



# Measuring Anxiety in Patients With Early-Stage Parkinson's Disease: Rasch Analysis of the State-Trait Anxiety Inventory

Hui-Jun Yang<sup>1</sup>, Joon-Ho Ahn<sup>2</sup>, Jungsun Lee<sup>3</sup>, Won Kee Lee<sup>4</sup>, Jiho Lee<sup>5</sup> and Yangho Kim<sup>5\*</sup>

<sup>1</sup> Department of Neurology, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, South Korea,

<sup>2</sup> Department of Psychiatry, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, South Korea,

<sup>3</sup> Department of Psychiatry, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, <sup>4</sup> Medical Research Collaboration Center, Kyungpook National University Hospital and School of Medicine, Kyungpook National University, Daegu, South Korea, <sup>5</sup> Department of Occupational and Environmental Medicine, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, South Korea

## OPEN ACCESS

### Edited by:

Davide Martino,  
King's College London,  
United Kingdom

### Reviewed by:

Cristian Falup-Pecurariu,  
Transilvania University of Braşov,  
Romania  
Veronica Andrea Bruno,  
University of Calgary, Canada

### \*Correspondence:

Yangho Kim  
yanghokm@ulsan.ac.kr

### Specialty section:

This article was submitted to  
Movement Disorders,  
a section of the journal  
Frontiers in Neurology

**Received:** 21 October 2018

**Accepted:** 14 January 2019

**Published:** 13 February 2019

### Citation:

Yang H-J, Ahn J-H, Lee J, Lee WK,  
Lee J and Kim Y (2019) Measuring  
Anxiety in Patients With Early-Stage  
Parkinson's Disease: Rasch Analysis  
of the State-Trait Anxiety Inventory.  
*Front. Neurol.* 10:49.  
doi: 10.3389/fneur.2019.00049

The State-Trait Anxiety Inventory (STAI), composed of two 20-item subscales (STAI-state and STAI-trait), has been increasingly used to assess anxiety symptoms in patients with Parkinson's disease (PD). However, the clinimetric attributes of the STAI under the statistical framework of the item-response theory (IRT) have not been fully elucidated within this population to date. We performed an IRT-based Rasch analysis of the STAI outcomes of patients with *de novo* PD from the Parkinson's Progression Markers Initiative database. The unidimensionality, Rasch model fit, scale targeting, separation reliability, differential item functioning, and response category utility of the STAI were statistically evaluated. A total of 326 (209 males, 117 females) patients without cognitive dysfunction were enrolled in our study. The original versions of the STAI-state and STAI-trait had acceptable separation reliability but lacked appropriate response category functioning, exhibited scale off-targeting, and several items demonstrated poor fit to the Rasch model. The response categories were reduced from four to three, and the rescored three-point TASI-trait demonstrated a marked improvement in clinimetric properties without a significant impact on unidimensionality and separation reliability. The rescored three-point version of the STAI-state required the additional removal of four misfitting items in order to improve the Rasch model fit. To our knowledge, this is the first study to assess the measurement properties based on the IRT of the STAI in patients with PD. Our Rasch analysis identified the components requiring possible amendments in order to improve the clinimetric attributes of the STAI.

**Keywords:** Parkinson's disease, item response theory, State-Trait Anxiety Inventory, anxiety, validity, reliability, Rasch analysis

## INTRODUCTION

Anxiety is a prevalent non-motor symptom that affects 12–57% of patients with Parkinson's disease (PD) (1–3). The structural evaluation of anxiety is crucial for the effective management of PD; therefore, the importance of the reliability and validity of the clinical rating scale used to assess the anxiety symptoms in patients with PD has been highlighted (3–5).

The classical test theory (CTT) is a popular statistical framework for evaluation of the reliability and validity of questionnaires, patient-reported outcomes, and rating scales in health-care studies (6–8). The CTT hypothesizes that an observed score is the sum of a true score (error-free score) and measurement error, and the true score in the CTT is estimated based on the mean value of repeatedly measured observed scale scores (6, 7). The CTT is relatively easy to interpret and has thus far been applied in the standardization of the diverse clinical rating scales; however, there are several known limitations of the CTT. Shortcomings of the CTT include dependency between the clinimetric properties of the rating scale and the patients' responses, an ordinal level of measurement rather than interval, and a lack of statistical assessment for polytomous response category function (8–10). The modern test theory or item response theory (IRT) was developed to complement the above limitations of the CTT (5, 7). According to the statistical framework of the IRT, the item and patient statistics are derived in a mutually independent manner (6, 7). The IRT calculates the probability of the patient's response to any particular item and converts the probability into a logit score that has properties of an interval scale (6, 9, 10). Moreover, the IRT can administer a series of fit statistics, average measures, and step calibration of the response categories, consequently allowing for the diagnostic analysis of polytomous category utility (10–13). The Rasch analysis, which is a widely used one-parameter IRT model approach, has previously been successfully utilized to validate several anxiety rating scales among patients with PD, including the Hospital Anxiety and Depression Scale-Anxiety subscale (HADS-A), Hamilton anxiety rating scale (HARS), and Beck Anxiety Inventory (BAI) (5, 14).

The State-Trait Anxiety Inventory (STAI), which was first developed in the 1960s by Spielberger et al. was designed to rate overall adult anxiety and is formulated as a four-point Likert scale (STAI "Form X") (15). A revision of the STAI, which reduced the overlap with depression and placed emphasis on better described state and trait anxiety factors, was published in 1983 (STAI "Form Y") (16). The STAI was applied in prior studies for quantification of anxiety symptoms in patients with PD (4, 5), and its reliability and validity based on the CTT have been demonstrated (17, 18). However, none of the previous studies have performed an IRT-based analysis of the clinimetric attributes of the STAI in patients with PD (3, 4). Moreover, concerns have been raised with regard to the suitability of the polytomous response structure including a multi-point Likert scale in the application in older adults, such as patients with PD (4, 19, 20). While the determination of the response category options is typically considered a priori in the CTT, Rasch analysis provides diagnostic statistics to evaluate if the empirically-determined polytomous response categories function as intended, and alternative rescoring can improve the overall measurement properties of scale (5, 10, 12).

The present study aims to address the measurement properties of the STAI in PD patients without cognitive dysfunction by conducting Rasch analysis with demonstration of the Rasch model fit, separation reliability, differential item functioning (DIF), and response category function of a multi-point Likert scale of STAI.

