



HHS Public Access

Author manuscript

J Psychopathol Behav Assess. Author manuscript; available in PMC 2019 December 01.

Published in final edited form as:

J Psychopathol Behav Assess. 2018 December ; 40(4): 636–644. doi:10.1007/s10862-018-9683-5.

The Contribution of Worry Behaviors to the Diagnosis of Generalized Anxiety Disorder

Timothy A. Brown and Esther S. Tung

Center for Anxiety & Related Disorders, Department of Psychological and Brain Sciences, Boston University, Boston, MA

Abstract

Worry behaviors (i.e., overt acts to avoid or cope with worry-induced distress) have been recognized as being important in the psychopathology and treatment of generalized anxiety disorder (GAD). This study evaluated the worry behaviors criterion proposed for *DSM-5* GAD, but was ultimately not adopted due to insufficient evidence. In 800 outpatients with emotional disorders (366 with GAD), most patients with GAD (92.6%) met the proposed worry behaviors criterion, which was at a rate significantly higher than other patient groups (e.g., patients with mood disorders). Patients who met the worry behaviors criterion had more severe GAD than patients who did not. The worry behaviors criterion, and 3 of its 4 constituent behaviors, were associated with no better than “fair” interrater reliability. Diagnostic reliability of GAD was not improved in cases where both interviewers agreed the worry behaviors criterion was met. The worry behaviors criterion significantly predicted *DSM-5* GAD holding core GAD features constant (e.g., excessive worry), but this contribution was weak and did not appreciably improve the classification accuracy of GAD diagnostic status. Mixed support was obtained for the discriminant validity of the worry behaviors criterion in relation to mood disorders. Raising the proposed threshold of the criterion (requiring 2 instead of 1 behaviors) did not result in a substantial improvement in reliability, prediction, and classification accuracy. Although additional research is warranted (e.g., importance of worry behaviors in the treatment and natural course of GAD), the results raise questions about the role of worry behaviors in the diagnostic classification of GAD.

Keywords

generalized anxiety disorder; worry behaviors; diagnostic criteria; *DSM-5*

Correspondence concerning this article should be addressed to Timothy A. Brown, Center for Anxiety & Related Disorders, Department of Psychology, Boston University, 648 Beacon Street, 6th floor, Boston, MA 02215-2013, tabrown@bu.edu, telephone: 617-353-9610.

Conflict of Interest: Timothy A. Brown and Esther S. Tung declare that they have no conflict of interest.

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.

Experiment Participants: Informed consent was obtained from all individual participants included in the study.

In *DSM-5*, the key diagnostic feature of generalized anxiety disorder (GAD) is excessive anxiety and difficult to control worry about a number of events or activities, occurring more days than not for at least six months (American Psychiatric Association, 2013). However, the diagnostic criteria for GAD have changed substantially since the category's inception in *DSM-III* (cf. Brown, Barlow, & Liebowitz, 1994), in part because of the disorder's chronically unsatisfactory diagnostic reliability (e.g., *DSM-III*: kappa = .47, Di Nardo, O'Brien, Barlow, Waddell, & Blanchard, 1983; *DSM-III-R*: ranges of kappas = .27 to .56, Di Nardo, Moras, Barlow, Rapee, & Brown, 1993; Mannuzza et al., 1989; Williams et al., 1992). Two significant changes to the definition of GAD were introduced in *DSM-IV* to improve diagnostic reliability (American Psychological Association, 1994): (1) the requirement that the individual finds the worry difficult to control, and (2) the reduction of the associated symptoms criterion from 18 to 6 symptoms (based on evidence that symptoms such as those involving autonomic hyperactivity are infrequently endorsed by patients with GAD and are not strongly related to worry; e.g., Brown, Marten, & Barlow, 1995). In addition, a diagnostic hierarchy rule was included stipulating that GAD should not be assigned if its features occurred exclusively during the course of a mood disorder or posttraumatic stress disorder. Results from a large-scale *DSM-IV* diagnostic reliability study (independent diagnostic interviews of 362 outpatients) indicated that the interrater agreement of GAD improved considerably (kappa = .65) relative to prior *DSM* definitions of the disorder (Brown, Di Nardo, Lehman, & Campbell, 2001). Although this improvement was encouraging, the diagnostic reliability of GAD continued to be among the lowest of the anxiety disorders (cf. Regier et al., 2013).

Several factors may contribute to the diagnostic reliability problems associated with GAD. An overarching issue pertains to concerns about GAD's discriminant validity. The diagnostic criteria for GAD overlap extensively with those of other anxiety and mood disorders (e.g., worry, associated symptoms such as restlessness and concentration difficulties). In fact, GAD has been regarded as the "basic" anxiety disorder because its core features of anxious apprehension and worry are present to varying degrees in all anxiety (and mood) disorders, and because these features are strongly related to dimensions that are thought to be higher-order vulnerability factors for the emotional disorders (e.g., neuroticism and negative affectivity; Brown & Barlow, 2002). Consequently, diagnostic comorbidity rates with GAD are routinely quite high, especially with the mood disorders (e.g., Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Kessler et al., 2008). The differential diagnosis of GAD may be more difficult because the disorder rarely presents in isolation. In a detailed analysis of the sources of diagnostic unreliability, Brown, Di Nardo, et al. (2001) found that, unlike other anxiety disorder categories, the majority (74%) of the diagnostic disagreements involving GAD were with other diagnoses (i.e., one interviewer assigned GAD while the other interviewer assigned a different diagnosis). In most instances (63%), these diagnostic disagreements were with the mood disorders, in accord with other evidence that the mood disorders may pose the greatest boundary problems with GAD (e.g., Brown, Campbell et al., 2001; Brown, Chorpita, & Barlow, 1998).

