Rasch Analysis of the Pain Catastrophizing Scale (PCS) in the Context of Workplace Bullying for a Sample of Nursing Faculty from Midwest U.S. Universities

Aryn C. Karpinski, Ph.D.*
Laura C. Dzurec, Ph.D., PMHCNS-BC**
Gail E. Bromley, Ph.D., RN, CNS**
Timothy W. Meyers, MSN, RN*
Shawn M. Fitzgerald, Ph.D.*

* College of Education, Health, and Human Services
** College of Nursing

August 3, 2012
10:55 AM
Introduction

- Workplace Bullying
  - Repeated, offensive intimidation and abuse of personal power
    (Einarsen, Hoel, & Notelaers, 2009)
  - Interactions are nearly invisible outside bully-victim dyads
    (Brannan, 2007)
  - Reporting is constrained
    (MacIntosh, Wuest, Gray, & Aldous, 2010)

It is important to find not only a clear definition of workplace bullying, but also the best possible way to measure it.
Paradoxes of Workplace Bullying

- Has defied clear definition for 30 years
  (e.g., Beale & Hoel, 2010; Field, 2003; Frankel, Leonard, & Denham, 2006; Jennifer, Cowie, & Ananiadou, 2003; MacIntosh et al., 2010; Wheeler, Halbesleben, & Shanine, 2010)

- Definition varies among victims, bullies, researchers, and others
  (Mikkelsen & Einarsen, 2001; Namie & Namie, 2009; Salin, 2003)
Catastrophization in Pain
(Edwards, Kronfli, Haythornthwaite, Smith, McGuire, & Page, 2008)

- Catastrophization is a “tendency to focus on and exaggerate the threat value of painful stimuli and negatively evaluate…(one’s) own ability to deal with pain.”
  (Keefe, Lipkus, Lefebvre, Hurwitz, Clipp, Smith, & Porter, 2003, p. 2)

- May model bullying victims’ experiences
  (Bromley & Dzurec, 2010; Dzurec & Bromley, 2012)
Catastrophization in Workplace Bullying

- Rumination, Magnification, and Helplessness (Dzurec & Bromley, 2012)

- Apparent in victim responses throughout an extensive body of international research
  (e.g., Baillien & DeWitte, 2009; Bromley & Dzurec, 2010; Einarsen et al., 2009; European Foundation for the Improvement of Living and Working Conditions, 2003; Hoel & Giga, 2006; Hutchinson, Vickers, Jackson, & Wilkes, 2006; Johnson, 2009; Lackner & Gurtman, 2004, 2005; Sullivan, Tripp, & Santor, 2000; Wheeler et al., 2010)
The objective of the current study is to define and measure the construct of victim responses to workplace bullying as pain catastrophization.

**Research Question:**

What are the psychometric properties of a measure of pain catastrophization in the context of workplace bullying?
Participants
  For the entire study, there were 154 nursing faculty members from Midwestern United States universities.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>M or n</th>
<th>SD or %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (n = 86)</td>
<td>51.6</td>
<td>9.3</td>
</tr>
<tr>
<td>Years Teaching (n = 88)</td>
<td>14.0</td>
<td>10.4</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>84</td>
</tr>
<tr>
<td>Faculty Rank</td>
<td>Non-Tenure Track</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Assistant Professor</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Associate Professor</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Full Professor</td>
<td>10</td>
</tr>
<tr>
<td>Frequency of Bullying</td>
<td>Daily</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Weekly</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Less than 3 Times/Year</td>
<td>4</td>
</tr>
<tr>
<td>Report Bullying</td>
<td>No</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>36</td>
</tr>
</tbody>
</table>
Materials

- The 13-item Pain Catastrophizing Scale (PCS) was used with permission.
  
  (Sullivan, Bishop, & Pivik, 1995)

- 5-Point Likert Scale from 0 ("Not at All") to 4 ("All the Time")

- Score Range = 0 to 52

- Rumination, Magnification, and Helplessness dimensions established through factor analysis
  
  - Coefficient α’s were .87, .60, and .79 for the three PCS subscales, respectively.
  
  - Coefficient α for the total instrument was .87.
**Methods**

**Pain Catastrophizing Scale**
(PCS; Sullivan, Bishop, & Pivik, 1995)

*Directions*: Please think about the worst case of workplace bullying you have ever encountered. Try to remember the emotional pain you experienced at the time (e.g., fear of losing your job, shame or embarrassment, a sense of personal rejection). In light of that memory, please answer the questions below as you reflect on past painful experiences and indicate the degree to which you experienced each of the 13 thoughts or feelings when experiencing pain on a 5-point scale: 0 (i.e., Not at All), 1 (i.e., To a Slight Degree), 2 (i.e., To a Moderate Degree), 3 (i.e., To a Great Degree), and 4 (i.e., All the Time).