## METHODS

The Parkinson's Progression Markers Initiative (PPMI) study is 5-years international multi-center prospective investigation that aims to identify potential biomarkers of disease progression using clinical, imaging, and biospecimen data from newly diagnosed patients with PD (21). The detailed design and procedure of the PPMI study have been previously published elsewhere (2, 21, 22). The study protocol was approved by the institutional review boards of all participating centers, and each patient was required to provide written informed consent before study participation. In our study, patients with untreated PD at baseline who met the selection criteria were recruited from the PPMI online database (<http://www.ppmi-info.org/data>), and the final database was accessed on March 31, 2018. The selection criteria in this Rasch analysis include the following: (1) complete baseline data of the clinical assessments including the STAI, Montreal cognitive assessment (MoCA), and dopamine transporter scans (21, 22); (2) MoCA total score above the cut-off value of mild cognitive impairment (MoCA score of  $\geq 26$ ) (23); and (3) modified Hoehn and Yahr (H-Y) stage  $\leq 2$ .

STAI "Form Y," consisting of two 20-item subscales, including the STAI state subscale (STAI-state) and STAI trait subscale (STAI-trait), was employed to quantify a range of anxiety symptoms associated with the disease (16). The STAI-state is used for measurement of temporary anxiety symptoms due to specific situations or particular objects. Responses in the STAI-state are formulated as a four-point Likert scale with the following category options: category 1 for "not at all," category 2 for "somewhat," category 3 for "moderately so," and category 4 for "very much so." The STAI-trait addresses innate and relatively stable personal tendencies to experience anxiety symptoms. Responses in the STAI-trait have the following four category options: category 1 for "almost never," category 2 for "sometimes," category 3 for "often," and category 4 for "almost always." Some *anxiety-absent* items (STAI-state items 1, 2, 5, 8, 10, 11, 15, 16, 19, and 20; STAI-trait items 21, 23, 26, 27, 30, 33, 34, 36, and 39) are negatively keyed, and these items are reverse-coded according to the instructions (16, 21). The score range for both the STAI-state and the STAI-trait is 20–80 points, and a higher score indicates a greater degree of anxiety. The disease severity of the patients was globally addressed by the H-Y staging and Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) parts I, II, and III. The 15-item Geriatric Depression Scale was used to evaluate the intensity of depressive symptoms, and the MoCA was employed to evaluate overall cognitive function (22).

The analyses of internal consistency and convergent validity based on the CTT were presented using IBM SPSS version 19 (IBM corporation, Somers, NY, USA). For internal consistency, Cronbach's  $\alpha$ -coefficient (should be  $>0.70$ ) and item-total correlation (should be  $>0.30$ ) were calculated (3, 4, 8). Convergent validity was assessed using the Spearman rank-order correlation with item 4 (anxiety symptoms) of the MDS-UPDRS part I, and the Spearman correlation coefficients of  $r_s >0.40$  was considered to indicate a moderate or greater correlation (7, 8).

The unidimensionality, Rasch model fit of items, person separation reliability, scale targeting, DIF, and response category utility were statistically evaluated by Rasch analysis (**Supplementary Figure 1**) based on Andrich's rating scale model using WINSTEPS version 4.0.1 (Winsteps Inc., Chicago, IL, USA) (10). Unidimensionality of the rating scale was assessed by a principal component analysis of the residuals (PCAR) wherein the Rasch factor is extracted. If the variance as explained by the Rasch factor was  $\geq 40\%$ , it was considered to support unidimensionality (6, 24). An eigenvalue for the first or second residual variances of  $\geq 3.0$ , or  $\geq 10\%$  of variance as explained by the first or second contrast, were thought to demonstrate the possibility of multidimensionality (25–27). The Rasch model fit indicating internal scale validity was investigated based on the infit (the information-weighted fit) and outfit (the outlier-sensitive fit) statistics (24). The mean-square (MnSq) value of infit and outfit was expected to be 1, and the acceptable range of the MnSq value was 0.5–1.5 (9, 13, 28). The measurement reliability of the STAI-state and STAI-trait were assessed by the person separation reliability (PSR), which measures the capacity of the discrimination among the patient groups with different levels of anxiety (7). The tentative criterion value of the PSR was 0.8, which indicates that the scale was able to discriminate the study patients into three strata of anxiety severity (i.e., mild, moderate, or severe) (22, 28). A person-item distribution map (Wright map) was used to visually inspect the STAI item measurement range with respect to the patient symptom severity distribution. The range of patient measures and item measures were displayed on the left and right side of the map, respectively. Test items with a higher logit score were located on the top of the map and considered measuring higher level of symptoms, while items on the bottom of the map were in the area of relatively lower patient anxiety. The difference in the person-item mean logit was used to address overall targeting of the rating scale (6, 9, 24). DIF was analyzed to determine whether STAI items function differently in relation to key demographic variables including gender (male vs. female) and age (aged  $<60$  vs. aged  $\geq 60$  years as older) (7, 12). DIF was considered to be significant if DIF contrast above 0.60 logits difference (29, 30). The utility of each response category of the four-point Likert scale was analyzed by the probability curves, outfit of the residual MnSq value, patient count, average measures, and step calibration for each response category (10, 31). The probability curves of the multi-point response options present the likelihood of patients selecting a certain response option on the STAI at various levels of the anxiety. Ordered thresholds on the probability curves, a minimum patient count of 10 in each category, an outfit MnSq value of  $<2.0$ , the hierarchical monotonic increase in average measures and step calibrations were required for appropriate category function (31, 32).

## RESULTS

A total of 326 patients with untreated PD including 209 men and 117 women (age range, 34–85 years) were enrolled. The

**TABLE 1 |** Demographic and clinical characteristics of 326 patients with Parkinson's disease.

Characteristics	Number	Percentage (%)	Mean	SD
Age, years			60.8	9.7
Gender				
Male	209	64.1		
Female	117	35.9		
Education, years			15.7	3.1
Disease duration, months			6.0	3.6
Hoehn and Yahr stage			1.6	0.5
Stage 1	147	45.1		
Stage 2	179	54.9		
MDS-UPDRS part 1			5.5	4.2
MDS-UPDRS part 2			6.0	4.3
MDS-UPDRS part 3			20.2	8.4
MoCA score			28.1	1.3
STAI-state score			33.1	10.6
STAI-trait score			32.4	9.6
GDS-15 score			5.3	1.4

SD, Standard Deviation; MDS-UPDRS, Movement Disorder Society—Unified Parkinson's Disease Rating Scale; STAI-state, State-Trait Anxiety Inventory state subscale; STAI-trait, State-Trait Anxiety Inventory trait subscale; MoCA, Montreal Cognitive Assessment; GDS-15, Geriatric Depression Scale 15-item short form.

median H-Y stage was 2. The mean score of the STAI-state was 33.1 (range, 0–45) and that of the STAI-trait was 32.4 (range, 0–45). **Table 1** presents the demographic and clinical data of the study patients.

