PAI-A Anxiety and Depression: IRT-Derived Scales and Their Construct Validity

Amelia D. Coffman, Lynne Steinberg, Carla Sharp, Leslie C. Morey

UNIVERSITY OF HOUSTON

Introduction

While the Personality Assessment Inventory – Adolescent (PAI-A; Morey, 2007) is well-known as an instrument for the assessment of psychopathology, there are few item-level and score level studies evaluating the structure and construct validity of the PAI-A. In two studies we examine the Anxiety (ANX) and Depression (DEP) subscales of the PAI-A. In Study 1, the items of the ANX and DEP scales were studied with IRT methods using the graded response model (Samejima, 1996). Item performance was examined separately for each scale in a community sample of adolescents. The goal in the IRT analysis was to make sure that all items are contributing to the measurement of the construct of interest, anxiety or depression. In Study 2, the construct validity of the IRT-derived item sets as well as the original full-length ANX and DEP scale scores was investigated using a sample of inpatient adolescents.

Study 1: Measurement

Data: PAI-A community standardization sample (N = 707, mean age = 15.00, SD = 2.00, 51.1% male; Morey, 2007).

Method
IRT item analysis was conducted to evaluate dimensionality and item performance. Items were evaluated with attention to the strength of the relation of the item response to the underlying construct (i.e., magnitude of the IRT slope parameters), redundancy in item content (evidenced as local dependence (LD) among pairs or triplets of items), and satisfactory model fit.

Results
ANX:
5 items were removed for low slope parameters, and redundancy among item pairs (detected with an LD chi-square index) .The unidimensional model fit well (M_2 (689) = 1000.34, p < .001; RMSEA = .03).

DEP:
5 items were removed for low slopes and redundancy. The unidimensional model fit well, (M_2 (689) = 985.07, p < .001; RMSEA = .02).
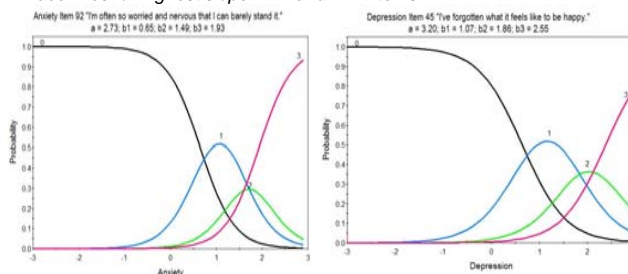
ANX and DEP
Although these scales were designed with cognitive, affective, and physiological subscales, the IRT analyses did not support this subscale structure.

Trace lines graphically show the probability of each response option across ranges of severity of the underlying construct (anxiety or depression). Figure 1 shows the trace lines for the ANX and the DEP item that had the highest slope parameters.

Figure 1.
Trace lines for highest slope ANX and DEP items

Anxiety item 92 "I'm often so worried and nervous that I can barely stand it."
a = 2.73; b1 = 0.65; b2 = 1.49; b3 = 1.93

Depression item 45 "I've forgotten what it feels like to be happy."
a = 3.20; b1 = 1.07; b2 = 1.86; b3 = 2.55

Study 2: Construct Validation of the Revised ANX and DEP scales

- Data: adolescents admitted to the adolescent program of a private tertiary care inpatient treatment facility (N = 169, mean age 15.43, SD = 1.44, 39.6% male).

Method
- Five instruments were used to validate each of the ANX and DEP scales of the PAI-A (Table 1)

Table 1.
Validating Instruments

ANX:
- Multidimensional Anxiety Scale for Children (MASC) total score
- Child Behavior Checklist (CBCL/6-18) Anxiety Problems subscale
- Youth Self-Report (YSR) Anxiety Problems subscale
- Youth report Diagnostic Interview Schedule for Children (Y-DISC) Generalized Anxiety Disorder subscale
- Parent report Diagnostic Interview Schedule for Children (P-DISC) Generalized Anxiety Disorder subscale

DEP:
- Beck Depression Inventory-II (BDI-II) total score
- Child Behavior Checklist (CBCL/6-18) Affective Problems subscale
- Youth Self-Report (YSR) Affective Problems subscale
- Youth report Diagnostic Interview Schedule for Children (Y-DISC) Major Depressive Disorder subscale
- Parent report Diagnostic Interview Schedule for Children (P-DISC) Major Depressive Disorder subscale

