Using the GAD-Q-IV to identify generalized anxiety disorder in psychiatric treatment seeking and primary care medical samples

Michael T. Moore, a Nicholas L. Anderson, b Jill M. Barnes, b Emily A.P. Haigh, c David M. Fresco a

a Jackson State University, United States
b Kent State University, United States
c University of Maine, United States

1. Introduction

Generalized Anxiety Disorder (GAD) has been described as the “basic” anxiety disorder, given its early onset, chronic course, and resistance to change (Brown, Barlow, & Liebowitz, 1994; Kessler, Berglund, Demler, Jin, & Walters, 2005). Perhaps owing to its chronicity, GAD is associated with a significant degree of functional impairment. For example, disability from GAD is comparable to the amount caused by depression or other mood disorders (Grant et al., 2005; Wittchen, Carter, Pfister, Montgomery, & Kessler, 2000). Further, GAD tends to go untreated for long periods of time in some individuals. In addition to the impairment and suffering caused by GAD, the disorder is also associated with high levels of comorbidity with other disorders. In terms of Axis I comorbidity, mood (e.g., major depressive disorder [MDD], dysthymic disorder) and other anxiety disorders (e.g., panic disorder, social phobia, specific phobia) are the most commonly occurring diagnoses (Borkovec, Abel, & Newman, 1995; Brown & Barlow, 1992; Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Grant et al., 2005; Kessler, Chiu, Demler, & Walters, 2005; Noyes & Russell, 2001; Rogers et al., 1996). Due to its frequency and cost, accurate and scalable methods of assessment for GAD have the potential to significantly benefit public health via the early initiation of treatment.

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 1994), a person who qualifies for a diagnosis of GAD must have excessive anxiety and worry that he or she finds difficult to control about a number of events or activities. This worry must occur more days than not for a period of at least six months. Significant distress or impairment must also be associated with worry in one or more important areas of the person’s life. Additionally, the person’s symptoms must not be due to the effects of a drug, a medical condition, or be better accounted for by another Axis I disorder. During this period of uncontrollable worry, the person must also experience three of six symptoms in order to classify for a diagnosis of GAD (restlessness/feeling keyed-up or on-edge, fatigue, difficulty concentrating, irritability, muscle tension, and sleep disturbance). While clinician-administered interviews are widely considered to be the gold standard in diagnostic assessment, they can be costly in both time and monetary cost to clients. Thus, a self-report measure, which could be used with a larger number of individuals to initially screen for the presence of GAD, could potentially streamline the process of identifying individuals who are likely to meet criteria for GAD. The Generalized Anxiety Disorder Questionnaire (GAD-Q, Roemer, Borkovec, Posa, & Borkovec, 1995) was originally designed as a self-report diagnostic measure of GAD based upon the revised third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R; APA, 1987). Roemer et al. (1995) found good agreement between the GAD-Q and a
diagnosis of GAD using the revised Anxiety Disorders Interview Schedule (Di Nardo & Barlow, 1988) in undergraduate student participants. Further, these researchers found that participants meeting GAD criteria using the GAD-Q showed similar levels of anxiety to individuals diagnosed with GAD in a clinical community sample.

The fourth edition of the Generalized Anxiety Disorder Questionnaire (GAD-Q-IV; Newman et al., 2002) is a revised version of the GAD-Q that was updated to reflect the diagnostic criteria of DSM-IV. Specifically, the GAD-Q-IV eliminated questions related to unrealistic worry and included items assessing excessive and uncontrollable worry. The GAD-Q-IV consists of nine questions. Five of the first six items use a dichotomous response format (yes/no) inquiring about: the presence or absence of worry, whether the worry is excessive and uncontrollable, whether the individual worries about minor matters, and whether the worry has lasted at least six months. Question 5 provides blanks where individuals can list discrete areas of worry. Item 7 lists the anxiety symptoms associated with GAD (e.g., restlessness, muscle tension) for which a respondent can endorse presence or absence. Items 8 and 9 inquire about impairment and distress associated with GAD features along a 9-point scale.