With regard to internal consistency according to the CTT, the Cronbach's  $\alpha$  coefficient of the STAI-state and STAI-trait was 0.932 and 0.922, respectively, both of which were higher than the threshold value of 0.70. The item-total correlation score for the STAI-state and STAI-trait was 0.377–0.764 and 0.446–0.726, respectively, and all items met the criterion for internal consistency (item-total correlation of  $>0.30$ ). Concerning the convergent validity, question four of the MDS-UPDRS part I exhibited moderate or greater correlation with the STAI-state ( $r_S = 0.427$ ,  $p < 0.000$ ) and STAI-trait ( $r_S = 0.523$ ,  $p < 0.000$ ).

Our initial Rasch analysis indicated that the acceptable separation reliability of the STAI-state and STAI-trait as assessed by PSR were 0.83 and 0.83, respectively, both of which were higher than the tentative criterion value of 0.80 (**Table 2**). However, three STAI-state items (item 9, 14, and 18) failed to meet the criterion value (between 0.5 and 1.5) of the infit or outfit MnSq values, indicating that the actual responses to these items did not match the expectation of the Rasch model (**Table 2**). DIF for gender was found for items 9 and 18 (greater in females). None of the STAI-state items exhibited age-related DIF. In the STAI-trait, four items (namely items 24, 32, 35, and 38) failed to demonstrate an acceptable outfit MnSq value for the Rasch model fit. While all STAI-trait items were free from DIF for age, items 25 and 35 exhibited gender-related DIF that was more severe in females. According to the person-item distribution map with the mean logit for patients set to 0, the mean logit for the STAI-state

**TABLE 2** | Psychometric validation-related statistics based on the Rasch model for the State-Trait Anxiety Inventory state and the State-Trait Anxiety Inventory trait subscales.

Rasch statistics of the clinical scales	Original 20-item scale (four-point)	Rescored 20-item scale (three-point)	Shortened 16-item scale (three-point)
<b>STAI-STATE</b>			
Rasch factor eigenvalue (variance %)	17.17 (46.2)	18.62 (48.2)	14.38 (47.3)
1st contrast eigenvalue (variance %)	2.90 (7.8)	2.91 (7.5)	2.88 (9.5)
2nd contrast eigenvalue (variance %)	1.81 (4.9)	1.92 (5.0)	1.56 (5.1)
Point-measure correlation, range	0.35~0.75	0.40~0.78	0.48~0.80
Items with infit MnSq <0.5 or >1.5 (%)	2 (10.0%)	0 (0.0%)	0 (0.0%)
Items with outfit MnSq <0.5 or >1.5 (%)	3 (15.0%)	4 (20.0%)	0 (0.0%)
Person separation reliability (index)	0.83 (2.20)	0.85 (2.35)	0.84 (2.27)
Item separation reliability (index)	0.97 (6.13)	0.98 (6.67)	0.98 (6.26)
Mean logit of the patient (standard deviation)	-2.03 (1.70)	-1.39 (1.85)	-1.22 (1.92)
<b>STAI-TRAIT</b>			
Rasch factor eigenvalue (variance %)	14.79 (42.5)	16.04 (44.5)	
1st contrast eigenvalue (variance %)	2.86 (8.2)	2.96 (8.2)	
2nd contrast eigenvalue (variance %)	1.62 (4.6)	1.51 (4.4)	
Point-measure correlation, range	0.46 ~ 0.72	0.50 ~ 0.73	
Items with infit MnSq <0.5 or >1.5 (%)	0 (0.0%)	0 (0.0%)	
Items with outfit MnSq <0.5 or >1.5 (%)	4 (20.0%)	0 (0.0%)	
Person separation reliability (index)	0.83 (2.18)	0.85 (2.35)	
Item separation reliability (index)	0.97 (5.29)	0.97 (5.68)	
Mean logit of the patient (standard deviation)	-2.09 (1.54)	-1.38 (1.69)	

STAI-state, State-Trait Anxiety Inventory state subscale; MnSq, mean squares; STAI-trait, State-Trait Anxiety Inventory trait subscale.

item was  $-2.03$ , indicating that item measurement range was highly distributed compared to the severity of anxiety symptoms in the study patients (Figure 1; Table 2). The person-item map (Figure 1) also displayed the hierarchy of the STAI-state items. Item 6 “I feel upset” and item 18 “I feel confused” were the hardest to endorse by patients with PD, and item 19 “I feel steady” the easiest. Similarly, item measurement distribution of the STAI-trait was targeted relatively higher than the degree of anxiety symptoms in the patient group with the mean difference in logit for items of  $-2.09$  (Figure 2; Table 2). Figure 2 indicated that STAI-trait item 25 “I feel like a failure” was the hardest to endorse, whereas item 26 “I feel rested” was the easiest.