The diagnostic reliability of GAD may be further compromised by the fact that its definition does not include overt symptoms. Brown, Di Nardo, et al. (2001) found that interrater agreement was highest for the *DSM-IV* anxiety disorders that included explicit behavioral

symptoms (e.g., situational avoidance, compulsions) in their diagnostic definitions (kappas ranged from .77 to .86 for the principal diagnoses of panic disorder with agoraphobia, social phobia, obsessive-compulsive disorder, and specific phobia). In this study, “difference in patient report” (e.g., patient gives different symptom reports in the two interviews) was the second most common source of GAD diagnostic unreliability (55% of disagreements), an inconsistency that may stem partly from the requirement that patients correctly recall and report their internal experiences in both interviews in order for the GAD diagnosis to be reliably assigned.

Anxiety-induced avoidance, or maladaptive avoidance behaviors, is included in the criteria of all *DSM-5* anxiety disorders except GAD (American Psychiatric Association, 2013). This may be because worry itself has been conceptualized as a form of avoidance, a cognitive process that is key to the maintenance of GAD (e.g., Dugas, Gagnon, Ladouceur, & Freeston, 1998; Newman & Llera, 2011). Although worry may indeed be construed as cognitive avoidance, individuals with GAD also engage in situational avoidance and safety behaviors such as excessive checking and reassurance-seeking (Coleman et al., 2011; Schut, Castonguay, & Borkovec, 2001; Townsend et al., 1999). Still, little research has been done to examine the nature and importance of overt behavioral avoidance in GAD (Mahoney et al., 2016). The available evidence suggests that behavioral avoidance may play an important role in the severity and treatment of GAD. For instance, a study of 56 patients who underwent psychological treatment of GAD found that higher levels of worry at long-term follow-up were predicted by post-treatment avoidance and the use of safety behaviors (Beesdo-Baum et al., 2012). Further research on behaviors associated with worry may have strong implications for the diagnosis and treatment of GAD.

Based on this initial evidence, the *DSM-5* Anxiety Disorders Workgroup proposed four worry behaviors for inclusion in the diagnostic criteria of GAD (Andrews et al., 2010). The four worry behaviors were: (1) marked avoidance of potentially negative events or activities, (2) marked time and effort preparing for possible negative outcomes of events or activities, (3) marked procrastination in behavior or decision-making due to worries, and (4) repeatedly seeking reassurance due to worries. It was proposed that at least one of these four behaviors be present in order for a *DSM-5* diagnosis of GAD to be assigned (Andrews et al., 2010). Ultimately, this worry behavior criterion and other proposals (e.g., revisions to the associated symptoms and duration criteria) were not adopted in *DSM-5* given the lack of sufficient evidence to support these changes (Mahoney et al., 2016). Consequently, the *DSM-5* diagnostic criteria for adult GAD are virtually the same as in *DSM-IV*, except for minor wording changes and the elimination of the diagnostic hierarchy rule with mood disorders.

To promote further research on the role of worry behaviors in GAD, the Anxiety and Related Disorders Interview Schedule for *DSM-5* (ADIS-5; Brown & Barlow, 2014) includes a section to assess the four worry behaviors that were proposed for *DSM-5* (Andrews et al., 2010). Using the ADIS-5, the present study provides the first large-scale, exploratory evaluation of these worry behaviors and the proposed worry behaviors criterion. In addition to the evaluation of the interrater reliability and prevalence of these behaviors in a large outpatient sample, of particular interest was the determination of the extent to which worry

behaviors and the proposed worry behaviors criterion contribute to the prediction of the GAD diagnosis beyond the extant features of the condition (e.g., excessive worry). The potential utility of the worry behaviors criterion was also evaluated with respect to its proposed cutoff (i.e., requirement of at least one of four behaviors), its ability to improve the diagnostic reliability of GAD, and its specificity to GAD in relation to mood disorders.

Method

Participants

The sample was 800 patients who presented for intake evaluation at an outpatient clinic specializing in the assessment and treatment of anxiety, mood, and related disorders. Women constituted the larger portion of the sample (59.3%); average age was 31.55 ($SD = 12.91$, range = 18 to 82). The sample was predominantly Caucasian (81.6%; African-American = 6.8%, Asian = 9.8%, Latino/Hispanic = 9.3%). The rates of current clinical disorders (collapsing across principal and additional diagnoses) that occurred most frequently in the sample were: generalized anxiety disorder ($n = 366$, 45.8%), social anxiety disorder (48.1%), mood disorders (i.e., major depression, persistent depressive disorder, other specified depressive disorder; 39.8%), panic disorder (16.1%), agoraphobia (15.5%), obsessive-compulsive disorder (13.6%), and specific phobia (16.1%). In addition, there were 78 (9.8%) cases who were assigned “other specific anxiety disorder” because they presented with clinically significant symptoms that did not quite meet the *DSM-5* diagnostic definition of GAD (e.g., one domain of excessive worry; cf. Lawrence & Brown, 2009).

Anxiety and Related Disorders Interview Schedule for DSM-5 (ADIS-5; Brown & Barlow, 2014)

The ADIS-5 is a semi-structured interview designed to ascertain reliable diagnosis of the *DSM-5* anxiety, mood, somatoform, obsessive-compulsive, trauma, and substance use disorders, and to screen for the presence of other conditions (e.g., eating and psychotic disorders). The ADIS-5 provides dimensional assessment of the key and associated features of disorders (0–8 ratings); these features are dimensionally rated regardless of whether a formal *DSM-5* diagnosis is under consideration. For each *DSM-5* diagnosis, interviewers assigned a 0–8 clinical severity rating (CSR) to indicate the degree of distress and lifestyle impairment associated with the disorder (0 = none, 8 = very severely disturbing/disabling). Disorders that met or exceeded the threshold for a formal *DSM-5* diagnosis were assigned CSRs of 4 (*definitely disturbing/disabling*) or higher (“clinical” diagnoses); “subclinical” diagnoses (e.g., disorders in partial remission) were given CSRs below 4. In patients with two or more current *DSM-5* disorders, the “principal” diagnosis was the condition associated with the highest level of distress and impairment (other co-occurring clinical diagnoses are referred to as “additional” diagnoses).

