



Repositorio Institucional de la Universidad Autónoma de Madrid

<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:

This is an **author produced version** of a paper published in:

Psychological Assessment 25.1 (2013): 279-287

DOI: <http://dx.doi.org/10.1037/a0030645>

Copyright: © 2012 American Psychological Association

El acceso a la versión del editor puede requerir la suscripción del recurso

Access to the published version may require subscription

Psychological Assessment

Psychometrical Assessment and Item Analysis of the General Health Questionnaire in Victims of Terrorism

David Delgado-Gomez, Jorge Lopez-Castroman, Victoria de Leon-Martinez, Enrique Baca-Garcia, Maria Luisa Cabanas-Arrate, Antonio Sanchez-Gonzalez, and David Aguado
Online First Publication, December 3, 2012. doi: 10.1037/a0030645

CITATION

Delgado-Gomez, D., Lopez-Castroman, J., de Leon-Martinez, V., Baca-Garcia, E., Cabanas-Arrate, M. L., Sanchez-Gonzalez, A., & Aguado, D. (2012, December 3). Psychometrical Assessment and Item Analysis of the General Health Questionnaire in Victims of Terrorism. *Psychological Assessment*. Advance online publication. doi: 10.1037/a0030645

Psychometrical Assessment and Item Analysis of the General Health Questionnaire in Victims of Terrorism

David Delgado-Gomez
Universidad Carlos III

Jorge Lopez-Castroman
Inserm, U1061 and IIS-Fundacion Jimenez Diaz

Victoria de Leon-Martinez
IIS-Fundacion Jimenez Diaz

Enrique Baca-Garcia
IIS-Fundacion Jimenez Diaz, New York State Psychiatric
Institute, and Columbia University

Maria Luisa Cabanas-Arrate
and Antonio Sanchez-Gonzalez
Fundación Instituto de Victimología

David Aguado
Instituto Ingeniería del Conocimiento

There is a need to assess the psychiatric morbidity that appears as a consequence of terrorist attacks. The General Health Questionnaire (GHQ) has been used to this end, but its psychometric properties have never been evaluated in a population affected by terrorism. A sample of 891 participants included 162 direct victims of terrorist attacks and 729 relatives of the victims. All participants were evaluated using the 28-item version of the GHQ (GHQ-28). We examined the reliability and external validity of scores on the scale using Cronbach's alpha and Pearson correlation with the State-Trait Anxiety Inventory (STAI), respectively. The factor structure of the scale was analyzed with varimax rotation. Samejima's (1969) graded response model was used to explore the item properties. The GHQ-28 scores showed good reliability and item-scale correlations. The factor analysis identified 3 factors: anxious-somatic symptoms, social dysfunction, and depression symptoms. All factors showed good correlation with the STAI. Before rotation, the first, second, and third factor explained 44.0%, 6.4%, and 5.0% of the variance, respectively. Varimax rotation redistributed the percentages of variance accounted for to 28.4%, 13.8%, and 13.2%, respectively. Items with the highest loadings in the first factor measured anxiety symptoms, whereas items with the highest loadings in the third factor measured suicide ideation. Samejima's model found that high scores in suicide-related items were associated with severe depression. The factor structure of the GHQ-28 found in this study underscores the preeminence of anxiety symptoms among victims of terrorism and their relatives. Item response analysis identified the most difficult and significant items for each factor.

Keywords: factor analysis, item response theory, terrorism victims, GHQ-28

In 2010, 11,604 terrorist attacks killed 13,186 and injured 30,665 persons globally (National Counter Terrorism Center, 2011). The indirect consequences of these attacks on mental health may be more frequent (10–100 psychiatric casualties for every physical injury), and they have more significance, cost and long-term effects wise, than previously thought (Ursano & Friedman, 2006). The families of the victims, independently of their own

exposure, may also present a high risk for the development of psychopathology (Hoven et al., 2009; Stoddard et al., 2011). The most frequent mental disorders in the aftermath of a terrorist attack are posttraumatic stress disorder (PTSD), depression, and other anxiety disorders (Henriksen, Bolton, & Sareen, 2010; Salguero, Cano-Vindel, Iruarizaga, Fernandez-Berrocal, & Galea, 2011). The aftereffects of the trauma exposure can affect victims and their

David Delgado-Gomez, Department of Statistics, Universidad Carlos III, Madrid, Spain; Jorge Lopez-Castroman, Inserm, U1061, Montpellier, France, and IIS-Fundacion Jimenez Diaz, Department of Psychiatry, CIBERSAM, Madrid, Spain; Victoria de Leon-Martinez, IIS-Fundacion Jimenez Diaz, Department of Psychiatry, CIBERSAM; Enrique Baca-Garcia, IIS-Fundacion Jimenez Diaz, Department of Psychiatry, CIBERSAM, and New York State Psychiatric Institute, and Department of Psychiatry, Columbia University; Maria Luisa Cabanas-Arrate and Antonio

Sanchez-Gonzalez, Instituto de Investigación en Psiquiatría y Psicoterapia, Fundación Instituto de Victimología, Madrid, Spain; David Aguado, Instituto Ingeniería del Conocimiento, Madrid, Spain.

David Delgado-Gomez and Jorge Lopez-Castroman have contributed equally to this work. The work by Jorge Lopez-Castroman was supported by a FondaMental Foundation grant.

Correspondence concerning this article should be addressed to Jorge Lopez-Castroman, IIS-Fundacion Jimenez Diaz, Department of Psychiatry, Avenida Reyes Catolicos 2, 28040 Madrid, Spain. E-mail: jlopez@fjd.es

relatives for years (Baca, Cabanas, & Baca-García, 2002). PTSD in particular often follows a chronic course in the general population (Breslau et al., 1998; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995) and after exposure to terrorist attacks (DiGrande, Neria, Brackbill, Pulliam, & Galea, 2011; Kawana, Ishimatu, & Kanda, 2001; North, Pfefferbaum, Kawasaki, Lee, & Spitznagel, 2011).

Although early detection of mental disorders in exposed populations may prevent future disabilities (Stoddard et al., 2011), the performance of screening instruments after a terrorist attack may present limitations. However, they have seldom been studied (Brewin, Fuchkan, Huntley, & Scragg, 2010). For instance, the diagnostic performance of screening instruments, such as the Trauma Screening Questionnaire, in an affected population was improved when the interval from the terrorist attacks was longer (Brewin et al., 2010). Moreover, there is a specific need for outreach to detect individuals with mental disorders after terrorist attacks (Brewin et al., 2008; Pfefferbaum, North, Flynn, Norris, & DeMartino, 2002).