The original GAD-Q can be used as either a continuous measure, by summing the responses to all items, or as a dichotomous measure based on an algorithm that tallied each of the criteria for GAD as laid out by the DSM-III-R. By contrast, Newman et al. (2002) devised a continuous scoring scheme for the GAD-Q-IV with potential range of scores from 0 to 12. The Yes/No items are coded 1 for Yes responses and 0 for No responses. Responses to Item 5 (number of discrete worries) are divided by 3 and responses to Item 7 (number of associated anxiety symptoms) are divided by 2. Responses for Items 8 (impairment) and 9 (distress) are divided by 4. Items that were skipped are coded as zero. The rationale for this weighted scoring scheme was to give priority to worry, which is regarded as the cardinal feature of GAD. In a sample of 143 college students who were also assessed with the lifetime version of the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV-L; Brown, Di Nardo, & Barlow, 1994), Newman et al. (2002) evaluated the ability of their continuous scoring scheme to reliably discriminate participants with a clinician-assessed diagnosis of GAD from participants without a GAD diagnosis. Participants in the study were college students who responded to an advertisement for free screening and possible referrals for help with anxiety problems. The sample included 30 participants who met criteria for GAD, 18 with panic disorder, 42 with social phobia, and 53 who did not meet criteria for any anxiety diagnoses. Using receiver operating characteristic (ROC) analysis, Newman et al. (2002) reported that a cut-score of 5.7 provided the best balance of specificity (.89), sensitivity (.83), positive predictive power (.68), and negative predictive power (.95), and correctly classified 88% of the participants. The performance of this cut score was compared to a dichotomous system based on the criteria for GAD listed in DSM-IV. Newman and her colleagues (2002) concluded that while the DSM-based method was associated with good specificity (.96), sensitivity was problematic (.67). Rodebaugh, Holaway, and Heimberg (2008) evaluated the factor structure and compared several possible methods of scoring the GAD-Q-IV, including both categorical and dimensional methods, the cut score identified by Newman and her colleagues, and a simple sum of all items. Results of both exploratory and confirmatory factor analysis indicated that the GAD-Q-IV was unfacorial. Interestingly, comparisons of the various scoring methods indicated that all, including the simple sum, resulted in scores that were equivalently correlated with a commonly-used self-report measure of worry.

Despite these successes for the GAD-Q-IV, several problems remain. For example, the continuous scoring scheme of Newman et al. (2002) creates a situation where respondents might earn a score of at least 5.7, but do not necessarily endorse all the required DSM-IV diagnostic criteria of GAD. Further, neither Newman et al. (2002), nor Rodebaugh et al. (2008) directly assess GAD in treatment seeking individuals, and thus, the performance of the GAD-Q-IV in an actual treatment setting has not been evaluated. Thus, the focus of the current study was to replicate and extend the findings of Newman et al. (2002) and Rodebaugh et al. (2008) with four important features. First, we sought to evaluate the ability of the GAD-Q-IV to differentiate between the presence/absence of GAD in a sample of individuals seeking psychotherapy at an outpatient psychological services clinic. Second, we sought to compare a self-report diagnosis of GAD based on affirmative endorsement of all of the DSM-IV criteria for GAD to the cut-score derived by Newman et al. (2002). Third, we sought to derive an optimal cut-score in two clinical samples: an outpatient sample seeking psychotherapy and primary care medical patients. The use of a primary care sample is particularly significant as GAD is both highly prevalent in the primary care setting (Üstün & Sartorius, 1995; Wittchen & Hoyer, 2001) and often goes unrecognized by primary care physicians (Weisberg, Dyck, Culpepper, & Keller, 2007; Wittchen et al., 2002). Fourth, among patients without GAD, we sought to identify which non-GAD diagnoses were most prone to be mistaken for GAD.