Diagnosticians were 5 clinical psychologists and 5 advanced clinical doctoral students. Before participating in the study, diagnosticians underwent extensive training and met strict certification criteria in the administration of the ADIS-5 (for a detailed description of these training and certification procedures, see Brown, Di Nardo, et al., 2001). A test-retest reliability study of the *DSM-IV* version of the ADIS indicated good-to-excellent interrater

agreement for current disorders (range of κ s = .67 to .86) except dysthymia (κ = .31; Brown, Di Nardo, et al., 2001). For purposes of estimating the interrater reliability of the diagnoses and dimensional ratings of the ADIS-5, a subset of the present sample ($n = 50$) was randomly selected to undergo two independent administrations of the ADIS-5; 31 (62.0%) of these reliability interviews involved cases who were ultimately assigned a consensus diagnosis of GAD. On average, the second interview occurred 9.20 days after the initial ADIS-5 ($SD = 4.06$, range = 2 to 17 days). All ADIS-5 intakes were presented in weekly staff meetings that entailed the presentation of interviewers' diagnoses, discussion of factors contributing to any diagnostic disagreements (in the case of reliability interviews), and establishment of consensus diagnoses and CSRs.

In the GAD section of the ADIS-5, interviewers inquired about each of the four worry behaviors proposed for inclusion in the *DSM-5* definition of GAD (i.e., avoidance, over-preparation, procrastination, reassurance seeking). After recording a description of the worry behavior, the behavior was rated on a 0–8 scale of frequency/severity (0 = *never*, 8 = *always/very severe*). As in other sections of the ADIS-5, these behaviors were recorded and rated regardless of whether a formal GAD diagnosis was under consideration. Along the lines of the disorder CSRs, if a given worry behavior was rated 4 or above (*sometimes/moderate*), it was deemed to meet the threshold for the proposed *DSM-5* criterion.

In addition to the four worry behaviors, two composite scores were created using dimensional ratings from the GAD section of the ADIS-5 (all 0–8 scales): (a) excessiveness of worry in eight areas (e.g., finances, minor matters); and (b) frequency and severity of the six symptoms comprising the *DSM-5* associated symptom criterion of GAD. Previous research using the *DSM-IV* version of the ADIS indicates that these composite scores have favorable interrater reliability (i.e., test-retest r s = .73 and .83 for worry excessiveness and associated symptoms, respectively; cf. Brown, Di Nardo, et al., 2001). In the present sample, the intraclass correlations for the 50 cases receiving two, independent administrations of the ADIS-5 were .75 and .79 for worry excessiveness and associated symptoms, respectively.

Self-Report Measures

Patients completed a battery of questionnaires as part of their intake assessment. From these measures, the Depression subscale from the briefer (21-item) version of the Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995) and the GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006) were selected as dimensional, self-report indicators of depression and GAD severity, respectively. The DASS-Depression scale is comprised of 7 items that measure symptoms believed to be specific to the construct of unipolar depression (e.g., hopelessness, anhedonia); respondents use a 0–3 scale to indicate the extent to which each symptom applied to them in the past week (range of scores = 0 to 21). The GAD-7 was designed to screen and measure the severity of GAD using 7 items that are closely aligned to the symptoms included in the *DSM* definition of GAD; respondents indicate on a 0–3 scale how often they were bothered by the symptom over the prior two weeks (range of scores = 0 to 21). Both the DASS and the GAD-7 are well-established and have been shown to possess favorable psychometric properties in clinical samples (e.g., Brown, Chorpita, Korotitsch, & Barlow, 1997; Rutter & Brown, 2017).

Results

Interrater Reliability of Worry Behaviors

Interrater reliability estimates for the worry behavior ratings are presented in Table 1. Intraclass correlations (ICCs) of the dimensional ratings of the four worry behaviors and the worry behavior composite score (summing the four behaviors) ranged from .39 (over-preparation) to .67 (avoidance) (all $ps < .01$). The interrater reliability of the binary outcomes (e.g., presence/absence of a clinically significant worry behavior) was estimated using kappa coefficients (κ). Following guidelines used in previous diagnostic reliability studies (e.g., Brown, Di Nardo, et al., 2001), the standards applied to the interpretation of κ were: excellent agreement ($\kappa \geq .75$), good agreement ($.60 < \kappa < .74$), fair agreement ($.40 < \kappa < .59$), and poor agreement ($\kappa < .40$) (cf. Fleiss, Nee, & Landis, 1979). As shown in Table 1, only avoidance behavior was associated with good interrater reliability ($\kappa = .61$); agreement was fair for over-preparation and procrastination ($\kappa = .43$ and $.52$, respectively) and poor for reassurance seeking ($\kappa = .32$). In accord with the *DSM-5* proposal, patients were classified as meeting the worry behaviors criterion if at least one behavior was present at a clinically significant level (i.e., rated 4 or higher on the ADIS-5). The proposed worry behaviors criterion was associated with fair interrater reliability ($\kappa = .44$). For comparison purposes, the reliability of the GAD diagnosis was also estimated. The diagnostic reliability of *DSM-5* GAD was good ($\kappa = .63$); the ICC of the ADIS-5 CSR for GAD was .62.

To evaluate whether the presence of worry behaviors fostered the interrater reliability of the GAD diagnosis, κ for GAD was re-estimated in cases where both interviewers agreed that the proposed worry behaviors criterion was met ($n = 31$). Although this result must be interpreted cautiously because of the smaller subsample, meeting the worry behaviors criterion did not improve the diagnostic reliability of GAD ($\kappa = .33$, 77% agreement).

