



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Behavior Therapy
and Experimental Psychiatry 34 (2003) 283–291

JOURNAL OF
behavior
therapy and
experimental
psychiatry

www.elsevier.com/locate/jbtep

Using the Penn State Worry Questionnaire to identify individuals with generalized anxiety disorder: a receiver operating characteristic analysis[☆]

David M. Fresco¹, Douglas S. Mennin², Richard G. Heimberg*,
Cynthia L. Turk³

*Department of Psychology, Adult Anxiety Clinic of Temple University, Weiss Hall, Temple University,
1701 North 13th Street, Philadelphia, PA 19122-6085, USA*

Received 1 June 2003; received in revised form 22 September 2003; accepted 23 September 2003

Abstract

We report on the use of the Penn State Worry Questionnaire (PSWQ) to identify individuals with generalized anxiety disorder (GAD). Fifty individuals with primary or secondary GAD and 114 individuals with social anxiety disorder (without GAD) completed the PSWQ. In receiver operating characteristic analyses, a score of 65 simultaneously optimized sensitivity and specificity in discriminating individuals with GAD from individuals with social anxiety disorder. Results support the use of the PSWQ in screening individuals likely to meet criteria for GAD who present for treatment at an anxiety disorders specialty clinic.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Generalized anxiety disorder; Worry; Sensitivity; Specificity; Assessment; Receiver operating characteristic analysis

[☆]Portions of this paper were presented at the annual meeting of the Association for Advancement of Behavior Therapy, Toronto, Ontario, Canada, November, 1999.

*Corresponding author. Tel.: +1-215-204-1575; fax: +1-215-204-5184.

E-mail address: heimberg@temple.edu (R.G. Heimberg).

¹Present address: Department of Psychology at Kent State University, Kent, OH, USA.

²Present address: Department of Psychology at Yale University, New Haven, CT, USA.

³Present address: Department of Psychology at La Salle University, Philadelphia, PA, USA.

1. Introduction

The central feature of DSM-IV generalized anxiety disorder (GAD) is chronic, excessive and uncontrollable worry (American Psychiatric Association, 1994). Although worry is common to all the anxiety disorders, it is present in greater amounts in GAD (Abel & Borkovec, 1995; Brown, Antony, & Barlow, 1992; Ladouceur, Blais, Freeston, & Dugas, 1998). Increasingly, studies have focused on the nature and function of worry, especially in the context of models of GAD (e.g., Borkovec, 1994; Dugas, Freeston, & Ladouceur, 1997; Wells & Carter, 1999).

The Penn State Worry Questionnaire (PSWQ) (Meyer, Miller, Metzger, & Borkovec, 1990) is the measure most frequently used to assess pathological worry in both clinical and non-clinical populations. The PSWQ is a 16-item inventory designed to capture the generality, excessiveness, and uncontrollability of pathological worry. It has been shown to have good internal consistency with samples consisting of older adults with GAD (Beck, Stanley, & Zebb, 1995), community subjects (Brown et al., 1992), and undergraduates (Meyer et al., 1990). It has also demonstrated good test–retest reliability over 8–10 weeks (Meyer et al., 1990).

The PSWQ is positively correlated with other self-report measures of worry (e.g., Beck et al., 1995; Davey, 1993; van Rijsoort, Emmelkamp and Vervaeke, 1999). Using the PSWQ to define groups of high, medium, and low worriers, high worriers generated more worrisome topics during an interview and reported spending more time worrying during the preceding week than did the other groups (Meyer et al., 1990). In another study, participants completed the PSWQ and asked three friends to do the same with instructions to make ratings regarding the target participant (e.g., “He/she is always worrying about something”). Self-ratings and aggregate peer ratings were significantly correlated (Stöber, 1998). As expected, the PSWQ has also been shown to discriminate patients with GAD from community controls and patients with other anxiety disorders (Brown et al., 1992). In summary, the reliability and validity of the PSWQ have been widely researched, and the instrument appears to have sound psychometric properties (see Molina & Borkovec, 1994, Turk, Heimberg, & Mennin, 2004, for more detailed reviews).

