



Screening Tool for Anxiety Disorders: Development and Validation of the Korean Anxiety Screening Assessment

Yeseul Kim¹, Yeonsoo Park², Gyeongcheol Cho³, Kiho Park¹, Shin-Hyang Kim¹,
Seung Yeon Baik², Cho Long Kim², Sooyun Jung¹, Won-Hye Lee⁴, Younyoung Choi⁵ ✉,
Seung-Hwan Lee^{2,6} ✉, and Kee-Hong Choi¹ ✉

¹Department of Psychology, Korea University, Seoul, Republic of Korea

²Clinical Emotion and Cognition Research Laboratory, Inje University Ilsan Paik Hospital, Goyang, Republic of Korea

³Department of Psychology, McGill University, Montreal, Quebec, Canada

⁴Department of Clinical Psychology, National Center for Mental Health, Seoul, Republic of Korea

⁵Department of Adolescent Psychology, Hanyang Cyber University, Seoul, Republic of Korea

⁶Department of Psychiatry, Inje University Ilsan Paik Hospital, Goyang, Republic of Korea

Objective This study evaluated the psychometric properties of the Korean Anxiety Screening Assessment (K-ANX) developed for screening anxiety disorders.

Methods Data from 613 participants were analyzed. The K-ANX was evaluated for reliability using Cronbach's alpha, item-total correlation, and test information curve, and for validity using focus group interviews, factor analysis, correlational analysis, and item characteristics based on item response theory (IRT). The diagnostic sensitivity and specificity of the K-ANX were compared with those of the Beck Anxiety Inventory (BAI) and Generalized Anxiety Disorder 7-item scale (GAD-7).

Results The K-ANX showed excellent internal consistency ($\alpha=0.97$) and item-total coefficients (0.92–0.97), and a one-factor structure was suggested. All items were highly correlated with the total scores of the BAI, GAD-7, and Penn State Worry Questionnaire. IRT analysis indicated the K-ANX was most informative as a screening tool for anxiety disorders at the range between 0.8 and 1.6 (i.e., top 21.2 to 5.5 percentiles). Higher sensitivity (0.795) and specificity (0.937) for identifying anxiety disorders were observed in the K-ANX compared to the BAI and GAD-7.

Conclusion The K-ANX is a reliable and valid measure to screen anxiety disorders in a Korean sample, with greater sensitivity and specificity than current measures of anxiety symptoms.

Psychiatry Investig 2018;15(11):1053-1063

Key Words Screening tool, Anxiety disorder, GAD, Item response theory, Validation.

INTRODUCTION

The US Preventive Services Task Force (USPSTF) reported that screening for mental disorders in primary care would reduce the prevalence rates by identifying patients with men-

tal disorders in advance and provide an optimal and timely treatment. Therefore, the Korean Ministry for Health and Welfare Affairs also placed greater emphasis in their policy for the prevention and early intervention of mental illnesses. The lifetime prevalence of anxiety disorders in South Korea

Received: June 4, 2018 **Revised:** August 16, 2018 **Accepted:** September 27, 2018

✉ **Correspondence:** Kee-Hong Choi, PhD

Department of Psychology, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea

Tel: +82-2-3290-2867, **Fax:** +82-2-3290-5390, **E-mail:** keehongchoi@gmail.com

✉ **Correspondence:** Seung-Hwan Lee, MD

Department of Psychiatry, Inje University Ilsan Paik Hospital, 170 Juhwa-ro, Ilsanseo-gu, Goyang 10380, Republic of Korea; Clinical Emotion and Cognition Research Laboratory, Inje University Ilsan Paik Hospital, 170 Juhwa-ro, Ilsanseo-gu, Goyang 10380, Republic of Korea

Tel: +82-31-910-7768, **Fax:** +82-31-910-7268, **E-mail:** lshpss@hanmail.net

✉ **Correspondence:** Younyoung Choi, PhD

Department of Adolescent Psychology, Hanyang Cyber University, 220 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

Tel: +82-2-2290-0367, **Fax:** +82-2-2290-0600, **E-mail:** 1120008@hycu.ac.kr

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

has increased from 6.9% in 2006 to 8.7% in 2011 and recently to 9.3% in 2016 according to the epidemiologic survey of psychiatric conditions.¹⁻³ There is therefore an increasing need for easily accessible and highly reliable and valid screening instruments for anxiety disorders. However, such screening tools are limited especially in South Korea owing to the paucity of psychometric validation studies and lack of consensus on common anxiety symptoms of Koreans.⁴⁻⁶ Importantly, with advances in psychometric methods, diagnostic sensitivity and specificity of screening tools for anxiety disorders are considered an area to be improved.⁷

Although some screening tools for anxiety disorders have already been developed, they seldom meet all of the required properties and have limited reliability and validity. For instance, the Beck Anxiety Inventory (BAI), one of the most widely used instruments in primary care settings, turned out to be a relatively poor screening measure for symptoms of clinically significant anxiety in a primary care population.⁸ The initial aim of the BAI was to distinguish anxiety symptoms from depressive symptoms of patients with depressive disorders, resulting in its heavy focus on somatic symptoms. This imbalance in assessment aspects inhibits the diagnostic screening capacity of the BAI.^{9,10} Various studies have suggested that anxiety disorders are not likely to be accompanied by a high score on the BAI.¹¹

The Generalized Anxiety Disorder 7-item scale (GAD-7) was developed for the clear purpose of screening for and assessing the severity of generalized anxiety disorder (GAD).¹² Since the GAD-7 identifies GAD effectively in a short time, it is widely used in primary care. According to a systematic review and meta-analysis, the GAD-7 appears to demonstrate relatively acceptable sensitivity and specificity for detecting any anxiety disorders at a cut-off point of 8 (sensitivity=0.77 to 0.91, specificity=0.74 to 0.83).¹³ Albeit its usefulness, the GAD-7 does not cover all of the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5)¹⁴ diagnostic domains, such as difficulty concentrating, sleep disturbance, and impairment of functioning.¹⁵ Especially in South Korea, the validation study of the GAD-7 was conducted only on a sample of migraine patients, thereby limiting the use of the same cutoff scores for the general population.¹⁶