Frequency of Worry Behaviors as a Function of DSM-5 Diagnostic Status

The sample rates of the four worry behaviors and the proposed *DSM-5* worry behavior criterion for patients with and without a current GAD diagnosis are presented in Table 2. The specific worry behaviors most commonly endorsed by patients with GAD were procrastination (67.2%) and reassurance seeking (57.4%), although nearly half of these patients also reported engaging in avoidance and over-preparation at a clinically significant level (49.7% and 47.5%, respectively). Whereas worry behaviors were present to a considerable degree in patients without GAD (as high as 27.4%, see Table 2), each behavior was significantly more common in GAD cases than in non-cases (ORs = 4.94, $ps < .001$). Moreover, 92.6% of patients with GAD met the proposed *DSM-5* worry behaviors criterion, which was at a rate significantly higher than patients without GAD (47.0%), OR = 14.16, $p < .001$.

Yet, it is noteworthy that 27 (7.4%) patients with GAD would not have been assigned this diagnosis if the worry behaviors criterion had been adopted in *DSM-5*. Compared to patients with GAD who met the worry behaviors criterion, these patients evidenced significantly less excessive worry ($Ms = 21.03$ and 16.19), GAD associated symptoms ($Ms = 23.50$ and 18.67), and overall GAD clinical severity ($Ms = 5.16$ and 4.63), as assessed by the ADIS-5

(all $ps < .001$). With regard to the ADIS-5 CSR, nearly half ($n = 13$, 48.1%) of the 27 patients who did not meet the worry behavior criterion were assigned GAD right at the diagnostic threshold (i.e., CSR = 4), compared to less than a quarter ($n = 73$, 21.5%) of the 339 patients with GAD who met this criterion.

To evaluate whether worry behaviors differentiate GAD from near-threshold presentations of GAD, the between-groups analyses were re-conducted changing the comparison group to the 79 patients who were assigned “other specified anxiety disorder” because their clinically significant symptoms did not quite meet the *DSM-5* definition of GAD (OSAD). As reflected in Table 2, it is noteworthy that at least 30% of patients without GAD who had a clinically significant worry behavior or met the worry behaviors criterion had OSAD; e.g., 28 of the 68 non-GAD cases (41.2%) who evidenced avoidance were assigned OSAD.¹ Although worry behaviors were more common in OSAD than in cases without GAD (endorsement percentages ranged from 29.5% to 51.3%), each behavior was still significantly more frequent in GAD (largest $p = .025$, Table 2). In addition, significantly fewer patients with OSAD met the proposed *DSM-5* worry behaviors criterion (79.5%) than patients with GAD (92.6%), OR = 3.24, $p < .001$.

Incremental Contribution of Worry Behaviors to the Prediction of Generalized Anxiety Disorder

Hierarchical logistic regressions were conducted to determine if the proposed *DSM-5* worry behaviors criterion, and each specific worry behavior, contributed uniquely to the prediction of GAD diagnostic status beyond the core symptom features of GAD (i.e., excessive of worry, associated symptoms).² The analyses were conducted in two steps (i.e., core GAD features entered in the first step) to examine the extent to which classification accuracy was improved by the addition of the worry behavior predictor. Classification accuracy was indexed by λ_p , which quantifies the proportional reduction in errors of prediction taking into account the base rate of the outcome (Menard, 2002). In each analysis, excessiveness of worry and GAD associated symptoms contributed significantly to the prediction of GAD caseness, with and without the worry behavior predictor in the equation (see Table 3). Without a worry behavior predictor in the equation, λ_p was .672 which indicates that the excessive worry and associated symptoms predictors resulted in a 67.2% reduction in classification errors relative to predicting every patient did not have GAD (“no GAD” was the modal category on the binary GAD diagnosis outcome).

As shown in Table 3, holding excessive worry and associated symptoms constant, the proposed *DSM-5* worry behaviors criterion contributed significantly to the prediction of GAD diagnostic status in the expected direction (OR = 2.70, likelihood ratio test value = 11.84, $p < .001$). However, the size of this effect was modest as reflected by the fact that this predictor uniquely reduced the deviance of the model ($-2LL$ of GAD status with no predictors) by only 1% ($R^2_L = .01$, Table 3). The tenuous clinical significance of this effect

¹The No GAD and OSAD groups presented in Table 2 are not mutually exclusive because the No GAD group was comprised in part by patients who were assigned OSAD.

²Because the ADIS-5 ratings of excessiveness of worry and uncontrollability of worry were highly correlated ($r = .92$), uncontrollability of worry was not included as a predictor in these analyses.

was further reflected by a negligible improvement in classification accuracy ($\lambda_p = .675$). Specifically, the addition of the proposed *DSM-5* worry behavior criterion resulted in just one fewer classification error relative to the prior step of the analysis that only included excessiveness of worry and associated symptoms (although the criterion led to two more correct classifications of GAD, there was one new false positive).

Avoidance, over-preparation, and reassurance seeking uniquely contributed to the prediction of GAD diagnostic status (p s < .05, .001, and .01, respectively), although the sizes of these contributions were small (all $R^2_L = .01$, Table 3). Procrastination did not significantly predict GAD when the core features of the disorder were held constant ($p = .075$).

