Research on generalized anxiety disorder (GAD) has significantly increased since it was first introduced as a residual diagnostic category in DSM-III (APA, 1980). Consistent with this progress, several self-report measures have been developed which assess symptomatology associated with GAD (e.g., Penn State Worry Questionnaire; Meyer, Miller, Metzger, & Borkovec, 1990).

The GAD-Q-IV (Newman, Zuellig, Kachin, Constantino, & Cashman, 2002) is a 9-item self-report measure designed to reflect the criteria of GAD as delineated in the DSM-IV (APA, 1994). Most items are dichotomous and measure excessiveness and uncontrollability of worry (e.g., “Do you experience excessive worry?”) and related physical symptoms (e.g., muscle tension). One item is open-ended and asks for a list of the most frequent worry topics. Two items are rated on a 0 (None) to 8 (Very Severe) scale and measure functional impairment and subjective distress. Thus, the measure assesses the DSM-IV criteria for GAD, with the exception of the exclusion criteria (criterion D & F; APA, 1994).

The GAD-Q-IV was initially designed as a screening device for GAD and is an updated version of the original GAD-Q (Roemer, Borkovec, Posas, & Borkovec, 1995), which was based on DSM-III-R (APA, 1987) criteria for GAD. The original version of this scale was scored by comparing individual items with DSM-III-R (APA, 1987) criteria for GAD. In an undergraduate sample, Newman et al. (2002) found that this scoring system resulted in 96% specificity and 67% sensitivity, suggesting that a third of participants who met GAD criteria were missed. Therefore, Newman et al. (2002) recommended using a dimensional scoring system (0-13) with a cut-off score demarking the presence or absence of GAD. Based on a Receiver Operating Characteristic analysis, Newman et al. (2002) suggested a cut-off score of 5.7 to achieve optimal balance between sensitivity and specificity. With this cut-off score, Newman et al. (2002) found good agreement between the GAD-Q-IV and clinician diagnosis based on the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV; T. Brown, DiNardo, & Barlow, 1994) or the Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L; DiNardo, Brown, & Barlow, 1994), kappa = .67, with 88% of clinician-diagnosed participants correctly classified. This dimensional scoring system has demonstrated good concurrent and discriminant validity and has good test-retest reliability (Newman et al., 2002).

However, a critique of this scoring system is that it allows for the possibility that an individual can be selected as meeting criteria for GAD without endorsing all of the diagnostic criteria. The 5.7 cut-off score may be over-diagnosing GAD, which is particularly pertinent for its utilization in research with analogue GAD samples. Turk et al. (in press) found that utilizing a 5.7 cut-off score in an unselected undergraduate sample resulted in 33% of the sample scoring positive for GAD, while the criteria matching system resulted in a more modest 14.5% of the sample as classifying for GAD. The criteria-matching system did not classify any participant as having GAD that the cutoff system did not.

The goal of the present study was to extend this research by examining the psychometric properties of the GAD-Q-IV and comparing the diagnostic accuracy of the two scoring systems among patients seeking treatment for worry and associated difficulties and a community control sample.

Method

Participants

The GAD group consisted of 31 treatment-seeking individuals with a primary diagnosis of GAD as defined by DSM-IV (APA, 1994). The control group consisted of 53 community control participants who had no current DSM-IV diagnosis, with the exception of two participants with specific phobia. The groups did not differ with regard to gender \( \chi^2 (1, N = 84) = .51, ns \), age \( t (82) = -1.62, ns \), or ethnic background \( \chi^2 (4, N = 84) = 4.82, ns \). Within the total sample, the mean age was 31.64 (SD = 11.86) and 59.5% of the sample was female. The total sample was primarily Caucasian (76.2%), while 19% were African American, 2.4% were Hispanic, 1.2% were Asian, and 1.2% indicated “other.”

Materials

Interview

- The Anxiety Disorders Interview Schedule for DSM-IV - Lifetime Version (ADIS-IV-L; DiNardo et al., 1994) is a semi-structured clinical interview that provides a thorough assessment for DSM-IV anxiety disorders but includes modules assessing mood disorders, substance abuse and dependence, and other disorders that overlap with anxiety disorders either conceptually or in terms of presenting symptoms (e.g., hypochondriasis).