2. Method

2.1. Participants

Participants included 163 adult clients (69% women) presenting for treatment at the psychological clinic of a large, Midwestern university in the United States serving approximately 24,000 students. Age of the clients ranged from 18 to 45 (M = 22.44, SD = 5.02). In terms of race, the sample was predominantly Caucasian (72%). The remainder of the sample was African American (8%), indicated another race (9%) or did not identify their race (11%). The clinic offers individual, couples, and family outpatient treatment to university students and staff, as well as residents of the community. Approximately 35–40% of the clients presenting at the clinic are children and adults from the surrounding community, whereas the remainder of the clients are students of the university. The most commonly occurring diagnoses were Major Depressive Disorder (43%), GAD (37%), Dysthymic Disorder (8%), and Specific Phobia (4%). In addition, 99 adults (85% women) were recruited from the waiting room of a primary care clinic located in a large tertiary care, not-for-profit teaching hospital in the Midwestern United States. Age of these participants ranged from 20 to 67 (M = 38.65, SD = 12.80). In terms of race, the sample was evenly split between Caucasian (48%) and African American (44%) participants. A small percentage of individuals indicated another race (8%). The most commonly occurring diagnoses were Major Depressive Disorder (22%), GAD (23%), Dysthymic Disorder (12%), and Specific Phobia (10%).

2.2. Measures

The fourth edition of the Generalized Anxiety Disorder Questionnaire (GAD-Q-IV; Newman et al., 2002) is a nine-item self-report questionnaire designed as an initial screen for the presence of GAD. Newman et al. (2002) high report two-week stability of GAD diagnoses based on the GAD-Q-IV (92%). In addition, they found excellent inter-rater agreement with a semi-structured diagnostic interview (Cohen’s κ = .67).

The Structured Clinical Interview for DSM-IV (SCID-IV; First et al., 1996) is a clinician-administered diagnostic interview for Axis I
disorders based on DSM-IV criteria. The SCID-IV is sensitive to both current and past mental health functioning and is well equipped to address issues of comorbidity and differential diagnosis. Primacy of diagnoses is typically assigned based on degree of functional impairment for each disorder. The length of the total interview varied by individual, as a short screen of yes or no questions was used to determine for which modules a particular individual should be assessed. Questions related to current (i.e., within the past month) and past (i.e., greater than 1 month ago) mood and anxiety disorders were asked of all individuals.

Interrater reliability (κ) for diagnosis of GAD using the SCID-IV ranges from .44 to .95 (Lobbestael, Leurgans, & Arntz, 2010; Segal, Kabacoff, Hersen, Van Hasselt, & Ryan, 1995; Skre, Onstad, Torgersen, & Kringlen, 1991; Williams et al., 1992; Zanarini & Frankenburg, 2001; Zanarini et al., 2000). In the present study, interrater reliability was established for the primary care sample by having three interviewers rate each of 10 randomly selected SCID-IV interviews. The original interviewer would audio record and rate the SCID-IV, then two other study interviewers who were blind to the original ratings would independently rate the SCID-IV. Kappa (κ) was calculated by comparing each of the two additional ratings to the original rating, yielding two κ values for each interview. For this sample, interrater reliability across interviewers for the SCID-IV ranged from 0.79 to 0.87 (average κ = 0.83) for all diagnoses and from 0.81 to 0.85 (average κ = 0.82) for GAD in particular. Any discrepancies in diagnoses were addressed in case conference to establish consensus.

2.3. Procedure

2.3.1. Outpatient psychiatric sample

As part of intake procedures at the outset of treatment, patients provided informed consent and completed a battery of measures that included the GAD–Q-IV. Participants were not reimbursed for their participation, and all clinic clients provided data as part of routine intake procedures.

2.3.2. Primary care medical patient sample

Participants were recruited from the waiting room of the family medical center prior to attending primary care appointments. Each participant was individually approached and screened for eligibility by a research assistant, and the majority of those approached elected to participate (approximately 85%). Participants provided informed consent and then completed a self-report packet containing demographic questions, the GAD–Q-IV, and other items not relevant to this study’s aims. Research assistants were available to answer any questions that arose. Participants were paid $25 for their time and effort.

Of the 243 participants who completed the GAD–Q-IV, 104 (42.2%) elected to complete the SCID-IV interview. The only significant differences between participants who completed the SCID-IV and participants who did not was that participants who were either engaged as homemakers or employed outside the home were less likely to complete the SCID-IV than participants who were retired, disabled, or unemployed. The SCID-IV was administered by one of 6 clinical psychology graduate students who were kept blind to participants’ GAD–Q-IV scores. Interviewers received extensive training in conducting the interview, including practicum training on administration techniques and 1 year of supervision rounds to establish accuracy of diagnoses. The average number of interviews conducted per interviewer was 26. Participants’ physicians were notified if their responses to either the self-report items or SCID-IV interview indicated the presence of GAD.

