

Psychometric Properties of the Worry Behaviors Inventory: Replication and Extension in a Large Clinical and Community Sample

Alison E. J. Mahoney*

*Clinical Research Unit for Anxiety and Depression, University of New South Wales at St Vincent's
Hospital, Sydney, NSW, Australia*

Megan J. Hobbs

*Clinical Research Unit for Anxiety and Depression, University of New South Wales at St Vincent's
Hospital, Sydney, NSW, Australia*

Jill M. Newby

Department of Psychology, University of New South Wales, Sydney, NSW, Australia

Alishia D. Williams

Department of Clinical Psychology, Utrecht University, The Netherlands

Gavin Andrews

*Clinical Research Unit for Anxiety and Depression, University of New South Wales at St Vincent's
Hospital, Sydney, NSW, Australia*

Background: The use of maladaptive behaviors by individuals with generalized anxiety disorder (GAD) is theoretically important and clinically meaningful. However, little is known about the specificity of avoidant behaviors to GAD and how these behaviors can be reliably assessed. **Aims:** This study replicated and extended the psychometric evaluation of the Worry Behaviors Inventory (WBI), a brief self-report measure of avoidant behaviors associated with GAD. **Method:** The WBI was administered to a hospital-based sample of adults seeking

* Correspondence to Dr Alison Mahoney, Clinical Research Unit for Anxiety and Depression, University of New South Wales at St Vincent's Hospital, Level 4 O'Brien Centre, 394–404 Victoria Street, Darlinghurst, NSW 2010, Australia. E-mail: Alison.Mahoney@svha.org.au

treatment for symptoms of anxiety and/or depression ($n = 639$) and to a community sample ($n = 55$). Participants completed measures of symptom severity (GAD, depression, panic disorder, health anxiety, and personality disorder), and measures of checking, reassurance-seeking and behavioral inhibition. Analyses evaluated the factor structure, convergent, divergent, incremental, and discriminant validity, as well the temporal stability and treatment sensitivity of the WBI. **Results:** The two-factor structure found in the preliminary psychometric evaluation of the WBI was replicated. The WBI was sensitive to changes across treatment and correlated well with measures of GAD symptom severity and maladaptive behaviors. The WBI was more strongly related to GAD symptom severity than other disorders. The WBI discriminated between clinical and community samples. **Conclusions:** The WBI provides clinicians and researchers with a brief, clinically meaningful index of problematic behaviors that may guide treatment decisions and contribute to our understanding of maintaining factors in GAD.

Keywords: generalized anxiety disorder, avoidance, behavior, psychometric

Introduction

Generalized anxiety disorder (GAD) involves excessive anxiety and worry about everyday concerns such as work, family and finances [American Psychiatric Association (APA), 2013]. To date, much of the research concerning GAD has focused on identifying the cognitive and somatic features of the disorder (Andrews et al., 2016). The extent to which maladaptive behaviors are relevant to the GAD classification is unclear. Behaviors are used to classify most of the other anxiety disorders in DSM-5 and were proposed for the DSM-5 classification of GAD (Andrews et al., 2010; APA, 2013). The proposed behaviors included excessive situational avoidance, preparatory activities, procrastination, and reassurance seeking. Although empirical evaluations demonstrate that individuals with GAD engage in these maladaptive behaviors (e.g. Beesdo-Baum et al., 2012; Coleman et al., 2011; Schut et al., 2001), the measurement and specificity of behavioral avoidance associated with GAD has been questioned (Starcevic et al., 2012).

Theoretical models of GAD and pathological worry propose that cognitive avoidance maintains the disorder (Borkovec et al., 2004; Dugas et al., 1998; Wells, 1999). Cognitive avoidance involves avoiding or reducing distressing internal experiences (e.g. thoughts and emotions) via mental strategies such as thought suppression and distraction (Sexton and Dugas, 2008). However, behavioral avoidance, i.e. avoiding or reducing distress by keeping away from distressing situations and activities, may be related to cognitive avoidance and may also play a significant role in the maintenance of pathological worry. While reducing maladaptive behaviors is recommended in contemporary cognitive treatment protocols for GAD, e.g. meta-cognitive therapy (MCT; Wells, 1995, 1999) and intolerance of uncertainty therapy (IUT; Robichaud, 2013), the relationships between pathological worry, anxiety, cognitive avoidance and behavioral avoidance require ongoing scrutiny. For instance, Beesdo-Baum et al. (2012) found that higher post-treatment cognitive and behavioral avoidance predicted greater worry at 6 and 12 months follow-up. An additional study found that while cognitive avoidance predicted subsequent rumination, worry and sadness, behavioral avoidance was only predictive of subsequent anxiety (Dickson et al., 2012). Further research is therefore needed to clarify the role of maladaptive behaviors in perpetuating GAD.

Before important diagnostic, theoretical and clinical questions can be addressed, a reliable and valid measure of maladaptive behaviors associated with GAD needs to be established. The limitations of existing scales led to the development of the Worry Behaviors Inventory (WBI; Mahoney et al., 2016), which was designed to provide a clinically meaningful index of maladaptive behaviors that are associated with GAD. Items were derived with reference to GAD diagnostic criteria, cognitive models of GAD, and the extant literature on maladaptive behaviors associated with GAD. The preliminary psychometric properties of the scale were examined in a large sample of adults seeking internet-delivered cognitive behavior therapy (iCBT) for their symptoms of anxiety and depression. The 10-item measure consisted of two related factors, labelled *Safety Behaviors* (e.g. items assessed checking, watching, planning, reassurance seeking and controlling others) and *Avoidance* (e.g. items assessed avoidance of decision making, worrying situations, people and activities). Given that the former study was the first to examine the psychometric properties of the WBI, additional studies are needed replicate its findings. Four facets of that study need to be replicated and/or extended. First, additional structural analyses are needed to examine the stability of the two-factor model. Given that face-to-face clinical settings continue to be the most common method of providing psychological therapies, further structural evaluations of the WBI among samples consisting of patients undertaking face-to-face psychotherapy would complement the existing data gathered from online treatment-seekers.