The ANX and DEP IRT-derived PAI-A scales were expected to exhibit the relationships and group mean differences found with the full-length PAI-A. Specifically, to evaluate construct validity:
(1) Correlations between the PAI-A ANX and DEP scale scores and the scores based on the continuously scored instruments (MASC, BDI-II, CBCL/6-18, YSR) were examined.
(2) Group mean differences on the PAI-A ANX and DEP scores between the Y-DISC and P-DISC categories "no" or "intermediate diagnosis" versus "positive diagnosis" were examined using t-tests.

Results
- Correlations were similar between the full-length and the IRT-revised DEP and ANX scores and scores from the validating measures (see Table 2).
- Using the full-length PAI-A scores, group mean differences comparing those with and without diagnoses based on the Youth and Parent DISC were significant for both ANX and DEP (see Table 3).
- Using the IRT-revised scores, group mean differences comparing those with and without diagnoses for the Youth DISC were significant for ANX and DEP. For the Parent DISC, group mean differences in scores were found for DEP, but not for ANX (see Table 3).

Table 2.
Correlation Coefficients with Continuous Anxiety and Depression Instruments

	Anxiety			Depression		
	MASC (N = 122)	YSR (N = 166)	CBCL (N = 162)	BDI-II (N = 128)	YSR (N = 166)	CBCL (N = 162)
PAI-A Full-length	0.73	0.65	0.24	0.84	0.79	0.38
PAI-A IRT-revised	0.73	0.66	0.24	0.84	0.79	0.37

Table 3.
Means for Categorical Anxiety and Depression Instruments

	Anxiety				Depression			
	Y-DISC		P-DISC		Y-DISC		P-DISC	
	No Dx (n = 147)	Yes Dx (n = 18)	No Dx (n = 138)	Yes Dx (no = 22)	No Dx (n = 104)	Yes Dx (n = 58)	No Dx (n = 82)	Yes Dx (no = 79)
PAI-A Full-length	39.22***	53.44***	40.06*	47.50*	36.35***	51.28***	39.44*	44.08*
PAI-A IRT-revised	26.90***	35.72***	27.54	31.41	24.62***	37.10***	27.46*	31.03*

Note. *** p < .001, ** p < .01, * p < .05

Discussion

This study extended previous classical test theory analyses of the PAI-A into an IRT framework for the ANX and DEP scales. Based on IRT item analysis, each original 18-item scale was reduced to 13 items. No specific effort was made to create equal length scales for ANX and DEP.

Evaluating the construct validity of the original, full-length PAI-A and the IRT-revised scale scores, all the other scale scores correlated well with the PAI-A ANX and DEP scale scores. The MASC, YSR, and BDI-II showed particularly strong correlations. The correlations for the youth-report YSR were higher compared to the parent report CBCL/6-18. For the full-length PAI-A scale scores, those with a positive diagnosis on the youth and parent DISC had higher scores compared to those with no/intermediate diagnosis. For the IRT-revised scores, the mean differences were found between the youth DISC categories for ANX and DEP; however, differences between the parent DISC categories were found only for DEP. Those with a diagnosis had higher scores compared to those with the no/intermediate diagnosis.

In sum, the full-length PAI-A ANX and DEP scores correlated well with scores from other measures of anxiety and depression, adding information regarding the construct validity of the original scale. Additionally, the IRT-derived scale functioned similarly to the full-length scale, for all analyses except the P-DISC group comparisons for ANX. Taken together, this evidence supports the use of the IRT-derived ANX and DEP scales as measures of anxiety or depression. Further research may consider the expansion of these methods to other PAI-A scales, or the use of the shortened ANX and DEP scales as individual instruments.

References
Morey, L. C. (2007). *The Personality Assessment Inventory – Adolescent professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.
Samejima, F. (1996). Evaluation of mathematical models for ordered polychotomous responses. *Behaviormetrika, 23*(1), 17-35. doi: 10.2333/bhmk.23.17