Participants chose to complete the SCID-IV interview at that time (16; 10.3”), at a future appointment (117; 75.5”), or at a future time over the telephone (22; 14.2”). Although the SCID-IV was not designed for telephone administration, it has been shown to be acceptably consistent (95% agreement) with in-person administration for determining a diagnosis of GAD (Cacciola, Alterman, Rutherford, McKay, & May, 1999). The average time between participant completion of the GAD–Q-IV and the SCID-IV interview was 27 days (SD = 39.46). The majority of participants (107; 69.0%) were able to complete the SCID-IV interview within 3 weeks of completing the GAD–Q-IV. We investigated whether the amount of time between GAD–Q-IV and SCID-IV completion affected the likelihood of agreement between the two measures as to GAD status. The correlation was not statistically significant (r = .08; p = .35), suggesting that the time delay between the two administrations did not introduce significant error into our results.

3. Results

3.1. Data analysis

The present study used ROC analysis (Fawcett, 2006; Kraemer, 1992) to identify an optimal GAD–Q-IV score to accurately classify individuals based upon their GAD status as assessed by the SCID-IV. Derived from logistic regression, ROC allows for the analysis of a continuous predictor (i.e., GAD–Q-IV score) on a dichotomous criterion (i.e., “gold standard” SCID-IV diagnosis) to allow clinical researchers to discriminate individuals with a characteristic from those without a characteristic. ROC output consists of a ROC curve in which the sensitivity is plotted against the specificity for each value of the test. Sensitivity is defined as the proportion of individuals who have the condition and who are predicted to have the condition (i.e., identifying someone as having the characteristic in question). Specificity, on the other hand, is the proportion of individuals who do not have the condition and who are predicted to not have the condition (i.e., identifying someone as not having the characteristic in question). Sensitivity and specificity values can range from 0 to 1, and a value of .5 represents the degree of classification accuracy that would be expected by chance. Positive Predictive Value (PPV) is defined as the proportion of individuals predicted to have the condition who actually have the condition (true positives). Negative Predictive Value (NPV) refers to the proportion of individuals predicted not to have the condition who actually do not have the condition (true negatives). Finally, overall predictive value is the percentage of cases that are correctly classified as true positives and true negatives for the disorder.

In ROC analysis, the area under the curve (AUC) is used to identify the probability that a test will correctly classify participants. In the present study, the AUC is defined as the probability that a randomly selected GAD case will score higher than a randomly selected non-GAD case on the GAD–Q-IV. Thus, AUC measures the overlap in distributions between cases and controls. AUC values range from 0 to 1 and are interpreted as a probability; AUC values greater than 0.50 are interpreted to indicate better than chance classification accuracy. Insofar as the upper left corner of ROC space indicates perfect classification, the point on the ROC closest to this location (i.e., the top of a convex curve) would represent the cut score which best balances sensitivity and specificity.

3.2. Outpatient psychiatric sample

ROC analysis was used to derive an optimal cut-score for the dimensionally scored GAD–Q-IV in a treatment seeking sample. Using a dimensional scoring system, the area under the curve was .86 (SE = .029; CI = .81–.92), p < .001, indicating that an individual with GAD is 86% more likely to have a higher score on the GAD–Q-IV than someone without GAD using this scoring system. A cut score of 7.67 provided the best balance of sensitivity (see Table 1)
Table 1
Accuracy in identifying GAD using three methods of scoring the GAD-Q-IV.