More recently, Gibbons and his colleagues developed a computerized adaptive test (CAT) based on item response theory (IRT) for anxiety disorders, named the Computerized Adaptive Testing-Anxiety Inventory (CAT-ANX). IRT provides an advanced psychometric algorithm considering both item characteristics and item response patterns of respondents.¹⁷ By computing an estimate of severity of anxiety disorders (i.e., latent trait) and an uncertainty estimate (i.e., standard error) of each item, the next item of maximizing in-

formation is automatically selected from the item pool. It is expected that an IRT based screening tool reflects not only item characteristics of anxiety disorders, but also specific item response patterns of the Korean population within smaller items. Gibbons et al.⁷ reported that the CAT-ANX demonstrated sensitivity and specificity of CAT-ANX both as 0.86 at a threshold of -0.85, and the tool is currently in clinical use. Likewise, we aimed to develop the IRT based Korean screening tool for anxiety disorders, K-ANX, which is composed of items with the best discriminative power for symptoms of GAD in Korean samples.

GAD was chosen out of all the other anxiety disorders because its central features such as excessive and uncontrollable worry and accompanying physiological symptoms, are common aspects shared by anxiety disorders.¹⁸ GAD also develops at a relatively early age and acts as a risk factor, or a gateway, for other anxiety disorders.¹⁹ For these reasons, GAD is now referred to as the basic anxiety disorder, and has become an optimal target for screening tools for general anxiety disorders.²⁰ Moreover, GAD is the most prevalent anxiety disorder in primary care,²¹ suggesting that it is essential to disseminate a reliable, valid and easily accessible anxiety disorders assessment tool that targets GAD symptoms which could be used in initial interviews that have time- and cost-constraints.

While developing a screening tool for GAD, there were three prerequisite considerations.¹⁵ First, a screening tool should contain or cover all specific diagnostic criteria of GAD. We considered all the GAD diagnostic criteria in the DSM-5 and the relevant dimensions of the RDoC (research domain criteria) for the current scale. Among the RDoC dimensions, sustained threat factors, which are the main symptom of GAD (i.e., wide-ranging anxiety and persistent anxiety symptoms), are considered and matched to the corresponding part of the DSM-5 domain. Second, considering the nature of the screening tool, the time and cost required were minimized for easy use and high accessibility. Third, despite the smaller number of items, we aimed to develop a tool with greater diagnostic sensitivity and specificity than previous tools.

The primary purpose of the current study was to evaluate the psychometric properties (i.e., reliability and validity) of the K-ANX to examine whether it would perform properly to screen GAD and other anxiety disorders in a Korean sample. Specifically, we examined the reliability of the K-ANX using Cronbach's α , item-total correlation, and standard error of measurement. Second, we collected validity evidence regarding content, construct, criterion, and response process. Lastly, the diagnostic sensitivity and specificity of the K-ANX were examined and compared with those of the BAI and GAD-7.

METHODS

Design and setting

The methodology of Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) was applied to evaluate the usefulness of the K-ANX as a screening tool.²² In the QUADAS-2, screening tools are evaluated based on patient selection, index test, reference standard, and flow and timing. In order to meet the criterion of the patient selection domain, participants for this research were recruited based on random sampling. In order to avoid bias in the index test phase, the K-ANX was administered while blinded to diagnostic decisions or reference standard information, and vice versa. As a reference test, the Mini International Neuropsychiatric Interview-Plus Version 5.0.0 (MINI), a simple structured diagnostic interview, was utilized.²³ The diagnostic interview was conducted by graduate students majoring in clinical psychology and their diagnostic decisions were supervised by licensed psychologists and psychiatrists. Lastly, anxiety assessment tools already widely used in the field were administered to verify the convergent validity of the K-ANX. This research was approved by the local Institutional Review Boards [1040548-KU-IRB-15-92-A-1(R-A-1)(R-A-2)(R-A-2), ISPAIK 2015-05-221-009]. All the subjects completed an informed consent form before participating in this study.

Participants

A total of 653 individuals were originally recruited for this study in two ways: by responding to an online recruiting advertisement or by introduction of potentially interested individuals by hospital staff. All participants voluntarily participated in this research. The inclusion criterion was not specific except that participants should be aged 19 years or older. Neither was the exclusion criterion specific, in order to maintain the benefit of random sampling. However, participants who did not finish the questionnaire or patients who could not answer the questions appropriately because of their medical or psychiatric symptoms were excluded from the analysis. Thus, responses from 613 participants were included in the final analysis.

Instruments

Korean Anxiety Screening Assessment

The K-ANX is a scale to screen anxiety disorders newly developed by the researchers of this study. In order to develop the K-ANX, a GAD item bank consisting of 273 items was constructed. Eleven items were selected considering the following criteria. First, correlation between individual items and total scores or other anxiety scales such as the BAI, GAD-7, and PSWQ were considered. Second, experts including

clinical psychologists, psychiatrists, and psychometricians discussed which items could appropriately reflect the construct of GAD. Third, analysis based on IRT was performed. Item characteristic curves for each item and item parameters such as difficulty and discrimination indices were considered. Item information curves were also utilized to determine which items gave the most information and lowest measurement error at a certain level of anxiety. Lastly, the final eleven items were composed to maximize diagnostic sensitivity and specificity while considering the diagnostic criteria of GAD from the DSM-5. The items were edited under the supervision of the National Institute of Korean Language.

Mini-international neuropsychiatric interview-plus version 5.0.0

The MINI is a structured interview that was developed for the diagnosis of major Axis 1 mental disorders from the International Classification of Diseases-10th Revision (ICD-10) and DSM-IV.²³ The Korean version of the M.I.N.I. was used in this study, which has been found to have a good level of diagnostic accuracy.²⁴ Excellent inter-rater reliability on the M.I.N.I. diagnoses was found (ICC=0.92).

Anxiety measures

Beck Anxiety Inventory (BAI)

The BAI⁹ is a scale to measure the severity of anxiety. This self-report questionnaire consists of 21 items, and subjects are asked to rate how uncomfortable they have been due to the symptoms they had during the past week using a 4-point Likert scale. In this study, the Korean version of the BAI translated and validated by Lee et al.²⁵ was used. In the present study, the BAI showed a high level of internal consistency ($\alpha=0.956$).

