J Psychopathol Behav Assess
DOI 10.1007/s10862-016-9571-9

Psychometric Properties of the Generalized Anxiety Disorder Scale-7 (GAD-7) in Outpatients with Anxiety and Mood Disorders

Lauren A. Rutter[1] · **Timothy A. Brown**[1]

Abstract This study examined the psychometric properties of the Generalized Anxiety Disorder Scale-7 (GAD-7) in a sample of 536 outpatients presenting at a specialty clinic for anxiety and mood disorders. A confirmatory factor analysis (CFA) was used to test the unidimensionality of the GAD-7. This model did not fit the data well. The CFA solution was respecified correlating residuals among items assessing somatic symptoms. This respecified model fit the data well. A series of multiple-groups CFAs determined that the measurement properties of the GAD-7 were invariant between sexes. Scale reliability estimates of the GAD-7 were favorable for the full sample, and for males and females. Sensitivity and specificity could not be balanced at any cut-point. Findings attest to the value of this instrument as a dimensional indicator of GAD severity rather than a screening tool for the presence or absence of the disorder in outpatients with anxiety and mood disorders.

Keywords GAD-7 · Anxiety · Validity · Psychometrics · Clinical sample

Generalized anxiety disorder (GAD) is classified in the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*) as chronic (lasting at least 6 months) excessive anxiety and worry about a number of events or activities that is difficult to control. GAD is one of the most common anxiety disorders, with lifetime prevalence rates estimated at 5.7 % (National Comorbidity Survey-Replication, 2011). GAD is associated with a variety of somatic complaints that can lead to costly medical testing and treatment in primary care settings (Roy-Bryne 1996). GAD has been linked to indirect costs at the workplace through lost productivity (Greenberg et al. 1999; Hoffman et al. 2006), and it may be a risk factor for the development of comorbid conditions (Roy-Bryne and Katon 1997) including other anxiety disorders (Grant et al. 2009), and alcohol-related disorders. Severe GAD pathology has been related to disability in areas of self-care, interpersonal functioning, and health care utilization (Ruiz et al. 2011). Increasing early recognition and providing appropriate treatment referrals could have wide-ranging benefits including reducing individual distress, disability status, overall healthcare usage, and the associated cost of GAD to society (Kertz et al. 2013).

The Generalized Anxiety Disorder Scale-7 (GAD-7) is a 7-item, self-rated scale developed by Spitzer et al. (2006) as a screening tool and severity indicator for GAD. It is easily scored and initially was created to increase recognition of GAD in primary care settings. The original validation of the GAD-7 in a large primary care sample revealed that the measure has good reliability, and good criterion, factorial, and procedural validity (based on a comparison of scores derived from the self-report scales with those derived from the clinician-administered versions of the same scales) (Spitzer et al. 2006). A cutoff score of 10 was identified as the optimal point for sensitivity (89 %) and specificity (82 %). The psychometric properties of the GAD-7 have also been evaluated in other primary care samples (Kroenke et al. 2007), a population-based sample (Löwe et al. 2008), psychiatric samples (Beard and Björgvinsson 2014; Kertz et al. 2013), Hispanic Americans (Mills et al. 2014), in addictions treatment (Delgadillo et al. 2012), and in different languages including Portuguese (Sousa et al. 2015) and in a Dutch web-based sample (Donker et al. 2011). While the literature has established strong psychometric support for the GAD-7 in a variety of contexts, more recently, Kertz et al. (2013) found that the measure did not perform well as a screener for GAD in

✉ Lauren A. Rutter
lrutter@bu.edu

[1] Center for Anxiety and Related Disorders, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215-2013, USA

Published online: 10 September 2016

patients presenting for treatment at a partial hospital program with a wide range of symptoms and high diagnostic comorbidity. These authors reported that while the GAD-7 demonstrated good sensitivity (.83), specificity was poor (.46). Otherwise, they concluded that the GAD-7 appears to be a valid measure of generalized anxiety symptoms in an acute psychiatric sample, on the basis of good internal consistency, convergent validity, and sensitivity to change.