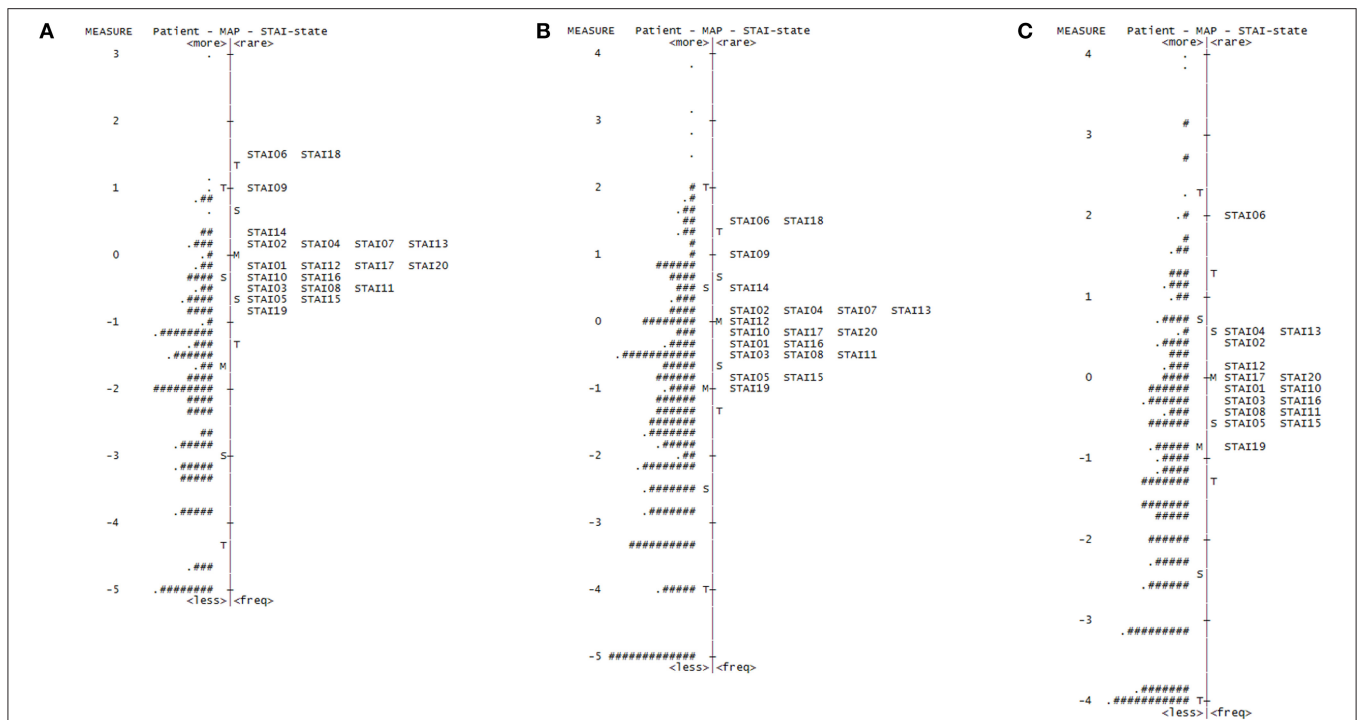
Concerning the response option utility, the category probability curves exhibited an ordered threshold for each response category in both the STAI-state and STAI-trait (Supplementary Figure 2). However, the outfit MnSq value of category 4 in the STAI-state exceeded the criterion value of 2.0. Category 4 of the STAI-trait also failed to meet the criteria (outfit MnSq value of  $<2.0$ ), and the average measure of category 3 ( $-0.20$ ) in the STAI-trait was higher than that of category 4 ( $-0.29$ ), not displaying a monotonic advance (Table 3). The response category functions of the four-point Likert scales were considered inappropriate, hence both of the STAI subscales were rescored by collapsing categories 3 and 4 (Supplementary Figure 2).

The rescored three-point version of the STAI-state indicated that the outfit MnSq value of four items (items 7, 9, 14, and 18) failed to demonstrate acceptable outfit MnSq values, while

the infit MnSq value for all rescored STAI-state items met the criteria. We also found notable gender-related DIF in items 9 and 18. After the above four misfitting items were removed, the infit, and outfit MnSq values of all 16 items were acceptable and free from DIF for gender and age. The revised 16-item STAI-state on a three-point Likert scale exhibited unidimensionality and the PSR (0.84) met the criteria. The range of item measurement was altered to better fit the distribution of the patients’ symptoms compared to that prior to revision, which was confirmed by a significant improvement in the mean logit from  $-2.08$  before revision to  $-1.39$  after revision (Figure 1, Table 2). Category function analysis of the revised three-point STAI-state with 16 items indicated that the outfit MnSq value for all response categories was  $<2.0$  and the average measure monotonically increased (Table 3).

The rescored three-point version of the STAI-trait presented apparent improvements in Rasch model fit, as all items of the rescored three-point STAI-trait had infit and outfit MnSq values between 0.5 and 1.5, and therefore all 20 items were retained (Table 2). A PCAR after the Rasch factor extraction ensured that the 20 items of the rescored STAI-trait were of a unidimensional construct. The PSR of the rescored STAI-trait (0.85) was adequate, and the category function analysis result was also satisfactory (Table 3). Only one item (item 35) exhibited a gender DIF contrast of  $-0.67$  and no items exhibited age-related DIF. The person-item map indicated that the item measurement range after rescored better reflected the distribution of the degree of the patients’ symptoms (Figure 2, Table 2), that is





**FIGURE 1 |** Person-item distribution plot (Wright map) of the STAI-state. Positive scores demonstrate higher levels of anxiety. M, mean of person or item distribution; S, 1 standard deviation from the mean; T, 2 standard deviations from the mean. **(A)** Wright map for the original 20-item State-Trait Anxiety Inventory state (STAI-state) subscale (four-point scale). “#” Indicates three persons and “.” indicates one to two persons. **(B)** Wright map for the rescored 20-item STAI-state (three-point scale). “#” Indicates two persons and “.” indicates one person. **(C)** Wright map for the shortened 16-item STAI-state (three-point scale). “#” Indicates three persons and “.” indicates one to two persons.

supported by the improvement in the mean logit score from  $-2.09$  prior to revision to  $-1.38$  after revision.