Specificity of the Worry Behaviors Criterion in Relation to Mood Disorders

Because the mood disorders pose the most significant diagnostic boundary problems with GAD (cf. Brown et al., 1994, Brown, Di Nardo, et al., 2001), an important question is whether worry behaviors bolster the differentiation of GAD from these neighboring conditions. The majority of patients with a current mood disorder met the proposed worry behaviors criterion (80.8%). However, 58.8% of these patients had a comorbid GAD diagnosis. In the subset of patients with a mood disorder without comorbid GAD ($n = 131$), the prevalence of the worry behaviors criterion was 62.6%. This rate was significantly lower than that seen in the 179 patients with GAD who did not have a co-occurring mood disorder (91.6%), $OR = 6.53$, $p < .001$. To determine if the worry behaviors criterion significantly contributed to the prediction of GAD caseness after also taking into account depression symptoms, the first analysis presented in Table 3 was re-conducted including the DASS-Depression scale in the initial step of the hierarchical logistic regression. The worry behaviors criterion remained a significant predictor of GAD diagnostic status when holding depression and core features of GAD constant ($OR = 2.61$, likelihood ratio test value = 10.13, $p < .001$), although the strength of this contribution continued to be weak ($R^2_L = .01$).³

The concurrent validity of the proposed worry behaviors criterion was further evaluated by correlating it with diagnostic and dimensional indicators of GAD and mood disorders. The criterion was more strongly correlated with GAD diagnostic status (tetrachoric $r = .79$) than with the presence of a current mood disorder ($r = .37$), as reflected by the significance of a test of the differential magnitude of these relationships, Steiger's (1980) $z = 13.00$, $p < .001$. However, this pattern of differential relationships was not found using dimensional, self-report measures of GAD and depression. The worry behaviors criterion was less strongly and equally correlated with the GAD-7 and DASS-Depression scales; r s = .29 and .26, respectively, $z = 0.89$, $p = .19$.

Raising the Threshold of the Worry Behavior Diagnostic Criterion

As previously noted, the worry behaviors criterion was proposed for *DSM-5* with the requirement that at least one of four behaviors is present. Because this cutoff is somewhat

³The worry behaviors criterion also remained a significant, albeit weak, predictor of GAD status when the presence/absence of a current mood disorder diagnosis was used as the depression covariate, $OR = 2.72$, $LRT = 12.01$, $p < .001$, $R^2_L = .01$.

arbitrary and without empirical basis, it was of interest to determine if the performance of the criterion could be improved by raising its diagnostic threshold. Thus, the analyses were reconducted using a criterion that required two or more clinically significant worry behaviors (i.e., rated 4 or higher on the ADIS-5). The interrater reliability of the revised criterion was slightly better than the original criterion ($\kappa = .50$, 74% agreement). Although more cases with GAD met this criterion than cases without GAD (72.4% and 22.4%, respectively; $OR = 9.12$, $p < .001$), an additional 20.2% of patients ($n = 74$) with a current *DSM-5* GAD diagnosis would not be given this disorder under the revised criterion. The revised criterion contributed significantly to the prediction of GAD diagnostic status after holding excessive worry and associated symptoms constant ($OR = 2.32$, likelihood ratio test value = 14.44, $p < .001$). Classification accuracy improved marginally ($\lambda_p = .683$) because three false positives produced by the original criterion were rectified by the revised criterion; both criteria correctly classified 85.2% of patients with GAD as having the disorder.

Discussion

The present findings indicate that although most patients with GAD meet the worry behaviors criterion proposed by the *DSM-5* Anxiety Disorders Workgroup (Andrews et al., 2010), the contribution of this criterion to the diagnosis of GAD is small. One of the key motivations for the proposed inclusion of a worry behaviors criterion was to improve GAD's diagnostic reliability. Interrater agreement of GAD has historically been modest, particularly in relation to anxiety disorders that are defined partly by overt behavioral symptoms (e.g., situational avoidance, safety behaviors). Although prior evidence suggests that overt symptoms foster diagnostic reliability (cf. Brown, Di Nardo, et al., 2001), the interrater agreement of the dimensional and categorical ratings of the individual worry behaviors and proposed worry behaviors criterion was no better than "fair" (e.g., range of κ s = .32 to .52), with the exception of behavioral avoidance ($\kappa = .61$, $ICC = .67$).⁴ This was despite the fact that the interrater reliability of the *DSM-5* GAD diagnosis was "good" ($\kappa = .63$) and at the same level of agreement previously obtained in our setting for *DSM-IV* GAD ($\kappa = .65$; Brown, Di Nardo, et al., 2001). The lack of variation in the interrater reliability of the GAD diagnosis is not surprising given that *DSM-IV* and *DSM-5* criteria are virtually identical, except for the removal of the diagnostic hierarchy rule with mood disorders. The current findings raise concerns about whether the addition of worry behaviors would improve the reliability and diagnostic boundary of GAD given that three of the four constituent symptoms and the criterion itself were not associated with favorable levels of interrater agreement. Perhaps the interrater reliability of the criterion could be increased by the refinement of the worry behaviors initially proposed by Andrews et al. (2010) (e.g., including specific examples of each type of worry behavior that are common in patients with GAD). Nonetheless, data from this study suggest that a more reliable worry behaviors criterion would not necessarily improve the diagnostic reliability of GAD. In the subset of

⁴Although dimensional approaches to disorder classification (e.g., Brown & Barlow, 2009) have been proposed in part to reduce measurement error associated with imposing categorical cut-offs on symptom features that operate in a continuous fashion (e.g., binary thresholds of symptom counts and severity in *DSM-5* diagnoses), it is interesting to note that the interrater reliability of the dimensional and categorical ratings of the worry behaviors did not appreciably differ (see Table 1).

double interview cases ($n = 31$) judged by both interviewers to meet the proposed worry behaviors criterion, the diagnostic reliability estimate of GAD was even lower ($\kappa = .33$).

