess the concurrent validity by comparing it to a second worry or anxiety scale. Nor did this study compare the PSWQ scores of those beginning SUD treatment who were diagnosed with GAD or other anxiety disorder and those without. This study did not address polysubstance use or co-occurring mental health conditions, which should be the focus of future work.

The identification of elevated mean PSWQ scores among individuals entering SUD treatment, which declines over time, suggests that while this population experiences heightened levels of worry compared to normative samples, it is not necessarily the symptom of an underlying anxiety disorder. Rather, elevated levels of worry at treatment commencement are likely related to, in part, SUD or the psychosocial stresses associated with entering SUD treatment. This study indicates that those initiating SUD treatment are distinct from the general population and suggests that measures of related constructs (e.g., psychosocial distress, resilience) must be assessed and normed for this population so that clinicians can properly assess their patients and create appropriate treatment plans and goals.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this

article: This study was supported by the National Institute on Drug Abuse T32 DA007209 (Strain, Weerts), and UH3 DA048734 (Huhn).

ORCID iD

Martin Hochheimer https://orcid.org/0000-0001-6644-4841

Supplemental Material

Supplemental material for this article is available online.

References

Acock, A. (2014). *A gentle introduction to stata. 4. izdaja*. Stata Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (*Vol. 5*). American Psychiatric Association.

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.

Behar, E., Alcaine, O., Zuellig, A. R., & Borkovec, T. (2003). Screening for generalized anxiety disorder using the Penn State Worry Questionnaire: A receiver operating characteristic analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *34*(1), 25–43.

Bergeria, C. L., Strickland, J. C., Huhn, A. S., Strain, E. C., & Dunn, K. E. (2021). A preliminary examination of the multiple dimensions of opioid craving. *Drug and Alcohol Dependence*, *219*, 108473. https://doi.org/10.1016/j.drugalcdep.2020.108473

Bluthenthal, R. N., Simpson, K., Ceasar, R. C., Zhao, J., Wenger, L., & Kral, A. H. (2020). Opioid withdrawal symptoms, frequency, and pain characteristics as correlates of health risk among people who inject drugs. *Drug and Alcohol Dependence*, 211, 107932. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2020.107932

Boschloo, L., Vogelzangs, N., van den Brink, W., Smit, J. H., Beekman, A. T., & Penninx, B. W. (2012). Predictors of the 2-year recurrence and persistence of alcohol dependence. *Addiction*, *107*(9), 1639–1640.

Brenes, G. A., Johnco, C., Mohlman, J., Wetherell, J. L., & Wuthrich, V. M. (2022). Evidence-based assessment of treatment outcomes for late-life generalized anxiety disorder using the Penn State Worry Questionnaire (PSWQ) and Penn State Worry Questionnaire—Abbreviated (PSWQ-A). *International Psychogeriatrics*, *34*(5), 489–501. https://doi.org/10.1017/S1041610221000351

Brown, S. A., & Schuckit, M. A. (1988). Changes in depression among abstinent alcoholics. *Journal of Studies on Alcohol*, *49*(5), 412–417.

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. https://doi.org/10.1016/S0005-7967(03)00059-7

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

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.

Chartoff, E. H., & Carlezon, W. A., Jr. (2014). Drug withdrawal conceptualized as a stressor. *Behavioural Pharmacology*, *25*(5 and 6), 337–339. https://journals.lww.com/behaviouralpharm/fulltext/2014/09000/drug_withdrawal_conceptualized_as_a_stressor.13.aspx

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*(2), 745–783. https://doi.org/10.1007/s10490-023-09871-y

Chorpita, B. F., Tracey, S. A., Brown, T. A., Collica, T. J., & Barlow, D. H. (1997). Assessment of worry in children and adolescents: An adaptation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *35*(6), 569–581. https://www.sciencedirect.com/science/article/pii/S0005796796001167?via%3Dihub

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319.

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, *31*(12), 1412.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.

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.

Davey, G. C. (1993). A comparison of three worry questionnaires. *Behaviour Research and Therapy*, *31*(1), 51–56.