Second, in the initial psychometric evaluation, the WBI correlated positively with GAD symptom severity, thereby providing support for the convergent validity of the scale. The WBI also correlated with measures of depression, panic disorder and social phobia symptom severity. Interestingly, the *Safety Behaviors* scale was more strongly related to GAD symptoms than symptoms of other disorders, while the *Avoidance* scale was as strongly related to GAD symptom severity as to depression, panic disorder and social phobia symptom severity. These associations require replication and extension given concerns about the specificity of maladaptive behaviors associated with GAD. It has been argued that the maladaptive behaviors that have been associated with GAD are also likely to characterize a number of conditions such as other anxiety disorders, depressive disorders and personality disorders (Starcevic et al., 2012). For example, situational avoidance forms part of the classification for panic disorder and avoidant personality disorder, and excessive reassurance seeking has been frequently observed in individuals with severe health anxiety (APA, 2013; Salkovskis and Warwick, 1986).

Third, convergent validity analyses should not be limited to examining associations between the WBI and symptom measures. Convergent validity analyses should be extended to examine factors that are theorized to maintain symptoms, for instance evaluating the associations between the WBI and existing measures of behavioral avoidance and behavioral inhibition. Additionally, the incremental validity of the WBI needs to be assessed in order to determine whether the WBI predicts unique variance in GAD symptom severity over and above existing measures of maladaptive behaviors and behavioral inhibition.

Finally, Mahoney et al. (2016) found that WBI scores discriminated between patients with and without a probable diagnosis of GAD. These analyses should be extended to examine the degree to which WBI scores discriminate between patients and community participants. These data are critical for identifying the clinical threshold of maladaptive behaviors, i.e. at what point the use of maladaptive behaviors is likely to be pathological. Furthermore, the temporal stability and treatment sensitivity of the WBI is currently unknown. These indices

are all important in clinical contexts where patients are screened and monitored over time and following interventions.

In order to establish the utility of the WBI across multiple clinical settings this study sought to replicate and extend the psychometric profile of the WBI in a large hospital-based out-patient sample and in a community sample. We predicted that the two-factor structure of the WBI would demonstrate adequate fit in the clinical sample. Consistent with previous findings, we predicted that the WBI subscales would demonstrate differential associations with measures of symptom severity. We expected that the *Safety Behaviors* subscale would be more strongly related to GAD symptom severity compared with symptoms of other disorders including depression, panic disorder, health anxiety and personality disorders. Alternatively, we predicted that the *Avoidance* subscale would be as strongly correlated with GAD symptom severity as symptoms of the other disorders. Investigating the incremental validity of the WBI, we anticipated that the scale would uniquely predict GAD symptom severity over and above measures of behavioral inhibition and checking behaviors. Lastly, we expected that the WBI would be sensitive to change such that WBI scores would be significantly reduced by cognitive behavior therapy.

Method

Participants

Clinical participants. 639 consecutive patients, referred to the Anxiety Disorders Clinic at St Vincent's Hospital, Sydney between 16 September 2013 and 30 December 2014 by their general practitioner, completed the WBI as part of their standardized intake assessment. On average, patients were 35.71 years of age ($SD = 12.38$, range = 18–80 years) and 56.5% were female. Most patients were born in Australia (79.8%) and described themselves as single/never married (54.0%) or married/*de facto* (36.6%), working full (36.5%) or part-time (24.1%), and completed a university degree (47.9%). 96.9% of patients completed modules of the Mini-Composite International Diagnostic Interview (Mini-CIDI; Sunderland et al., 2012), which assessed for GAD, social phobia, panic disorder, agoraphobia, major depressive disorder (MDD) and dysthymic disorder (further details are provided below). Of these, 16.3% did not meet diagnostic criteria for any of the assessed disorders, 41.0% met criteria for one disorder, 25.2% for two disorders, 11.5% for three disorders, 4.0% for four disorders, and 1.6% for five disorders. Proportions meeting criteria for each diagnosis were: GAD = 40.1%, social phobia = 39.7%, panic disorder = 5.85%, agoraphobia = 17.1%, MDD = 24.9%, and dysthymic disorder = 22.5%.

A subset of the clinical sample [$n = 107$, 57.9% female, mean age (SD) = 35.63 (12.17) years] completed the WBI prior to treatment commencing and formed the sample that was used to examine the temporal stability of the scale. A proportion of this subsample [$n = 48$, 54.2% female, mean age (SD) = 39.81 (15.35) years] completed the WBI a third time after completing an online CBT course that formed part of their treatment at the hospital.

Community participants. An unselected sample of 55 adult participants was recruited from the community via advertisements on local community noticeboards. On average, community participants were 44.37 years of age ($SD = 15.1$, range = 26–70 years). Most were female (70.9%), born in Australia (76.4%), and working full (42.2%) or part-time (31.5%).

Measures

Patients completed the Mini-CIDI, Generalized Anxiety Disorder 7-item (GAD-7), Penn State Worry Questionnaire-Ultra Brief Version (PSWQ-3), Patient Health Questionnaire-9, Panic Disorder Severity Scale Self-Report Version, Whiteley Index-7, Self-Report Standardized Assessment of Personality-Abbreviated Scale; WBI, Obsessive Compulsive Inventory-Revised: Checking Subscale, and Behavioral Inhibition System Scale.

Community participants completed the GAD-7, PSWQ-3, WBI and Threat-related Reassurance Seeking Scale.

Mini-Composite International Diagnostic Interview. The Mini-CIDI (Sunderland et al., 2012) is a computerized, self-report, short-form version of the World Mental Health Composite International Diagnostic Interview (WMH-CIDI 3.0), which is a structured interview for DSM-IV common mental disorders (Kessler et al., 2004). The Mini-CIDI has excellent concordance with the CIDI 3.0 for each of the disorders that were assessed in this study (Sunderland et al., 2012).

Measures of symptom severity

Generalized Anxiety Disorder 7-item. The GAD-7 is a seven-item self-report screener for a probable diagnosis of GAD and of GAD symptom severity over the past two weeks (Spitzer et al., 2006). A total score ≥ 10 indicates a probable GAD diagnosis (sensitivity = 89%, specificity = 82%; Spitzer et al., 2006). Studies support a unidimensional structure, reliability ($\alpha = .92$, $r = .83$ over one week) and validity (e.g. convergent/divergent validity with the measures of anxiety, depression, self-esteem; criterion validity with respect to diagnosis) of the GAD-7 (Löwe et al., 2008; Spitzer et al., 2006). Internal consistency in the current samples was $\alpha = .89$ (clinical sample) and $\alpha = .74$ (community sample).