Theoretical developments in the field of GAD have given rise to the frequent use of screened samples to test initial predictions of new models of worry and GAD (e.g., Freeston, Ladouceur, Rheaume, & Leger, 1998; Lyonfields, Borkovec, & Thayer, 1995; Mennin, Heimberg, Turk, & Fresco, 2004). However, minimal research has been conducted to provide guidance in using the PSWQ to identify participants with self-reported GAD or to prescreen individuals likely to subsequently meet criteria for clinician-assessed GAD (but see, Behar, Alcaine, Zuellig, & Borkovec, 2003, for an exception). Two important criteria that speak to the ability of a scale to classify individuals into groups are sensitivity (the likelihood of having positive test results among individuals with a positive diagnosis) and specificity (the likelihood of having negative test results in individuals without the diagnosis) (Kraemer, 1992).

Receiver operating characteristic analysis (ROC) (Kraemer, 1992) allows researchers and clinicians to evaluate the ability of tests to discriminate individuals

with a characteristic from individuals without the characteristic. In ROC analysis, one obtains an ROC curve in which the sensitivity is plotted against the specificity for each value of the test. ROC analysis is a non-parametric test, and the most common way to index the probability that a test will correctly classify participants is with the area under the curve (AUC). The AUC is defined as the probability that a randomly selected case will score higher than a randomly selected control on the test variable. The AUC is a measure of the overlap in distributions between cases and controls and is indexed from 0 to 1. Values greater than 0.50 are interpreted as a probability greater than chance. One particular strength of ROC analysis is that the test is robust even when representation of cases and controls is unequal in the sample (Rice & Harris, 1995).

Behar et al. (2003) represents the first examination of the PSWQ using ROC analysis. In their first study, Behar et al. (2003) obtained PSWQ scores from 159 treatment-seeking participants with clinician-assessed GAD as well as 113 non-anxious controls. Although the AUC was not reported, the authors conclude that a PSWQ cut-score of 45 provided the best balance of sensitivity (0.99) and specificity (0.98). A strength of this first study was the use of clinician assessment to determine the presence of GAD as well as the absence of other current psychopathology. However, another important test of the PSWQ is to find GAD in a more heterogeneous sample that includes individuals likely to experience problems with other anxiety and mood disorders. Thus, in their Study 2, Behar et al. (2003) utilized a large sample of unselected college students ($N = 2449$) who completed the PSWQ as well as the Generalized Anxiety Disorders Questionnaire for DSM-IV (Newman et al., 2002) (to determine GAD caseness). They determined that a PSWQ cut-score of 62, provided the best balance of sensitivity (0.86) and specificity (0.75). Not surprisingly, in a sample that included participants who met criteria for self-reported diagnoses of post-traumatic stress disorder, social anxiety disorder, and/or moderate depression, the task of the PSWQ to detect GAD was made more difficult, yet it performed very well. However, one limitation of this study was that assessment of caseness of GAD, PTSD, social anxiety disorder, and a moderate level of depression was done solely with self-report measures. Further, the sample did not consist of participants seeking treatment for their emotional problems. Building upon the important findings of Behar et al. (2003), the present study used ROC analyses to examine the ability of the PSWQ to correctly identify patients with DSM-IV GAD in a sample of patients who sought treatment at an anxiety disorders specialty clinic.

2. Method

2.1. Participants

The sample of 164 participants consisted of 28 patients (19 women) who met criteria for GAD (without social anxiety disorder), 22 patients (10 women) with a principal diagnosis of social anxiety disorder who also met criteria for secondary diagnosis of GAD, and 114 patients (65 women) who met criteria for social anxiety

disorder (without GAD) by structured diagnostic interview. Analyses, discussed below, were conducted twice: (1) attempting to identify the 28 patients with primary GAD from a sample made up of these patients and the 114 patients with social anxiety disorder, and (2) attempting to identify the 50 patients with either primary GAD or GAD secondary to social anxiety disorder from a sample made up of these patients and the 114 patients with social anxiety disorder.