Self-Report Measures

- The Penn State Worry Questionnaire (PSWQ; Meyer et al., 1990) is a 16-item measure of the generality, uncontrollability, and pervasiveness of worry. The validity of the PSWQ has been supported by an analysis indicating that the measure distinguished individuals with GAD from individuals with other anxiety disorders (Brown, Antony, & Barlow, 1992).
- The Worry Domains Questionnaire (WDQ; Tallis, Eysenck, & Mathews, 1992) is a 25-item measure that assesses different content areas in which a person worries. The total score of the WDQ provides a general indication of worry frequency.
- The Intolerance of Uncertainty Scale (IU; Freeston, Rheume, Letarte, Dugas, & Ladouceur, 1994) is a 27-item measure of uncertainty, emotional and behavioral reactions to ambiguous situations, implications of being uncertain, and attempts to control the future.
• The Social Interaction Anxiety Scale (SIAS; Mattick & Clark, 1998) is a 20-item self-report measure that assesses anxiety experienced in interactions in dyads and groups.

• The Social Phobia Scale (SPS; Mattick & Clark, 1998) is a 20-item self-report measure that assesses anxiety and distress experienced in situations where the person may be observed by others.

• The Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979) is a 21-item instrument that assesses the affective, cognitive, behavioral, somatic, and motivalional components of depression.

• The Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988) is a 21-item instrument that assesses the severity of anxiety symptoms.

Procedure
A psychologist or doctoral level graduate student interviewed all participants with the ADIS-IV-L (DiNardo et al., 1994). Following the interview, participants were asked to complete a battery of self-report questionnaires, which included the above measures and the GAD-Q-IV. The GAD-Q-IV was altered from its original version so that all participants were instructed to answer all of the nine items no matter whether question 6 was endorsed. (On the GAD-Q-IV, “skip out” instructions ask participants to skip questions 7-9 if item 6 is not endorsed). This allowed us to assess internal consistency and item-total correlations, which have not been previously assessed as a result of this “skip out” instruction. Individuals with a diagnosis of GAD were offered open cognitive-behavioral treatment. Community control participants were paid $40 to complete the assessment.

Results
Using the dimensional scoring system reported by Newman et al. (2002), the GAD-Q-IV achieved high internal consistency ( = .93), with GAD participants (M = 10.35, SD = 1.33), as measured by the ADISIV-L, exhibiting a significantly higher GAD-Q-IV dimensional score than non-anxious controls (M = 2.60, SD = 2.00), t(82) = -19.21, p < .001. Item-total correlations showed that all items were significantly correlated with the total score. Coefficients varied between 0.40 to 0.94, with a mean of 0.81.

This dimensional score was also utilized to assess the GAD-Q-IV’s convergent and discriminant validity. Correlations between the GAD-Q-IV dimensional score and the above self-report measures are presented in Table 1.

As hypothesized, the GAD-Q-IV dimensional score was highly correlated with both worry measures (PSWQ, r = .92; WDQ, r = .74). Further, tests of dependent correlations revealed that the GAD-Q-IV was more highly correlated with constructs associated with worry and GAD than with measures of social anxiety (see Table 2). Specifically, the GAD-Q-IV was more highly correlated with the PSWQ than with the SPS (r = .58, t(81) = 8.61, p < .001) and the SIAS (r = .65, t(81) = 7.78, p < .001). The GAD-Q-IV was also more correlated with the IU (r = .77) and the WDQ (r = .74), than with the SPS (t(81) = 3.14, p < .01; t(81) = 2.37, p < .05; respectively). While the GAD-Q-IV was more correlated with the IU than the SIAS (t(81) = 2.84, p < .05), it was not more correlated with the WDQ than the SIAS (t(81) = 1.60, ns). Further, it was highly correlated with the BAI (r = .77) and BDI (r = .76).

As two scoring systems have been reported in the literature, the sensitivity and specificity of the GAD-Q-IV in comparison to ADIS-IV-L interview was assessed with each of these scoring systems.

Newman et al. (2002) Scoring System
The cut-off score of 5.7 suggested by Newman et al. (2002) revealed that 50 of the 53 (94.3%) non-anxious community participants were correctly classified as not having a diagnosis of GAD with the GAD-Q-IV. Within the GAD sample, all of the participants were correctly classified as having GAD based on the GAD-Q-IV. These classifications correspond with 94.3% specificity and 100% sensitivity. Scoring the GAD-Q-IV to include the “skip out” instructions led to similar results, with 96.2% specificity and 96.8% sensitivity.