Penn State Worry Questionnaire-Ultra Brief Version. The PSWQ-3 is a three-item measure of trait worry (Berle et al., 2011). Evidence of internal consistency ($\alpha = .85$) and validity (e.g. convergent/divergent validity via correlations with measures of anxiety, depression, social anxiety; treatment sensitivity; discriminant validity) has been reported (Berle et al., 2011; Kertz et al., 2014). Current internal consistency was $\alpha = .82$ (clinical sample) and $\alpha = .78$ (community sample).

Patient Health Questionnaire-9. The PHQ-9 is a nine-item self-report screener for probable MDD in the past two weeks (Kroenke et al., 2001). Evidence of a one- or two-factor structure, reliability ($\alpha = .86$, $r = .84$ over 48 h), and validity (including convergent/divergent validity via correlations with measures of depression, health, disability; criterion validity via sensitivity/specificity with respect to diagnosis; treatment sensitivity) has been provided (Beard et al., 2016; Kroenke et al., 2010). Current $\alpha = .87$.

Panic Disorder Severity Scale Self-Report Version. The seven-item PDSS-SR measures symptoms of panic disorder in the past week (Shear et al., 1997). Evidence supports the reliability ($\alpha = .92$, ICC over 2 days = .81), validity (e.g. correlations with the interviewer-administered PDSS; treatment sensitivity), and single-factor structure of the measure (Houck et al., 2002; Shear et al., 2001). Current $\alpha = .91$.

Whiteley Index-7. The seven-item WI-7 is a self-report screener for somatization disorder and health anxiety with analyses indicating a one-dimensional structure (Fink et al., 1999). Evidence of internal consistency ($\alpha = .68$), convergent validity (e.g. correlations with measures of somatization symptoms) and discriminant validity (e.g. distinguishes between those with and without somatoform disorders) has been provided (Conradt et al., 2006; Fink et al., 1999). Current $\alpha = .79$.

Self-Report Standardised Assessment of Personality-Abbreviated Scale. The SAPAS-SR is an eight-item, self-report screener for personality disorders with total score ≥ 4 indicating a probable diagnosis (Germans et al., 2008). Evidence of reliability ($r = .89$ over 2 weeks, and $\alpha = .45$ which is thought to appropriately reflect the heterogeneous construct of 'personality disorder'), criterion validity (e.g. sensitivity/specificity with respect to diagnosis) and predictive validity (predicts treatment outcome) has been provided (Germans et al., 2008; Goddard et al., 2015). Current $\alpha = .63$.

Measures of maladaptive behaviors and behavioral inhibition

Worry Behaviors Inventory. The WBI is a 10-item, self-report measure that assesses how often respondents typically use maladaptive behaviors to control or prevent worry about everyday concerns (see Appendix). Evidence of the internal consistency (total $\alpha = .86$, *Safety Behaviors* subscale $\alpha = .85$, *Avoidance* subscale $\alpha = .75$) and validity (e.g. convergent/divergent validity and discriminant validity) has been provided (Mahoney et al., 2016).

Obsessive Compulsive Inventory-Revised: Checking Subscale. The three-item checking subscale of the OCI-R assesses excessive checking behaviors over the preceding month (Foa et al., 2002). Evaluations support the reliability ($\alpha = .83$; $r = .74$ over 1–2 weeks) and validity (e.g. convergent validity with measures of checking; divergent validity with indices of depression; discriminant validity via distinguishing between those with and without OCD) of the scale (Abramowitz and Deacon, 2006; Foa et al., 2002). Current $\alpha = .84$.

Threat-related Reassurance Seeking Scale. The TRSS is an eight-item transdiagnostic measure of reassurance seeking (Cougle et al., 2012). Analyses support the two-factor structure, reliability ($\alpha = .85$ and $.86$, and $r = .80$ and $.79$ over one month for the General and Evaluative Threat scales, respectively) and convergent validity (e.g. correlations with measures of reassurance seeking and anxiety symptoms) of the TRSS (Cougle et al., 2012). Internal consistency in the community sample was $\alpha = .89$ and $.93$ (General and Evaluative Threats scales, respectively).

Behavioral Inhibition System Scale. The BIS is a seven-item self-report measure of temperamental harm avoidance, sensitivity to threat, and proneness to anxiety (Carver and White, 1994). Evaluations support the unidimensional structure, reliability ($\alpha = .74$, $r = .66$ over 8 weeks) and validity (e.g. convergent/divergent validity with measures of harm avoidance, negative affectivity and novelty seeking; criterion validity by predicting nervousness during experimental tasks) of the BIS (Carver and White, 1994). Current $\alpha = .61$.

Procedure

Prior to their initial in-person assessment at the out-patient hospital clinic, patients completed the following self-report assessments via the internet: Mini-CIDI, GAD-7, PSWQ-3, PHQ-9, PDSS-SR, WI-7, SAPAS-SR, WBI, OCI-R-Checking, BIS and WHODAS-II.

A subset of patients ($n = 107$) completed the WBI a second time prior to commencing an online CBT programme (2–4 weeks after initial administration) as part of their treatment at the clinic. This sample provided an index of the test–retest reliability of the WBI. A portion of this subsample ($n = 48$) completed the WBI a third time after completing their online CBT programme, thereby providing an index of treatment sensitivity. The online cognitive behavior therapy (iCBT) programmes contained six lessons and were completed over 12 weeks. In the test–retest reliability sample, 35 participants (32.7%) enrolled in the GAD iCBT course and 72 (67.3%) participants enrolled in the iCBT course for GAD and depression. In the treatment sensitivity sample, 15 participants (31%) completed the GAD iCBT course and 33 participants (69%) completed the course for GAD and depression. Both iCBT courses are efficacious and effective, and have been described in detail elsewhere (Newby et al., 2013, 2014; Mewton et al., 2012; Robinson et al., 2010). In brief, the programmes involve (a) psycho-education about the nature of GAD, (b) arousal reduction skills, (c) cognitive restructuring to shift unhelpful thought patterns (including those associated with cognitive distortions, intolerance of uncertainty and metacognition), (d) graded exposure and behavioral experiments to reduce cognitive avoidance and maladaptive behaviors, and (e) relapse prevention. The two iCBT courses lead to comparable effect size reductions in GAD symptom severity, functional impairment and psychological distress (Newby et al., 2017).