<table>
<thead>
<tr>
<th></th>
<th>DSM-based algorithm</th>
<th>Cut score = 5.7</th>
<th>Cut score = 7.67</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outpatient PRIMARY CARE</td>
<td>Outpatient PRIMARY CARE</td>
<td>Outpatient PRIMARY CARE</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>.81</td>
<td>.89</td>
<td>.81</td>
</tr>
<tr>
<td>Specificity</td>
<td>.78</td>
<td>.65</td>
<td>.76</td>
</tr>
<tr>
<td>PPV</td>
<td>.66</td>
<td>.60</td>
<td>.62</td>
</tr>
<tr>
<td>NPV</td>
<td>.88</td>
<td>.91</td>
<td>.89</td>
</tr>
<tr>
<td>% correctly classified</td>
<td>79</td>
<td>74</td>
<td>78</td>
</tr>
</tbody>
</table>

Note: PPV = positive predictive value; NPV = negative predictive value.

and specificity. Based on this cut score, the GAD-Q-IV correctly classified 43 of 53 true GAD individuals and 82 of 110 true non-GAD individuals. The cut score of 5.7 derived by Newman et al. (2002) identified GAD cases with higher sensitivity (see Table 1), but lower specificity. Using a cut-score of 7.67, 26 individuals would be falsely diagnosed with GAD, compared to 36 using the Newman et al. (2002) approach. Finally, the items on the GAD-Q-IV, which reflect the DSM-IV criteria for GAD were used to create an alternative scoring algorithm to Newman et al. (2002). Positive endorsement of Items 1–4 and 6, in addition to endorsement of two or more worry topics in Item 5, three or more symptoms in Item 7 and a score of four or higher on Items 8 or 9 was used to indicate probable diagnosis of GAD. This DSM-based algorithm provided a similar balance of sensitivity and specificity (see Table 1) to the optimal dimensional cut score identified above.

Finally, in service of a more fine-grained analysis of false positives, an investigation was made into cases that did not have GAD (according to the SCID-IV), but which the GAD-Q-IV misclassified as having GAD (using either of the two cut scores or the DSM-based algorithm). Twenty-five individuals were misclassified as having GAD using the cut score of 5.7. Of these cases, the most common diagnoses were Major Depressive Disorder (MDD; 40%), Anxiety Disorder Not Otherwise Specified (AnxNOS; 20%), and dysthyemic disorder (DYS; 12%). Note that these values will not add up to 100% as individuals could have multiple diagnoses. Of these 25 misclassified cases, 28% met criteria for both a mood and anxiety disorder, 25% met criteria for a mood disorder only, 25% met criteria for neither a mood nor an anxiety disorder, 13% met criteria for an anxiety disorder only, and 9% did not meet criteria for any diagnosis. Eighteen individuals were misclassified as having GAD using the cut-score of 7.67. Of these 18 cases, the most common diagnoses were MDD (33%), AnxNOS (22%), Adjustment Disorder (17%), and DYS (11%). Of the cases misclassified as GAD, 28% met criteria for both a mood and anxiety disorder, an equal number met criteria for neither a mood nor an anxiety disorder, 22% met criteria for an anxiety disorder only, 17% met criteria for a mood disorder only, and 6% did not meet criteria for any diagnosis. The DSM-based algorithm misclassified 2 fewer participants as compared to the cut score of 7.67. Of these 16 cases, the most common diagnoses were MDD (44%), AnxNOS (19%), DYS (19%), and Adjustment Disorder (6%). Of the cases misclassified as GAD, 37% met criteria for both a mood and anxiety disorder, 25% met criteria for a mood disorder only, 19% met criteria for neither a mood nor an anxiety disorder, 13% met criteria for an anxiety disorder only, and 6% did not meet criteria for any diagnosis.

3.3. Primary care medical patient sample

ROC analysis was again used in a primary care medical patient sample. Using a dimensional scoring system, the area under the curve was .85 (SE = .041; CI = .76–.93), p < .001, indicating that an individual with GAD is 85% more likely to have a higher score on the GAD-Q-IV than someone without GAD using this scoring system. The performance of the two dimensional cut scores and the DSM-based scoring algorithm were compared in a sample obtained from a primary care medical clinic. Using the dimensional cut score of 5.7 from Newman et al. (2002), the GAD-Q-IV correctly classified 24 of 27 true GAD individuals and 45 of 72 true non-GAD individuals. Using the dimensional cut score of 7.67, identified as the optimal value in the outpatient sample above, once again provided the optimal balance of sensitivity and specificity as compared to the lower cut score of Newman et al. (2002) (see Table 1). Based upon this higher threshold, 24 of 27 true GAD individuals and 52 of 72 true non-GAD individuals were correctly classified. The DSM-based scoring algorithm fared the best of the three scoring methods when compared with the SCID-IV diagnoses (see Table 1). Using this method, the GAD-Q-IV correctly classified 18 of 27 true GAD individuals and 63 of 72 true non-GAD individuals.