Surprisingly, the GAD-7 has not been examined in outpatient samples of patients with anxiety and mood disorders. This omission is notable, as the GAD-7 is being increasingly used in anxiety disorders research and clinical practice (Dear et al. 2011). Anxiety disorders are commonly treated in outpatient settings, and GAD is one of the most common anxiety disorders. Therefore, determining the psychometric properties of a brief measure that was designed to identify probable cases of GAD and assess symptom severity seems important, given our specialized sample. The GAD-7 has been validated in a variety of samples, but its psychometric properties have yet to be studied in a diagnostically diverse sample of patients with emotional disorders. The primary aim of this study was to evaluate the psychometric properties of the GAD-7 in a large sample of patients with a variety of anxiety and mood disorders. This is the first study of the GAD-7 in an outpatient sample specialty clinic for anxiety and mood disorders, where GAD is prevalent both as a principal and an additional diagnosis. While not all patients seeking treatment at the clinic were diagnosed with GAD, most of the sample presented with worries and anxiety as a part of their pathology. By focusing on this type of sample, this study also expands on previous work (e.g., Kertz et al. 2013) by examining the discriminant validity of the GAD-7 with closely neighboring conditions including obsessive-compulsive disorder (OCD) and depression.

Based on prior research (e.g., Löwe et al. 2008), we expected that the latent structure of the GAD-7 would be unidimensional. We evaluated the measurement invariance, scale reliability, and sensitivity and specificity of the GAD-7. We expected that the measure would be invariant between sexes, the scale reliability would be favorable, and the measure would be sensitive to detecting GAD. Based on previous work (e.g., Kertz et al. 2013), we did not expect the GAD-7 would show good specificity in our sample. Finally, we predicted that the GAD-7 would be more strongly related to measures of anxiety than to measures of depression and OCD.

Method

Participants

Participants were 536 patients who presented for assessment and treatment at the Center for Anxiety and Related Disorders (CARD) at Boston University. Patients completed the GAD-7 along with a battery of self-report questionnaires and a semi-structured interview as a part of their intake assessment. Women constituted the larger proportion of the sample (62.7 % female). The average age was 31.49 (*SD* = 12.05, range = 18–80). The sample was predominantly Caucasian (82.3 %; African American, 8.2 %, Asian, 8.8 %, Pacific Islander, 0.4 %, American Indian/Alaskan, 0.4 %). Diagnoses were established using the Anxiety Disorders Interview Schedule for *DSM-IV:* Lifetime version (ADIS-IV-L; Di Nardo et al. 1994), a semi-structured interview designed to establish a diagnosis of *DSM-IV* anxiety, mood, somatoform, and substance use disorders, and to screen for other disorders (e.g., psychotic disorders). The sample breakdown of current clinical disorders (collapsing across principal and additional diagnoses) was: panic disorder with or without agoraphobia (22.6 %), social phobia (47.6 %), specific phobia (17.7 %), GAD (30.0 %), GAD ignoring *DSM-IV* diagnostic hierarchy rules with mood disorders (38.2 %), OCD (14.7 %), posttraumatic stress disorder (3.5 %), major depressive disorder (19.6 %), dysthymia (5.4 %), agoraphobia without a history of panic (1.5 %), anxiety disorder not otherwise specified (11.6 %), and depressive disorder not otherwise specified (3.2 %).

Measures

ADIS-IV-L (Di Nardo et al. 1994)

The ADIS-IV-L was administered by trained Ph.D.-level psychologists and advanced doctoral students in clinical psychology who underwent extensive training to meet strict certification criteria (see Brown et al. 2001, for details). Certified interviewers complete training by: (1) observing interviews, (2) conducting collaborative interviews, and (3) conducting interviews under observation of a senior interviewer. After each interview, trainee and senior interviewer discussed current and lifetime diagnoses. The criteria for ADIS-IV-L certification requires that trainee's diagnoses match the senior interviewers' diagnoses within three of five consecutive interviews. The ADIS-IV-L has been shown to have good or excellent diagnostic reliability for most anxiety disorders (ks = .67–.86; Brown et al. 2001).