1. I worry all the time about whether the pain will end.
2. I feel I can't go on.
3. It's terrible and I think it's never going to get any better.
4. It's awful and I feel that it overwhelms me.
5. I feel I can't stand it anymore.
6. I become afraid that the pain will get worse.
7. I keep thinking of other painful events.
8. I anxiously want the pain to go away.
9. I can't seem to keep it out of my mind.
10. I keep thinking about how much it hurts.
11. I keep thinking about how badly I want the pain to stop.
12. There's nothing I can do to reduce the intensity of the pain.
13. I wonder whether something serious may happen.

- **Helplessness**
- **Magnification**
- **Rumination**
Methods

- Procedure
  - Demographic questions and the PCS were posted on a survey-hosting website.

  - Deans from 26 baccalaureate and higher degree nursing programs were contacted to allow their faculty members to participate.

  - On the survey, respondents were asked to indicate whether or not they had experienced or observed workplace bullying. Only those answering “Yes” were allowed to continue.
Methods

- Data Analysis
  - Rasch Analysis – Rating Scale Model
    (Andrich, 1978)
Results

Descriptive Information – PCS

- At the item level, participants used every Likert category option available for all the items.

- Use of the Likert categories “To a Great Degree” and “All the Time” (i.e., coded 3 and 4) were not endorsed as frequently as the others.
  - Lowest Item Mean: Item 2 – “I feel I can’t go on”
    - $M = .85$, $SD = 1.22$
  - Highest Item Mean: Item 9 – “I can’t seem to keep it out of my mind”
    - $M = 1.47$, $SD = 1.35$
  - All the correlations between the items were strong ($r \geq .601$ for all) and statistically significant ($p < .001$).
Rasch Analysis

- Summary Statistics

Person Summary Statistics

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>S.D.</th>
<th>Max.</th>
<th>Min.</th>
<th>REAL RMSE</th>
<th>MODEL RMSE</th>
<th>S.E. of Person Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Score</td>
<td>16.9</td>
<td>14.9</td>
<td>51.0</td>
<td>1.0</td>
<td>.60</td>
<td>.57</td>
<td>.23</td>
</tr>
<tr>
<td>Count</td>
<td>12.8</td>
<td>.4</td>
<td>13.0</td>
<td>11.0</td>
<td>2.25</td>
<td>2.26</td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>-1.33</td>
<td>2.33</td>
<td>4.64</td>
<td>-4.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEASURE</td>
<td>.52</td>
<td>.23</td>
<td>1.03</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFIT</td>
<td>MNSQ</td>
<td>ZSTD</td>
<td>OUTFIT</td>
<td>MNSQ</td>
<td>ZSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>.0</td>
<td>.97</td>
<td>-1.1</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPARATION</td>
<td>3.77</td>
<td>PERSON RELIABILITY</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPARATION</td>
<td>4.00</td>
<td>PERSON RELIABILITY</td>
<td>.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Item Summary Statistics

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>S.D.</th>
<th>Max.</th>
<th>Min.</th>
<th>REAL RMSE</th>
<th>MODEL RMSE</th>
<th>S.E. of Item Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Score</td>
<td>140.2</td>
<td>18.7</td>
<td>168.0</td>
<td>98.0</td>
<td>.16</td>
<td>.15</td>
<td>.13</td>
</tr>
<tr>
<td>Count</td>
<td>114.5</td>
<td>2.6</td>
<td>116.0</td>
<td>106.0</td>
<td>.41</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>MEASURE</td>
<td>.00</td>
<td>.44</td>
<td>.98</td>
<td>-1.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFIT</td>
<td>MNSQ</td>
<td>ZSTD</td>
<td>OUTFIT</td>
<td>MNSQ</td>
<td>ZSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>-1.1</td>
<td>.97</td>
<td>-1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPARATION</td>
<td>2.56</td>
<td>ITEM RELIABILITY</td>
<td>.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPARATION</td>
<td>2.71</td>
<td>ITEM RELIABILITY</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For both persons and items, the mean square infit and outfit are expected to be 1.0, and the standardized infit and outfit are expected to be 0.0.

Low person separation (< 2; Reliability < .8) implies that the instrument may not be sensitive enough to distinguish between high and low performers.

Low item separation (< 3; Reliability < .9) implies that the sample is not large enough to confirm the item difficulty hierarchy (construct validity).

The mean of the item logit position is always 0.0. If the person mean is negative compared to the item mean, this indicates that the items were potentially too “difficult” for the sample.