All patients in the clinical sample were informed and provided electronic informed consent that their data would be collected, their pooled data analysed and published in scientific journals. Patients could opt out of the use of their data for these purposes via email with no impact on their receipt of treatment.

Community participants completed the WBI, GAD-7, PSWQ-3 and TRSS online. The human ethics committee at St Vincent's Hospital, Sydney approved all study procedures involving community participants (LNR/13/SVH/325).

Analyses

Factor structure. Confirmatory factor analysis (CFA) examined the model fit of the two-factor model in the clinical data. This model was examined in the MPlus v5.12 software package (Muthén and Muthén, 1998–2009) using a weighted least squares mean and variance estimator that is suitable for ordinal variables. Model fit was assessed with reference to the comparative fit index (CFI), Tucker–Lewis index (TLI), and the root mean square error of approximation (RMSEA). Hu and Bentler (1998) suggest that CFI and TLI values $> .90$ indicate acceptable fit and values $\geq .95$ indicate good fit, while RMSEA values $< .05$ indicate good fit. MacCallum et al. (1996) recommend that RMSEA values between .08 and .10 indicate mediocre fit, and Browne and Cudeck (1993) propose that RMSEA values $> .10$ indicate poor fit.

Changes in WBI scores over time. Paired t tests and Pearson correlation coefficients were used to assess the temporal stability of the WBI in the clinical subset that completed the WBI on two occasions. Paired sample t -tests with associated Cohen's d effect sizes were used to

estimate the treatment sensitivity of the WBI in the clinical subset that completed the WBI following their iCBT programme.

Validity. Pearson correlation coefficients between WBI scores and symptoms of GAD (GAD-7, PSWQ-3), and measures of maladaptive behaviors and behavioral inhibition (BIS, OCI-R-Checking, TRSS) were estimated to provide indices of convergent validity. Correlations between WBI scores and measures of related but theoretically distinct constructs (PHQ-9, PDSS-SR, WI-7 and SAPAS-SR) provided indices of divergent validity. Steiger's z was used to examine differences between the magnitudes of overlapping correlation coefficients to assess whether the WBI had stronger associations with measures of GAD symptom severity than with symptoms of other disorders (Diedenhofen and Musch, 2015; Steiger, 1980).

Independent samples t tests and receiver operating characteristic (ROC) analyses were used to indicate the extent to which the WBI scores discriminated between community participants and clinical participants with and without a probable GAD diagnosis (defined as individuals with a diagnosis of GAD according to the Mini-CIDI and scoring ≥ 10 on the GAD-7). Analyses also compared WBI scores between patients with and without with a probable GAD diagnosis.

Finally, to examine the incremental validity of the WBI in the clinical sample, a hierarchical regression was conducted to determine if the WBI demonstrated a unique relationship with GAD symptom severity over and above behavioral inhibition and checking behaviors. In this regression, the GAD-7 was the dependent variable. The BIS and OCI-R-Checking subscale were entered in Step 1, and the WBI was entered in Step 2.

Results

Symptom measure descriptives

In the clinical sample, mean scores for the symptom measures were GAD-7 = 12.49 (5.36), PSWQ-3 = 11.61 (2.83), PHQ-9 = 13.37 (6.67), PDSS-SR = 11.10 (6.59), WI-7 = 3.72 (2.24) and SAPAS-SR = 3.80 (2.03). The community sample's mean scores were GAD-7 = 2.98 (2.32), PSWQ-3 = 4.98 (2.21), TRSS-General Threat = 10.22 (5.42) and TRSS-Evaluative Threat = 10.11 (6.19). On all measures, the clinical sample's means fell within clinical ranges while community means fell within non-clinical ranges (Berle et al., 2011; Cogle et al., 2012; Fink et al., 1999; Germans et al., 2008; Shear et al., 2001; Spitzer et al., 2006; Wittkamp et al., 2007).

Factor structure

The two-factor structure was examined in the clinical sample. CFI and TLI indicated acceptable to good fit, while RMSEA did not (CFI = .93, TLI = .96 and RMSEA = .11).

Reliability

Internal consistency. In the clinical sample, the internal consistency of the WBI was good ($\alpha = .83, .82$ and $.70$ for the total, *Safety Behaviors* and *Avoidance* subscales, respectively). In the community sample $\alpha = .81, .77$ and $.67$ for the total, *Safety Behaviors* and *Avoidance* subscales, respectively. In the clinical sample, the WBI total correlated strongly with the *Safety*

Table 1. Correlations between self-report measures completed by the clinical sample

	GAD-7	PSWQ-3	OCI-R-Checking	PDSS	WI-7	PHQ-9	BIS
GAD-7	1.0						
PSWQ-3	.67	1.0					
OCI-R-Checking	.27	.30	1.0				
PDSS	.55	.45	.21	1.0			
WI-7	.35	.29	.27	.30	1.0		
PHQ-9	.65	.53	.22	.47	.29	1.0	
BIS	.33	.37	.14	.17	.09	.31	1.0
SAPAS-SR	.25	.28	.19	.21	.22	.29	.18

GAD-7, Generalized Anxiety Disorder 7-item scale; PSWQ-3, Penn State Worry Questionnaire-3 item version; OCI-R, Obsessive Compulsive Inventory-Revised; PDSS, Panic Disorder Severity Scale; WI-7, Whitley Index-7 item; PHQ-9, Patient Health Questionnaire-9; BIS, Behavioral Inhibition System scale; SAPAS-SR, Self-Report Standardized Assessment of Personality-Abbreviated Scale.

Behaviors subscale ($r = .96, p < .001$) and moderately with the *Avoidance* subscale ($r = .73, p < .001$) subscales. The subscales correlated moderately with each other ($r = .53, p < .001$). Similar correlations were found in the community sample.

Temporal stability and treatment sensitivity. For the test-retest sample, correlations between time 1 and 2 scores indicated temporal stability ($r = .89, .89$ and $.74$ for WBI total, *Safety Behaviors* and *Avoidance*, respectively, all p values $< .01$). WBI scores did not significantly differ from time 1 to time 2 (WBI total mean (SD) = 17.18 (8.03) vs 17.23 (7.52), $t(106) = .16, p = .88$; *Safety Behaviors* mean (SD) = 11.25 (6.35) vs 11.34 (5.75), $t(106) = .30, p = .89$; *Avoidance* mean (SD) = 5.93 (2.73) vs 5.90 (2.66), $t(106) = -.15, p = .88$).

Supporting the treatment sensitivity of the WBI, WBI scores significantly reduced from pre- to post-iCBT (WBI total mean (SD) = 18.77 (9.12) vs 12.58 (7.78), $t(47) = 6.77, p < .001, d = .73$; *Safety Behaviors* mean (SD) = 12.35 (6.90) vs 8.02 (5.81), $t(47) = 6.35, p < .001, d = .68$; *Avoidance* mean (SD) = 6.42 (2.89) vs 4.56 (2.61), $t(47) = 5.24, p < .001, d = .68$).

Validity

Convergent and divergent validity. Correlations between study measures are provided in Table 1 (all p values $< .01$, except $p < .05$ for the WI-7/BIS and SAPAS-SR/BIS correlations). Moderate, positive correlations were found between the WBI total score and measures of GAD symptoms, checking behaviors, and behavioral inhibition (see Table 2, r values = .32 to .51, all p values $< .001$). The WBI total was also significantly correlated to measures of anxiety, depression and personality pathology (r values = .18 to .46, all p values $< .001$; Table 2). Supporting the divergent validity of the WBI, tests of significant differences between the magnitude of overlapping correlations suggested that the WBI total score was more strongly correlated with the GAD-7 than the PDSS ($z = 3.61, p < .001$), PHQ-9 ($z = 3.08, p < .001$), WI-7 ($z = 4.72, p < .001$), and SAPAS-SR ($z = 4.85, p < .001$). This pattern of results was identical for the *Safety Behaviors* and *Avoidance* subscales with one exception; the *Avoidance* scale was as strongly related to the GAD-7 as it was to the PHQ-9 ($z = .00, p = 1.00$). This pattern of results was identical when the PSWQ-3 replaced the GAD-7 in these analyses.

Table 2. Correlations between WBI scores and validity measures in the clinical sample

	WBI total	<i>Safety Behaviors</i>	<i>Avoidance</i>
Convergent validity			
GAD-7	.48	.41	.46
PSWQ-3	.50	.42	.50
OCI-R-Checking	.51	.51	.32
BIS	.41	.34	.41
Divergent validity			
WI-7	.29	.29	.18
PDSS	.36	.31	.33
PHQ-9	.39	.30	.46
SAPAS-SR	.27	.23	.25

WBI, Worry Behaviours Inventory; GAD-7, Generalized Anxiety Disorder 7-item scale; PSWQ-3, Penn State Worry Questionnaire-3 item version; OCI-R-Checking, Obsessive Compulsive Inventory-Revised- Checking Subscale; BIS, Behavioral Inhibition System scale; WI-7, Whitely Index- 7 item; PDSS, Panic Disorder Severity Scale; PHQ-9, Patient Health Questionnaire-9; SAPAS-SR, Self-Report Standardized Assessment of Personality-Abbreviated Scale.

Table 3. Descriptive statistics for the WBI in the clinical and community samples

	n	WBI total score Mean (SD)	<i>Safety Behaviors</i> score Mean (SD)	<i>Avoidance</i> score Mean (SD)
Community	55	10.07 (5.99)	6.76 (4.39)	3.31 (2.37)
Clinical	639	19.03 (8.39)	12.38 (6.58)	6.65 (2.87)
With probable GAD	192	21.43 (8.10)	14.12 (6.57)	7.31 (2.71)
Without probable GAD	447	18.00 (8.31)	11.63 (6.45)	6.37 (2.90)

WBI, Worry Behaviors Inventory; *SD*, standard deviation; GAD, generalized anxiety disorder.

In the community sample, moderate correlations were found between the WBI total score and the GAD-7 ($r = .32, p < .01$), PSWQ-3 ($r = .49, p < .001$), TRSS-General Threat ($r = .65, p < .001$) and TRSS-Evaluative Threat ($r = .46, p < .001$). Similar results were found for the WBI subscales.

Discriminant validity. WBI scores were higher in the clinical sample compared with the community sample (see Table 3, total score $t(692) = 7.75, p < .001, d = 1.23$; *Safety Behaviors* $t(692) = 6.21, p < .001, d = 1.00$; *Avoidance* $t(692) = 8.39, p < .001, d = 1.27$). ROC analyses suggested that WBI scores discriminated between clinical and community participants {total WBI score area under the curve [AUC (95% CI)] = 0.80 (0.75–0.86); *Safety Behaviors* = 0.75 (0.70–0.81); *Avoidance* = 0.81 (0.75–0.86)}. The operating characteristics of the WBI are shown in Table 4. Sensitivity and specificity for the discrimination between the clinical and community samples were maximized at cut scores of 13 for the WBI total score, and 10 and 5 for the *Safety Behaviors* and *Avoidance* subscales, respectively.

Table 4. Operating characteristics between the WBI and a probable diagnosis of generalized anxiety disorder at varying cut-scores

Total WBI			Safety Behaviors			Avoidance		
Cut-score*	Sensitivity (%)	Specificity (%)	Cut-score*	Sensitivity (%)	Specificity (%)	Cut-score*	Sensitivity (%)	Specificity (%)
12	80	64	8	73	60	3	91	42
13	76	71	9	68	67	4	83	62
14	71	73	10	64	75	5	78	75
15	67	78	11	58	80	6	68	78

WBI, Worry Behaviors Inventory. *Cut-score: the actual WBI score is greater than or equal to the cut-score shown.

Table 5. Summary of hierarchical regression analysis for variables predicting GAD-7 scores ($n = 631$)

Predictor variables	ΔR^2	B	SE B	β	t	Part r
Step 1	.16***					
BIS		.57	.07	.29	7.90***	.29
OCI-R-Checking		.38	.06	.23	6.14***	.23
Step 2	.10***					
BIS		.30	.07	.15	4.01***	.14
OCI-R-Checking		.08	.07	.05	1.12	.04
WBI		.25	.03	.40	9.02***	.31

BIS, Behavioral Inhibition System scale; OCI-R-Checking, Obsessive Compulsive Inventory-Revised Checking subscale; WBI, Worry Behaviours Inventory; Δ , change. *** $p < .001$.