Generalized Anxiety Disorder Questionnaire (GAD-7; Spitzer et al. 2006)

The GAD-7 is a 7-item self-report scale developed to assess the defining symptoms of GAD. Items are rated on a 4-point Likert-type scale (0 = *not at all* to 3 = *nearly every day*). GAD-7 items describe some of the most salient diagnostic features of GAD (i.e., *feeling nervous, anxious, or on edge* and *worrying too much about different things*). Scores range from 0 to 21 with higher scores indicating more severe GAD symptoms.

Research has suggested that the GAD-7 is a valid screening tool for GAD in a primary care setting and for assessing its severity in clinical practice and research (Spitzer et al. 2006). The average GAD-7 score was 11.60 (*SD* = 5.44) in our sample.

Penn State Worry Questionnaire (PSWQ; Meyer et al. 1990)

The PSWQ consists of 16 items rated on a 5-point Likert-type scale (1 = *not at all typical of me* to 5 = *very typical of me*). The PSWQ assesses the extent to which worry is pervasive, excessive, and uncontrollable (e.g., *I am always worrying about something*). The PSWQ has excellent internal consistency ($\alpha = .91$; Meyer et al. 1990) and good convergent and discriminant validity for GAD compared to other anxiety disorders and community controls (e.g., Brown, Antony, & Barlow, 1992). The average PSWQ score was 64.35 (*SD* = 11.78) in our sample (Cronbach's α in this sample = .91).

Depression Anxiety Stress Scales (DASS; Lovibond and Lovibond 1995)

The DASS consists of 21 items rated on a 4-point Likert-type scale (0 = *did not apply to me at all* to 3 = *applied to me very much, or most of the time*). The DASS assesses levels of depression (DASS-D), general anxiety (DASS-A), and general tension/negative affect symptoms (DASS-S) over the past 2 weeks. The three-factor structure of the DASS has been substantiated by both exploratory and confirmatory factor analyses in a variety of samples (e.g., Brown et al., 1997; Lovibond and Lovibond 1995). Descriptive statistics for the subscales in our sample are as follows: DASS-D *M* = 7.50, *SD* = 5.53, DASS-A *M* = 6.82, *SD* = 4.66, DASS-S *M* = 9.85, *SD* = 4.92. Internal consistencies of the DASS subscales in the present sample were .91, .80. , and .85 for DASS-D, DASS-A, and DASS-S, respectively.

Obsessive-Compulsive Inventory—Revised (OCI-R; Foa et al. 2002)

The OCI-R consists of 18 items rated on a 4-point Likert-type scale (0 = *not at all* to 3 *extremely*). Based on factor analytic evidence, the OCI-R is defined by six subscales (obsessing, checking, neutralizing, hoarding, ordering, and washing) and a total score. The OCI-R total score and its subscales have been used to effectively differentiate people with and without clinical levels of OCD. The instrument has also performed well in discriminating OCD from other anxiety disorders (Abramowitz and Deacon 2006). Total OCI-R scores evidence excellent internal consistency (Cronbach's $\alpha = .90$; Foa et al. 2002). The average OCI-R score was 17.60 (*SD* = 12.72) in our sample (Cronbach's α in this sample = .90).

Data Analysis

A latent variable software program using maximum likelihood fitting functions (Mplus 7.11, Muthén and Muthén 2008–2014) was used to analyze data. Goodness of fit for confirmatory factor analysis (CFA) models was evaluated using the root mean square error of approximation (RMSEA) and its 90 % confidence interval, the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the standardized root mean square residual (SRMR). Acceptable model fit was defined as follows: RMSEA (close to or below .08, 90 % CI < .08), CFI (close to or above .95), TLI (close to or above 0.95), and SRMR (close to or below .08; Hu and Bentler 1999). Multiple indices were used because they provide different information for evaluating model fit (i.e., absolute fit, fit adjusting for model parsimony, fit relative to a null model). When used together, these indices provide a more conservative and reliable evaluation of the model fit (cf. Brown 2015). In the instances of nested models, comparative fit was evaluated using χ^2 difference tests (χ^2_{diff}) and by the interpretability of the solutions. Additionally, SPSS 20.0 was used to calculate receiver operating characteristics curves to determine the classification accuracy of the GAD-7 in predicting a current GAD diagnosis.