Again, a comparison was made between the conditions misclassified as GAD by the three methods for scoring the GAD-Q-IV. In the primary care sample, the cut score of 5.7 (n = 29) and 7.67 (n = 21) misclassified many more individuals as GAD than the DSM-based method (n = 11). For the cut score of 5.7, the most common diagnoses misclassified as GAD were a substance misuse diagnosis (24%), DYS (21%), MDD (21%), SPEC (14%), a bipolar diagnosis (14%), and Panic Disorder (with or without Agoraphobia; 10%). Of the cases misclassified as GAD by the cut score of 5.7, 34% reported a mood disorder only, 24% reported no diagnosis, 17% reported both a mood and an anxiety disorder, 14% reported an anxiety disorder only, and 11% reported neither a mood nor an anxiety disorder. For the cut score of 7.67, the most common diagnoses misclassified as GAD were SPEC (24%), MDD (19%), a substance misuse diagnosis (14%), a bipolar diagnosis (14%), DYS (10%), and Panic Disorder (with or without Agoraphobia; 10%). Of the cases misclassified as GAD by the cut score of 7.67, 33% did not meet criteria for any diagnosis, 29% met criteria for a mood disorder only, 14% met criteria for an anxiety disorder only, 14% met criteria for both a mood and an anxiety disorder, and 10% met criteria for neither a mood nor an anxiety disorder. For the DSM-based method, the most common diagnoses were a bipolar diagnosis (27%), a substance misuse diagnosis (27%), MDD (18%), and SPEC (18%). Of the cases misclassified as GAD by the DSM-based method, 27% reported no diagnosis, 27% reported a mood disorder only, 10% reported an anxiety disorder only, 18% reported both a mood and an anxiety disorder, and 18% reported neither a mood nor an anxiety disorder.

4. Discussion

The primary focus of this paper was to evaluate the ability of the GAD-Q-IV to differentiate between the presence/absence of GAD in both a sample of individuals seeking outpatient psychotherapy at a general psychology clinic and medical treatment in a primary care setting. We sought to compare the cut score (5.7) proposed by Newman et al. (2002) to an optimal cut score (7.67) derived from ROC analyses, as well as a scoring algorithm meant to closely approximate DSM-IV criteria for GAD. Based upon the present results, use of a cut score of 5.7 would seem to overestimate the number of affirmative GAD diagnoses in both an outpatient clinical
and primary care setting. By contrast, our optimal cut score and the
DSM-based algorithm performed similarly in an outpatient ther-
apy setting. The DSM-based algorithm provided the best balance
of sensitivity to specificity in the primary care setting and would
seem to be the most generalizable of the three scoring methods for
the GAD-Q-IV. However, it should be noted that under conditions
where false negatives are particularly costly (i.e., missing any cases
with GAD), the superior sensitivity of the lower cut score identi-
fied by Newman and her colleagues may outweigh the inevitable
trade-off in an increased number of false positives.

Comparison of various methods of classification cannot exist
without a consideration of variation in base rates of the condition
in question (Finn, 2009). Both Newman et al. (2002) and Rodebaugh
et al. (2008) evaluated the performance of the GAD-Q-IV in under-
graduate samples, where the base rate of GAD is likely to be lower
than in the samples utilized in the current study. It would be incor-
rect to conclude that one cut score will be ideal for all potential
settings where the GAD-Q-IV might be used. The results of Newman
et al. (2002) and Rodebaugh et al. (2008) demonstrate the validity
of the lower cut score in an undergraduate population while our
results suggest that a higher cut score represents the best bal-
ance of sensitivity and specificity in both psychiatric and medical
treatment-seeking populations.