One instrument that has been successfully used to assess the psychopathology of terrorism victims and their relatives is the General Health Questionnaire (GHQ; Baca et al., 2002; Boscarino, Figley, & Adams, 2004). The GHQ, in all its versions, is a self-administered Likert-type questionnaire frequently used to detect nonpsychotic mental disorders. The results of the GHQ might depend on the characteristics of the sample. For instance, the GHQ provided greater sensitivity and specificity in a clinical sample of 382 adult females than in a comparable sample of 154 males (Hobbs, Ballinger, Greenwood, Martin, & McClure, 1984). Differences were also observed among 260 primary care respondents: respondents younger than 40 years of age had 3 times more risk of being misclassified with the GHQ than respondents older than 40 years of age (de Jesus Mari & Williams, 1986). Moreover, Tarnopolsky et al. indicated that validity coefficients could differ from one population to another. Therefore, it is necessary to analyze the reliability and external validity of the scores of the GHQ and its factor structure in any population under study (Tarnopolsky, Hand, McLean, Roberts, & Wiggins, 1979). To date, the GHQ has been validated in several populations varying in age (Baksheev, Robinson, Cosgrave, Baker, & Yung, 2011; Costa et al., 2006), sex (Romans-Clarkson, Walton, Herbison, & Mullen, 1989; Tran, Tran, & Fisher, 2012), nationality (Politi, Piccinelli, & Wilkinson, 1994; Quek, Low, Razack, & Loh, 2001), and work status (Banks et al., 1980), however not in terrorism victims. Moreover, Tarnopolsky et al. supported the validation of the GHQ on any particular population because validity coefficients could differ from one setting to another (Tarnopolsky et al., 1979).

The reliability, factor structure, and external validity represent the classical test theory in psychological assessment. This traditional assessment can be complemented with the item response theory (IRT), which is not commonly used to evaluate clinical instruments (Thomas, 2011) and, to our knowledge, has never been applied to the GHQ. The IRT analyses provide indices of discrimination and difficulty for each item, and therefore a deeper understanding of the questionnaire (P. J. Ferrando, 2001; Ploubidis & Frangou, 2011).

In this article, the psychometric properties of the GHQ as a screening questionnaire in a population of terrorism victims and their relatives were evaluated. Although the questionnaire has been

used before to assess terrorism victims and their relatives (Baca et al., 2002), its psychometric properties have never been analyzed in this population. Thus, the first goal of this study was to analyze the reliability and external validity of scores of the 28-item version of the GHQ (GHQ-28) and its factor structure in this population. The second goal was to estimate the parameters of each item of the GHQ-28 using the IRT. The GHQ-28 items were then sorted according to the IRT. The item characteristic curves (ICCs) provided a visual understanding of the item responses. In conclusion, the IRT allowed a more robust estimate of the latent variable, given that the item parameters are independent of the sample.

Method

Participants

The sample was composed of 891 participants: 162 participants had experienced a terrorist attack, and 729 participants were relatives of the victims. The sample was obtained from interviews with victims of terrorist attacks in Spain and their immediate families (first-degree relatives). The average time from the terrorist attacks to the interviews was 13.5 years ($SD = 6.6$). This sample is part of the Phoenix Project (Baca et al., 2002), which was funded by the Association of Victims of Terrorism and designed to study the long-term consequences of terrorist attacks. The types of the terrorist attacks were mostly bombings (58.7%) or shootings (36.2%). All participants effectively completed the GHQ-28 and the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970).

The mean age of the sample was 38.7 ± 14.8 years. Although there was a majority of females in the total sample (60.6%), the victims were composed of mostly males (69.7% males). The victims and their relatives were mostly married (50.3%) or single (34.4%). Regarding the level of education in the sample, 28.2% reported primary studies or less, 46.0% reported secondary studies, and 25.8% reported university studies. A majority of the victims (74.9%) were policemen or army soldiers.

Measures

The GHQ-28. The GHQ-28 is a Likert-type scale composed of 28 items selected from the 60-item GHQ version by a principal component analysis (Goldberg & Hillier, 1979). Each item receives a discrete score ranging from 0 to 3. The 28 items were then divided into four groups. Items in each group measure a different trait. These four traits were somatic, anxiety and depression symptoms, and social dysfunction (see Table 2). Three different approaches have appeared in the literature to evaluate the questionnaire. They differ in the way the score of each item is calculated. In the classical GHQ approach, values 0 and 1 score a point value of 0, whereas values 2 and 3 score a point value of 1. In the Likert approach, values are not transformed. There also exists a third uncommon approach, called the c-GHQ, where the score of each item relies on whether the item is formulated positively or negatively (Richard, Lussier, Gagnon, & Lamarche, 2004). In this article, the Likert approach was used. It has been observed that the Likert method is the preferable method as it produces less skewed distributions and is more suitable for correlation analysis and intergroup comparisons (Banks et al., 1980). The Spanish version of the questionnaire was handed to the participants (Lobo & Muñoz, 1996).

The STAI. The STAI (Spielberger et al., 1970) is a Likert-type self-report questionnaire composed of 40 items. The first 20 items measure anxiety as a state, whereas the last 20 items measure anxiety as a trait with a single score for each. Previous literature has demonstrated a high reliability and validity for STAI scores in different samples (Quek, Low, Razack, Loh, & Chua, 2004; Zhang & Gao, 2012).

Procedure

Reliability, factor analysis, and external validity. The reliability of the GHQ-28 was estimated using Cronbach's alpha. The adjusted correlations of each item with the GHQ-28 were also calculated. An exploratory factor analysis using varimax rotation was conducted to discover the structure of the GHQ-28. A maximum likelihood estimator was used for the exploratory factor analysis. Coefficients obtained with this method were similar to those obtained with other estimation methods, such as principal components and generalized least squares, except for an inversion of the second and third factor. However, the second and third factor presented very similar eigenvalues using any of the aforementioned methods (data not shown). In order to determine the factors to be retained, a parallel analysis was conducted following Hayton, Allen, and Scarpello (2004). The Bartlett method was used to calculate the factor scores. Finally, the external validity was analyzed in comparison with a questionnaire measuring anxiety (STAI) in agreement with Sánchez-López and Dresch (2008). We estimated the correlation between the STAI (trait and state scores) and the GHQ-28 (global and factor scores).