The three groups (GAD only, social anxiety disorder only, and GAD secondary to social anxiety disorder) did not differ significantly on any demographic characteristic. The final sample had a mean age of 33.5 years ($SD = 11$). The racial/ethnic composition of the sample was 78.9% White, 12.9% African American, 4.3% Asian, 3.1% Latino, and 1.2% Other. With regard to marital status, 68.4% of the sample was single; 15.6% were married; and 16% were divorced, separated, or widowed. The sample was highly educated, with 35.2% reported having completed college, another 21.1% with a post-graduate degree, 36.4% with some college, and 7.3% with a high school education or less.

2.2. Assessment

PSWQ (Meyer et al., 1990) is a 16-item inventory designed to assess the pathological worry characteristic of GAD.

The Anxiety Disorders Interview Schedule for DSM-IV: lifetime version (ADIS-IV-L; DiNardo, Brown, & Barlow, 1994) assesses current and lifetime anxiety disorders, unipolar and bipolar mood disorders, substance abuse and dependence, and disorders that are similar to the anxiety disorders either conceptually or in terms of presenting symptoms (e.g., hypochondriasis). Screening questions are provided for other major disorders (e.g., psychosis). In a sample of 362 patients, Brown, DiNardo, Lehman, and Campbell (2001) reported a kappa of 0.67 for a principal diagnosis of GAD and a kappa of 0.77 for a principal diagnosis of social anxiety disorder. Interviewers in the current study were either clinical psychologists or doctoral students in clinical psychology who had participated in the rigorous training program suggested by Brown et al. (2001). Specifically, trainees first studied the users' guide and observed at least three live interviews, which were conducted by a senior interviewer. While observing live interviews, the trainee made ratings and assigned diagnoses, which were later discussed with the senior interviewer. Thereafter the trainee conducted a minimum of three interviews with the senior interviewer in the room. The trainee conducted the interview, but the senior clinician observed and interjected questions as needed. The trainee and senior interviewer derived diagnoses independently. Before the trainee could act as a diagnostic interviewer, the trainee must have agreed with the senior interviewer on three consecutive interviews on the principal diagnosis and on the presence of all additional current and lifetime diagnoses. In a sample of 60 patients from the current study who were also interviewed by a second assessor who administered either the social anxiety disorder or GAD module of the ADIS-IV-L, there was 100% agreement with the original principal diagnosis ($\text{kappa} = 1.0$).

3. Procedure

Patients seeking treatment for worry or social anxiety reported to the clinic for initial assessment with the ADIS-IV-L. They were sent home with a packet of self-report instruments that included the PSWQ.

3.1. ROC analysis procedure

ROC analyses in the current study were conducted using STATA version 7 (StataCorp, 2001). Specifically, the continuous worry score was regressed onto the dichotomous GAD status variable using a logistic regression procedure. The predicted probabilities (a log odds transformation of the worry scores) from this analysis were then saved as a new variable. After solving the regression equation, the *lroc* command was implemented to derive the AUC and ROC curve. To test whether the AUC was significantly better than chance, the *brier* procedure was used. This command generates several decompositions of the Brier mean probability score based on a binary outcome (GAD Status) and a forecast variable (predicted probability of worry scores). Finally, optimal cut-scores were obtained using the *roctab* procedure, which generates sensitivity, specificity, and percent correctly classified for each value of the forecast variable. In each analysis, we elected to present three cut-scores: optimal sensitivity (the score that optimized sensitivity without reducing specificity to less than chance); optimal specificity (the score that optimized specificity without reducing sensitivity to less than chance); and optimal sensitivity and specificity (the score that produced the best balance of sensitivity and specificity).

4. Results

Table 1 displays the means and standard deviations for the PSWQ total score. GAD patients and comorbid patients scored significantly higher than patients with social anxiety disorder on the PSWQ total score. The two GAD groups did not differ from one another. Given the differences in variances between participants with and

Table 1
Means (standard deviations) on the penn state worry questionnaire (PSWQ) in the clinical samples

	GAD patients (<i>n</i> = 28)	Social anxiety patients (<i>n</i> = 114)	Comorbid patients (<i>n</i> = 22)	<i>F</i> (2,164)
PSWQ				
Total score	68.11 ^a (7.33)	56.35 ^b (14.84)	68.55 ^a (6.67)	14.68

Omnibus test is statistically significant ($p < 0.00001$); means with different superscripts are significantly different from each other as determined by Bonferroni-corrected pairwise *t*-tests. Comorbid patients received a primary diagnosis of social anxiety disorder and a secondary diagnosis of generalized anxiety disorder.

without GAD on PSWQ total score, analyses were repeated with a Kruskal–Wallis non-parametric test and follow-up Mann-Whitney Wilcoxon pairwise tests. The results were comparable to the univariate ANOVA.