Results

Factor Structure

A one-factor CFA was conducted to evaluate whether the latent structure of the GAD-7 was unidimensional. This model did not fit the data well, χ^2 (14) = 122.39, $p < .001$, RMSEA = 0.12 (90 % CI = 0.10, 0.14), SRMR = .05, CFI = .94, TLI = 0.91. Evaluation of localized areas of strain (i.e., large modification indices) showed evidence of correlated residuals for items 4, 5, and 6, each of which assess the somatic symptoms of GAD (i.e., trouble relaxing, feeling restless/unable to sit still, being easily annoyed or irritable). The CFA solution was respecified by freely estimating the error covariances of these items. The revised model fit the data well, χ^2 (11) = 35.92, $p < .001$, RMSEA = 0.07 (90 % CI = 0.04, 0.09), SRMR = .02, CFI = .99, TLI = 0.97. There were no salient areas of strain in the solution, as indicated by small modification indices and standardized residuals. Thus, in subsequent models, the correlated errors between items 4, 5, and 6 were freely estimated. The GAD-7 items had factor loadings ranging from .48 to .90, as shown in Table 1 (all *p*s < .001). The correlated residuals of items 4, 5, and 6 ranged from .11 to .32 (all *p*s < .05).

Measurement Invariance Between Sexes

A series of multiple-groups CFAs was conducted to determine if the measurement properties (i.e., latent structure, factor

Table 1 Latent structure of the Generalized Anxiety Disorder Scale-7: Factor loadings from confirmatory factor analyses using the full sample ($N = 536$) and males ($n = 200$) and females ($n = 336$)

GAD-7 Item	Full sample	Males	Females
1. Feeling nervous, anxious or on edge	.79	.78	.80
2. Not being able to stop or control worrying	.90	.88	.92
3. Worrying too much about different things	.85	.85	.84
4. Trouble relaxing	.73	.77	.70
5. Being so restless that it is hard to sit still	.52	.54	.50
6. Becoming easily annoyed or irritable	.48	.48	.48
7. Feeling afraid as if something awful might happen	.60	.54	.61

GAD-7 Generalized Anxiety Disorder Scale – 7; all factor loadings, $p < .001$

loadings, indicator intercepts) of the GAD-7 were invariant between female and male participants (*n*s = 336 and 200, respectively). For women, the respecified CFA fit the data well, χ^2 (11) = 20.96, $p < .05$, RMSEA = 0.05 (90 % CI = 0.01, 0.09), SRMR = .02, CFI = .99, TLI = 0.98. For men, the model also fit the data well, χ^2 (11) = 23.02, $p < .05$, RMSEA = 0.07 (90 % CI = 0.03, 0.12), SRMR = .03, CFI = .98, TLI = 0.97. The equal form fit the data well indicating that the revised one-factor measurement model was acceptable for both sexes, χ^2 (22) = 44.03, $p < .01$, RMSEA = 0.06 (90 % CI = 0.03, 0.09), SRMR = 0.03, CFI = .99, TLI = 0.98. As seen in Table 1, for females, factor loadings ranged from .48 to .92. For males, factor loadings ranged from .48 to .88. Given the support for equal form, the next analysis addressed metric invariance by holding the factor loadings to equality in the male and female solutions. This restriction did not significantly degrade the fit of the model, χ^2_{diff} (6) = 3.46, *ns*; χ^2 (28) = 47.49, $p < .05$, RMSEA = 0.05 (90 % CI = 0.02, 0.08), SRMR = 0.03, CFI = .99, TLI = 0.98, indicating that factor loadings were equivalent in males and females. The next analysis examined the scalar invariance of the GAD-7 by placing equality constraints on the intercepts of the 7 items. These constraints did not result in degradation of model fit, χ^2_{diff} (6) = 9.02, *ns;* χ^2 (34) = 56.51, $p < .01$, RMSEA = 0.05 (90 % CI = 0.03, 0.07), SRMR = 0.03, CFI = .99, TLI = 0.99. Collectively, these results indicated that the measurement properties of the GAD-7 were equivalent between sexes (i.e., a given observed score on the GAD-7 reflects the same degree of GAD symptom severity in males and females).