Each of the four worry behaviors was significantly more common in patients with GAD than in patients without GAD and patients with OSAD (Table 2). In addition, the vast majority (92.6%) of patients with GAD met the proposed *DSM-5* worry behaviors criterion, which was at a rate significantly higher than was observed in patients without GAD (47.0%) and patients with OSAD (79.5%). The finding that worry behaviors differentiate GAD from relevant patient groups was bolstered by multivariate analyses showing that the worry behaviors criterion and all but one worry behavior significantly contributed to the prediction of GAD caseness beyond the core features of GAD (Table 3). However, the clinical significance of this contribution could be questioned because the worry behavior predictors were associated with very small effect sizes (i.e., R^2_L no higher than .01) and led to trivial improvements in the classification accuracy of GAD (i.e., produced just one fewer classification error).

Worry behaviors were very common in GAD, yet 7.4% of patients would not have been assigned GAD if the criterion was adopted in *DSM-5*. These patients were significantly less symptomatic than patients with GAD who met the proposed criterion (on measures of excessive worry, GAD associated symptoms, and the GAD clinical severity rating). These results suggest that the inclusion of the worry behaviors criterion would raise the severity threshold of the GAD diagnosis; i.e., many cases with less severe clinical presentations would no longer meet *DSM* diagnostic criteria for GAD. Although it is possible that the worry behaviors criterion would foster the differentiation of GAD from nonpathological worry, this addition would have the less desirable consequence of increasing the prevalence of the OSAD diagnosis. Specifically, our findings indicate that nearly 1 out of every 13 outpatients (7.4%) with clinically significant (albeit less severe) symptoms meeting the current definition of a GAD would be relegated to a residual category (OSAD) if the worry behaviors criterion was implemented by *DSM*. It might be argued that an elevation in GAD's diagnostic threshold could improve diagnostic reliability by reducing the number of disagreements due to threshold issues (i.e., instances where diagnosticians agree on the presence of the key and associated features of the disorder, but disagree on whether the symptoms cause sufficient interference and distress). However, the fact that the interrater reliability of GAD was not better in cases where both interviewers concurred on the presence of the worry behaviors criterion suggests that the criterion would not reconcile this or other sources of diagnostic unreliability.⁵

A key dilemma in *DSM* classification is that the definition of GAD does not include specific features that differentiate it from closely neighboring conditions such as mood disorders (Starcevic, Portman, & Beck, 2012). Although worry behaviors were common in patients with mood disorders (62.6% met the worry behaviors criterion), these features were significantly more prevalent in patients with GAD. Of note, the prevalence of the worry behaviors criterion in patients with GAD was not affected by mood disorder comorbidity;

⁵It is also noteworthy that "threshold" issues were a relatively less common reason for GAD disagreements (9%) in a study examining the sources of diagnostic unreliability of *DSM-IV* anxiety and mood disorders (Brown, Di Nardo, et al., 2001).

92.6% of all patients with GAD met this criterion compared to 91.6% of patients with GAD without a comorbid mood disorder. At the diagnostic level, the worry behaviors criterion was more strongly correlated with GAD than with mood disorders, although this differential relationship was not found using dimensional, self-report indicators of GAD and depression. This discrepancy may have been due to source effects in the diagnostic-level analyses (i.e., GAD diagnoses and worry behaviors ratings were made by the same interviewer) and possible psychometric fallibility of the GAD-7 as a measure of GAD severity (i.e., because the item content of the GAD-7 is more heavily weighted on nonspecific, associated symptoms, it is as strongly correlated with measures of general distress as with measures of worry and GAD; cf. Rutter & Brown, 2017). Collectively, the results support the notion that worry behaviors are more strongly associated with GAD. However, the considerable prevalence of these features in mood disorders and other diagnostic groups indicates that worry behaviors are not unique to GAD. Thus, it is not clear to what extent worry behaviors can advance the differential diagnosis of GAD.

If the four worry behaviors proposed by Andrews et al. (2010) were included in the *DSM* definition of GAD, the current results support a criterion requiring the presence of at least one behavior. A cutoff requiring two or more behaviors resulted in only marginal improvements in interrater reliability and classification accuracy, but would have the deleterious effect of reassigning an additional 20.2% of patients with *DSM-5* GAD to the residual diagnosis of OSAD; in total, over 1 in 4 (27.6%) patients with *DSM-5* GAD would no longer meet criteria for this diagnosis. Of course, these and other study findings are based specifically on the four behaviors proposed by the *DSM-5* Anxiety Disorders Workgroup. In view of evidence of less than favorable interrater reliability and weak predictability of the four individual behaviors, the question arises as to whether future research could produce a more psychometrically sound and more valid list of behaviors, and ultimately a diagnostically and clinically useful worry behaviors criterion. Along these lines, further consideration should be given to how a worry behaviors criterion could be refined to be more specific to GAD. Contributing to the current lack of diagnostic specificity is the possibility that patients with disorders other than GAD endorse worry behaviors for reasons related to their disorder, and not because of excessive worry; e.g., procrastination may be prevalent in mood disorders as a consequence of low energy and a loss of interest in usual activities. Finer-grained analysis of extant questionnaires of the worry behaviors construct (cf. Mahoney et al., 2016) may have heuristic value to improving the validity and specificity of a worry behaviors criterion (e.g., translating particularly strong questionnaire items to potential diagnostic criteria), as well as qualitative research in patient samples that produces a wider array of behaviors for further study.