4.1. ROC: patients with GAD and patients with social anxiety disorder

The first set of ROC analyses assessed the ability of the PSWQ total score to correctly identify 28 patients with primary GAD (without social anxiety disorder) from a sample that also included 114 social anxiety disorder patients without GAD. The analysis revealed a strong ROC curve for the PSWQ total score ($AUC = 0.74$, $p < 0.00001$) that was significantly better than chance in classifying individuals with GAD versus social anxiety disorder. See Fig. 1.

Cut scores optimizing sensitivity, optimizing specificity, and simultaneously optimizing both sensitivity and specificity were derived for values of the PSWQ total score. When optimizing sensitivity (cut score = 57) for the PSWQ total score, 25 of 28 GAD patients (89.29%) and 57 of 114 patients with social anxiety disorder (50%) were correctly classified. Overall, 57.75% of the sample was correctly classified. When optimizing specificity (cut score = 69), 14 of 28 GAD patients (50.00%) and 87 of 114 patients with social anxiety disorder (76.32%) was correctly classified. Overall, 71.13% of the sample was correctly classified. Finally, when sensitivity and specificity were both optimized (cut score = 65), 19 of 28 GAD patients (67.86%) and 71 of 114 patients with social anxiety disorder (62.28%) were correctly classified. Overall, 63.38% of the sample was correctly classified.

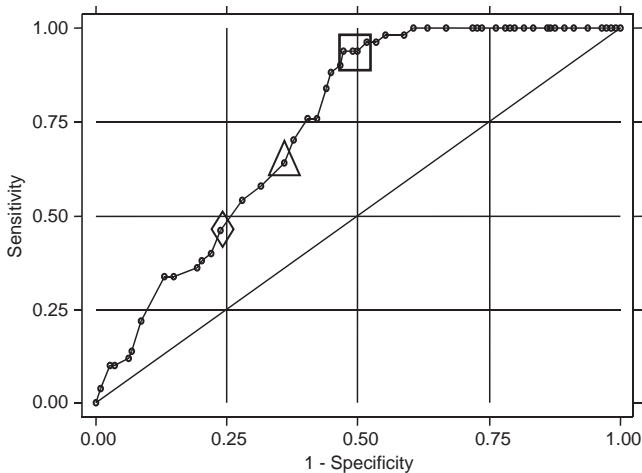


Fig. 1. Receiver operating characteristic analysis curve for the PSWQ total score in a sample of patients with GAD and social anxiety disorder. The open diamond denotes the point that optimizes specificity without dropping sensitivity below chance. The open triangle denotes the point that simultaneously optimizes sensitivity and specificity. The open square denotes the point that optimizes sensitivity without dropping specificity below chance.

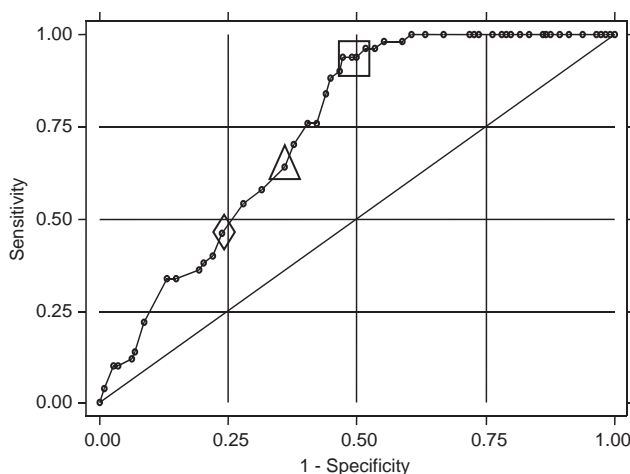


Fig. 2. Receiver operating characteristic analysis curve for the PSWQ total score in a sample of patients with primary or secondary GAD and social anxiety disorder. The open diamond denotes the point that optimizes specificity without dropping sensitivity below chance. The open triangle denotes the point that simultaneously optimizes sensitivity and specificity. The open square denotes the point that optimizes sensitivity without dropping specificity below chance.