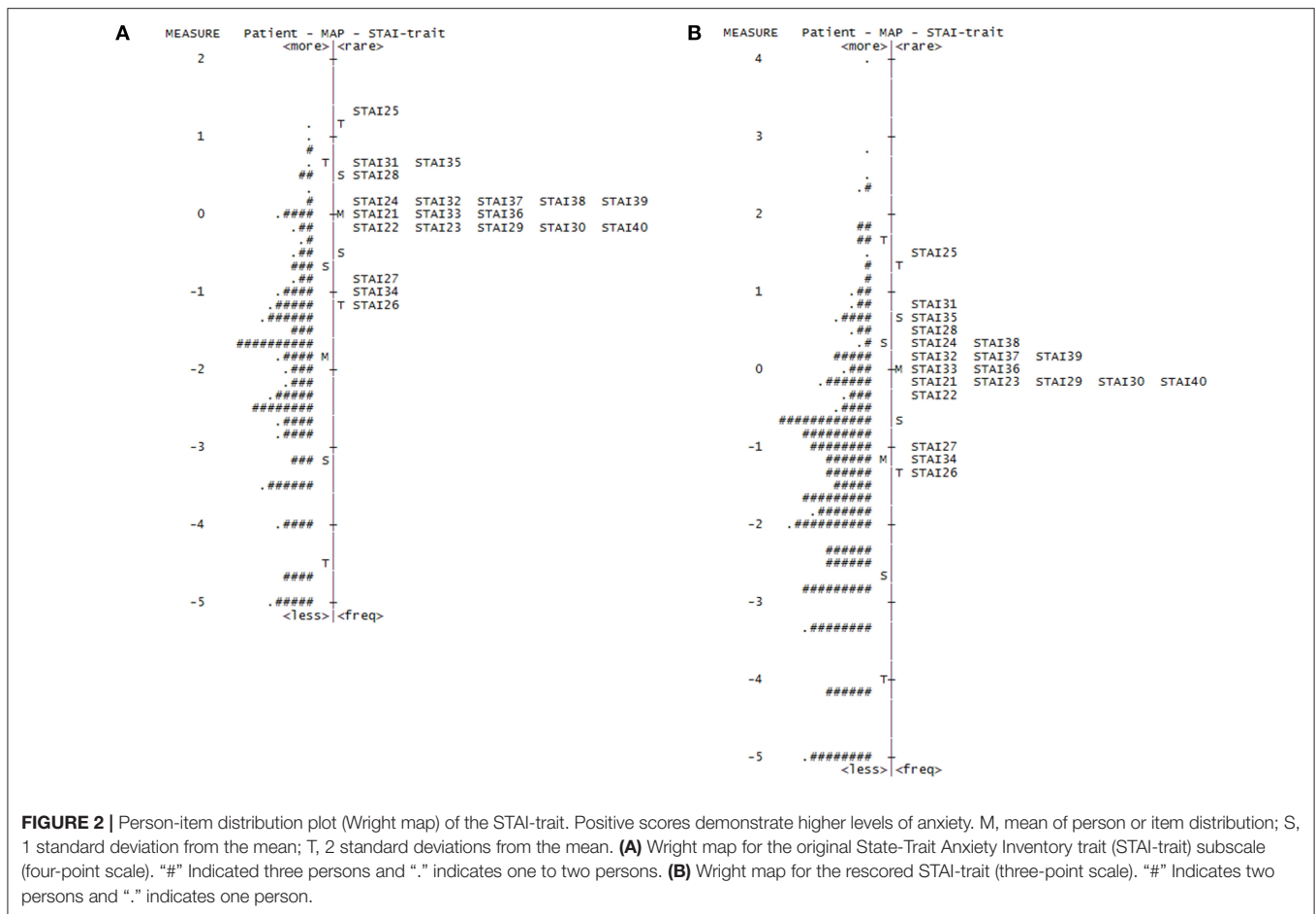
We evaluated the reliability and validity of both of the revised subscales using the CTT. The rescored STAI-state (16 items) and the rescored STAI-trait (20 items) yielded Cronbach’s  $\alpha$  coefficients of 0.936 and 0.927, respectively, with item-total correlations in the range of 0.518–0.795 and 0.507–0.795. All values were higher than the thresholds. We confirmed a moderate or greater correlation with item 4 (anxiety symptoms) of the MDS-UPDRS part I for both the rescored STAI-state (16 items) and STAI-trait (20 items) according to the  $r_s$  values of 0.423 ( $p < 0.000$ ) and 0.435 ( $p < 0.000$ ), respectively.

## DISCUSSION

This study explored the clinimetric attributes of STAI in non-demented *de novo* PD patients using the CTT and IRT (3, 7). Our result based on the CTT indicated that the STAI-state and STAI-trait displayed good internal consistency according to Cronbach’s  $\alpha$  coefficient and the item-total correlation. There was a moderate correlation with the MDS-UPDRS anxiety item, supporting the convergent validity in patients with PD. The current findings are in line with previous CTT studies in the PD population reporting that the STAI is significantly correlated with the HADS-A, HARS, and Geriatric Anxiety Inventory (GAI) (17, 18).

The initial Rasch analysis suggested that both the STAI-state and STAI-trait exhibited good separation reliability in patients with PD. However, the person-item distribution maps indicated that item difficulties of both STAI subscales were more highly distributed than the level of anxiety in the study patients, which indicated the inability of the questions to capture low level of anxiety and that the questions were off-targeted in PD (6, 9). The result of the present study was similar to a previous Rasch analysis in a PD cohort using other anxiety rating scales, including the HADS-A, BAI, and HARS, which were found to be more appropriate for patients with moderate or severe anxiety symptoms (5, 14). Off-targeting of the above scales was likely due to the fact that they were originally developed for assessment of anxiety in patients with more severe symptoms (4, 15, 16).

Testing with Rasch analysis also suggested possible issues in the function of the four-point polytomous response categories, and indicated that some of the items did not fit the Rasch model in the original STAI-state and STAI-trait (11, 27). To date, there have been several studies assessing the clinimetric properties of STAI in non-PD samples using Rasch analysis (33–37). Tenenbaum et al. found that the STAI-state and STAI-trait comprise some items that do not fit the Rasch model, and concluded that item deletion may be needed to refine the STAI-state and -trait scales in a non-PD population (33). Kaipper et al. found that both the 20-item STAI-state and the 20-item STAI-trait did not fit the Rasch model in the measurement of



**FIGURE 2 |** Person-item distribution plot (Wright map) of the STAI-trait. Positive scores demonstrate higher levels of anxiety. M, mean of person or item distribution; S, 1 standard deviation from the mean; T, 2 standard deviations from the mean. **(A)** Wright map for the original State-Trait Anxiety Inventory trait (STAI-trait) subscale (four-point scale). “#” Indicated three persons and “.” indicates one to two persons. **(B)** Wright map for the rescored STAI-trait (three-point scale). “#” Indicates two persons and “.” indicates one person.

anxiety levels in surgical patients scheduled for elective operation. The authors reported that the shorter version of the STAI-state after removal of seven items and the shortened STAI-trait after removal of eight items have acceptable fit to the Rasch model (34). However, these two studies used the first version of the STAI (STAI “Form X”) in a relatively younger population (15, 16). In addition, the reduced number of rating scale items can result in decreased scale reliability, and some studies demonstrated that several shorter forms of the STAI were associated with reduced separation reliability (35, 38).