The clinical sample was divided into patients with and without a probable diagnosis of GAD (defined as a diagnosis of GAD according to the Mini-CIDI and scoring ≥ 10 on the GAD-7). Patients with a probable diagnosis of GAD had significantly higher WBI scores than those without GAD (see Table 3, total $t(637) = 4.82, p < .001, d = .42$; *Safety Behaviors* $t(637) = 4.45, p < .001, d = .38$; *Avoidance* $t(637) = 3.82, p < .001, d = .33$), but ROC analyses suggested that discrimination between these groups was sub-optimal [AUC (95% CI) = 0.61 (0.57–0.66) for total; 0.60 (0.56–0.65) for *Safety Behaviors*; 0.59 (0.54–0.64) for *Avoidance*].

Incremental validity. WBI total scores predicted symptoms of GAD over and above behavioral inhibition and checking behaviors (see Table 5). This pattern of results was identical when (i) WBI subscales were entered in Step 2 instead of the total score, and (ii) the PSWQ-3 replaced the GAD-7 as the criterion variable.

Discussion

Understanding how people with GAD engage in maladaptive behaviors is important from both a theoretical and clinical perspective because the use of such behaviors may contribute to the maintenance of the disorder. Before the role of maladaptive behaviors in maintaining GAD can be appropriately scrutinized, an empirically supported measure of these behaviors is needed.

The aim of this study was to replicate and extend Mahoney et al.'s (2016) initial psychometric evaluation of the WBI in order to provide clinicians and researchers with a reliable and valid index of the maladaptive behaviors that are associated with GAD.

Confirmatory factor analysis supported the two-factor structure of the WBI. Consistent with the initial structural examinations of the WBI (Mahoney et al., 2016), the CFI and TLI indicated acceptable to good fit, however, the RMSEA was slightly over the threshold of the *a priori* cut-point of 0.10. The internal consistency of the WBI was good in both the clinical and community samples, although the internal consistency of the *Avoidance* subscale was fair ($\alpha = .70$ and $.67$). This partly reflects the small number of items in the *Avoidance* scale, and future investigations could evaluate whether adding further items to the *Avoidance* scale from the original WBI item pool would enhance its internal consistency without weakening its validity. To complement the current findings of the traditional psychometric properties of the WBI, future research may also examine the psychometric properties of the WBI items within an item-response theory framework. Such analyses would provide additional details about the relationship between the behaviors that are associated with GAD, GAD symptom severity and the likelihood of engaging in each maladaptive behavior. Additionally, the small size of the community sample precluded examination of the factor structure in a non-clinical population. Nevertheless, the two-factor structure of the WBI has now been observed in two large clinical samples – one drawn from a face-to-face treatment clinic and the other from an online clinic (see Mahoney et al., 2016). These are the most common settings where individuals with GAD undertake CBT and are therefore most likely to benefit from the use of the WBI to identify maladaptive behaviors that can be addressed in treatment.

The WBI demonstrated temporal stability and treatment sensitivity. The online CBT programmes lead to moderate effect size reductions in WBI scores (d range = 0.68–0.73) which are commensurate with previously reported treatment outcomes for these programmes (Mewton et al., 2012; Newby et al., 2017). Although replication of our findings is needed, the current estimates of temporal stability and treatment sensitivity suggest that the WBI can be reliably used to identify problematic behaviors associated with GAD and to monitor patients' progress through treatment. It is currently unclear whether reductions in worry across treatment cause reductions in maladaptive behaviors or whether these relationships are more complex. Future research is now needed to understand the mechanisms that are responsible for the reduction of maladaptive behaviors during CBT and whether these reductions predict treatment outcomes.

The WBI discriminated well between the clinical and community samples. The WBI scores of the clinical participants were about twice as high as those of the community participants. Sensitivity and specificity for discriminating between these groups were maximized at cut scores of 13 for the total score, and 10 and 5 for the *Safety Behaviors* and *Avoidance* subscales, respectively. Clinicians can use these thresholds as a guide for identifying pathological levels of maladaptive behaviors when screening and assessing patients as well as identifying when patients are approaching non-clinical levels of maladaptive behaviors during treatment. Our findings supporting the incremental validity of the WBI also attest to the clinical utility of the scale. We found that the WBI predicted symptoms of GAD over and above existing measures of checking and behavioral inhibition. Compared with existing measures, the WBI indexes a wider range of clinically meaningful maladaptive behaviors that clinicians can use to inform their case formulation and treatment plan. The WBI is also brief, which is particularly important in settings where there is limited time to administer measures.

The current study found evidence in support of the convergent and divergent validity of the WBI. The WBI total and subscales correlated moderately with measures of GAD symptom severity, safety-seeking behaviors (reassurance seeking, checking behaviors) and behavioral inhibition. Although the WBI also significantly correlated with symptoms of other disorders (depression, panic disorder, health anxiety, and personality disorders), the WBI was invariably more strongly related to GAD symptom severity. The one exception was that the *Avoidance* subscale was as strongly related to the GAD-7 as it was to the PHQ-9. These results are similar to the findings of the initial psychometric evaluation of the WBI (Mahoney *et al.*, 2016). It is unsurprising that the WBI correlates with symptoms of multiple mental disorders because of the risk factors and clinical correlates that are shared by these disorders (e.g. Goldberg *et al.*, 2009). Nevertheless, our results suggest that the maladaptive behaviors assessed by the WBI demonstrate relative specificity to GAD symptoms compared with symptoms of other disorders. We also found that patients with a probable diagnosis of GAD reported significantly higher WBI scores than patients without a probable diagnosis of GAD. However, ROC analyses suggested that the WBI's ability to discriminate between these two groups was suboptimal. This again suggests that while the maladaptive behaviors indexed by the WBI may be most strongly related to GAD, these behaviors are not exclusively related to GAD but are associated with other anxiety disorders and depression.