4.2. ROC: patients with primary or secondary GAD and patients with social anxiety disorder

The second set of ROC analyses assessed the ability of the PSWQ total score to correctly identify the 50 patients with either primary or secondary GAD from a sample that also included 114 social anxiety disorder patients without GAD. Analyses again revealed strong ROC curves for the PSWQ total score ($AUC = 0.74$, $p < 0.00001$) that was significantly better than chance in classifying individuals with GAD versus social anxiety disorder. See Fig. 2.

Cut scores optimizing sensitivity, optimizing specificity, and simultaneously optimizing both sensitivity and specificity were derived for values of the PSWQ total score. When optimizing sensitivity (cut score = 57) for the PSWQ total score, 47 of 50 GAD patients (94%) and 57 of 114 patients with social anxiety disorder (50%) were correctly classified. Overall, 63.41% of the sample was correctly classified. When optimizing specificity (cut score = 68), 25 of 50 GAD patients (50.00%) and 82 of 114 patients with social anxiety disorder (71.93%) were correctly classified. Overall, 66.46% of the sample were correctly classified. Finally, when sensitivity and specificity were both optimized (cut score = 65), 32 of 50 GAD patients (67.86%) and 73 of 114 patients with social anxiety disorder (64.84%) were correctly classified. Overall, 63.41% of the sample was correctly classified.

5. Discussion

The PSWQ total score reliably distinguished GAD patients from individuals with social anxiety disorder, whether or not GAD was the primary or secondary

diagnosis. Results from the current study offer strong support for the use of the PSWQ for initial screening and identification of individuals with GAD who present for treatment at an anxiety disorders specialty clinic.

Both sets of ROC analyses, which used a sample of patients with a principal diagnosis of social anxiety disorder as the comparison group, offered a stringent test for the PSWQ given the high degree of overlap between the features, symptoms, and concerns of individuals with GAD and social anxiety disorder. By definition, social anxiety disorder is characterized by significant interpersonal concerns regarding negative evaluation by others. However, evidence is mounting that interpersonal difficulties are also an important concern among patients with GAD (Borkovec, Shadick, & Hopkins, 1991; Breitholtz, Westling, & Öst, 1998). Thus, it is impressive that the PSWQ total score performed as well as it did.

Findings from the current study complement those of Behar et al. (2003), particularly their second study where PSWQ scores were obtained in large student sample of participants who endorsed the presence of PTSD, social anxiety disorder, and moderate depression. However, one additional complementary study would be to obtain a large, non-treatment seeking sample with participants assessed with clinician interview to determine the ability of the PSWQ to find GAD among other anxiety and mood disorders.

ROC analysis allows one to derive optimal cut scores that can facilitate screenings or assessments based on the needs of the research or clinical endeavor. The cut scores presented herein, which optimized sensitivity, specificity, or both, were selected so that a variety screening scenarios could be satisfied. For example, if identifying everyone with GAD in a population is critical (as might be the case when screening in a limited population for a study of new methods of treatment of GAD), then a cut score with high sensitivity would be best. However, in a case in which having a homogeneous sample of participants with GAD is needed and missing some true positives is acceptable, then a cut score with high specificity would be ideal. This might be the case in an fMRI or PET scan study of GAD, where false positives would result in great expense. In the absence of similar concerns, however, we recommend the use of cut scores that optimize both sensitivity and specificity.

References

- Abel, J. L., & Borkovec, T. D. (1995). Generalizability of DSM-III-R generalized anxiety disorders to proposed DSM-IV criteria and cross-validation of proposed changes. *Journal of Anxiety Disorders, 9*, 303–315.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Beck, J. G., Stanley, M. A., & Zebb, B. J. (1995). Psychometric properties of the Penn State Worry Questionnaire. *Journal of Clinical Geropsychology, 1*, 33–42.
- Behar, E., Alcaine, O., Zuellig, A. R., & Borkovec, T. D. (2003). Screening for generalized anxiety disorder using the Penn State Worry Questionnaire: A receiver operating characteristics analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 34*, 25–43.
- Borkovec, T. D. (1994). The nature, functions, and origins of worry. In G. C. L. Davey, & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 5–33). Chichester, England: Wiley.

- Borkovec, T. D., Shadick, R. N., & Hopkins, M. (1991). The nature of normal and pathological worry. In R. M. Rapee, & D. H. Barlow (Eds.), *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression* (pp. 29–51). New York, NY: Guilford Press.
- Breitholtz, E., Westling, B. E., & Öst, L.-G. (1998). Cognitions in generalized anxiety disorder and panic disorder patients. *Journal of Anxiety Disorders, 12*, 567–577.
- Brown, T. A., Antony, M. M., & Barlow, D. H. (1992). Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. *Behaviour Research and Therapy, 30*, 33–37.
- Brown, T. A., DiNardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of DSM-IV anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 110*, 49–58.
- Davey, G. C. (1993). A comparison of three worry questionnaires. *Behaviour Research and Therapy, 31*, 51–56.
- DiNardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV: Lifetime version (ADIS-IV-L)*. Albany, NY: Graywind Publications.
- Dugas, M. J., Freeston, M. H., & Ladouceur, R. (1997). Intolerance of uncertainty and problem orientation in worry. *Cognitive Therapy and Research, 21*, 593–606.
- Freeston, M. H., Ladouceur, R., Rheaume, J., & Leger, E. (1998). Applications of cognitive models of OCD in clinical practice. In E. Sanavio (Ed.), *Behavior and cognitive therapy today: Essays in honor of Hans J. Eysenck* (pp. 117–126). Oxford, England: Anonima Romana.
- Kraemer, H. C. (1992). *Evaluating medical tests*. Newbury Park, CA: Sage Publications.
- Ladouceur, R., Blais, F., Freeston, M. H., & Dugas, M. J. (1998). Problem solving and problem orientation in generalized anxiety disorder. *Journal of Anxiety Disorders, 12*, 139–152.
- Lyonfields, J. D., Borkovec, T. D., & Thayer, J. F. (1995). Vagal tone in generalized anxiety disorder and the effects of aversive imagery and worrisome thinking. *Behavior Therapy, 26*, 457–466.
- Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2004). *Preliminary evidence for an emotion dysregulation model of generalized anxiety disorder*. submitted for publication.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*, 487–495.
- Molina, S., & Borkovec, T. D. (1994). The Penn state worry questionnaire: psychometric properties and associated characteristics. In G. C. L. Davey, & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 265–283). Chichester, England: Wiley.
- Newman, M. G., Zuellig, A. R., Kachin, K. E., Constantino, M. J., Przeworski, A., Erickson, T., & Cashman-McGrath, L. (2002). Preliminary reliability and validity of the GAD-Q-IV: A revised self-report diagnostic measure of generalized anxiety disorder. *Behavior Therapy, 33*, 215–233.
- Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737–748.
- StataCorp. (2001). *Stata Statistical Software: Release 7.0*. College Station, TX: Stata Corporation.
- Stöber, J. (1998). Reliability and validity of two widely used worry questionnaires: Self-report and self-peer convergence. *Personality and Individual Differences, 24*, 887–890.
- Turk, C. L., Heimberg, R. G., & Mennin, D. S. (2004). Assessment of worry and generalized anxiety disorder. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized anxiety disorder: Advances in research and practice* (pp. 219–247). New York: Guilford Press.
- van Rijsoort, S., Emmelkamp, P., & Vervaeke, G. (1999). The Penn state worry questionnaire and the worry domains questionnaire: Structure, reliability and validity. *Clinical Psychology and Psychotherapy, 6*, 297–307.
- Wells, A., & Carter, K. (1999). Preliminary tests of a cognitive model of generalized anxiety disorder. *Behaviour Research and Therapy, 37*, 585–594.