Recently, Davey et al. evaluated STAI “Form Y” in 322 ophthalmology patients (mean age  $\pm$  SD: 61  $\pm$  19 years) by Rasch analysis, indicating that response category reduction (combining categories 3 and 4) can improve the fit to the Rasch measurement model (36). Fernández-Blázquez et al. also proposed that recoding the four original polytomous response options (0, 1, 2, 3) of STAI “Form Y” to a dichotomous structure (0, 1) can be useful in clinical settings for adults aged >69 years (35). There have been substantial debates regarding whether multi-point response options should be used in the rating scales assessing elderly populations, such as patients with PD. A greater number of response categories in an ordinal scale allow for the capture of detailed information and distinction of minor clinical differences. However, an excessive number of response categories

can induce confusion and fatigue in older patients, particularly if the investigator cannot make a clear distinction between each category or if the overall cognitive function of the patient is impaired (6, 9, 12). The recently developed GAI is one instance of employment of a dichotomous response category (“Agree” or “Disagree”) to indicate the level of anxiety in older adults (17).

We evaluated whether a reduction in the number of STAI response options would improve the measurement properties of the questionnaire in patients with PD. In the first instance, the rescored three-point STAI-trait demonstrated considerable improvement not only in polytomous category function, but also in Rasch model fit and scale targeting. The mean logit score was reduced on the person-item distribution map, indicating improvement in scale target deviation. Analysis of the gender-related DIF showed that item 35 “I feel inadequate” was easier for females to endorse. However, the DIF contrast was only marginally above the criterion, indicating that the gender DIF of item 35 was of minor concern. The benefit of reduction in the number of response options and conversion of the scale to a three-point Likert scale improved the polytomous response category function, scale targeting, and Rasch model fit without hindering the unidimensionality and separation reliability.

In contrast, although the rescored three-point STAI-state demonstrated an improvement in response category function

**TABLE 3 |** Threshold ordering test statistics for the State-Trait Anxiety Inventory state and the State-Trait Anxiety Inventory trait subscales.

Response categories	Count (%)	Average measure	Step calibration	Outfit MnSq
<b>STAI-STATE (ORIGINAL 20-ITEM SCALE)</b>				
Category 1 (not at all)	3501 (64)	-2.62	None	0.93
Category 2 (somewhat)	2003 (12)	-1.15	-1.47	0.83
Category 3 (moderately so)	766 (10)	-0.16	0.22	0.95
Category 4 (very much so)	249 (8)	0.10	1.25	2.81
<b>STAI-STATE (RESCORED 20-ITEM SCALE)</b>				
Category 1	3501 (64)	-2.11	None	0.95
Category 2	2003 (12)	-0.42	-0.87	0.86
Category 3	1015 (16)	0.83	0.87	1.62
<b>STAI-TRAIT (SHORTENED 16-ITEM SCALE)</b>				
Category 1	3361 (50)	-2.03	None	0.96
Category 2	2239 (33)	-0.35	-0.96	0.85
Category 3	1135 (17)	0.91	0.96	1.34
<b>STAI-TRAIT (ORIGINAL 20-ITEM SCALE)</b>				
Category 1 (almost never)	3512 (54)	-2.66	None	0.98
Category 2 (sometimes)	2129 (33)	-1.39	-1.61	0.84
Category 3 (often)	705 (11)	-0.20	0.19	0.76
Category 4 (almost always)	173 (3)	-0.29	1.41	2.91
<b>STAI-TRAIT (RESCORED 20-ITEM SCALE)</b>				
Category 1	3512 (54)	-2.06	None	1.05
Category 2	2129 (33)	-0.63	-0.94	0.85
Category 3	878 (13)	0.79	0.94	1.31

MnSq, mean squares; STAI-state, State-Trait Anxiety Inventory state subscale; STAI-trait, State-Trait Anxiety Inventory trait subscale.

and scale targeting, the number of items that did not fit the Rasch model (four misfitting items) or presenting gender-related DIF (two items) was similar to that of the original STAI-state. Therefore, additional modification (removal of items 7, 9, 14, and 18) was required in order to resolve the Rasch model misfit and gender DIF. Our revisions with removal of inappropriate items of the STAI-state subscale are in line with those of previous works proposing several shorter forms of the STAI-state composed of six or seven items (39–41). Although only four inappropriate items were removed from the STAI-state in the present study, future investigations in patients with PD should consider the assessment of clinimetric properties of several shorter versions of the STAI-state that have already been developed for clinical use.

As described above, the current analysis found that some of the original STAI items (STAI-state item 9 “I feel frightened” and item 18 “I feel confused,” STAI-trait item 25 “I feel like a failure” and item 35 “I feel inadequate”) were relatively less severe indicators of anxiety for male than for female. These gender-related DIFs were not virtually observed in prior Rasch analysis among non-PD population (34, 37). The discrepancy with the findings of previous reports could be due to the use of different versions of the instruments or to the different samples examined. The previous studies applied the first version of the STAI in relatively younger non-PD samples (34, 37). Moreover, STAI-state item 14 “I feel indecisive” and item 18 “I feel confused” exhibited high misfit to the Rasch model. This finding of misfitting items also differed from prior Rasch analysis

studies of the STAI in non-PD samples (35, 36). While confusion or indecisiveness could be related to the overall anxiety levels of patients, they can also result from other non-motor symptoms such as apathy, fatigue, or depression in PD (42, 43).

The proposed amendment of the collapse of the four-category option to a three-category system in our study involves the application of odd numbers of response categories to offer other than even. An even or odd number of response options each have certain strengths and shortcomings to be considered. The rating scales that have an even number of response categories such as the original STAI can lead the patient to provide more distinct answers instead of simply providing a midpoint option in their odd counterparts (44). However, odd numbers of categories with midpoint options are known to resolve the questionnaire bias caused by forcing the respondents to provide an answer, and further studies are needed to elucidate the optimal number of response categories (44, 45).