Several implications for the conceptualization and classification of GAD can be drawn from this study. Supporting cognitive models of GAD, participants with a probable diagnosis of GAD reported using maladaptive behaviors to manage their worry more frequently than patients without probable GAD and community participants. Moreover, CBT significantly reduced the use of these behaviors. It is now important to understand how maladaptive behaviors interact with the more established maintaining factors of GAD such as cognitive avoidance, intolerance of uncertainty and metacognitive beliefs, and whether there is a causal relationship between maladaptive behaviors and symptoms of GAD. When behavioral criteria were considered for the DSM-5 classification of GAD, scant data were available for determining which behaviors were specific to GAD and at which threshold behaviors became pathological. This study found that the maladaptive behaviors indexed by the WBI discriminated well between patients and community participants, and while these behaviors were not specific to GAD, they were more strongly related to symptoms of GAD than symptoms of other disorders. That said, our findings must be interpreted within their context. This study relied exclusively on self-report data. We did not conduct clinician-rated structured diagnostic assessments and we did not assess all disorders where maladaptive behaviors associated with worry may be prevalent (e.g. post-traumatic stress disorder and separation anxiety disorder). As such, many questions remain about the diagnostic specificity of maladaptive behaviors in GAD. It also remains to be seen if including maladaptive behaviors in the GAD classification would enhance the reliability, validity and clinical utility of the diagnosis.

Conclusion

The WBI provides clinicians and researchers with a brief, self-report index of clinically meaningful maladaptive behaviors that are associated with GAD. The psychometric performance of the scale has now been evaluated in two large clinical samples with findings supporting the reliability and validity of the scale.

Acknowledgements

Financial support: A.M. is supported by an Australian Government Research Training Program Scholarship. J.N. is supported by a NHMRC Early Career Research Fellowship (1033787).

Ethical statement: All authors have abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the APA. Data from the clinical sample were collected within the Quality Assurance Activities of St Vincent's Hospital, Sydney Australia. All patients were informed and provided electronic informed consent that their data would be collected, their pooled data analysed and published in scientific journals. Patients could opt out of the use of their data for these purposes via email with no impact on their receipt of treatment. The human ethics committee at St Vincent's Hospital, Sydney approved all study procedures involving community participants (LNR/13/SVH/325).

Conflicts of interest: Alison Mahoney, Megan Hobbs, Jill Newby, Alishia Williams and Gavin Andrews have no conflicts of interest with respect to this publication.

References

- Andrews, G., Hobbs, M. J., Borkovec, T. D., Beesdo, K., Craske, M. G., Heimberg, R. G. et al. (2010). Generalised worry disorder: a review of DSM-IV generalised anxiety disorder and options for DSM-V. *Depression and Anxiety*, 27, 134–147.
- Andrews, G., Mahoney, A.E.J., Hobbs, M.J. and Genderson, M. (2016). *Treatment of Generalized Anxiety Disorder: Therapist Guides and Patient Manual*. Oxford: Oxford University Press.
- Abramowitz, J. S. and 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.
- American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual for Mental Disorders* (5th edition). Washington, DC.
- Beard, C., Hsu, K. J., Rifkin, L. S., Busch, A. B. and Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, 193, 267–273.
- Beesdo-Baum, K., Jenjahn, E., Hoffer, M., Lueken, U., Becker, E. S. and Hoyer, J. (2012). Avoidance, safety behavior, and reassurance-seeking in generalized anxiety disorder. *Depression and Anxiety*, 29, 948–957.
- Berle, D., Starcevic, V., Moses, K., Hannan, A., Milicevic, D. and Sammut, P. (2011). Preliminary validation of an ultra-brief version of the Penn State Worry Questionnaire. *Clinical Psychology and Psychotherapy*, 18, 339–346.
- Borkovec, T. D., Alcaine, O. M. and Behar, E. (2004). Avoidance theory of worry and generalized anxiety disorder. In R. G. Heimberg, C. L. Turk and D. Mennin (eds), *Generalized Anxiety Disorder: Advances in Research and Practice*. New York, NY: Guilford Press.
- Browne, M. W. and Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Long (eds), *Testing Structural Equation Models*, pp. 136–162. Beverly Hills, CA: Sage.
- Carver, C. S. and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319–333.
- Coleman, S. L., Pieterfesa, A. S., Holaway, R. M., Coles, M. E. and Heimberg, R. G. (2011). Content and correlates of checking related to symptoms of obsessive compulsive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders*, 25, 293–301.

- Conradt, M., Cavanagh, M., Franklin, J. and Rief, W. (2006). Dimensionality of the Whiteley Index: Assessment of hypochondriasis in an Australian sample of primary care patients. *Journal of Psychosomatic Research*, 60, 137–143.
- Cogle, J. R., Fitch, K. E., Finchman, F. D., Riccardi, C. J., Keough, M. E. and Timpano, K. R. (2012). Excessive reassurance seeking and anxiety pathology: tests of incremental associations and directionality. *Journal of Anxiety Disorders*, 26, 117–125.
- Dickson, K. S., Ciesla, J. A. and Reilly, L. C. (2012). Rumination, worry, cognitive avoidance, and behavioral avoidance: Examination of temporal effects. *Behavior Therapy*, 43, 629–640.
- Diedenhofen, B. and Musch, J. (2015). cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10, e0121945.
- Dugas, M. J., Gagnon, F., Ladouceur, R. and Freeston, M. H. (1998). Generalized anxiety disorder: a preliminary test of a conceptual model. *Behavior Research and Therapy*, 36, 215–226.
- Fink, P., Ewald, H., Jensen, J., Sorensen, L., Engberg, M., Holm, M. and Munk-Jorgensen, P. (1999). Screening for somatization and hypochondriasis in primary care and neurological in-patients: A seven-item scale for hypochondriasis and somatization. *Journal of Psychosomatic Research*, 46, 261–273.
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G. et al. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychological Assessment*, 14, 485–496.
- Germans, S., Van Heck, G. L., Moran, P. and Hodiament, P. G. (2008). The Self-report Standardized Assessment of Personality-abbreviated Scale: Preliminary results of a brief screening test for personality disorders. *Personality and Mental Health*, 2, 70–76.
- Goddard, E., Wingrove, J. and Moran, P. (2015). The impact of comorbid personality difficulties on response to IAPT treatment for depression and anxiety. *Behavior Research and Therapy*, 73, 1–7.
- Goldberg, D. P., Andrews, G., Krueger, R. F. and Hobbs, M. J. (2009). Emotional disorders: Cluster 4 of the proposed meta-structure for DSM-V. *Psychological Medicine*, 39, 2043–2059.
- Houck, P. R., Spiegel, D. A., Shear, M. K. and Rucci, P. (2002). Reliability of the self-report version of the panic disorder severity scale. *Depression and Anxiety*, 15, 183–185.
- Hu, L. T. and Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453.
- Kertz, S. J., Lee, J. and Björqvinnson, T. (2014). Psychometric properties of abbreviated and ultra-brief versions of the Penn State Worry Questionnaire. *Psychological Assessment*, 14, 1146–1154.
- Kessler, R. C., Abelson, J., Demler, O., Escobar, J. I., Gibbon, M. and Guyer, M. E. (2004). Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *International Journal of Methods in Psychiatric Research*, 3, 122–139.
- Kroenke, K., Spitzer, R., and Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16, 606–613.
- Kroenke, K., Spitzer, R., Williams, J. B. and Löwe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *General Hospital Psychiatry*, 32, 345–359.
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W. et al. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46, 266–274.
- MacCallum, R. C., Browne, M. W. and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1, 130–149.
- Mahoney, A., Hobbs, M. J., Newby, J. M., Williams, A. D., Sunderland, M. and Andrews, G. (2016). The Worry Behaviors Inventory: assessing the behavioral avoidance associated with generalized anxiety disorder. *Journal of Affective Disorders*, 203, 256–64.

- Mewton, L., Wong, N. and Andrews, G.** (2012). The effectiveness of internet cognitive behavioral therapy for generalized anxiety disorder in clinical practice. *Depression and Anxiety*, 29, 843–849.
- Muthén, L. K. and Muthén, B. O.** (1998–2009). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Newby, J. M., Mackenzie, A., Williams, A. D., McIntyre, K., Watts, S., Wong, N. et al.** (2013). Internet cognitive behavioral therapy for mixed anxiety and depression: a randomized controlled trial and evidence of effectiveness in primary care. *Psychological Medicine*, 43, 2635–2648.
- Newby, J. M., Mewton, L. and Andrews, G.** (2017). Transdiagnostic versus disorder-specific internet-delivered cognitive behavioral treatment for anxiety and depression in primary care. *Journal of Anxiety Disorders*, 46, 25–34.
- Newby, J. M., Mewton, L., Williams, A. D. and Andrews, G.** (2014). Effectiveness of transdiagnostic internet cognitive behavioral treatment for mixed anxiety and depression in primary care. *Journal of Affective Disorders*, 165, 45–52.
- Robichaud, M.** (2013). Cognitive behavior therapy targeting intolerance of uncertainty: application to a clinical case of generalized anxiety disorder. *Cognitive and Behavioral Practice*, 20, 251–263.
- Robinson, E., Titov, N., Andrews, G., McIntyre, K., Schwencke, G. and Solley, K.** (2010). Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS ONE*, 5, e10942.
- Salkovskis, P. M. and Warwick, H. M.** (1986). Morbid preoccupations, health anxiety and reassurance: a cognitive-behavioral approach to hypochondriasis. *Behavior Research and Therapy*, 24, 597–602.
- Schut, A. J., Castonguay, L. G. and Borkovec, T. D.** (2001). Compulsive checking behaviors in generalised anxiety disorder. *Journal of Clinical Psychology*, 57, 705–715.
- Sexton, K. A. and Dugas, M. J.** (2008). The Cognitive Avoidance Questionnaire: validation of the English translation. *Journal of Anxiety Disorders*, 22, 355–370.
- Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomka, D. E., Woods, S. W. et al.** (1997). Multicenter collaborative Panic Disorder Severity Scale. *American Journal of Psychiatry*, 154, 1571–1575.
- Shear, M. K., Rucci, P., Williams, J., Frank, E., Grochocinski, V., Vander Bilt, J. et al.** (2001). Reliability and validity of the Panic Disorder Severity Scale: replication and extension. *Journal of Psychiatric Research*, 35, 293–296.
- Spitzer, R. L., Kroenke, K., Williams, J. B. and Löwe, B.** (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166, 1092–1097.
- Starcevic, V., Portman, M. E. and Beck, A. T.** (2012). Generalized anxiety disorder: between neglect and an epidemic. *Journal of Nervous and Mental Disorders*, 200, 664–667.
- Steiger, J. H.** (1980). Tests for comparing elements of a correlation matrix. *Psychology Bulletin*, 87, 245–251.
- Sunderland, M., Slade, T. and Andrews, G.** (2012). Developing a short-form structured diagnostic interview for common mental disorders using signal detection theory. *International Journal of Methods in Psychiatric Research*, 21, 247–257.
- Wells, A.** (1995). Meta-cognition and worry: a cognitive model of generalized anxiety disorder. *Behavioral and Cognitive Psychotherapy*, 23, 301–320.
- Wells, A.** (1999). A metacognitive model and therapy for generalized anxiety disorder. *Clinical Psychology and Psychotherapy*, 6, 86–95.
- Wittkamp, K. A., Naeije, L., Schene, A. H., Huyser, J. and van Weert, H. C.** (2007). Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *General Hospital Psychiatry*, 29, 388–395.

Appendix: Worry Behaviors Inventory

People can worry about everyday things, like their family, finances, work, health, the future, or minor things like being on time. This questionnaire concerns the sorts of things people do to prevent, control, or avoid worrying. Please select the number that best describes what you do generally.

0 1 2 3 4
None of the time A little of the time Some of the time Most of the time All of the time

1 I try to control what other people do or think (e.g. stop others doing things, tell others to be careful, give advice, over-protect others, do things for others)	0	1	2	3	4
2 I make plans 'just in case'	0	1	2	3	4
3 I keep a close watch for anything bad that could happen	0	1	2	3	4
4 I avoid situations or people that worry me	0	1	2	3	4
5 I seek reassurance from sources of information (e.g. personal records, internet, reviews, books)	0	1	2	3	4
6 I check to make sure nothing bad has happened or that everything is OK	0	1	2	3	4
7 I delay making decisions about things or get other people to make decisions for me	0	1	2	3	4
8 I avoid saying or doing things that worry me	0	1	2	3	4
9 I over-plan activities (e.g. prepare for all possible bad outcomes, have a plan B, plan every step of an activity)	0	1	2	3	4
10 I repeatedly check that things have been done properly	0	1	2	3	4

Items 4, 7 and 8 form the *Avoidance* scale. Remaining items form the *Safety Behaviors* scale.