Generalized Anxiety Disorder 7-item scale (GAD-7)

The GAD-7¹² screens for patients with GAD and assesses the severity of the symptoms with 7 items. Participants are asked how often they have been disturbed by anxiety symptoms during the past two weeks using a 4-point Likert scale. The Korean version of the GAD-7 is freely available on the Patient Health Questionnaire web-page.²⁶ The GAD-7 showed good internal consistency in the present study ($\alpha=0.932$).

Penn State Worry Questionnaire (PSWQ)

The PSWQ²⁷ is a 16-item self-report questionnaire that measures frequency and intensity of chronic and uncontrollable worry, which is the main symptom of GAD. The PSWQ consists of questions that ask about various aspects of anxiety, and each item is assessed on a 5-point Likert scale. The Kore-

an version of the PSWQ validated by Lim et al.²⁸ was utilized in this research. The PSWQ showed good internal consistency in the present study ($\alpha=0.826$).

Data analysis

Descriptive statistics were calculated including the means, standard deviations, skewness, kurtosis, and item-total correlations. The psychometric properties were computed regarding reliability and validity. Reliability values were evaluated by Cronbach's alpha, item-total correlation as internal consistency, and the standard measurement errors across different anxiety levels using IRT.²⁹ The sources of validity evidence were collected based on the Standard for Educational and Psychological Testing provided by AERA, APA, and NCME.³⁰ They recommend that a validation study should collect several sources of evidence. In this study, evidence based on content (i.e., content validity) was collected through a domain analysis by focus group interviews (FGIs) with patient groups and mental health professionals. Evidence based on internal structure (i.e., construct validity) was collected by exploratory and confirmatory factor analysis using Mplus.³¹ Evidence based on relationships with other variables (criterion validity) was evaluated by computing the correlation values with the BAI, GAD-7, and PSWQ. Also, evidence based on response process was assessed by estimating item characteristics across the different levels of anxiety using IRT by IRTpro.³² Lastly, the diagnostic sensitivity and specificity values of the K-ANX, BAI, GAD-7, and PSWQ were computed and compared.

RESULTS

Demographic information

The demographic characteristics and diagnostic information of the participants is provided in Table 1. Data from 613 participants including 194 males (31.6%) and 419 females (68.4%) were analyzed in this study. The average age of participants was 49.6 (SD=39.5) and the average education level of the participants was 14.8 years (SD=2.71). The psychiatric symptoms of all participants by the MINI examination were as follows: major depressive disorder (MDD) (n=40, 6.5%), depressive disorders (DD) (n=90, 14.7%), generalized anxiety disorder (GAD) (n=38, 6.2%), anxiety disorders (AD) (n=98, 16.0%), and comorbidity of DD and AD (n=41, 6.7%). In terms of suicidal tendency, 505 (82.4%) had no suicidal tendency, 62 (10.1%) had a low tendency, 38 (6.2%) were intermediate, and 8 (1.3%) had a high tendency. Lastly, 13 (2.1%) were diagnosed with bipolar type 1 and 4 (0.7%) with type 2.

Reliability of the K-ANX

Table 2 shows descriptive statistics for each K-ANX item.

The mean of each item score of the healthy group was lower than that of the anxiety disorders group. Cronbach's α was 0.97 and the item-total correlations ranged from 0.92 to 0.97. Further, test information curve (TIC) based on IRT were estimated to report the standard error of measurement across the different anxiety levels. Figure 1 shows the TIC and standard error of measurement curve for the K-ANX. The TIC indicates a certain range of anxiety levels (X axis) where the highest information and the lowest standard error measurement (Y axis) were provided. This result implies that the K-ANX offers the most informative diagnostic decisions with the highest reliability and the lowest standard error of measurement at an anxiety trait around 1.2SD (between 0.8SD and 1.6SD) above the mean.

Table 1. Demographic characteristics and diagnostic information of participants

		Mean	SD
Age		49.6	39.5
Educational years		14.8	2.71
		Categories	N
Marital Status	Married	224	36.5
	Single	358	58.4
	Others	31	5.1
Gender	Male	194	31.6
	Female	358	68.4
Diagnosed disorders	DD ¹	90	14.7
	MDD ²	40	6.5
	AD ³	98	16.0
	GAD ⁴	38	6.2
	Comorbidity ⁵	41	6.7
Suicidality ⁶	N/A (0)	505	82.4
	Low (1)	62	10.1
	Intermediate (2)	38	6.2
	High (3)	8	1.3
Bipolar disorders	N/A (0)	596	97.2
	Bipolar type I ⁷ (1)	13	2.1
	Bipolar type II ⁸ (2)	4	0.7
Total		613	100

¹DD: depressive disorders (including MDD, PDD, R/O MDD), ²MDD: major depressive disorder (including current episode, partial remission), ³AD: anxiety disorders (including GAD, social phobia, specific phobia, agoraphobia, panic disorder), ⁴GAD: generalized anxiety disorder, ⁵Comorbidity: depressive disorders and anxiety disorders occurring in the same person, ⁶Suicidality: the present level of current suicide risk and lifetime suicide attempts (low risk: 1–5, intermediate: 6–9, and high risk: > 9), ⁷Bipolar type I: characterized by at least one episode of mania, ⁸Bipolar type II: characterized by at least one episode of hypomania and at least one episode of one major depression). SD: standard deviation

Table 2. Descriptive statistics for each item

Item	Healthy group				Anxiety disorders group			
	Mean	SD	Skewness	Kurtosis	Mean	SD	Skewness	Kurtosis
Item 1	0.26	0.663	2.997	9.463	1.38	1.403	0.531	-1.203
Item 2	0.43	0.804	2.145	4.301	1.67	1.368	0.219	-1.366
Item 3	0.29	0.640	2.742	8.729	1.36	1.341	0.652	-0.837
Item 4	0.44	0.710	1.743	3.292	1.51	1.237	0.476	-0.622
Item 5	0.37	0.692	2.433	7.346	1.49	1.379	0.672	-0.825
Item 6	0.65	0.981	1.825	3.386	1.04	1.331	1.022	-0.324
Item 7	0.49	0.734	1.662	3.053	1.62	1.288	0.328	-1.026
Item 8	0.42	0.846	2.343	5.778	1.73	1.517	0.427	-1.210
Item 9	0.43	0.769	1.889	3.191	1.59	1.299	0.226	-1.175
Item 10	0.23	0.578	3.463	5.135	1.24	1.340	0.878	-0.466
Item 11	0.34	0.763	2.711	7.980	1.60	1.525	0.470	-1.195