Although the present study was indeed restricted to the four behaviors proposed by Andrews et al. (2010), it provided the first large-scale evaluation of the worry behaviors criterion that was considered for *DSM-5*. The study's large sample ($N = 800$) allowed for a strong evaluation of the prevalence of these behaviors in a clinical sample, as well as their predictability and specificity to GAD. However, the number of cases receiving two, independent administrations of the ADIS-5 was considerably smaller ($n = 50$) which raises concern about the role of sampling error in the observed interrater reliability estimates of the worry behaviors and worry behaviors criterion (as well as the extent to which these estimates

generalize to different sets of diagnostic interviewers). It is encouraging that the diagnostic reliability estimate for GAD in this study was virtually the same as the estimate obtained in our large-scale evaluation of *DSM-IV* emotional disorders (Brown, Di Nardo, et al., 2001), but the interrater agreement of worry behaviors should be further investigated in larger samples and with different interviewers and instruments. Moreover, the cross-sectional nature of this inquiry precluded some potentially informative analyses (e.g., temporal stability of worry behaviors over a longer interval). Although the current findings suggest that worry behaviors contribute only modestly to the diagnostic classification of GAD, these features may nonetheless be quite salient to treatment planning, treatment response, and the natural course of the disorder (e.g., represent important treatment targets, have a significant role in prognosis; Beesdo-Baum et al., 2012). Thus, continued study should not be limited to diagnostic issues but should also involve evaluation of the role that worry behaviors play in the maintenance and treatment of GAD.

Acknowledgments

The authors thank Anthony J. Rosellini for his comments on this manuscript.

Funding: This research was funded by a grant from the National Institute of Mental Health (R01 MH039096; PI: Brown).

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 1994.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. Washington, DC: Author; 2013.
- Andrews G, Hobbs MJ, Borkovec TD, Beesdo K, Craske MG, Heimberg RG, Rapee RM, Ruscio AM, Stanley MA. 2010; Generalized worry disorder: A review of DSM-IV generalized anxiety disorder and options for DSM-V. *Depression and Anxiety*. 27:134–147. [PubMed: 20058241]
- Beesdo-Baum K, Jenjahn E, Höfler M, Lueken U, Becker ES, Hoyer J. 2012; Avoidance, safety behavior, and reassurance seeking in generalized anxiety disorder. *Depression and Anxiety*. 29:948–957. [PubMed: 22581482]
- Brown, TA, Barlow, DH. Classification of anxiety and mood disorders. In: Barlow, DH, editor *Anxiety and its disorders: The nature and treatment of anxiety and panic*. 2. New York: Guilford Press; 2002. 292–327.
- Brown, TA, Barlow, DH. *Anxiety and Related Disorders Interview Schedule for DSM-5 (ADIS-5)*. New York: Oxford University Press; 2014.
- Brown TA, Barlow DH, Liebowitz MR. 1994; The empirical basis of generalized anxiety disorder. *American Journal of Psychiatry*. 151:1272–1280. [PubMed: 8067480]
- Brown TA, Campbell LA, Lehman CL, Grisham JR, Mancill RB. 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. [PubMed: 11727948]
- Brown TA, Chorpita BF, Barlow DH. 1998; Structural relationships among dimensions of the DSM-IV anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*. 107:179–192. [PubMed: 9604548]
- Brown TA, Chorpita BF, Korotitsch W, Barlow DH. 1997; Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy*. 35:79–89. [PubMed: 9009048]
- Brown TA, Di Nardo PA, Lehman CL, Campbell LA. 2001; Reliability of DSM-IV anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology*. 110:49–58. [PubMed: 11261399]

- Brown TA, Marten PA, Barlow DH. 1995; Discriminant validity of the symptoms constituting the DSM-III-R and DSM-IV associated symptom criterion of generalized anxiety disorder. *Journal of Anxiety Disorders*. 9:317–328.
- Coleman SL, Pieterfesa AS, Holaway RM, Coles ME, Heimberg RG. 2011; Content and correlates of checking related to symptoms of obsessive-compulsive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders*. 25:293–301. [PubMed: 21242055]
- Di Nardo PA, Moras K, Barlow DH, Rapee RM, Brown TA. 1993; Reliability of DSM-III-R anxiety disorder categories using the Anxiety Disorders Interview Schedule-Revised (ADIS-R). *Archives of General Psychiatry*. 50:251–256. [PubMed: 8466385]
- Di Nardo PA, O'Brien GT, Barlow DH, Waddell MT, Blanchard EB. 1983; Reliability of the DSM-III anxiety disorder categories using a new structured interview. *Archives of General Psychiatry*. 40:1070–1074. [PubMed: 6625856]
- Dugas MJ, Gagnon F, Ladouceur R, Freeston MH. 1998; Generalized anxiety disorder: A preliminary test of a conceptual model. *Behaviour Research and Therapy*. 36:215–226. [PubMed: 9613027]
- Fleiss JL, Nee JCM, Landis JR. 1979; Large sample variance of kappa in the case of different sets of raters. *Psychological Bulletin*. 86:974–977.
- Kessler RC, Gruber M, Hettema JM, Hwang I, Sampson N, Yonkers KA. 2008; Comorbid major depression and generalized anxiety disorders in the National Comorbidity Survey follow-up. *Psychological Medicine*. 38:365–374. [PubMed: 18047766]
- Lawrence AE, Brown TA. 2009; Differentiating generalized anxiety disorder from anxiety disorder not otherwise specified. *Journal of Nervous and Mental Disease*. 12:879–886.
- Lovibond PF, Lovibond SH. 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*. 33:335–342. [PubMed: 7726811]
- Mahoney AE, Hobbs MJ, Newby JM, Williams AD, Sunderland M, Andrews G. 2016; The Worry Behaviors Inventory: Assessing behavioral avoidance associated with generalized anxiety disorder. *Journal of Affective Disorders*. 203:256–264. [PubMed: 27314812]
- Mannuzza S, Fyer AJ, Martin LY, Gallops MS, Endicott J, Gorman JM, Liebowitz MR, Klein DF. 1989; Reliability of anxiety assessment: I. Diagnostic agreement. *Archives of General Psychiatry*. 46:1093–1101. [PubMed: 2589923]
- Menard, S. *Applied logistic regression analysis*. 2. Thousand Oaks, CA: Sage Publications; 2002.
- Newman MG, Llera SJ. 2011; A novel theory of experiential avoidance in generalized anxiety disorder: A review and synthesis of research supporting a contrast avoidance model of worry. *Clinical Psychology Review*. 31:371–382. [PubMed: 21334285]
- Regier DA, Narrow WE, Clarke DE, Kraemer HC, Kuramoto SJ, Kuhl EA, Kupfer DJ. 2013; DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry*. 170:59–70. [PubMed: 23111466]
- Rutter LA, Brown TA. 2017; Psychometric properties of the Generalized Anxiety Disorder Scale - 7 (GAD-7) in outpatients with anxiety and mood disorders. *Journal of Psychopathology and Behavioral Assessment*. 39:140–146. [PubMed: 28260835]
- Schut AJ, Castonguay LG, Borkovec TD. 2001; Compulsive checking behaviors in generalized anxiety disorder. *Journal of Clinical Psychology*. 57:705–715. [PubMed: 11344459]
- Spitzer RL, Kroenke K, Williams JB, Löwe B. 2006; A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine*. 166:1092–1097. [PubMed: 16717171]
- Starcevic V, Portman ME, Beck AT. 2012; Generalized anxiety disorder: Between neglect and epidemic. *Journal of Nervous and Mental Disease*. 200:664–667. [PubMed: 22850300]
- Steiger JH. 1980; Tests for comparing elements of a correlation matrix. *Psychological Bulletin*. 87:245–251.
- Townsend MH, Weissbecker KA, Barbee JG, Peña JM, Snider LM, Tynes LL, Tynes SF, Boudoin C, Green-Liebovitz MI, Winstead D. 1999; Compulsive behaviors in generalized anxiety disorder and obsessive-compulsive disorder. *Journal of Nervous and Mental Disease*. 187:697–699. [PubMed: 10579600]

Williams JBW, Gibbon M, First MB, Spitzer RL, Davies M, Borus J, Howes MJ, Kane J, Pope HG, Rounsaville B, Wittchen H. 1992; The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite test-retest reliability. *Archives of General Psychiatry*. 49:630–636. [PubMed: 1637253]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Interrater Reliability of Generalized Anxiety Disorder and Worry Behaviors

	κ	% agreement	ICC
Generalized Anxiety Disorder			
<i>DSM-5</i> diagnosis	.63	82.0	
ADIS-5 clinical severity rating			.62
Worry Behaviors			
Avoidance	.61	82.0	.67
Over-preparation	.43	72.0	.39
Procrastination	.52	76.0	.52
Reassurance seeking	.32	66.0	.34
Composite score			.53
Proposed <i>DSM-5</i> criterion	.44	78.0	

Note. $n = 50$, ADIS-5 = Anxiety and Related Disorders Interview Schedule for *DSM-5* κ = kappa, ICC = intraclass correlation correlation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2
Frequencies (and Percentages) of Worry Behaviors and the Proposed DSM-5 Worry Behavior Criterion Across DSM-5 Diagnostic Groups

Worry Behaviors	Diagnostic Group				OR (GAD-OSAD)
	GAD (n = 366)	No GAD (n = 434)	OSAD (n = 78)	OR (GAD-No GAD)	
Avoidance	182 (49.7)	68 (15.7)	28 (35.9)	5.32 ^{***}	1.77 [*]
Over-preparation	174 (47.5)	63 (14.5)	23 (29.5)	5.34 ^{***}	2.17 ^{**}
Procrastination	246 (67.2)	119 (27.4)	40 (51.3)	5.43 ^{***}	1.95 ^{**}
Reassurance seeking	210 (57.4)	93 (21.4)	29 (37.2)	4.94 ^{***}	2.27 ^{***}
Proposed DSM-5 criterion	339 (92.6)	204 (47.0)	62 (79.5)	14.16 ^{***}	3.24 ^{***}

Note. GAD = generalized anxiety disorder, No GAD = patients with diagnoses other than GAD, OSAD = other specific anxiety disorder cases below the DSM-5 GAD diagnostic threshold; OR = odds ratio; the No GAD and OSAD groups are not mutually exclusive (see Footnote 2).

* $p < .05$

** $p < .01$

*** $p < .001$

Table 3

Final Step of Hierarchical Logistic Regressions Predicting GAD Diagnostic Status From GAD Symptom Features and Worry Behaviors

Predictor	B	SE _B	OR	LRT	R ² _L
Proposed DSM-5 Criterion					
Excessive worry	0.22	0.02	1.24	167.91	.15
Associated symptoms	0.10	0.02	1.11	49.16	.04
Proposed DSM-5 criterion	0.99	0.29	2.70	11.84	.01
Worry Behaviors					
Excessive worry	0.23	0.02	1.25	196.49	.18
Associated symptoms	0.11	0.02	1.11	52.39	.05
Avoidance	0.46	0.23	1.58	4.01	.00
Excessive worry	0.23	0.02	1.25	184.04	.17
Associated symptoms	0.11	0.01	1.12	59.69	.05
Over-preparation	0.88	0.24	2.41	14.02	.01
Excessive worry	0.23	0.02	1.25	188.47	.17
Associated symptoms	0.11	0.02	1.12	55.42	.05
Procrastination	0.39	0.22	1.48	3.17	.00
Excessive worry	0.22	0.02	1.25	184.61	.17
Associated symptoms	0.11	0.02	1.12	57.75	.05
Reassurance seeking	0.64	0.22	1.89	8.41	.01

Note. GAD = generalized anxiety disorder, OR = odds ratio, LRT = likelihood ratio test (critical values: 3.84, 6.64, and 10.83 for $\alpha = .05, .01, .001$, respectively), R^2_L = proportion of the model deviance ($-2LL$) uniquely reduced by the predictor.