There are limitations to the study that influence the generalizability of our findings. One such limitation of this study is its questionnaire-based design, which did not include patient interviews or clinical diagnosis of anxiety according to the Diagnostic and Statistical Manual of Mental Disorders criteria (1, 16). However, our study did not aim to diagnose specific anxiety disorders, and such patient interviews may not be appropriate for retrospective investigations such as this. Second, the subjects in this study were limited to *de novo* PD patients without any cognitive impairment at the time of PPMI study

registration (21, 22). Our study was cross-sectional, using the PPMI online database among early and unmedicated patients, thus there was a limitation in the number of advanced stage PD patients. Floor effects of both subscales can be expected when utilizing the STAI in patients with early-stage PD with mild non-motor manifestations including neuropsychiatric and cognitive symptoms. Notably, anxiety levels appear to be too high in some early PD patients (17, 18). Given that the PPMI cohort has a relatively short duration after diagnosis ( $6.0 \pm 3.6$  months), it is possible that initial anxiety level may be partly explained as a psychological reaction to the diagnosis of their disease in some early-stage PD population (2, 46). Third, convergent validity was only tested against the MDS-UPDRS part 1 anxiety item and not against the other previously validated clinical rating scales for the measurement of anxiety in patients with PD. Fourth, our study did not assess the inter-rater or test-retest reliability. The test-retest reliability was demonstrated to be good for the STAI-trait but was not satisfactory for the STAI-trait in non-PD patients (3, 4). No such information is provided for PD patients and further studies are necessary to address its inter-rater and test-retest reliability.

## CONCLUSIONS

In conclusion, the rating scales that evaluate and quantify the severity of anxiety symptoms are particularly relevant for patients with PD. The STAI is one of the widely-used questionnaires for anxiety in clinical practice; however, the STAI has not been clinimetrically tested in patients with PD, based on modern test theories such as the Rasch measurement model. Our Rasch analysis complements previous findings based on the CTT and identified areas for possible amendments. Future studies should assess the clinimetric properties of the original version

of STAI in broader populations including late-stage PD patients, and the harmonized amendments can be tested to assess the interpretation of items and ease of comprehension. We expect that the present study will substantially help to evaluate anxiety symptoms in patients with PD in both clinical research and clinical practice.

## AUTHOR CONTRIBUTIONS

YK and H-JY participated in study design. J-HA, JuL, and WL interpreted data and contributed analytical tools. WL and H-JY acquired data and performed statistical analysis. J-HA, JuL, WL, and JiL contributed to writing of the manuscript and revised the manuscript for content. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the Ulsan University Hospital Research Grant (UUH-2017-03).

## ACKNOWLEDGMENTS

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database ([www.ppmi-info.org/data](http://www.ppmi-info.org/data)). For up-to-date information on the study, visit [www.ppmi-info.org](http://www.ppmi-info.org).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fneur.2019.00049/full#supplementary-material>

## REFERENCES

- Pontone GM, Williams JR, Anderson KE, Chase G, Goldstein SA, Grill S, et al. Prevalence of anxiety disorders and anxiety subtypes in patients with Parkinson's disease. *Mov Disord.* (2009) 24:1333–8. doi: 10.1002/mds.22611
- Rutten S, van der Ven PM, Weintraub D, Pontone GM, Leentjens AFG, Berendse HW, et al. Predictors of anxiety in early-stage Parkinson's disease - results from the first two years of a prospective cohort study. *Parkinsonism Relat Disord.* (2017) 43:49–55. doi: 10.1016/j.parkreldis.2017.06.024
- Leentjens AF, Dujardin K, Marsh L, Martinez-Martin P, Richard IH, Starkstein SE, et al. Anxiety rating scales in Parkinson's disease: critique and recommendations. *Mov Disord.* (2008) 23:2015–25. doi: 10.1002/mds.22233
- Dissanayaka NNW, Torbey E, Pachana NA. Anxiety rating scales in Parkinson's disease: a critical review updating recent literature. *Int Psychogeriatr.* (2015) 27:1777–84. doi: 10.1017/S1041610215000885
- Forjaz MJ, Martinez-Martin P, Dujardin K, Marsh L, Richard IH, Starkstein SE, et al. Rasch analysis of anxiety scales in Parkinson's disease. *J Psychosom Res.* (2013) 74:414–9. doi: 10.1016/j.jpsychores.2013.02.009
- Cappelleri JC, Lundy JJ, Hays RD. Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures. *Clin Ther.* (2014) 36:648–62. doi: 10.1016/j.clinthera.2014.04.006
- Forjaz MJ. How to evaluate validation data. In: Sampaio C, Goetz CG, Schrag A, editors. *Rating Scales in Parkinson's Disease: Clinical Practice and Research.* New York, NY: Oxford University Press (2012). p. 16–41.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* (2010) 19:539–49. doi: 10.1007/s11136-010-9606-8
- da Rocha NS, Chachamovich E, Fleck MPD, Tennant A. An introduction to Rasch analysis for Psychiatric practice and research. *J Psychiatr Res.* (2013) 47:141–8. doi: 10.1016/j.jpsychires.2012.09.014
- Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas.* (2002) 3:85–106.
- Pesudovs K, Noble BA. Improving subjective scaling of pain using Rasch analysis. *J Pain* (2005) 6:630–6. doi: 10.1016/j.jpain.2005.04.001
- Chung H. The Rasch model: an alternative method for analyzing ordinal data. *J Coaching Development* (2005) 7:133–41.
- Hagell P, Whalley D, McKenna SP, Lindvall O. Health status measurement in Parkinson's disease: validity of the PDQ-39 and Nottingham Health Profile. *Mov Disord* (2003) 18:773–83. doi: 10.1002/mds.10438
- Forjaz MJ, Rodriguez-Blázquez C, Martinez-Martin P. Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord.* (2009) 24:526–32. doi: 10.1002/mds.22409