Item 1: “(for the last 2 weeks) It was hard to endure anxiety even though I tried to tolerate it rationally;” Item 2: “(for the last 2 weeks) I could not control or stop worrying;” Item 3: “(for the last 2 weeks) I was nervous and restless;” Item 4: “(for the last 2 weeks) I was too tired to think of anything else;” Item 5: “(for the last 2 weeks) I could not concentrate on anything;” Item 6: “(for the last 2 weeks) All the stimuli from the surroundings bothered me;” Item 7: “(for the last 2 weeks) My head was heavy and my neck was stiff;” Item 8: “(for the last 2 weeks) I was tired all morning because I could not sleep well;” Item 9: “(for the last 2 weeks) I had difficulty in work and social life because I was anxious and irritated;” Item 10: “(for the last 2 weeks) I felt that my chest has been suddenly oppressed;” Item 11: “(for the last 2 weeks) I was nervous or tensed.” SD: standard deviation

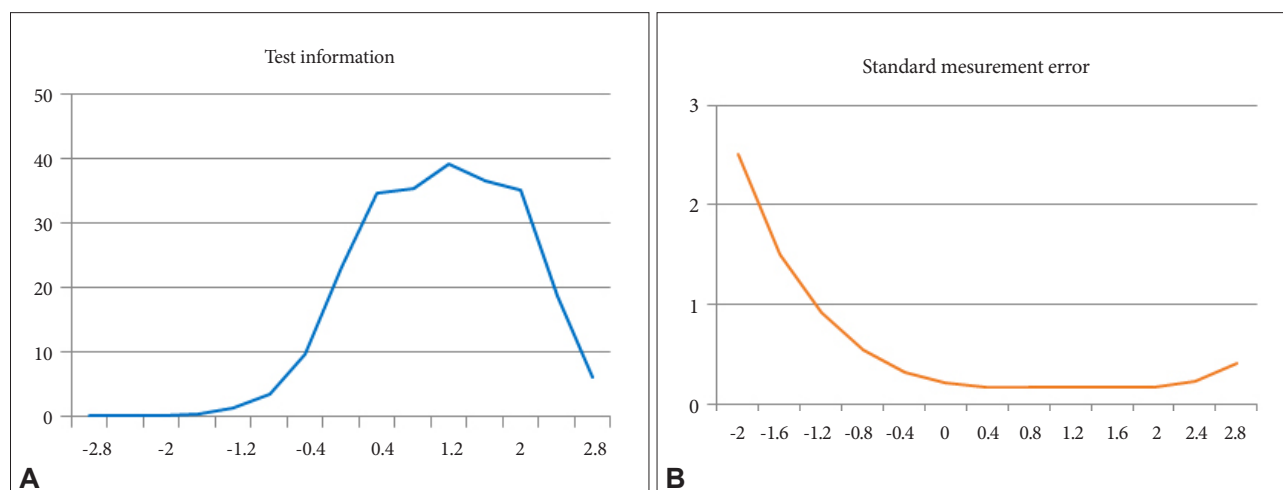


Figure 1. Test information curve (TIC) of the K-ANX (A); standard error of measurement of the K-ANX (B). K-ANX: Korean Anxiety Screening Assessment.

Content validity

Domain analysis regarding core diagnostic domains of GAD based on the DSM-5 was conducted by mental health professionals including clinical psychologists and psychiatrists. The K-ANX was evaluated to include all necessary GAD diagnostic domains of the DSM-5 which are excessive anxiety and worry, difficulty in controlling worry, restlessness, fatigue, difficulty in concentrating, irritability, muscle tension, sleep disturbance, and impaired functioning. Table 3 shows the domain elements that the K-ANX and other instruments cover.¹⁵

Factor structure of the K-ANX

Exploratory factor analysis (EFA) was conducted to explore the internal structure, while confirmatory factor analysis (CFA) was conducted to confirm EFA results and other theoretical candidate models considering DSM-5 and RoDOC. A principal axis factor analysis with a varimax rotation was conducted. An examination of the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy suggested that the sample was factorable (KMO=0.901) and that Bartlett's sphericity test was significant ($p<0.001$), which indicated that the data were suitable for factor analysis. The scree plot sug-

Table 3. DSM-5 GAD diagnostic domains of the K-ANX and other instruments

	1	2	3	4	5	6	7	8	9
K-ANX	•	•	•	•	•	•	•	•	•
HARS	•		•	•	•		•	•	•
STAI-X	•	•	•	•		•			
SAS	•		•	•		•	•	•	
STAI-Y	•	•	•	•		•			
BAI	•		•			•	•		
PSWQ	•	•							
GAD-Q-IV	•	•	•	•	•	•	•	•	•
GAD-7	•	•	•	•		•	•		

1. Excessive anxiety&worry, 2. Difficult to control worry, 3. Restlessness, 4. Fatigue, 5. Difficulty concentrating, 6. Irritability, 7. Muscle tension, 8. Sleep disturbance, 9. Impairment of functioning. Adapted from Kim et al. Kor J Clin Psychol 2016;35:635.¹⁵ GAD: generalized anxiety disorder, K-ANX: Korean Anxiety Screening Assessment, HARS: Hamilton Anxiety Rating Scale, STAI-X: State-Trait Anxiety Inventory-X, SAS: Zung's Self-Rating Anxiety Scale, STAI-Y: State-Trait Anxiety Inventory-Y, BAI: Beck Anxiety Inventory, PSWQ: Penn State Worry Questionnaire, GAD-Q-IV: Generalized Anxiety Disorder Questionnaire-IV, GAD-7: Generalized Anxiety Disorder 7-item scale

gested a one factor solution; the first eigenvalue was over 8 and the second eigenvalue was below one (Figure 2). The one factor solution accounted for 86.3% of the variance. All factor loadings were above 0.90 on the first factor (Table 4). This result implies that all items of the K-ANX possess a dominant domain explained by anxiety disorders. Also, based on the results of the CFA, the model fit statistics of the one-factor model showed an appropriate fit except for the RMSEA (Table 5) (the criterion are CFI and TLI above 0.90; RMSEA below 0.08; SRMR below 0.10).³³⁻³⁶

Correlations between the K-ANX and other instruments

This study used the BAI, GAD-7, and PSWQ to evaluate criterion validity. As displayed in Table 6, the total score of the K-ANX had high positive correlations with total scores on the GAD-7 ($r=0.821$), PSWQ ($r=0.653$), and BAI ($r=0.821$).