Scale Reliability

Scale reliability was computed using the unstandardized parameter estimates from the revised one-factor model (cf. Raykov 2001). This method accounts for the limitations of Cronbach's alpha, a misestimator of reliability except for the rare situation when all elements of a multiple-item measure are tau equivalent and free of nonrandom measurement error (Raykov 2001, 2004). The scale reliability was favorable for the full sample ($\rho = .85$), and for males and females (both ρs = .85).

Concurrent Validity

To evaluate the convergent and discriminant validity of the GAD-7, the DASS, PSWQ, and OCI-R were brought into the revised CFA model as covariates of the GAD-7 factor. We predicted that the GAD-7 would evidence stronger correlations with the anxiety measures (i.e., DASS-S, DASS-A, and PSWQ) than with indicators of depression (DASS-D) and OCD (OCI-R). Measurement error of these single indicators was adjusted for in the analyses by imposing constraints on error variances using previously published reliability estimates. The differential magnitude of the correlations of the GAD-7 factor with DASS-D, DASS-A, DASS-D, PSWQ, and OCI-R was evaluated using Steiger's *z* tests. As shown in Table 2, the magnitude of the convergent validity *r*s was moderate (*r*s = .52–.68) whereas the discriminant validity *r*s were somewhat smaller (*r*s = .42 and .47). We found that the GAD-7 was more strongly correlated with the DASS-S than the OCI-R ($z = 7.15$, $p < .001$). Similarly, the GAD-7 was more strongly correlated with the PSWQ than the OCI-R ($z = 6.58$, $p < .001$). The GAD-7 was more strongly correlated with the DASS-S than the DASS-D ($z = 6.98$, $p < .001$). Finally, the GAD-7 was more strongly correlated with the PSWQ than the DASS-D ($z = 5.12$, $p < .001$). In summary, consistent with prediction, the GAD-7 factor was more strongly related to measures of anxiety and general negative affect/distress than measures of depression and OCD.

Sensitivity and Specificity Analyses

Cut scores on the GAD-7 were examined to determine whether a score was able to predict those who currently

Table 2 Concurrent and discriminant validity of GAD-7 with measures of anxiety and depression

	PSWQ	DASS-S	DASS-A	DASS-D	OCI-R
GAD-7	$.66_a$	$.68_a$	$.52_b$	$.48_{b,\ c}$	$.42_c$

GAD-7 Generalized Anxiety Disorder Scale-7; *OCI-R* Obsessive Compulsive Inventory-Revised; *PSWQ* Penn State Worry Questionnaire; *DASS-D* Depression Anxiety and Stress Scales – Depression Scale; *DASS-A* Depression Anxiety and Stress Scales – Anxiety Scale; *DASS-S* Depression Anxiety and Stress Scales – Stress Scale. Correlations with different subscripts are significantly different in magnitude ($p < .001$) as indicated by Stieger's *z* tests

met diagnostic criteria for GAD (collapsing principal and additional diagnoses and ignoring *DSM* hierarchy rules for mood disorders) from those who did not. A receiver operating characteristics curve with the GAD-7 total score as a continuous variable and diagnostic status as the categorical outcome variable, and corresponding sensitivity and specificity values were generated. Estimates of positive predictive value and negative predictive value were calculated. In previous analyses of this issue using a primary care sample, Spitzer et al. (2006) found that a GAD-7 cutoff score of 10 maximized sensitivity and specificity (89 and 82 %, respectively). Our results showed that a cut score of 10 was associated with good sensitivity (79.5 %), but poor specificity (44.7 %). Given the poor performance of the 10 cutoff score, the classification accuracy of alternative cutoff scores was examined. The areas under the curve ranged from .65 to .74. As shown in Table 3, no cut scores demonstrated adequately balanced sensitivity and specificity. Our findings were similar to those of Kertz et al. (2013), as this investigation also showed good sensitivity (.83), but poor specificity (.46).