Ediati, A., & Utari, A. (2019). Assessing worry in children: Psychometric evaluation of the Indonesian version of Penn State Worry Questionnaire (PSWQ). *Program Magister Psikologi Universitas Ahamd Dahlan*, *8*(1), 54–66.

Ellis, J. D., Rabinowitz, J. A., Wells, J., Liu, F., Finan, P. H., Stein, M. D., Antoine, D. G. II, Hobelmann, G. J., & Huhn, A. S. (2022). Latent trajectories of anxiety and depressive symptoms among adults in early treatment for nonmedical opioid use. *Journal of Affective Disorders*, *299*, 223–232.

Fatséas, M., Denis, C., Lavie, E., & Auriacombe, M. (2010). Relationship between anxiety disorders and opiate dependence—A systematic review of the literature: Implications for diagnosis and treatment. *Journal of Substance Abuse Treatment*, *38*(3), 220–230. https://doi.org/10.1016/j.jsat.2009.12.003

Fey, C. F., Hu, T., & Delios, A. (2023). The Measurement and communication of effect sizes in management research. *Management and Organization Review*, *19*(1), 176–197.

Fortune, D. G., Richards, H. L., Griffiths, C. E. M., & Main, C. J. (2005). Worry and pathological worry in patients with psoriasis: Cross sectional and longitudinal analyses of the Penn State Worry Questionnaire (PSWQ) in four samples of

patients. *Journal of Clinical Psychology in Medical Settings*, *12*(2), 143–152. https://doi.org/10.1007/s10880-005-3274-9

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. https://doi.org/10.1016/S0005-7967(00)00113-3

Garey, L., Olofsson, H., Garza, T., Rogers, A. H., Kauffman, B. Y., & Zvolensky, M. J. (2020). Directional effects of anxiety and depressive disorders with substance use: A review of recent prospective research. *Current Addiction Reports*, *7*, 344–355.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*, 115–135.

Jesse, S., Bråthen, G., Ferrara, M., Keindl, M., Ben-Menachem, E., Tanasescu, R., Brodtkorb, E., Hillbom, M., Leone, M. A., & Ludolph, A. C. (2017). Alcohol withdrawal syndrome: mechanisms, manifestations, and management. *Acta Neurologica Scandinavica*, *135*(1), 4–16. https://doi.org/https://doi.org/10.1111/ane.12671

Johnco, C., Wuthrich, V. M., Brenes, G. A., Wetherell, J. L., & Mohlman, J. (2022). Evidence-based assessment of treatment outcomes for late-life generalized anxiety disorder using the Penn State Worry Questionnaire (PSWQ) and Penn State Worry Questionnaire–Abbreviated (PSWQ-A). *International Psychogeriatrics*, *34*(5), 489–501.

Kelly, J. F., & Hoeppner, B. B. (2013). Does alcoholics anonymous work differently for men and women? A moderated multiple-mediation analysis in a large clinical sample. *Drug Alcohol Depend*, *130*(1–3), 186–193. https://doi.org/10.1016/j.drugalcdep.2012.11.005

Kelly, W. E. (2008). Anxiety and stress as contributory factors in pathological and nonpathological worry. *Psychology Journal*, *5*(3), 147–157.

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. https://doi.org/10.1037/a0037251

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Koob, G. F. (2015). The dark side of emotion: The addiction perspective. *European Journal of Pharmacology*, *753*, 73–87.

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: A neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760–773.

Lai, H. M. X., Cleary, M., Sitharthan, T., & Hunt, G. E. (2015). Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990–2014: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, *154*, 1–13. https://doi.org/10.1016/j.drugalcdep.2015.05.031

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, *70*(4), 810.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30.

McHugh, R. K. (2015). Treatment of co-occurring anxiety disorders and substance use disorders. *Harvard Review of Psychiatry*, *23*(2), 99–111.

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. https://doi.org/10.1016/0005-7967(90)90135-6

Molina, S., & Borkovec, T. D. (1994). The Penn State Worry Questionnaire: Psychometric properties and associated characteristics. In G. C. L. Davey & F. Tallis (Eds.) *Worrying: Perspectives on theory, assessment and treatment*. (pp. 265–283). John Wiley & Sons.

Motooka, H., Tanaka-Matsumi, J., & Hayashi, K. (2009). The reliability and validity of a Japanese version of The Penn State Worry Questionnaire (PSWQ): A self-report inventory of "worry". *Japanese Journal of Counseling Science*, *42*(3), 247–255.

Nieminen, P. (2022). Application of standardized regression coefficient in meta-analysis. *BioMedInformatics*, *2*(3), 434–458. https://www.mdpi.com/2673-7426/2/3/28

Nitschke, J. B., Heller, W., Imig, J. C., McDonald, R. P., & Miller, G. A. (2001). Distinguishing dimensions of anxiety and depression. *Cognitive Therapy and Research*, *25*, 1–22.

Oliveira, J. T., Faustino, D., Freitas, F., Gonçalves, M. M., Ribeiro, E., Gonçalves, S., & Machado, P. P. P. (2023). Penn State Worry Questionnaire in emotional disorders: Validation and normative data for Portuguese population. *British Journal of Guidance & Counselling*, *51*(2), 251–261. https://doi.org/10.1080/03069885.2021.1897969

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. https://doi.org/10.1111/j.1467-9450.2006.00518.x

Puccinelli, C., Cameron, D. H., Ouellette, M. J., McCabe, R. E., & Rowa, K. (2023). Psychometric properties of the Penn State Worry Questionnaire-Past Week (PSWQ-PW) in an anxiety and related disorders sample. *Journal of Psychopathology and Behavioral Assessment*, *45*(2), 549–557.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rabinowitz, J. A., Ellis, J. D., Wells, J., Strickland, J. C., Maher, B. S., Hobelmann, J. G., & Huhn, A. (2023). Correlates and consequences of anxiety and depressive symptom trajectories during early treatment for alcohol use. *Alcohol*, *108*, 44–54.

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. 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. https://doi.org/10.5209/rev_SJOP.2011.v14.n1.41

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, *25*(1), 6–14.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/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 non-clinical and clinical samples. *Anales de Psicología/Annals of Psychology*, *34*(3), 451–457.

Schuckit, M. A. (2014). Recognition and management of withdrawal delirium (delirium tremens). *New England Journal of Medicine*, *371*(22), 2109–2113. https://doi.org/10.1056/NEJMra1407298

Schuckit, M. A. (2006). Comorbidity between substance use disorders and psychiatric conditions. *Addiction*, *101*, 76–88.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420.

Stöber, J., & Bittencourt, J. (1998). Weekly assessment of worry: An adaptation of the Penn State Worry Questionnaire for monitoring changes during treatment. *Behaviour Research and Therapy*, *36*(6), 645–656. https://doi.org/10.1016/S0005-7967(98)00031-X

Stoeber, J. (1995). Besorgnis: Ein Vergleich dreier Inventare zur Erfassung allgemeiner Besorgnis [Worry: A comparison of three inventories to assess general worries]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *16*, 50–63.

Tallis, F., Eysenck, M., & Mathews, A. (1992). A questionnaire for the measurement of nonpathological worry. *Personality and Individual Differences*, *13*(2), 161–168.

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. https://doi.org/10.1002/(SICI)1099-0879(199910)6:4<297::AID-CPP206>3.0.CO;2-E

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, *44*(1), 119–134. https://doi.org/10.1007/s11747-015-0455-4

Vorspan, F., Mehtelli, W., Dupuy, G., Bloch, V., & Lépine, J.-P. (2015). Anxiety and substance use disorders: Co-occurrence and clinical issues. *Current Psychiatry Reports*, *17*, 1–7.

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. https://doi.org/10.1016/j.janxdis.2014.07.005

Xie, S.-S., Xiao, H.-W., & Lin, R.-M. (2023). Abbreviated version of Penn State Worry Questionnaire for Chinese adolescents: Age, gender and longitudinal invariance. *Frontiers in Psychiatry*, *14*, 1086592. https://doi.org/10.3389/fpsyt.2023.1086592