Item characteristics based on IRT

IRT was applied to evaluate item characteristics. The Graded Response Model (GRM),³⁷ as one of the IRT models, deals with ordered polytomous categories including rating scale and Likert scale. It estimates the discrimination parameter (α) and the boundary parameters between two adjacent categories (b_j) for each item. The former indicates how sensitive of each probability of response is according to the levels of the latent trait. The latter indicates the level of the latent trait needed to have a 50% probability of endorsing j category or higher, which reflects the fact that the actual intervals between adjacent categories are not exactly the same in Likert scales. These values were evaluated to see if each item worked well (Table 7). Also, the plot of the probability of response as a function of anxiety level for each option for each item was

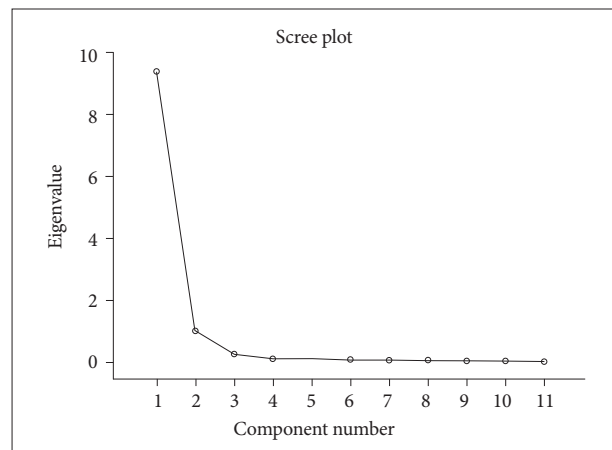


Figure 2. Scree plot of the K-ANX. K-ANX: Korean Anxiety Screening Assessment.

Table 4. Factor loadings of EFA and unstandardized and standardized coefficients of CFA

Items	EFA	CFA		
	Factor loading	Unstandardized coefficient	Standard error	Standardized coefficient
Item 1	0.963	1.000	0.018	0.954
Item 2	0.968	1.129	0.016	0.970
Item 3	0.974	0.973	0.017	0.967
Item 4	0.961	0.990	0.016	0.961
Item 5	0.975	1.034	0.02	0.973
Item 6	0.904	1.128	0.018	0.961
Item 7	0.958	1.028	0.019	0.959
Item 8	0.976	1.212	0.017	0.974
Item 9	0.975	1.080	0.017	0.975
Item 10	0.952	0.897	0.017	0.940
Item 11	0.978	1.145	0.018	0.972

EFA: exploratory factor analysis, CFA: confirmatory factor analysis

evaluated (Figure 3). This plot indicates the probability of a response in the given category across different anxiety levels. Furthermore, the TIC of the K-ANX was evaluated (Figure 1). The TIC provides the knowledge of how an instrument will behave in estimating person locations and permits the design of an instrument with specific estimation properties. Therefore, the concept of total information function can be used to design instruments with specific psychometric properties. According to the TIC of the K-ANX, the maximum information and the lowest standard error of measurement were provided at an anxiety level of 1.2 (between 0.8 and 1.6). This means that the K-ANX performs well in discriminating and screening participants at an anxiety level around 1.2 standard deviations above the mean.

Table 5. Model fit statistics of CFA

Model	χ^2 (df)	TLI	CFI	RMSEA [90% CI]	SRMR
One Factor	2311.832 (35)	0.891	0.901	0.101 [0.059, 0.313]	0.019

CFA: confirmatory factor analysis, TLI: Tucker-Lewis Index, CFI: Comparative Fit Index, RMSEA: root mean square error of approximation, SRMR: standardized root mean square residual

Table 6. Correlations among K-ANX, GAD-7, PSWQ, and BAI

	K-ANX	GAD-7	PSWQ	BAI
K-ANX	1	0.876*	0.653*	0.821*
GAD-7		1	0.689*	0.763*
PSWQ			1	0.556*
BAI				1

* $p < 0.001$. K-ANX: Korean Anxiety Screening Assessment, GAD-7: Generalized Anxiety Disorder 7-item scale, PSWQ: Penn State Worry Questionnaire, BAI: Beck Anxiety Inventory

Table 7. Parameter estimates for a graded response model

Items	a	se	b ₁	se	b ₂	se	b ₃	se	b ₄	se
Item 1	5.54	0.71	0.16	0.05	0.84	0.07	1.34	0.10	1.68	0.11
Item 2	4.57	0.52	-0.02	0.06	0.51	0.06	1.03	0.09	1.51	0.11
Item 3	3.47	0.44	-0.26	0.08	0.59	0.09	1.34	0.13	1.85	0.18
Item 4	3.75	0.52	0.11	0.07	0.79	0.11	1.35	0.17	1.84	0.22
Item 5	3.51	0.50	0.11	0.08	0.76	0.10	1.19	0.14	1.73	0.21
Item 6	4.18	0.61	0.23	0.07	0.95	0.12	1.40	0.17	1.93	0.23
Item 7	4.62	0.58	0.37	0.06	1.03	0.09	1.48	0.11	2.10	0.17
Item 8	3.72	0.45	0.32	0.07	0.94	0.09	1.36	0.11	1.83	0.15
Item 9	3.27	0.47	0.14	0.07	0.92	0.12	1.57	0.20	2.23	0.28
Item 10	4.47	0.82	0.73	0.09	1.07	0.11	1.50	0.16	2.79	0.94
Item 11	2.23	0.36	0.21	0.10	0.99	0.16	1.73	0.26	2.33	0.41

a: represents the item discrimination, b₁-b₄: parameters represent category boundary location

Comparison of diagnostic sensitivity and specificity with other instruments

The diagnostic sensitivity and specificity of the K-ANX, BAI, and GAD-7 were computed and compared. The sensitivity tells us how well a test identifies and rules in individuals with disease. However, efforts to increase sensitivity excessively may lead to some individuals without the disease also testing positive; consequently, there is the need to consider specificity. The specificity reports the proportion of individuals without disease correctly having negative results from the test. The optimal cutoff points were determined by maximizing both diagnostic sensitivity and specificity. The sensitivity and specificity of the K-ANX for identifying anxiety disorders were 0.795 and 0.937 respectively at a cutoff score of 25, and 0.869 and 0.972 for identifying GAD at a cutoff score of 27 (Table 8). It was observed that these values were higher than those of other widely used anxiety screening measures. Meanwhile, conventionally used cut-off points of the measures that were presented in the original article but have not been validated in Korea showed unfavorable results in this sample: an imbalance of sensitivity and specificity. By comparison, individuals with anxiety disorders are expected to obtain higher scores on the K-ANX, whereas individuals without the disorders would obtain lower scores.

DISCUSSION

Despite the Ministry for Health & Welfare Affairs' emphasis on the early intervention and prevention of mental illnesses, relatively few studies have examined the validity and reliability of anxiety screening tools, impeding the early identification of anxiety disorders. Moreover, no anxiety screening instrument that has been developed up to date appropriately reflects the unique characteristics of the Korean culture. Therefore, the aim of this study was to develop a Korean

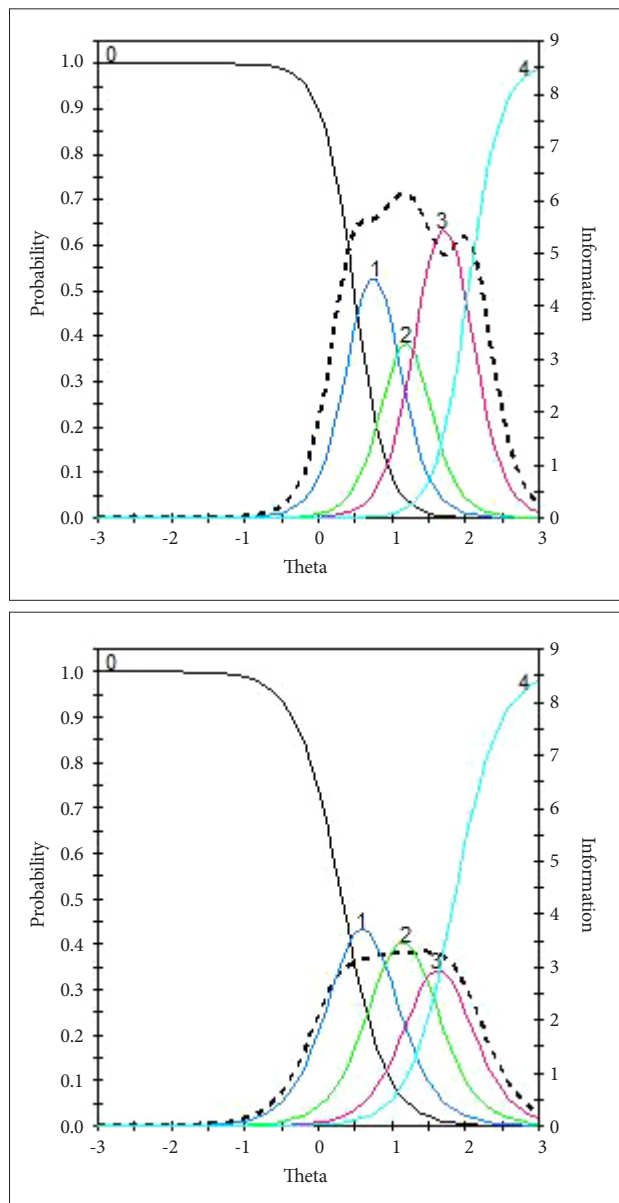


Figure 3. Exemplary category response function (solid lines) and item information function (dotted lines) for item 1 (left) and item 10 (right). The item category response curves for the rest of the items can be provided category response curves at the request of readers.

screening instrument for anxiety disorders, the K-ANX, which incorporates all nine core diagnostic criteria of GAD based on the DSM-5, and characteristics relevant to the Korean culture and recent anxiety symptoms.

First, the final items of the K-ANX reflected all of the diagnostic criteria of GAD as suggested by the DSM-5. The K-ANX further includes three diagnostic domains that the GAD-7 does not cover, namely difficulty in concentrating, sleep disturbance, and impairment of functioning. Among these, ‘sleep disturbance’ had been included reflecting the opinions of clinicians that patients with GAD often complain

Table 8. Diagnostic sensitivity and specificity of the K-ANX, BAI, and GAD-7

Assessment	Identified disorder	Cutoff point	Sensitivity	Specificity
K-ANX	Anxiety disorders	≥25*	0.795	0.937
	GAD	≥27*	0.869	0.972
BAI	Anxiety disorders	≥9*	0.694	0.783
		≥16†	0.398	0.901
		≥9*	0.868	0.744
GAD-7	Anxiety disorders	≥16†	0.605	0.883
		≥5*	0.704	0.775
		≥9†	0.500	0.901
		≥8*	0.763	0.849
		≥9†	0.684	0.871

*optimal cutoff point which maximizes both diagnostic sensitivity and specificity in the present sample, †conventionally used cutoff point. K-ANX: Korean Anxiety Screening Assessment, GAD: generalized anxiety disorder, GAD-7: Generalized Anxiety Disorder 7-item scale, BAI: Beck Anxiety Inventory

of sleep problems. The ‘impairment of functioning’ domain is essential to the diagnosis of most mental illnesses including GAD. While the BAI focuses heavily on somatic symptoms and the GAD-7 places more emphasis on worry, the K-ANX is a balanced scale that considers both worry and body sensations. In addition, because the K-ANX covers all domains, it would be suitable for assessing the patients’ conditions more objectively if it was used as a severity rating application for patients with anxiety disorders.

Moreover, items of the K-ANX were culturally relevant. For instance, one item on somatic symptoms from the K-ANX, “(for the last 2 weeks) I felt that my chest has been suddenly oppressed,” better distinguished Koreans with anxiety disorders from Koreans without such disorders than other somatic symptoms items that appear in previous anxiety disorder assessments such as the BAI. Interestingly, our results indicate that somatic symptoms were less appropriate to assess anxiety disorders in Koreans. Possible explanations might be their intimate association with concomitant depressive disorders as well as anxiety disorders, and the tendency of Koreans to report more subjective discomfort, vague fears, or worries.^{6,38,39} Hence, the heavy reliance on somatic symptoms (e.g., BAI) appears to reduce the diagnostic specificity of an anxiety disorder screening measure.

Second, the exploratory and confirmatory factor analysis, which was conducted to evaluate the structural properties of the K-ANX, supported a one-factor model. In addition, the factor loadings of all items were above 0.50, which signifies a stable factor structure⁴⁰ and provides evidence for construct validity, suggesting that all of the items represent a dominant do-

main that can be explained by a single construct, anxiety. Our results suggest that the K-ANX might be a more reliable screening tool for anxiety disorders than the BAI, which has been identified to have two or four factors depending on the study.^{9,41} Its multiple factors (i.e., “somatization” and “panic”) can explain why the BAI is more efficient in screening those with and without panic disorder rather than anxiety disorders.¹¹

Third, the reliability of the K-ANX was supported by an excellent internal consistency and high item inter-relatedness. Internal consistency refers to the degree to which the items measure a single construct and is related to the item-total correlation. The internal consistency (Cronbach's α) of the K-ANX was 0.97, indicating excellent reliability. However, this value has to be interpreted with caution, considering that an extremely high alpha might indicate that items are too inter-related and therefore some of them are redundant.⁴²

Fourth, the total score of the K-ANX highly and positively correlated with the total scores of the BAI, GAD-7, and PSWQ at a significance level of $p < 0.001$, providing support for criterion validity. Specifically, the highest correlations of the K-ANX with the GAD-7, followed by that with the BAI (0.821) and the PSWQ (0.653), imply that the K-ANX was most appropriate for screening probable cases of GAD and overall anxiety disorders similar to the GAD-7, even though the K-ANX was also associated with cognitive components of GAD (i.e., PSWQ) and more complicated anxiety components (i.e., BAI).

Fifth, our IRT analysis indicates that each item provides high information values across different anxiety levels, and the K-ANX offered the most informative diagnostic decisions with the highest reliability and the lowest standard error of measurement at the range between 0.8 and 1.6 (i.e., top 21.2 to 5.5 percentiles) level of anxiety. While we have mentioned that a Cronbach's alpha too high might suggest redundancy, the IRT result indicates that all items have their own informative value.

Finally, the K-ANX was revealed to be a more sensitive and specific screening tool than the BAI or GAD-7. Sensitivity refers to the probability of a positive response in the presence of an illness, whereas specificity refers to the probability of a negative response in the absence of an illness.^{43,44} Sensitivity and specificity of diagnostic screening tools are critical to reduce medical costs by minimizing type 1 and 2 errors, and to help provide appropriate interventions to those who are in most need of professional care or treatment.^{45,46} Thus, higher sensitivity and specificity of the K-ANX will have benefits in clinical and research settings. Meanwhile, there were several combinations of items that yielded even higher sensitivity and specificity than this K-ANX item combination. In order to develop a balanced screening tool in terms of content, sophisticated efforts were made to extract the

items with strong discriminative power from each domain and combine them without significantly lowering the high level of sensitivity and specificity. As in this second-year study, high sensitivity and specificity are expected to be maintained in the third-year of study, and if so, the K-ANX could be a valid test with fewer false positives or false negatives than the BAI or GAD-7.

Regardless of the significant implications, there are some limitations. First, the gender of our study was disproportionate (male to female ratio=0.462). Based on differential item functioning (DIF) analysis using IRT, two items detected the uniform DIF and one item detected the non-uniform DIF with respect to gender. Two items, “(for the last 2 weeks) All the stimuli from the surroundings bothered me,” and “(for the last 2 weeks) I was too tired to think of anything else,” measure different levels of anxiety in the two groups. Specifically, male has lower probability of responding to the items than female at the same level of anxiety (Male=0.1, Female=-0.1375; Male=0.2925, Female=-0.225). Another item, “(for the last 2 weeks) I was nervous or tensed,” responds more sensitively to changes in anxiety level in male group than in female group (M=8.13, F=3.65). Even though three items were detected as differently functioning items in terms of gender at the item level analysis, the item-total correlation, factor structure, and reliability index did not display any gender effect at the test level analysis. Therefore, a total score of the K-ANX would be interpretable regardless of the gender effect existing at the item level. However, future study should replicate the current findings in larger samples with more balanced gender ratios, and develop a scoring system considering gender effect. Second, given that the purpose of this study was to construct an item bank and to select the final items with high reliability, validity, and diagnostic sensitivity and specificity, data were collected and analyzed using the linking method to allow subjects to respond to as many pilot items as possible before selecting the final items. Therefore, there is a methodological caution in that the estimated item responses were used in part. Third, our study focused on developing and examining the psychometric properties of the K-ANX. Thus, subsequent studies should provide information on its implementation and utilization (e.g., precautions in execution, comparison of test results according to various groups, interannual testability, interpretation of scores, etc.), which is crucial in disseminating a screening tools.

Nevertheless, the K-ANX is a promising novel screening tool for anxiety disorders. It was developed in accordance with the policy to provide early diagnosis, screening, and treatment of anxiety disorders in Korea. The K-ANX was constructed with a small number of optimal items that possess high information value for screening of anxiety disor-

ders in Koreans. The measure includes all diagnostic criteria based on the DSM-5 and demonstrated higher sensitivity and specificity than the BAI and GAD-7. Therefore, with further examination of the psychometric properties in another sample, the K-ANX is expected to be used to detect anxiety disorders and provide appropriate early intervention for potential anxiety patients, thereby reducing social costs associated with treatment and time off work.

Acknowledgments

This study was supported by Korea Mental Health Technology R&D Project under the Korean Ministry of Health and Welfare (MOHW) (HM15C1169).