Table 3 Operating characteristics of the GAD-7 at different cutoffs

GAD-7 score	Sensitivity %	Specificity %	PPV%	NPV%
0	100	0		
1	100	1.6	38.7	100
2	100	4.0	39.3	100
3	100	7.8	40.2	100
4	98.5	12.1	41.0	92.9
5	97.5	17.4	42.3	91.8
6	95.0	23.3	43.5	88.2
7	90.5	30.1	44.6	83.6
8	86.5	34.8	45.2	80.6
9	83.5	40.1	46.4	79.6
10	79.5	44.7	47.2	77.8
11	76.0	51.6	49.4	77.6
12	68.0	60.2	51.5	75.2
13	64.5	65.5	53.8	74.8
14	58.0	71.7	56.0	73.3
15	49.0	74.8	54.7	70.3
16	44.5	78.9	56.7	69.6
17	35.5	86.6	62.3	68.4
18	28.5	89.8	63.3	66.9
19	19.5	93.2	63.9	65.1
20	10.9	97.2	69.0	63.5
21	6.0	97.8	63.2	62.6

GAD-7 Generalized Anxiety Disorder Scale – 7; *PPV* Positive Predictive Value; *NPV* Negative Predictive Value

Discussion

This study is the first to evaluate the psychometric properties of the GAD-7 in an outpatient sample of patients presenting with anxiety and mood disorders. While the GAD-7 has demonstrated strong psychometric properties in primary care settings (Spitzer et al. 2006) and in a population-based sample (Löwe et al. 2008), the results of a more recent validation study in an acute psychiatric sample (Kertz et al. 2013) led us to reexamine the GAD-7 in a clinical outpatient sample of patients with emotional disorders. We found that while the psychometric properties of the GAD-7 were strong, the measure may better serve as a dimensional indicator of GAD severity than a screening tool for the presence or absence of GAD in clinical samples, consistent with the findings of Kertz et al. (2013) and Beard and Björgvinsson (2014).

Our findings are consistent with Kertz et al. (2013), who found that the one higher-order GAD factor did not fit the data well, but a revised model covarying the error terms on items 4, 5, and 6 improved the fit. Our findings expand upon those of Beard and Björgvinsson (2014) who suggested a two-factor structure for the GAD-7. A key difference between the current analyses and those conducted in Beard and Björgvinsson (2014) is that the latter relied on exploratory factor analysis (EFA) to explore the latent structure of the GAD-7. Statistical identification restrictions in traditional EFA prevent the estimation of indicator error covariances (Brown 2015). Thus, what would otherwise be represented as method effects in CFA often emerge as additional, yet substantively trivial, factors in EFA (see Brown 2015, and Brown 2003, for an applied example). Our CFA results indicated that each of the GAD-7's items had salient loadings on a general factor, but error covariances were necessary to account for the additional covariance among items 4, 5, and 6, the items that represent the associated symptoms of GAD. Given the evidence for a single, general factor, an implication of our findings is that all items can be summed to form a composite score as an index of GAD severity in clinics that specialize in treating anxiety and mood disorders.

Our hypothesis that the GAD-7 would show good convergent and discriminant validity was supported. This finding from our clinical sample is consistent with the literature, which has also shown that the GAD-7 has good convergent and discriminant validity with other measures of anxiety and depression, respectively (see Kertz et al. 2013; Spitzer et al. 2006). Of note, some psychometric studies on the GAD-7 reported only convergent validity (Donker et al. 2011; Mills et al. 2014) and therefore, this study expands on previous work by providing new analyses of discriminant validity, particularly with OCD, given the high comorbidities and phenotypic overlap between GAD and OCD (see Abramowitz and Foa 1998; Brown et al. 2001).