Examination of the two cut scores identified here has several
research and clinical implications. For instance, researchers who
are attempting to screen for GAD potential diagnoses might want to
“cast a wider net” by using the cut score identified by Newman et al.
(2002). However, if the goal is to more accurately identify probable
GAD diagnoses from outpatient or primary medical care clinics (e.g.,
creating an analog GAD group in a comparison study), the cut score
of 7.67 identified here might be a more acceptable representation
of true GAD diagnoses. Alternatively, the higher cut score would be
useful, due to its lower false positive rate, in circumstances where
there is a high cost to incorrect identification of potential GAD cases
(e.g., where potentially ineligible research participants will have to
be paid for a lengthy diagnostic interview).

Recently, a brief measure for assessing GAD was published, the
GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006). Unlike the
GAD-Q-IV, the GAD-7 was developed as a brief self-report instrument
to assess probable cases of GAD. However, the GAD-7 was not de-
veloped to include all DSM-IV criteria for GAD. As such, an individual
who is deemed a “probable GAD” on the GAD-7 may not meet the
DSM-IV definition of GAD. Initial findings for the utility of this mea-
sure as a screening tool for probable cases of GAD seem promising,
but little research has been published to evidence that the GAD-
7 shows more clinical utility than the GAD-Q-IV. The cut score
of the current investigation improves upon other GAD measures, such
as the GAD-7, by increasing the clinical utility of the GAD-Q-IV to
increase the likelihood that an individual endorsed as a probably
GAD is likely to endorse all necessary items to qualify as a true GAD
diagnosis. Therefore, a cut score of 7.67 on the GAD-Q-IV may be a
better clinically useful tool over the GAD-7.

4.1. Limitations

Several limitations must be addressed in the present study.
Because the GAD-Q-IV is a self-report measure, we are limited
by the nature of self-report data and cannot establish whether par-
ticipants’ responses are accurate. This study is also limited by the
nature of the samples which consisted primarily of Caucasians.
Because of this, the results from this study may not generalize to
other populations. Finally, the current investigation utilized the
DSM-IV definition of GAD. The fifth edition (DSM-5; APA, 2013)
was recently published and, while the core criteria for GAD were
unchanged, one revision was the removal of the stipulation that
symptoms cannot “occur exclusively during a Mood Disorder, a
Psychotic Disorder, or a Pervasive Developmental Disorder” (APA,
2000, p. 476). While this change is unlikely to significantly alter
the results, as it is not reflected in the GAD-Q-IV, future revision
of the GAD-Q-IV and subsequent replication of the current inves-
tigation, will be necessary to reflect DSM-5. One potential area for
future research is to investigate the results of the GAD-Q-IV using
Newman and colleague’s (2002) cut score and the cut score estab-
lished in the present study and compare this to probable diagnoses
using the GAD-7. Results would allow direct comparison of compet-
ing GAD diagnoses. Nevertheless, future research should continue
to use and evaluate the GAD-Q-IV as a valid measure of GAD in
clinical and research samples.

4.2. Conclusions

The GAD-Q-IV is a useful tool for assessing self-report GAD
probable diagnoses. The results of this study suggest that a brief,
self-report instrument can adequately screen for GAD in a vari-
ety of settings. Given that the GAD-Q-IV is quickly and easily
administered, its adoption as a screening instrument can improve
the identification of GAD in primary care: the setting where it is
both most likely to be first encountered (Üstün & Sartorius, 1995;
Wittchen & Hoyer, 2001) and to go unrecognized (Weisberg, Dyck,
Culpepper, & Keller, 2007; Wittchen et al., 2002). Improvement in
the identification of GAD can serve to direct clients to treatment
who might otherwise not have gotten it and, hopefully, reduce the
burden of this disorder.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental
bid conditions in generalized anxiety disorder. Journal of Consulting and Clinical
Psychology, 63, 479–483.
atations for treatment and DSM-IV. Journal of Consulting and Clinical Psychology,
60, 835–844.
Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a
Corp.
Comparability of telephone and inperson structured clinical interview for DSM-
(ADIS-R): Albany: Phobia and Anxiety Disorders Clinic, State University of New
York.
882–891.
Finn, S. E. (2009). Incorporating base-rate information in daily clinical decision
disorders in the US: results from the national epidemiologic survey on alcohol
prevalence and age-of-onset distributions of DSM-IV disorders in the national
and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey


