PROMIS Item Banks for Emotional Distress: Depression, Anxiety, and Anger

Paul A. Pilkonis, Seung W. Choi, Steven P. Reise, Angela M. Stover, William T. Riley, and David Cella on behalf of the PROMIS Cooperative Group

NIH Roadmap Initiative: U01 AR052155
A Critical Distinction: Item Banks vs. Static Scales

- **Item Banks**
  - Item Response Theory (IRT)
  - Computerized adaptive testing
  - Customized short forms

- **Static Scales**
  - Classical Test Theory (CTT)
  - A single form of the test
  - Lack of comparability across different scales
Psycho-metric Testing

Item Bank v 1.0
(IRT-calibrated items reviewed for reliability, validity, and sensitivity)

Short Form Instruments

New Items

Item Pool

Content Expert Review
Focus Groups
Cognitive Testing
Secondary Data Analysis

Questionnaire administered to large representative sample

Probability of Response

Theta

Psychometric Testing

CAT
<table>
<thead>
<tr>
<th>Subdomain</th>
<th>Core Citations</th>
<th>Number of Measures</th>
<th>Number of Items (Initial Pool)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>198</td>
<td>75</td>
<td>567</td>
</tr>
<tr>
<td>Anxiety</td>
<td>173</td>
<td>145</td>
<td>465</td>
</tr>
<tr>
<td>Anger</td>
<td>213</td>
<td>82</td>
<td>446</td>
</tr>
<tr>
<td>Total</td>
<td>584</td>
<td>302</td>
<td>1,478</td>
</tr>
</tbody>
</table>
Conceptual Framework: Depression

» 1 Domain
» 6 Subdomains
  » Mood
  » Cognition
  » Behavior
  » Somatic Complaints
  » Suicidality
  » Perceived need for treatment
» 23 Factors
» 46 Facets ("Bins")
Legacy Scale Selection

- Benchmarking to current gold standard measures
- Convergent and discriminant validity
- Emotional Distress:
  - **Depression**: Center for Epidemiologic Studies Depression Scale (CESD)
  - **Anxiety**: General Distress (Anxiety) Subscale from the Mood and Anxiety Symptom Questionnaire (MASQ)
  - **Anger**: Aggression Questionnaire (AQ)
Wave 1 Sample

- N = 14,829 (Block testing 14,057, Full-bank 782)
- General population n = 6,995
- Clinical population n = 7,844
  - Online panel respondents who answered “yes” to the question: “Has a doctor or health care provider ever told you that you have depression?”
- Clinical sample in Pittsburgh: 250 psychiatric outpatients
- Scale setting sub-sample matching 2000 census on gender, age, race, ethnicity, and education
  - Used to establish U.S. population norms
### Full Sample Demographics

<table>
<thead>
<tr>
<th></th>
<th>Anxiety</th>
<th>Depression</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Testing Format</strong></td>
<td>Full-Bank: 788</td>
<td>Full-Bank: 782</td>
<td>Full-Bank: 858</td>
</tr>
<tr>
<td></td>
<td>Block: 14,048</td>
<td>Block: 14,057</td>
<td>Block: 14,043</td>
</tr>
<tr>
<td><strong>Population Type</strong></td>
<td>General: 6,992</td>
<td>General: 6,995</td>
<td>General: 7,061</td>
</tr>
<tr>
<td></td>
<td>Clinical: 7,844</td>
<td>Clinical: 7,844</td>
<td>Clinical: 7,840</td>
</tr>
<tr>
<td><strong>Age Mean (SD)</strong></td>
<td>54.098 (16.474)</td>
<td>54.093 (16.465)</td>
<td>54.063 (16.434)</td>
</tr>
<tr>
<td><strong>Gender (%)</strong></td>
<td>Male: 48.02</td>
<td>Male: 48.00</td>
<td>Male: 48.11</td>
</tr>
<tr>
<td><strong>Ethnicity (%)</strong></td>
<td>Hispanic: 8.39</td>
<td>Hispanic: 8.42</td>
<td>Hispanic: 8.39</td>
</tr>
<tr>
<td><strong>Race (%)</strong></td>
<td>White: 82.95</td>
<td>White: 82.94</td>
<td>White: 82.95</td>
</tr>
<tr>
<td></td>
<td>Black: 8.20</td>
<td>Black: 8.22</td>
<td>Black: 8.18</td>
</tr>
<tr>
<td></td>
<td>Asian: 0.53</td>
<td>Asian: 0.53</td>
<td>Asian: 0.52</td>
</tr>
<tr>
<td></td>
<td>Native: 0.64</td>
<td>Native: 0.63</td>
<td>Native: 0.61</td>
</tr>
<tr>
<td><strong>Education (%)</strong></td>
<td>HS or less: 17.96</td>
<td>HS or less: 17.97</td>
<td>HS or less: 17.73</td>
</tr>
</tbody>
</table>
Statistical and Psychometric Analyses

• Preliminary analyses
  – Assess data quality, identify sparse cells, monotonic responses, internal consistency reliability

• Assessment of dimensionality
  – Exploratory factor analysis, confirmatory factor analysis, fit indices, root mean square error of approximation, local dependence

• Estimation of IRT parameters
  – Graded response model
Statistical and Psychometric Analyses

- Assessment of fit to IRT model (IRTFIT)
- Differential item functioning
  - Gender
  - Age (<65 and ≥65)
  - Education (≤high school and >high school)
Final Item Bank Content

Emphasis on cognitive and affective content
(internalized psychological distress)

Final number of items in banks:

- **Depression**: 28 (cognition 17, mood 9, behavior 1, suicidality 1)
- **Anxiety**: 29 (mood 10, cognition 9, somatic complaints 8, need for treatment 1)
- **Anger**: 29 (cognition 13, behavior 8, mood 7, need for treatment 1)
Short Form Selection

• CAT Selection Rank
• Expected Information
• Content Consideration
• Final Number of SF Items: 7-8
Test Information Function: Depression

![Test Information Function Diagram](image-url)
Test Information Function: Anxiety

![Graph showing Test Information Function for Anxiety](image-url)
# Correlations: Depression Measures

<table>
<thead>
<tr>
<th></th>
<th>Bank (28)</th>
<th>CAT (8)</th>
<th>SF (8)</th>
<th>CESD (20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank</td>
<td>.98</td>
<td>.96</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td></td>
<td>.97</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>Short Form</td>
<td></td>
<td></td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td>CESD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Correlations: Anxiety Measures

<table>
<thead>
<tr>
<th></th>
<th>Bank (29)</th>
<th>CAT (7)</th>
<th>SF (7)</th>
<th>MASQ (11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank</td>
<td>0.97</td>
<td>0.96</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td></td>
<td>0.97</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Short Form</td>
<td></td>
<td></td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>MASQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Correlations: Anger Measures

<table>
<thead>
<tr>
<th></th>
<th>Bank (29)</th>
<th>CAT (8)</th>
<th>SF (8)</th>
<th>AQ (12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank</td>
<td></td>
<td>.97</td>
<td>.96</td>
<td>.57</td>
</tr>
<tr>
<td>CAT</td>
<td></td>
<td></td>
<td>.98</td>
<td>.54</td>
</tr>
<tr>
<td>Short Form</td>
<td></td>
<td></td>
<td></td>
<td>.54</td>
</tr>
<tr>
<td>AQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Preliminary Validity Evidence

- **Content Validity**
  - Focus group participants mentioned over 90% of depression facets (43/46 bins) from a priori list
  - Strong evidence for content validity of depression conceptual framework (Kelly et al., 2010)

- **Convergent Validity**
  - Correlations with legacy measures
Wave 2 Validation Study: Pain and Depression

- Target N = 115 patients with major depressive disorder who recently started treatment
- Target N = 185 with back and leg pain scheduled for pain-relief injection
- Baseline, 1 month follow-up, 3 month follow-up
- PROMIS CATs, legacy measures, diagnostic interview, medical records
- Analysis of relationship between pain and depression
Demographic Characteristics: Depression Sample

- Depression Sample \((n = 111)\)
  - Mean age = 35 (SD = 12), range = 18-59
  - 27% Male
  - 5% Hispanic
  - 79% White, 14% African American, 3% American Indian, and 3% Asian American (2% unknown)
  - 25% high school or less, 60% some college or college degree, and 14% advanced degree.
  - 33% income <$20,000, 32% income $20,000-$49,999, 16% income $50,000-$99,999, and 12% income >$100,000 (6% unknown)
Demographic Characteristics: Pain Sample

- Pain Sample \( (n = 229) \)
  - Mean age = 54 (SD = 15), range = 19-86
  - 44% Male (3% unknown)
  - 4% Hispanic (5% unknown)
  - 84% White, 4% African American, 4% American Indian, 2% Native Hawaiian, and 1% Asian American (5% unknown)
  - 19% high school or less, 58% some college or college degree, and 19% advanced degree (3% unknown)
  - 13% income <$20,000, 21% income $20,000-$49,999, 38% income $50,000-$99,999, and 23% income >$100,000 (5% unknown)
### Descriptive Statistics: Depression Sample

<table>
<thead>
<tr>
<th>Measures</th>
<th>Baseline</th>
<th>T2 (1 Month)</th>
<th>T3 (3 Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Anger</td>
<td>56.5</td>
<td>9.3</td>
<td>111</td>
</tr>
<tr>
<td>AQ</td>
<td>30.9</td>
<td>8.8</td>
<td>110</td>
</tr>
<tr>
<td>Anxiety</td>
<td>63.4</td>
<td>6.9</td>
<td>111</td>
</tr>
<tr>
<td>MASQ</td>
<td>25.9</td>
<td>8.2</td>
<td>110</td>
</tr>
<tr>
<td>Depression</td>
<td>63.9</td>
<td>7.7</td>
<td>111</td>
</tr>
<tr>
<td>CES-D</td>
<td>31.0</td>
<td>9.8</td>
<td>111</td>
</tr>
<tr>
<td>HRSD</td>
<td>19.2</td>
<td>4.6</td>
<td>111</td>
</tr>
</tbody>
</table>
### Convergent Validity: Depression Sample

<table>
<thead>
<tr>
<th></th>
<th>Anger</th>
<th>Anxiety</th>
<th>Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>r</strong></td>
<td>.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AQ</strong></td>
<td></td>
<td>.70</td>
<td></td>
</tr>
<tr>
<td><strong>MASQ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CES-D</strong></td>
<td></td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td><strong>HRSD</strong></td>
<td></td>
<td>.42</td>
<td></td>
</tr>
</tbody>
</table>

All sig at .01 (2-tailed)
## Descriptive Statistics: Pain Sample

<table>
<thead>
<tr>
<th>Measures</th>
<th>Baseline</th>
<th>T2 (1 Month)</th>
<th>T3 (3 Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Anger</td>
<td>53.4</td>
<td>8.4</td>
<td>218</td>
</tr>
<tr>
<td>AQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>55.4</td>
<td>8.8</td>
<td>219</td>
</tr>
<tr>
<td>MASQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>53.2</td>
<td>8.8</td>
<td>218</td>
</tr>
<tr>
<td>CES-D</td>
<td>18.0</td>
<td>11.2</td>
<td>209</td>
</tr>
<tr>
<td>HRSD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Responsiveness for Depression Sample: PROMIS Anger vs. AQ

![Graph showing the comparison of effect size for Anger and AQ over time points 1, 2, and 3.](image-url)
Responsiveness for Depression Sample: PROMIS Anxiety vs. MASQ

![Graph showing the comparison between Anxiety and MASQ effect sizes over time points 1, 2, and 3. The graph displays a line for Anxiety and another for MASQ, illustrating their respective changes over time.](image-url)
Responsiveness for Depression Sample: PROMIS Depression vs. CES-D vs. HRSD

![Graph showing the comparison between Depression, HRSD, and CESD over time points 1, 2, and 3.]
Next Steps …

- Development of additional item banks
  - Alcohol use and abuse
  - Externalizing disorders
What We Learned from Prior Work on Alcohol Use

• **Item content:** Use of criteria for alcohol abuse and dependence, i.e., severe symptoms and consequences.

• **Item parameters:** High threshold parameters, indicating that it is difficult for respondents to endorse such items in most samples.

• **Item bank:** A peaked but narrow information curve at the upper range of severity (1-2 SDs above the mean).
Test Information Curve: Alcohol
Conceptual Framework: Alcohol Use

8 Sub-domains

- Consumption
- Craving and Control
- Triggers
- Positive Consequences
- Negative Consequences
- Positive Expectancies
- Negative Expectancies
- General Attitudes

» 105 Facets ("Bins")
Demographic Characteristics: Clinical Sample

- Clinical sample ($n = 407$)
  - Mean age = 37 (SD = 12), range = 18-70
  - 61% Male
  - 8% Hispanic
  - 72% White, 24% African American, 8% American Indian, 2% Native Hawaiian, and 1% Asian American
  - 63% high school or less, 36% some college or college degree, and 2% advanced degree.
  - 75% income <$20,000, 18% income $20,000-$49,999, 6% income $50,000-$99,999, and 1% income >$100,000
Demographic Characteristics: General Population (Internet) Sample

• YouGov Polimetrix sample ($n = 1,000$)
  – Mean age = 45 (SD = 16), range = 18-86
  – 54% Male
  – 10% Hispanic
  – 88% White, 10% African American, 4% American Indian, 2% Asian American, and 1% Native Hawaiian
  – 28% high school or less, 57% some college or college degree, and 15% advanced degree.
  – 9% income <$20,000, 31% income $20,000-$49,999, 40% income $50,000-$99,999, and 21% income >$100,000
Exploratory Factor Analysis: Total Sample

• 5-factor structure
  – Consumption (CONS): included consumption, general attitudes, craving and control, and triggers
  – Negative Consequences (NECO)
  – Positive Consequences (POCO)
  – Negative Expectancies (NEXP)
  – Positive Expectancies (PEXP)
Confirmatory Factor Analysis

• The final sets of items after CFA:
  – Consumption: 46 items
  – Negative Consequences: 31 items
  – Positive Consequences: 21 items
  – Negative Expectancies: 13 items
  – Positive Expectancies: 13 items

• Total item pool = 124 items
Test Information Curve - CONS
Test Information Curve - NECO
Test Information Curve - POCO
Test Information Curve - NEXP
Test Information Curve - PEXP
Convergent and Discriminant Validity

• Correlations with the AUDIT total score
  – Consumption = .89
  – Negative Consequences = .78
  – Positive Consequences = .50
  – Positive Expectancies = .46
  – Negative Expectancies = .23
Next Steps ...

- Use of the PROMIS measures for depression, anxiety, anger, and sleep in the DSM-5 field trials
- Transdiagnostic, dimensional assessment
PROMIS Website

http://www.nihPROMIS.org

The site includes a link to Assessment Center, where PROMIS measures can be accessed.