REFERENCES

- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005;62:593-602.
- Cho MJ, Ham BJ, Hong JP, Lee DW, Lee JY, Jung IW. The Epidemiological Survey of Psychiatric Illnesses in Korea 2006. Seoul: Ministry of Health & Welfare; 2007.
- Cho MJ, Sung S, Shin S, Kim J, Jeon S, Kim M. The epidemiological survey of mental disorders in Korea. Seoul: Ministry of Health and Welfare 2011. Available at: <http://www.mohw.go.kr/>. Accessed March 23, 2018.
- Zimmerman M, Chelminski I, Young D, Dalrymple K. A clinically useful anxiety outcome scale. *J Clin Psychiatry* 2010;7:534-542.
- Zimmerman M, Chelminski I, Young D, Dalrymple K, Walsh E, Rosenstein L. A clinically useful self-report measure of the DSM-5 anxious distress specifier for major depressive disorder. *J Clin Psychiatry* 2014; 75:601-607.
- Yook SP, Kim ZS. A clinical study on the Korean version of Beck Anxiety Inventory: comparative study of patient and non-patient. *Kor J Clin Psychol* 1997;16:185-197.
- Gibbons RD, Weiss DJ, Pilkonis PA, Frank E, Moore T, Kim JB, et al. Development of the CAT-ANX: a computerized adaptive test for anxiety. *Am J Psychiatry* 2014;171:187-194.
- McQuaid JR, Stein MB, McCahil, M, Laffaye C, Ramel W. Use of brief psychiatric screening measures in a primary care sample. *Depress Anxiety* 2000;12:21-29.
- Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988; 56:893-897.
- Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care Res (Hoboken)* 2011;63 (Suppl 11): S467-S472.
- Leyfer OT, Ruberg JL, Woodruff-Borden J. Examination of the utility of the Beck Anxiety Inventory and its factors as a screener for anxiety disorders. *J Anxiety Disord* 2006;20:444-458.
- Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166: 1092-1097.
- Plummer F, Manea L, Treppe, D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. *Gen Hosp Psychiatry* 2016;39:24-31.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington, DC: American Psychiatric Publishing; 2013.
- Kim SH, Jung S, Park K, Jaekal E, Lee SH, Choi Y, et al. Development of the Korean screening tool for anxiety disorders: review of current anxiety scales and development of preliminary item pools. *Kor J Clin Psychol* 2016;35:630-644.
- Seo JG, Park SP. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. *J Headache Pain* 2015;16:97.
- Embretson SE, Reise SP. Item Response Theory for Psychologists. Mahway, NJ: Lawrence Erlbaum Associates; 2000.
- Barlow DH. Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic. New York: Guilford Press; 2004.
- Brown TA, Barlow DH, Liebowitz MR. The empirical basis of generalized anxiety disorder. *Am J Psychiatry* 1994;151:1272-1280.
- Fresco DM, Mennin DS, Heimberg RG, Turk CL. Using the Penn State Worry Questionnaire to identify individuals with generalized anxiety disorder: A receiver operating characteristic analysis. *J Behav Ther Exp Psychiatry* 2003;34:283-291.
- Wittchen, HU. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depress Anxiety* 2002;16:162-171.
- Whiting PF, Rutjes AW, Westwood ME, Mallet S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-536.
- Hergueta T, Baker R, Dunbar GC. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59:22-33.
- Yoo SW, Kim YS, Noh JS, Oh KS, Kim CH, NamKoong K, et al. Validity of Korean version of the Mini-International Neuropsychiatric Interview. *Anxiety Mood* 2006;2:50-55.
- Lee HK, Lee EH, Hwang ST, Hong SH, Kim JH. Psychometric properties of the beck anxiety inventory in the community-dwelling sample of Korean adults. *Korean J Clin Psychol* 2016;35:822-830.
- Pfizer Inc. The Korean version of the GAD-7. Patient Health Questionnaire (PHQ) screeners. Available at: <http://www.phqscreener.com>. Accessed May 1, 2018.
- Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. *Behav Res Ther* 1990;28:487-495.
- Lim YJ, Kim YH, Lee EH, Kwon SM. The Penn State Worry Questionnaire: Psychometric properties of the Korean version. *Depress Anxiety* 2008;25:E97-E103.
- Lord FM. Applications of Item Response Theory to Practical Testing Problems. New York: Routledge; 2012.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. AERA, APA, & NCME. Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association; 2014.
- Muth n LK, Muth n BO. Mplus. Statistical Analysis with Latent Variables. User's Guide (7th Ed). Los Angeles, CA: Muth n&Muth n; 2012.
- Cai L, Thissen D, du Toit SHC. IRTPRO for Windows [Computer Software]. Lincolnwood, IL: Scientific Software International; 2011.
- Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238-246.
- Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull* 1980;88:588-606.
- Browne MW, Cudek R. Alternative Ways of Assessing Model fit. In: Bollen KA, Long JS, Editors. Testing Structural Equation Models. Newbury Park, CA: Sage, 1993, p. 136-162.
- Medsker GJ, Williams LJ, Holahan PJ. A review of current practices for evaluating causal models in organizational behavior and human resources management research. *J Manage* 1994;20:439-464.
- Samejima F. Estimation of Latent Ability Using a Response Pattern of Graded Scores. *ETS Res Rep Series* 1968;1:i-169.
- Han EK, Cho YR., Park SH, Kim HR, Kim SH. Factor structure of the Korean version of the Beck Anxiety Inventory: an application of confirmatory factor analysis in psychiatric patients. *Kor J Clin Psychol*

- 2003;22:261-270.
39. Kroenke K. Patients presenting with somatic complaints: epidemiology, psychiatric co-morbidity and management. *Int J Methods Psychiatr Res* 2003;12:34-43.
 40. Costello AB, Osborne JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract Assess Res Eval* 2005;10:1-9.
 41. Beck AT, Steer RA. *Manual for the Beck Anxiety Inventory*. San Antonio, TX: Psychological Corporation; 1990.
 42. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53-55.
 43. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;16:981-991.
 44. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008;149:889-897.
 45. Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatr* 2007;96:338-341.
 46. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr* 2007;96:644-647.