15. Spielberger CD, Gorsuch RL, Lushene R. *Test Manual for the State-Trait Anxiety Inventory (Self Evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychologists Press (1970).
16. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. *Manual for the State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Consultant Psychologist Press (1983).
17. Matheson SE, Byrne GJ, Dissanayaka NN, Pachana NA, Mellick GD, O'Sullivan JD, et al. Validity and reliability of the Geriatric Anxiety Inventory in Parkinson's disease. *Australas J Ageing* (2012) 31:13–6. doi: 10.1111/j.1741-6612.2010.00487.x
18. Mondolo F, Jahanshahi M, Granà A, Biasutti E, Cacciatori E, Di Benedetto P. Evaluation of anxiety in Parkinson's disease with some commonly used rating scales. *Neuro Sci.* (2007) 28:270–5. doi: 10.1007/s10072-007-0834-9
19. Slaughter JR, Slaughter KA, Nichols D, Holmes SE, Martens MP. Prevalence, clinical manifestations, etiology, and treatment of depression in Parkinson's disease. *J Neuropsychiatry Clin Neurosci.* (2001) 13:187–96. doi: 10.1176/jnp.13.2.187
20. Byrne GJ, Pachana NA, Goncalves DC, Arnold E, King R, Khoo SK. Psychometric properties and health correlates of the geriatric anxiety inventory in Australian community-residing older women. *Aging Ment Health* (2010) 14:247–54. doi: 10.1080/13607861003587628
21. Marek K, Jennings D, Lasch S, Siderowf A, Tanner C, Simuni T, et al. The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol.* (2011) 95:629–35. doi: 10.1016/j.pneurobio.2011.09.005
22. Kletzel SL, Hernandez JM, Miskiel EF, Mallinson T, Pape TL. Evaluating the performance of the Montreal Cognitive Assessment in early stage Parkinson's disease. *Parkinsonism Relat Disord.* (2017) 37:58–64. doi: 10.1016/j.parkreldis.2017.01.012
23. Dalrymple-Alford JC, MacAskill MR, Nakas CT, Livingston L, Graham C, Crucian GP, et al. The MoCA: well-suited screen for cognitive impairment in Parkinson disease. *Neurology* (2010) 75:17–25. doi: 10.1212/WNL.0b013e3181fc29c9
24. Tennant A, Conaghan PG. Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum.* (2007) 57:1358–62. doi: 10.1002/art.23108
25. Ramakrisnan P, Jaafar A, Yatim NFBM. Towards an Understanding of user satisfaction measurement in Online Discussion Site (ODS) design using Rasch analysis. *Int J Education Res.* (2013) 1:125–36.
26. Eakman AM. Measurement characteristics of the engagement in meaningful activities survey in an age-diverse sample. *Am J Occup Ther.* (2012) 66:e20–9. doi: 10.5014/ajot.2012.001867
27. Ishar NIM, Roslin RM. Customer Voice Retaliation (CVR) construct verification: a rasch analysis approach. *Proc Econ Finan* (2016) 37:214–20. doi: 10.1016/S2212-5671(16)30116-2
28. Linacre JM. Prioritizing misfit indicators. *Rasch Meas Trans.* (1995) 9:422–3.
29. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Med Care* (2003) 41:582–92. doi: 10.1097/01.MLR.0000062554.74615.4C
30. Conrad KJ, Smith EV. International conference on objective measurement: applications of Rasch analysis in health care. *Med Care* (2004) 42(Suppl. 1):1–6. doi: 10.1097/01.mlr.0000103527.52821.1c
31. Andrich D. A rating formulation for ordered response categories. *Psychometrika* (1978) 43:357–74.
32. Janssen KC, Phillipson S, O'Connor J, Johns MW. Validation of the epworth sleepiness scale for children and adolescents using rasch analysis. *Sleep Med.* (2017) 33:30–5. doi: 10.1016/j.sleep.2017.01.014
33. Tenenbaum G, Furst D, Weingarten G. A statistical reevaluation of the STAI anxiety questionnaire. *J Clin Psychol.* (1985) 41:239–44.
34. Kaipper MB, Chachamovich E, Hidalgo MPL, da Silva Torres IL, Caumo W. Evaluation of the structure of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *J Psychosom Res.* (2010) 68:223–33. doi: 10.1016/j.jpsychores.2009.09.013
35. Fernández-Blázquez MA, Ávila-Villanueva M, López-Pina JA, Zea-Sevilla MA, Frades-Payo B. Psychometric properties of a new short version of the State-Trait Anxiety Inventory (STAI) for the assessment of anxiety in the elderly. *Neurologia* (2015) 30:352–8. doi: 10.1016/j.nrl.2013.12.015
36. Davey CJ, Harley C, Elliott DB. Levels of state and trait anxiety in patients referred to ophthalmology by primary care clinicians: a cross sectional study. *PLoS ONE* (2013) 8:e65708. doi: 10.1371/journal.pone.0065708
37. Quintão S, Delgado AR, Prieto G. Validity study of the Beck Anxiety Inventory (Portuguese version) by the Rasch Rating Scale model. *Psicol Reflex Crit.* (2013) 26:305–10. doi: 10.1590/S0102-79722013000200010
38. Vianya-Estopa M, Elliott DB, Barrett BT. An evaluation of the Amblyopia and Strabismus Questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci.* (2010) 51:2496–503. doi: 10.1167/iovs.09-4381
39. Court H, Greenland K, Margrain TH. Measuring patient anxiety in primary care: Rasch analysis of the 6-item Spielberger State Anxiety Scale. *Value Health* (2010) 13:813–9. doi: 10.1111/j.1524-4733.2010.00758.x
40. Lincoln N, Moreton B, Turner K, Walsh D. The measurement of psychological constructs in people with osteoarthritis of the knee: a psychometric evaluation. *Disabil Rehabil.* (2017) 39:372–84. doi: 10.3109/09638288.2016.1146356
41. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol.* (1992) 31:301–6.
42. Dissanayaka NN, White E, O'sullivan JD, Marsh R, Pachana NA, Byrne GJ. The clinical spectrum of anxiety in Parkinson's disease. *Mov Disord.* (2014) 29:967–75. doi: 10.1002/mds.25937
43. Broen MPG, Leentjens AFG, Hinkle JT, Moonen AJH, Kuijff ML, Fischer NM, et al. Clinical markers of anxiety subtypes in Parkinson disease. *J Geriatr Psychiatry Neurol.* (2018) 31:55–62. doi: 10.1177/0891988718757369
44. Simms LJ, Watson D. The construct validation approach to personality scale construction. In: Robins RW, Fraley RC, Krueger RF, editors. *Handbook of Research Methods in Personality Psychology*. New York, NY: Guilford Press (2007). p. 240–58.
45. Krosnick JA, Fabrigar LR. Designing rating scales for effective measurement in surveys. In: Lyberg L, Biemer PP, Collins M, de Leeuw E, Dippo CS, Schwarz N, Trewin D, editors. *Survey Measurement and Process Quality*. New York, NY: Wiley (1997). p. 141–64.
46. Pachana NA, Egan SJ, Laidlaw K, Dissanayaka N, Byrne GJ, Brockman S, et al. Clinical issues in the treatment of anxiety and depression in older adults with Parkinson's disease. *Mov Disord.* (2013) 28:1930–4. doi: 10.1002/mds.25689

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Yang, Ahn, Lee, Lee, Lee and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.