The graded response model. The IRT is based on the assumption that the answer given by an individual to a specific item depends on an individual's latent variable and the properties of the item. Changes to the latent variable will modify the response probability. Moreover, it is assumed that the latent variable under study follows a standard Gaussian distribution. This assumption allows the determination of the percentile at which an individual can be placed in a given population. Specifically, let i denote an item of an item bank, and let n_i be the number of possible choices for the i th item. Let k be the number of examinees. The probability that the k th examinee with a given latent variable value θ^k chooses the answer $g \in \{0, \dots, n_i - 1\}$ on an item i is represented by $P_{i(g)}(\theta^k)$. Several models have been developed in order to define this probability (van der Linden & Hambleton, 1997).

Among the different item response models, we used Samejima's (1969) graded response for the two following reasons: (a) Responses are ordered, and (bi) specific parameters for discrimination and difficulty are obtained for each item, allowing us to differentiate participants with the same score on two different items. This model calculates the probabilities of answering each of the item choices given the trait level θ^k of the examinee. The probability that the k th individual will choose the category g or higher on the i th item is expressed by,

$$P_{i(g)}^* = \frac{e^{a_i(\theta^k - b_{i(g)})}}{1 + e^{a_i(\theta^k - b_{i(g)})}},$$

where a_i is the discrimination parameter, $b_{i(g)}$ indicates the difficulty specified for the answer g , and the difficulty parameters

satisfy that $b_{i(g)} < b_{i(g+1)}$. The probability $P_{i(g)}$ is computed as a difference of the cumulative probabilities for adjacent answers:

$$P_{i(g)} = P_{i(g)}^* - P_{i(g+1)}^*.$$

This formula requires the probabilities when $g = 0$ and $g = n_i$ for all the items:

$$P_{i(0)}^*(\theta^k) = 1, P_{i(n_i)}^*(\theta^k) = 0.$$

The IRT term *difficulty* cannot be directly applied to psychological assessment with the GHQ-28 as there are no correct or incorrect answers to the test and is substituted by the term *location* henceforward. The discrimination and location parameters that characterize each item of a questionnaire are estimated by the marginal maximum likelihood technique using the answers of a sample of examinees. The item location identifies the value of the latent variable where the probability of choosing or not choosing a specific answer to an item is the same. The higher the discrimi-

Table 1

Adjusted Item-Scale Correlation and Cronbach's Alphas After Removing Each Item in the GHQ-28

Item	Adjusted item-scale correlation	Cronbach's α if the item is eliminated
1	.603	0.954
2	.647	0.954
3	.766	0.953
4	.684	0.953
5	.589	0.954
6	.649	0.954
7	.615	0.954
8	.673	0.954
9	.657	0.954
10	.774	0.953
11	.769	0.953
12	.695	0.953
13	.773	0.953
14	.762	0.953
15	.288	0.956
16	.565	0.955
17	.548	0.955
18	.513	0.955
19	.514	0.955
20	.453	0.955
21	.584	0.954
22	.665	0.954
23	.729	0.953
24	.702	0.953
25	.602	0.954
26	.782	0.953
27	.689	0.954
28	.623	0.954
Internal consistency		
Total sample	$\alpha = 0.955$	Std $\alpha = 0.954$
Victims	$\alpha = 0.962$	Std $\alpha = 0.961$
Relatives	$\alpha = 0.951$	Std $\alpha = 0.950$

Note. Internal consistency of the GHQ-28 as measured with Cronbach's alphas (Standardized) for the total sample (victims and their relatives). GHQ-28 = 28-item version of the General Health Questionnaire; Std = Standardized.

Table 2

Eigenvalues, Percentage of Explained Variance for Each Factor, and Maximum Likelihood Estimate of Factor Loadings for the GHQ-28

Factors	1	2	3
Eigenvalues and percentage of explained variance before rotation			
Eigenvalue	12.32	1.79	1.40
% Variance	43.99	6.39	5.01
% Cumulative variance	43.99	50.39	55.40
Eigenvalues and percentage of explained variance after rotation			
Eigenvalue	7.95	3.85	3.70
% Variance	28.40	13.75	13.24
% Cumulative variance	28.40	42.16	55.40
Factor loadings (Item: Have you recently . . .)			
Somatic			
1. been feeling perfectly well and in good health?	0.53	0.34	0.09
2. been feeling in need of a good tonic?	0.61	0.25	0.16
3. been feeling run down and out of sorts?	0.70	0.32	0.20
4. felt that you are ill?	0.62	0.28	0.18
5. been getting any pains in your head?	0.61	0.13	0.13
6. been getting a feeling of tightness or pressure in your head?	0.66	0.15	0.17
7. been having hot or cold spells?	0.61	0.16	0.15
Anxiety			
8. lost much sleep over worry?	0.71	0.15	0.16
9. had difficulty in staying asleep once you are off?	0.67	0.19	0.14
10. felt constantly under strain?	0.79	0.20	0.22
11. been getting edgy and bad-tempered?	0.79	0.19	0.23
12. been getting scared or panicky for no good reason?	0.64	0.20	0.26
13. been satisfied with the way you've carried out your task?	0.71	0.25	0.29
14. felt constantly under strain?	0.76	0.21	0.24
Social Dysfunction			
15. been managing to keep yourself busy and occupied?	0.14	0.40	0.01
16. been taking longer than usual to do things?	0.34	0.52	0.14
17. felt on the whole you were doing things well?	0.22	0.71	0.15
18. been satisfied with the way you've carried out your tasks?	0.18	0.72	0.13
19. felt that you are playing a useful part in things?	0.16	0.60	0.29
20. felt capable of making decisions about things?	0.09	0.58	0.27
21. been able to enjoy your normal day-to-day activities?	0.36	0.47	0.22
Depression			
22. been thinking of yourself as a worthless person?	0.35	0.45	0.45
23. felt that life is entirely hopeless?	0.43	0.43	0.46
24. felt that life isn't worth living?	0.38	0.33	0.60
25. thought of the possibility that you might do away with yourself?	0.23	0.16	0.85
26. found at times you couldn't do anything because your nerves were too bad?	0.65	0.28	0.38
27. found yourself wishing you were dead and away from it all?	0.33	0.30	0.72
28. found that the idea of taking your own life kept coming into your mind?	0.25	0.19	0.83