Our results examining the sensitivity and specificity of the GAD-7 showed that a cut score of 10 was associated with good sensitivity (79.5 %), but poor specificity (44.7 %). Due to the shared features of anxiety disorders (see Brown and Barlow 2009), cutoffs are less likely to apply to clinical samples that are characterized by high general distress. Given our findings, it is recommended that the GAD-7 not be used to screen for GAD in a clinical sample of outpatients with depression and anxiety. In our sample, the GAD-7 did not provide sufficient specific information to indicate the presence of a GAD diagnosis, and is likely capturing the high levels of negative affect and severe distress that are present in this clinical, treatment-seeking sample.[1]

Despite our strengths in sampling and methodology, this study is not without limitations. Our sample was primarily Caucasian (82.3 %), limiting generalizability to other racial groups. Future studies should explore the GAD-7 in a more diverse outpatient sample. Additionally, the present results may not generalize to other types of clinical samples (patients with principal substance abuse disorders, personality disorders, psychotic disorders).

Notwithstanding the limitations in generalizability, the collective results of our study show that the GAD-7 has strong psychometric properties in terms of reliability and validity, and a unidimensional latent structure that is invariant between male and female patients. Future studies should investigate the convergent and discriminant validity of the GAD-7 with respect to other criteria (e.g., behavioral, biological, information-processing) that are relevant to the psychopathology of GAD.

Compliance with Ethical Standards

Funding This study was funded by the National Institute of Mental Health (Grant MH039096) to the second author. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

Conflict of Interest Lauren A. Rutter declares that she has no conflict of interest. Timothy A. Brown declares that he has no conflict of interest.

[1] To address this issue, we tested the relationship between the GAD-7 and high negative affect/neuroticism using the Neuroticism scale from the NEO Five Factor Inventory (NFFI-N). Although the NFFI-N was moderately correlated with the GAD-7 factor ($r = .56$), inferential tests of the differential magnitude of relationships indicated that the GAD-7 factor was more strongly correlated with the DASS-S ($r = .68$, $z = 4.50$, $p < .001$) and the PSWQ ($r = .66$, $z = 3.93$, $p < .001$) than the NFFI-N. Although providing support for the convergent validity of the GAD-7, the results also indicated that the GAD-7 factor was more strongly correlated with the NFFI-N than with the DASS-D ($r = .47$, $z = 2.97$, $p < .01$) and the OCD-R ($r = .42$, $z = 3.40$, $p < .001$) in accord with the notion that GAD is more closely related to neuroticism than other emotional disorders (e.g., major depressive disorder, OCD). The test of the differential magnitude of correlations involving the GAD-7 and NFFI-N with the DASS-A was not significant ($r = .52$, $z = 0.85$).

Experiment Participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Abramowitz, J. S., & Deacon, B. J. (2006). Psychometric properties and construct validity of the Obsessive-Compulsive Inventory—Revised: replication and extension with a clinical sample. *Journal of Anxiety Disorders, 20*, 1016–1035.

Abramowitz, J. S., & Foa, E. B. (1998). Worries and obsessions in individuals with obsessive–compulsive disorder with and without comorbid generalized anxiety disorder. *Behaviour Research and Therapy, 36*, 695–700.

Beard, C., & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders, 28*, 547–552.

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: multiple factors or method effects? *Behaviour Research and Therapy, 41*, 1411–1426.

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

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the *DSM-IV* anxiety and mood disorders: implications for assessment and treatment. *Psychological Assessment, 21*, 256–271.

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), 33–7.

Brown T. A., Chorpita B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy, 30* 35(1),79–89.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the *DSM-IV* anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110*, 585–599.

Dear, B. F., Titov, N., Sunderland, M., McMillan, D., Anderson, T., Lorian, C., & Robinson, E. (2011). Psychometric comparison of the Generalized Anxiety Disorder Scale-7 and the Penn State Worry Questionnaire for measuring response during treatment of generalised anxiety disorder. *Cognitive Behaviour Therapy, 40*, 216–227.

Delgadillo, J., Payne, S., Gilbody, S., Godfrey, C., Gore, S., Jessop, D., & Dale, V. (2012). Brief case finding tools for anxiety disorders: validation of GAD-7 and GAD-7 in addictions treatment. *Drug and Alcohol Dependence, 125*, 37–42.

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorder Interview Schedule for DSM-IV: Lifetime version (ADIS-IV-L)*. New York: Oxford University Press.

Donker, T., van Straten, A., Marks, I., & Cuijpers, P. (2011). Quick and easy self-rating of generalized anxiety disorder: validity of the Dutch web-based GAD-7, GAD-2 and GAD-SI. *Psychiatry Research, 188*, 58–64.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The obsessive-compulsive inventory: development and validation of a short version. *Psychological Assessment, 14*, 485–496.

Grant, B. F., Goldstein, R. B., Chou, S. P., Huang, B., Stinson, F. S., Dawson, D. A., & Compton, W. M. (2009). Sociodemographic and psychopathologic predictors of first incidence of *DSM-IV* substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Molecular Psychiatry, 14*, 1051–1066.

Greenberg, P. E., Sisitsky, T., Kessler, R. C., Finkelstein, S. N., Berndt, E. R., Davidson, J. R. T., & Fyer, A. J. (1999). The economic burden of anxiety disorders in the 1990s. *Journal of Clinical Psychiatry, 60*, 427–435.

Hoffman, D. L., Dukes, E. M., & Wittchen, H.-U. (2006). Human and economic burden of generalized anxiety disorder. *Depression and Anxiety, 25*, 72–90.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Kertz, S., Bigda-Peyton, J., & Björgvinsson, T. (2013). Validity of the Generalized Anxiety Disorder-7 Scale in an acute psychiatric sample. *Clinical Psychology and Psychotherapy, 20*, 456–464.

Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine, 146*, 317–325.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*, 335–343.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care, 46*, 266–274.

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*, 487–495.

Mills, S. D., Fox, R. S., Malcarne, V. L., Roesch, S. C., Champagne, B. R., & Sadler, G. R. (2014). The psychometric properties of the Generalized Anxiety Disorder-7 Scale in Hispanic Americans with English or Spanish language preference. *Cultural Diversity and Ethnic Minority Psychology, 20*, 463–468.

Muthén, L. K., & Muthén, B. O. (2008–2014). *Mplus 6.0* [Computer software]. Los Angeles: Author.

National Comorbidity Survey (NCS) (2011). NCS-R lifetime prevalence estimates. Available from: http://www.hcp.med.harvard.edu/ncs/ftpdir/NCS-R_Lifetime_Prevalence_Estimates.pdf. Accessed 22 August 2015

Raykov, T. (2001). Estimation of congeneric scale reliability via covariance structure models with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology, 54*, 315–323.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35*, 299–331.

Roy-Byrne, P. P. (1996). Generalized anxiety and mixed anxiety-depression: association with disability and health care utilization. *Journal of Clinical Psychiatry, 57*, 86–91.

Roy-Byrne, P. P., & Katon, W. (1997). Generalized anxiety disorder in primary care: the precursor/modifier pathway to increased health care utilization. *Journal of Clinical Psychiatry, 58*, 34–40.

Ruiz, M. A., Zamorano, E., García-Campayo, J., Pardo, A., Freire, O., & Rejas, J. (2011). Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. *Journal of Affective Disorders, 128*, 277–286.

Sousa, T. V., Viveiros, A., Chai, M. V., Vicente, F. L., Jesus, G., Carnot, M. J., & Ferreira, P. L. (2015). Reliability and validity of the Portuguese version of the Generalized Anxiety Disorder (GAD-7) scale. *Health and Quality of Life Outcomes, 13*, 1–8.

Spitzer, R. L., Kroenke, K., Williams, Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine, 166*, 1092–1097.