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. GAD-Q-IV</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PSWQ</td>
<td>.92**</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. WDQ</td>
<td>.74**</td>
<td>.78**</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. IU</td>
<td>.77**</td>
<td>.79**</td>
<td>.88**</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. SIAS</td>
<td>.65**</td>
<td>.68**</td>
<td>.69**</td>
<td>.71**</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. SPS</td>
<td>.58**</td>
<td>.58**</td>
<td>.55**</td>
<td>.57**</td>
<td>.80**</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>7. BDI</td>
<td>.76**</td>
<td>.73**</td>
<td>.70**</td>
<td>.74**</td>
<td>.47**</td>
<td>.48**</td>
<td>1.0</td>
</tr>
<tr>
<td>8. BAI</td>
<td>.77**</td>
<td>.73**</td>
<td>.73**</td>
<td>.71**</td>
<td>.62**</td>
<td>.67**</td>
<td>.78**</td>
</tr>
</tbody>
</table>

** p < .001 (two-tailed)
Criteria Matching Scoring System

Utilizing the criteria matching scoring system based on the original GAD-Q-IV, 51 of the 53 (96.2%) community control participants were correctly classified as not having GAD. Within the GAD sample, 24 of 31 participants were correctly classified as having GAD based on the GAD-Q-IV. These classifications correspond with a specificity of 96.2% and a sensitivity of 77.4%.

Discussion

The findings from this study suggest that the GAD-Q-IV can be an effective way of detecting the presence or absence of GAD in a quick and low-cost manner. In this clinical sample, the GAD-Q-IV dimensional score was internally consistent and has excellent convergent and divergent validity. The GAD-Q-IV was generally more correlated with measures of worry and associated constructs than with measures of social anxiety. Further, the dimensional score significantly distinguished GAD participants from community control participants. By deleting the “skip out” instructions, we were able to demonstrate that the GAD-Q-IV was internally consistent and has good item-total correlations.

The dimensional scoring system suggested by Newman et al. (2002) was very accurate in determining GAD, with 100% sensitivity and 94.3% specificity. Scoring the GAD-Q-IV to include the “skip out” instructions also resulted in high sensitivity (96.8%) and specificity (96.2%). The more conservative scoring approach of criteria matching utilized with the original version of the GAD-Q-IV had a specificity of 96.2% and a sensitivity of 77.4% against a diagnostic interview.

Thus, this more liberal scoring system suggested by Newman et al. (2002) was found to be more sensitive, without significantly compromising specificity. However, this may be an artifact of the sample utilized in this study, especially since most (51 out of 53) participants in the community control sample had no current Axis I diagnoses as measured by the ADIS-IV-L. Future studies are needed that include samples with other Axis I diagnoses (particularly anxiety disorders) and GAD-Q-IV scores that fall throughout the distribution to determine whether this scoring system compromises specificity. We are currently in the process of conducting a study to assess the psychometric properties of the GAD-Q-IV in comparison to diagnostic interview with individuals who range across the distribution of GAD-Q-IV dimensional scores. In conclusion, the results of this study suggest that the GAD-Q-IV can be utilized with confidence as an initial screener for detecting GAD.

Table 2. Dependent Correlations with the GAD-Q-IV between measures of worry and measures of social anxiety.

<table>
<thead>
<tr>
<th></th>
<th>PSWQ vs. SPS</th>
<th>PSWQ vs. SIAS</th>
<th>WDQ vs. SPS</th>
<th>WDQ vs. SIAS</th>
<th>IU vs. SPS</th>
<th>IU vs. SIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t(81)</td>
<td>d</td>
<td>t(81)</td>
<td>d</td>
<td>t(81)</td>
<td>d</td>
</tr>
<tr>
<td>GAD-Q-IV</td>
<td>8.61**</td>
<td>1.91</td>
<td>7.78**</td>
<td>1.73</td>
<td>2.37*</td>
<td>.53</td>
</tr>
<tr>
<td></td>
<td>2.84*</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: ** p < .001 (two-tailed), *p < .05 (two-tailed). d: Cohen’s (1977) d effect size conventions, small = .20, medium = .50, large = .80.

References