Note. GHQ-28 = 28-item version of the General Health Questionnaire. Bold type identifies .40 or larger loads.

nation value, the easier the discrimination is between the distinct levels of the latent variable. These parameters are graphically represented in the ICCs. The ICCs represent the probability of each item response conditioned on the specific values of the latent variable. The slope of the curve of the location values determines the discrimination of an item. An example of ICC with three possible answers, and a discriminative parameter (a_i) of 2.5, and location parameters (b_i) of -1 and 1 are displayed in Figure 2. In this article, the ICCs have been summarized in an expected average score for each item, in accordance with:

$$\sum_{k=0}^{n_i} k \int_{-\infty}^{\infty} P_{i(k)}(\theta) f(\theta) d\theta,$$

where $f(\theta)$ is the density function of a standard Gaussian distribution. The average score allows us to sort the items by their inherent location for a given latent variable. For instance, a person

with a higher latent variable would obtain higher scores in low-average score items than a person with a smaller latent variable.

All procedures were performed using the ltm R statistical package.

Results

Reliability

Table 1 shows the Cronbach's alphas (and the standardized alpha) for the total sample, the victims and their relatives together with the adjusted item-scale correlation, and the Cronbach's alphas after removing each item. The calculated Cronbach's alphas were superior to 0.95 in all cases. The adjusted item-scale correlation ranged from .45 to .78 for all the items, except Item 15 in which the item-scale correlation was .29.

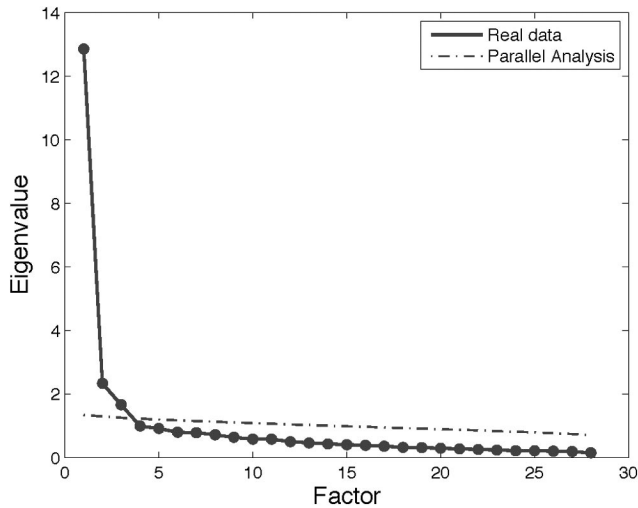


Figure 1. Plot of actual versus randomly generated eigenvalues.

Factor Analysis

Three factors were retained by the parallel analysis. They also showed an associated eigenvalue higher than 1 (see Figure 1). The eigenvalues of these factors together with the corresponding percentage of explained variance and the factor loadings are displayed in Table 2. The highest loadings of the first factor were found in the first 14 items (somatic and anxiety symptoms) and Item 26 (depression symptoms). The highest loadings of the second factor were located mainly in the items measuring social dysfunction symptoms. Finally, with the exception of Item 26, the highest loadings of the third factor were found in the items indicating depression symptoms. The loading of Item 22 was very similar for the first and third factor, but it was classified within the third factor in agreement with previous reports. When comparing the victims of terrorism and their relatives with regards to the factor scores, we found significant differences for the first ($p = .046$), second ($p <$

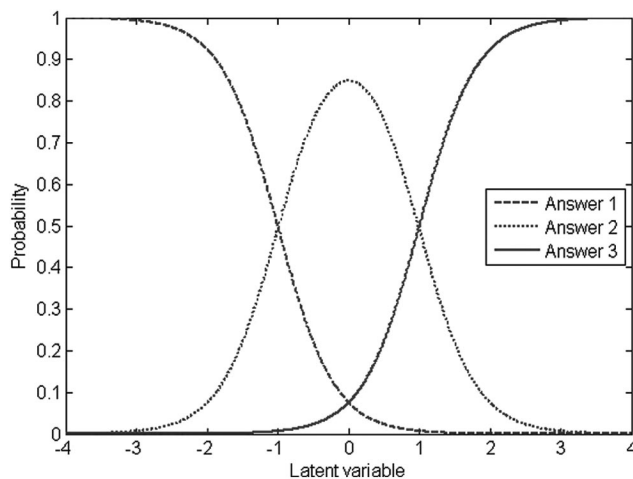


Figure 2. Example of an item characteristic curve with three possible answers: The discriminative parameter (a) is 2.5, and the location parameters (b_i) are -1 and 1 .

Table 3

Correlation of the GHQ Global Score and the Obtained Factors With the State and Trait Global Scores of Anxiety, and Cronbach's Alphas for STAI Internal Validity

GHQ	STAI	
	State score $r(p)$	Trait score $r(p)$
Global score	.80 (<.0001)	.79 (<.0001)
Factor I	.67 (<.0001)	.65 (<.0001)
Factor II	.37 (<.0001)	.33 (<.0001)
Factor III	.35 (<.0001)	.39 (<.0001)
STAI Internal validity	State α	Trait α
Victims of terrorism	0.9596	0.9337
Relatives	0.9449	0.9189
Total	0.9492	0.9229

Note. GHQ = General Health Questionnaire; STAI = State-Trait Anxiety Inventory; r = Pearson's coefficient.

.001), and third factor ($p = .028$). Factor structure and internal validity of samples separated by gender or class (victims or family members) showed no differences.

External validity. Table 3 shows the correlation of the GHQ-28 global score and the obtained factors with the state and trait global scores of anxiety according to the STAI. The first factor had the greatest correlation with the STAI resulting in a state score of 0.67 and a trait STAI score of 0.65.

IRT-based analyses. The factor analysis indicates that the GHQ-28 measures three dimensions or latent variables. We conducted an IRT-based analysis for each. On the basis of the factor analysis, items numbering from 1 to 14 and Item 26 were associated with the first latent variable, items numbering from 15 to 21 were associated with the second latent variable, and the remaining items were associated with the third latent variable. Table 4 shows the discrimination and location parameters for each item. The ICCs of the items with the highest and lowest expected average value for each dimension are displayed in Figure 2. We compared the average latent variables for each factor according to the IRT between the victims of terrorism and their relatives using t tests. We found significant differences ($p < .001$) between the groups in the three factors.

Discussion

In this study, we analyzed the psychometric properties, including an item analysis using IRT, of the GHQ-28 in a sample of terrorism victims and their relatives. The reliability of the scores of GHQ-28 was high, as demonstrated by an alpha value of over 0.955 in the total sample (Bland & Altman, 1997). Moreover, with the exception of Item 15 ("Have you recently felt that you are playing a useful part in things?"), all items exhibited a reasonably good adjusted item-scale correlation with values ranging from .45 to .78. The factor structure and internal validity did not differ when separating the sample by gender or class (victims or family members). Previous literature supports a continuous increase of risk for mental disorders depending on the degree to which participants are affected by a terrorist act, independently of being directly affected or through a familiar bond (Baca, Cabanas, Perez-Rodriguez, &

Table 4

Discrimination and Location Parameters for Each Item Sorted by Its Expected Average Score

	Item (Have you recently . . .)	A	B1	B2	B3	Av. score
Factor 1						
12	been getting scared or panicky for no good reason?	2.11	0.15	1.34	2.29	1.48
7	been having hot or cold spells?	1.66	0.01	1.31	2.71	1.57
26	found at times you couldn't do anything because your nerves were too bad?	2.42	-0.11	1.04	2.17	1.61
6	been getting a feeling of tightness or pressure in your head?	1.86	-0.19	0.95	2.19	1.62
4	felt that you are ill?	1.88	-0.29	1.00	2.25	1.66
13	been satisfied with the way you've carried out your task?	2.70	-0.33	0.72	1.85	1.70
14	felt constantly under strain?	3.03	-0.41	0.68	1.81	1.73
2	been feeling in need of a good tonic?	1.78	-0.48	0.96	2.49	1.76
9	had difficulty in staying asleep once you are off?	1.90	-0.82	0.55	1.83	1.80
3	been feeling run down and out of sorts?	2.54	-0.71	0.67	1.94	1.84
5	been getting any pains in your head?	1.56	-0.93	0.85	2.43	1.85
11	been getting edgy and bad-tempered?	3.35	-0.81	0.41	1.56	1.86
8	lost much sleep over worry?	2.10	-0.98	0.51	1.81	1.87
10	felt constantly under strain?	3.33	-0.70	0.44	1.75	1.89
1	been feeling perfectly well and in good health?	1.52	-2.72	0.78	2.68	2.13
Factor 2						
15	been managing to keep yourself busy and occupied?	0.94	-1.91	2.27	3.89	1.85
20	felt capable of making decisions about things?	1.73	-1.43	1.31	2.76	1.93
19	felt that you are playing a useful part in things?	1.79	-1.66	1.24	2.49	1.95
18	been satisfied with the way you've carried out your tasks?	3.22	-1.11	0.94	2.18	1.95
17	felt on the whole you were doing things well?	3.52	-1.40	1.17	2.35	1.98
21	been able to enjoy your normal day-to-day activities?	1.56	-1.98	0.90	2.47	2.02
16	been taking longer than usual to do things?	1.65	-2.54	0.88	2.50	2.09
Factor 3						
27	found yourself wishing you were dead and away from it all?	3.50	0.84	1.59	2.62	1.27
28	found that the idea of taking your own life kept coming into your mind?	2.46	0.85	1.62	2.70	1.28
25	thought of the possibility that you might do away with yourself?	2.55	0.73	1.51	2.44	1.31
24	felt that life isn't worth living?	3.21	0.47	1.32	2.11	1.36
22	been thinking of yourself as a worthless person?	2.15	0.51	1.31	2.19	1.36
23	felt that life is entirely hopeless?	2.45	0.41	1.27	2.10	1.38

Note. A = discrimination parameter; B1, B2, and B3 = location parameters of the three possible responses to each item; Av. score = Average score of the item.

Baca-Garcia, 2004). These reasons lead us to analyze victims of terrorism and their family members together.

The factor analysis identified three factors. The highest loadings of the first factor appeared in the group of anxiety and somatic symptoms. The highest loadings of the second and third factor were associated with the social dysfunction and depression items, respectively. Previous reports identified similar factors (Aderibigbe, Riley, Lewin, & Gureje, 1996; Gibbons, Flores de Arévalo, & Mónico, 2004; Vallejo, Jordan, Diaz, Comeche, & Ortega, 2007; Werneke, Goldberg, Yalcin, & Ustun, 2000). Notably, the highest loadings of the third factor correspond to symptoms of severe depression associated with suicidal behavior (see Items 24, 25, 27, and 28 in Table 2). This finding goes in line with the increased risk of suicide that has been reported in a sample of victims after the 11 March terrorist attacks in Madrid (Conejo-Galindo et al., 2008). Finally, the highly significant correlation between the STAI scores and the three factors of the GHQ-28 indicates a good external validity.

After varimax rotation, the first factor exhibited more variance (28.4%) than the other two jointly (13.8% and 13.2% for the second and third factor, respectively). Consequently, the GHQ-28 could be interpreted as a one-dimensional measure of psychiatric disorders (Banks et al., 1980) or, on the contrary, as a multidimensional measure in which all factors are needed to establish an

individual's profile (Huppert, Walters, Day, & Elliott, 1989). A closer analysis of the first factor may favor a one-dimensional view in this sample. In support of this view, the highest loadings of the first factor were found in the anxiety items. The predominance of anxiety symptoms seems reasonable in a population who is particularly at risk for anxiety disorders due to their traumatic experiences (Stoddard et al., 2011). Moreover, the highly significant correlation between both the trait and state scores of the STAI and the three factors of the GHQ-28 supports the central role of anxiety in this population. Somatic symptoms, which are frequently associated with underlying anxiety (Romera et al., 2010; Zhu et al., 2012), also presented high loadings in the first factor.

The IRT analysis demonstrates the importance of items within each identified factor. Interestingly, we found a certain gradient in the first factor that could reflect the severity of anxiety symptoms. We found that the five most discriminative items were Items 12 and 26, both related with subjective feelings of anxiety, and Items 7, 6, and 4, related with somatic symptoms of anxiety. On the contrary, the least discriminative items are those questions that address anxiety symptoms indirectly. For instance, Item 1 ("good health"), Item 10 ("under strain"), or Item 8 ("lost sleep"). Within the second factor (social dysfunction), the most discriminative items appeared to be closely related with the capacity to commit

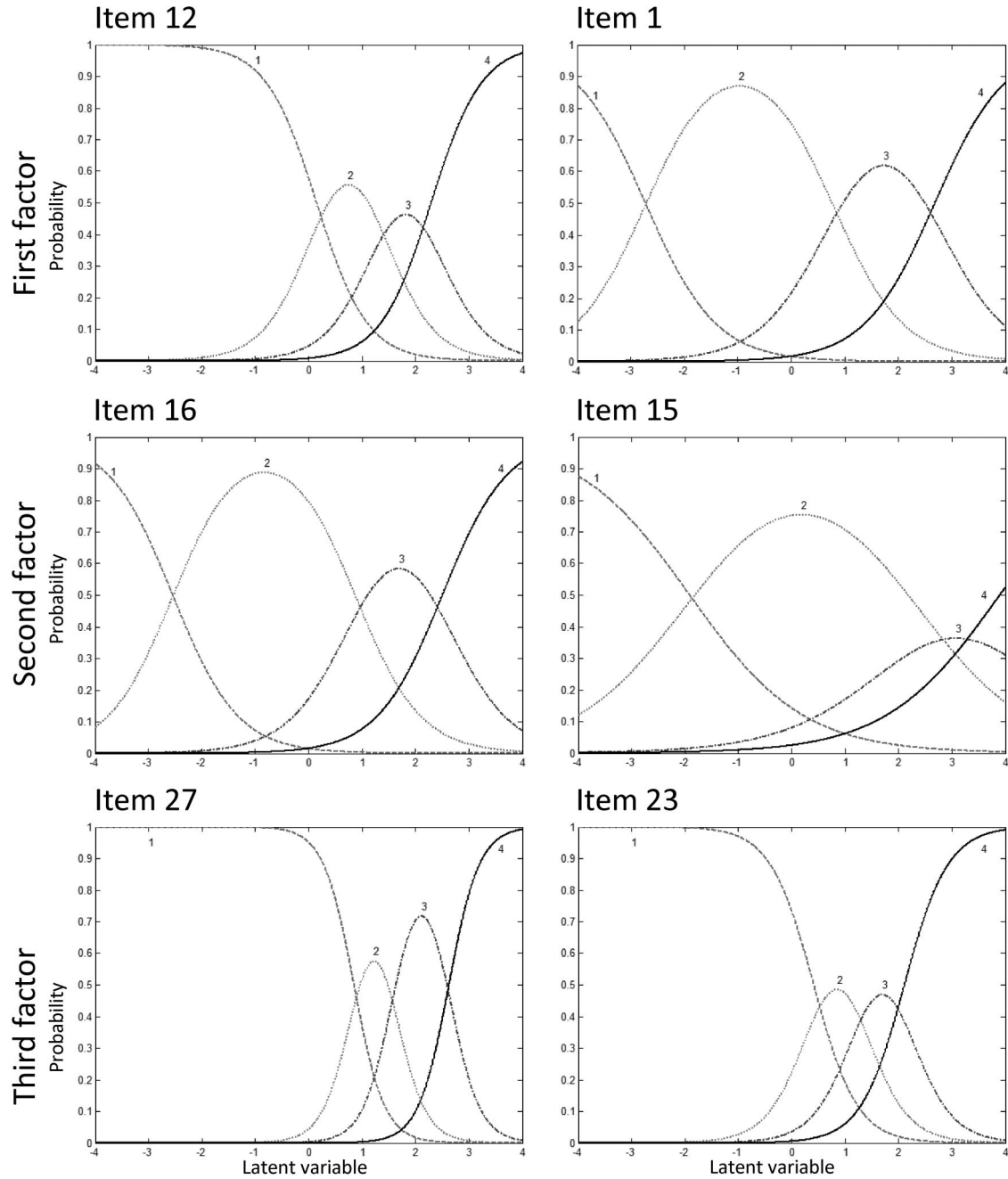


Figure 3. Item characteristic curves of the items with the lowest and highest expected average score of each factor.

oneself to any task (Item 15) and to making decisions (Item 20). Finally, the most discriminative items among the depression symptoms seem to be the items regarding passive or active ideas of suicide (Items 25, 27, and 28). On the contrary, low self-esteem and hopelessness (Items 22 and 23), usual descriptors of depression, appeared to be the least discriminative.

The ICCs (see Figure 3) show that items included in the same factor may have very different discrimination powers. For instance, it can be observed that the latent variable in Item 1 has a

wider range of values of location for each answer that can be chosen. This implies that small differences in the latent variable will not result in different response options. Thus, Item 1 is very discriminative and can provide accurate information on the actual value of the latent variable (the anxiety factor). Item 12 is the least discriminative item in this factor. Small changes in the value of the latent variable may change the participant's response and therefore provides less valuable information. In the third factor, the small differences between the ICCs of the items indicate a similar

capacity to discriminate the changes in the latent variable (depression).

This study presents several strengths. We examined a representative sample of victims of terrorism and their relatives. The traditional analyses of GHQ-28 were enhanced with a politomic model of item response. However, the validity of our results could have been confirmed with other sources, such as a clinical assessment or other screening scales. We had no information of the diagnoses assigned to the participants in the sample; therefore, gender and age effects could not be analyzed regarding the diagnoses of mental disorders. The time gap between the terrorist attacks and the GHQ-28 assessment is another limitation. However, several studies have demonstrated the long-term effects of terrorist attacks on direct victims and their relatives (Baca et al., 2002; DiGrande et al., 2011; L. Ferrando et al., 2011; Kawana et al., 2001; North et al., 2011). It appears symptoms may be more stable 1 year after the traumatic event rather than in the immediate aftermath (Brewin, 2005).

The analysis of the GHQ-28 applied to this sample of victims of terrorism and their relatives demonstrated a high reliability and a good external validity, as well as a three-factor structure. The first factor had the highest loadings in anxiety symptoms, suggesting that anxiety could be the most frequent psychopathological finding in this population. A selection of the most discriminative items, based on the IRT, could be used to improve the screening for detection of mental disorders.

References

- Aderibigbe, Y. A., Riley, W., Lewin, T., & Gureje, O. (1996). Factor structure of the 28-item general health questionnaire in a sample of antenatal women. *International Journal of Psychiatry in Medicine*, 26, 263–269. doi:10.2190/3XAV-M1BC-DA2B-DCMF
- Baca, E., Cabanas, M. L., & Baca-García, E. (2002). Terrorist attacks and short-long time psychiatric morbidity. *Actas Espanolas de Psiquiatria*, 30, 85–90.
- Baca, E., Cabanas, M. L., Perez-Rodriguez, M. M., & Baca-García, E. (2004). Mental disorders in victims of terrorism and their families. *Medicina Clinica*, 122, 681–685.
- Baksheev, G. N., Robinson, J., Cosgrave, E. M., Baker, K., & Yung, A. R. (2011). Validity of the 12-item General Health Questionnaire (GHQ-12) in detecting depressive and anxiety disorders among high school students. *Psychiatry Research*, 187, 291–296. doi:10.1016/j.psychres.2010.10.010
- Banks, M. H., Clegg, C. W., Jackson, P. R., Kemp, N. J., Stafford, E. M., & Wall, T. D. (1980). The use of the general health questionnaire as indicator of mental health in occupational studies. *Journal of Occupational Psychology*, 53, 187–194. doi:10.1111/j.2044-8325.1980.tb00024.x
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal*, 314, 572. doi:10.1136/bmj.314.7080.572
- Boscarino, J. A., Figley, C. R., & Adams, R. E. (2004). Compassion fatigue following the September 11 terrorist attacks: A study of secondary trauma among New York City social workers. *International Journal of Emergency Mental Health*, 6, 57–66.
- Breslau, N., Kessler, R. C., Chilcoat, H. D., Schultz, L. R., Davis, G. C., & Andreski, P. (1998). Trauma and posttraumatic stress disorder in the community: The 1996 Detroit area survey of trauma. *Archives of General Psychiatry*, 55, 626–632. doi:10.1001/archpsyc.55.7.626
- Brewin, C. R. (2005). Systematic review of screening instruments for adults at risk of PTSD. *Journal of Traumatic Stress*, 18, 53–62. doi:10.1002/jts.20007
- Brewin, C. R., Fuchkan, N., Huntley, Z., & Scragg, P. (2010). Diagnostic accuracy of the Trauma Screening Questionnaire after the 2005 London bombings. *Journal of Traumatic Stress*, 23, 393–398.
- Brewin, C. R., Scragg, P., Robertson, M., Thompson, M., d'Ardenne, P., & Ehlers, A. (2008). Promoting mental health following the London bombings: A screen and treat approach. *Journal of Traumatic Stress*, 21, 3–8. doi:10.1002/jts.20310
- Conejo-Galindo, J., Medina, Ó., Fraguas, D., Terán, S., Sainz-Cortón, E., & Arango, C. (2008). Psychopathological sequelae of the 11 March terrorist attacks in Madrid: An epidemiological study of victims treated in a hospital. *European Archives of Psychiatry and Clinical Neuroscience*, 258, 28–34. doi:10.1007/s00406-007-0758-7
- Costa, E., Barreto, S. M., Uchoa, E., Firmo, J. O., Lima-Costa, M. F., & Prince, M. (2006). Is the GDS-30 better than the GHQ-12 for screening depression in elderly people in the community? The Bambui Health Aging Study (BHAS). *International Psychogeriatrics*, 18, 493–503. doi:10.1017/S1041610205002954
- de Jesus Mari, J., & Williams, P. (1986). Misclassification by psychiatric screening questionnaires. *Journal of Chronic Diseases*, 39, 371–378. doi:10.1016/0021-9681(86)90123-2
- DiGrande, L., Neria, Y., Brackbill, R. M., Pulliam, P., & Galea, S. (2011). Long-term posttraumatic stress symptoms among 3,271 civilian survivors of the September 11, 2001, terrorist attacks on the World Trade Center. *American Journal of Epidemiology*, 173, 271–281. doi:10.1093/aje/kwq372
- Ferrando, L., Galea, S., Sainz Cortón, E., Mingote, C., García Camba, E., Fernández Liria, A., & Gabriel, R. (2011). Long-term psychopathology changes among the injured and members of the community after a massive terrorist attack. *European Psychiatry*, 26, 513–517. doi:10.1016/j.eurpsy.2010.07.009
- Ferrando, P. J. (2001). The measurement of neuroticism using MMQ, MPI, EPI and EPQ items: A psychometrical analysis based on item response theory. *Personality and Individual Differences*, 30, 641–656. doi:10.1016/S0191-8869(00)00062-3
- Gibbons, P. F., Flores de Arévalo, H., & Mónico, M. (2004). Assessment of the factor structure and reliability of the 28 item version of the General Health Questionnaire (GHQ-28) in El Salvador. *International Journal of Clinical and Health Psychology*, 4, 389–398.
- Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139–145. doi:10.1017/S0033291700021644
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7, 191–205. doi:10.1177/1094428104263675
- Henriksen, C. A., Bolton, J. M., & Sareen, J. (2010). The psychological impact of terrorist attacks: Examining a dose-response relationship between exposure to 9/11 and Axis I mental disorders. *Depression and Anxiety*, 27, 993–1000. doi:10.1002/da.20742
- Hobbs, P., Ballinger, C. B., Greenwood, C., Martin, B., & McClure, A. (1984). Factor analysis and validation of the general health questionnaire in men: A general practice survey. *British Journal of Psychiatry*, 144, 270–275. doi:10.1192/bjp.144.3.270
- Hoven, C. W., Duarte, C. S., Wu, P., Doan, T., Singh, N., Mandell, D. J., . . . Cohen, P. (2009). Parental exposure to mass violence and child mental health: The First Responder and WTC Evacuee Study. *Clinical Child and Family Psychology Review*, 12, 95–112. doi:10.1007/s10567-009-0047-2
- Huppert, F. A., Walters, D. E., Day, N. E., & Elliott, B. J. (1989). The factor structure of the General Health Questionnaire (GHQ-30). A reliability study on 6317 community residents. *British Journal of Psychiatry*, 155, 178–185. doi:10.1192/bjp.155.2.178

- Kawana, N., Ishimatsu, S., & Kanda, K. (2001). Psycho-physiological effects of the terrorist sarin attack on the Tokyo subway system. *Military Medicine*, 166, 23–26.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry*, 52, 1048–1060. doi:10.1001/archpsyc.1995.03950240066012
- Lobo, A., & Muñoz, P. E. (1996). *Versiones en lengua española validadas. Cuestionario de Salud General GHQ (General Health Questionnaire). Guía para el usuario de las distintas versiones* [Validated versions in Spanish language. General Health Questionnaire. A user's guide for the different versions]. Barcelona, Spain: Masson.
- National Counter Terrorism Center. (2011). *2010 Report on Terrorism*. Retrieved from www.nctc.gov
- North, C. S., Pfefferbaum, B., Kawasaki, A., Lee, S., & Spitznagel, E. L. (2011). Psychosocial adjustment of directly exposed survivors 7 years after the Oklahoma City bombing. *Comprehensive Psychiatry*, 52, 1–8. doi:10.1016/j.comppsy.2010.04.003
- Pfefferbaum, B., North, C. S., Flynn, B. W., Norris, F. H., & DeMartino, R. (2002). Disaster mental health services following the 1995 Oklahoma City bombing: Modifying approaches to address terrorism. *CNS Spectrums*, 7, 575–579.
- Ploubidis, G. B., & Frangou, S. (2011). Neuroticism and psychological distress: To what extent is their association due to person-environment correlation? *European Psychiatry*, 26, 1–5. doi:10.1016/j.eurpsy.2009.11.003
- Politi, P. L., Piccinelli, M., & Wilkinson, G. (1994). Reliability, validity and factor structure of the 12-item General Health Questionnaire among young males in Italy. *Acta Psychiatrica Scandinavica*, 90, 432–437. doi:10.1111/j.1600-0447.1994.tb01620.x
- Quek, K. F., Low, W. Y., Razack, A. H., & Loh, C. S. (2001). Reliability and validity of the General Health Questionnaire (GHQ-12) among urological patients: A Malaysian study. *Psychiatry and Clinical Neurosciences*, 55, 509–513. doi:10.1046/j.1440-1819.2001.00897.x
- Quek, K. F., Low, W. Y., Razack, A. H., Loh, C. S., & Chua, C. B. (2004). Reliability and validity of the Spielberger State-Trait Anxiety Inventory (STAI) among urological patients: A Malaysian study. *Medical Journal of Malaysia*, 59, 258–267.
- Richard, C., Lussier, M. T., Gagnon, R., & Lamarche, L. (2004). GHQ-28 and cGHQ-28: Implications of two scoring methods for the GHQ in a primary care setting. *Social Psychiatry and Psychiatric Epidemiology*, 39, 235–243. doi:10.1007/s00127-004-0710-3
- Romans-Clarkson, S. E., Walton, V. A., Herbison, G. P., & Mullen, P. E. (1989). Validity of the GHQ-28 in New Zealand women. *Australian and New Zealand Journal of Psychiatry*, 23, 187–196. doi:10.3109/00048678909062135
- Romera, I., Fernandez-Perez, S., Montejo, A. L., Caballero, F., Caballero, L., Arbesu, J. A., . . . Gilaberte, I. (2010). Generalized anxiety disorder, with or without co-morbid major depressive disorder, in primary care: Prevalence of painful somatic symptoms, functioning and health status. *Journal of Affective Disorders*, 127, 160–168. doi:10.1016/j.jad.2010.05.009
- Salguero, J. M., Cano-Vindel, A., Iruarizaga, I., Fernandez-Berrocá, P., & Galea, S. (2011). Trajectory and predictors of depression in a 12-month prospective study after the Madrid March 11 terrorist attacks. *Journal of Psychiatric Research*, 45, 1395–1403. doi:10.1016/j.jpsychires.2011.05.012
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34(4, Pt. 2).
- Sánchez-López, M. P., & Dresch, V. (2008). The 12-Item General Health Questionnaire (GHQ-12): Reliability, external validity and factor structure in the Spanish population. *Psicothema*, 20, 839–843.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Stoddard, F. J., Jr., Gold, J., Henderson, S. W., Merlino, J. P., Norwood, A., Post, J. M., . . . Katz, C. L. (2011). Psychiatry and terrorism. *Journal of Nervous and Mental Disease*, 199, 537–543. doi:10.1097/NMD.0b013e318225ee90
- Tarnopolsky, A., Hand, D. J., McLean, E. K., Roberts, H., & Wiggins, R. D. (1979). Validity and uses of a screening questionnaire (GHQ) in the community. *British Journal of Psychiatry*, 134, 508–515. doi:10.1192/bjp.134.5.508
- Thomas, M. L. (2011). The value of item response theory in clinical assessment: A review. *Assessment*, 18, 291–307. doi:10.1177/1073191110374797
- Tran, T. D., Tran, T., & Fisher, J. (2012). Validation of three psychometric instruments for screening for perinatal common mental disorders in men in the north of Vietnam. *Journal of Affective Disorders*, 136, 104–109. doi:10.1016/j.jad.2011.08.012
- Ursano, R. J., & Friedman, M. J. (2006). Mental health and behavioral interventions for victims of disasters and mass violence systems, caring, planning, and needs. In E. C. Ritchie, M. J. Friedman, & P. J. Watson (Eds.), *Interventions following mass violence and disasters: Strategies for mental health practice* (pp. 405–414). New York, NY: Guilford Press.
- Vallejo, M. A., Jordan, C. M., Diaz, M. I., Comeche, M. I., & Ortega, J. (2007). Psychological assessment via the Internet: A reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *Journal of Medicine Internet Research*, 9, e2. doi:10.2196/jmir.9.1.e2
- van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York, NY: Springer.
- Werneke, U., Goldberg, D. P., Yalcin, I., & Ustun, B. T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, 30, 823–829. doi:10.1017/S0033291799002287
- Zhang, J., & Gao, Q. (2012). Validation of the trait anxiety scale for state-trait anxiety inventory in suicide victims and living controls of Chinese rural youths. *Archives of Suicide Research*, 16, 85–94. doi:10.1080/13811118.2012.641440
- Zhu, C., Ou, L., Geng, Q., Zhang, M., Ye, R., Chen, J., & Jiang, W. (2012). Association of somatic symptoms with depression and anxiety in clinical patients of general hospitals in Guangzhou, China. *General Hospital Psychiatry*, 34, 113–120. doi:10.1016/j.genhosppsych.2011.09.005

Received March 23, 2012

Revision received September 14, 2012

Accepted September 17, 2012 ■