This is similar to Coefficient α.
Rasch Analysis

- Response Scale Analysis

**SUMMARY OF CATEGORY STRUCTURE.** Model="R"

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>OBSERVED</th>
<th>OBSVD</th>
<th>SAMPLE</th>
<th>INFIT</th>
<th>OUTFIT</th>
<th>STRUCTURE</th>
<th>CATEGORY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LABEL</td>
<td>COUNT</td>
<td>AVGRE</td>
<td>EXPECT</td>
<td>MNSQ</td>
<td>CALIBRATN</td>
<td>MEASURE</td>
</tr>
<tr>
<td>0 0</td>
<td>635</td>
<td>43</td>
<td>-3.39</td>
<td>-3.38</td>
<td>1.00</td>
<td>.99</td>
<td>NONE</td>
</tr>
<tr>
<td>1 1</td>
<td>346</td>
<td>23</td>
<td>-1.71</td>
<td>-1.70</td>
<td>.97</td>
<td>.92</td>
<td>-2.17</td>
</tr>
<tr>
<td>2 2</td>
<td>195</td>
<td>13</td>
<td>-.19</td>
<td>-2.27</td>
<td>.86</td>
<td>.82</td>
<td>-.39</td>
</tr>
<tr>
<td>3 3</td>
<td>165</td>
<td>11</td>
<td>1.01</td>
<td>1.01</td>
<td>.96</td>
<td>1.04</td>
<td>.52</td>
</tr>
<tr>
<td>4 4</td>
<td>148</td>
<td>10</td>
<td>2.70</td>
<td>2.75</td>
<td>1.31</td>
<td>1.28</td>
<td>2.03</td>
</tr>
<tr>
<td>MISSING</td>
<td>19</td>
<td>1</td>
<td>-.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Observed Count** is the number of times the category was selected across all items and all persons. **Observed Average** is the average of logit positions modeled in the category. It should increase by category value.

The probability curves display the likelihood of category selection (Y-axis) by the person-minus-item measure (X-axis). If all categories are utilized appropriately, each category value will be the most likely at some point on the continuum.

Structure Calibration is the logit calibrated difficulty of the step. This is demonstrated in the probability curve figure where the transition points between one response scale category and the next are the values in the table. These values are expected to increase with each category.
**Results**

- **Rasch Analysis**
  - **Item Misfit Diagnostics**

**Measure** is the logit position of the item, with Model S.E. being the standard error of measurement for the item.

Fit statistics (mean square infit and outfit) show the amount of distortion of the measurement system. Values less than 1.0 indicate observations are too predictable (i.e., redundancy, data overfit the model). Values greater than 1.0 indicate unpredictability (i.e., unmodeled noise). No definitive rules exist regarding what is considered acceptable or unacceptable fit. For the current study, infit and outfit values of less than 2.0 were considered acceptable.

A point measure correlation (PT-Measure) is a correlation between the observations on an item and the corresponding person measures. Below .15 indicates a potentially misfitting item.
Results

- Rasch Analysis
  - Item/Person Map (Variable Map)

This is the map of persons and items. The distribution of person positions is on the left side of the vertical line and items on the right. "M" marks the person and item mean; "S" is one standard deviation away from the mean; and "T" is two standard deviations away from the mean.

There were a few persons whose position was above where items were measuring, but a considerable amount of persons whose position was below where the items were measuring. This could indicate that the measure is too difficult for the intended sample.

The items are approximately normally distributed (i.e., although slightly leptokurtic). The majority of the items are compressed between -1 and +1 on the scale. There are no large gaps between items, and several items appeared at the exact same location on the vertical ruler (i.e., 13, 3, 4, 5, 8) indicating redundancy.
Although the internal consistency of the PCS was high (Coefficient $\alpha = .97$), there were a few concerns with some items and the measure as a whole.

Overall, the results support the use of the PCS to gauge response to workplace bullying (*upon further revision*).

- This may support diagnosis and remedying of problems before they become major impediments to productivity and individual well-being.
Discussion

- Concerns with the Items
  - Items 13, 3, 4, 5, and 8 were located at the same position on the vertical ruler indicating redundancy. This suggests a need to review these items to make sure they are not too similarly worded or structured.
    - Items 3, 4, and 5 were categorized as Helplessness items.
  - Unstandardized and standardized infit and outfit measures indicated that items 7, 10, 11, and 13 were potentially problematic in terms of being too predictable or unpredictable (i.e., construct-irrelevant variance).
Concerns with the Measure

- Items spanned from -1 to +1 logits, which was much narrower than the persons range of -5 to +5 logits, indicating that the range of the trait being measured may need to be extended.

- Every Likert category option was used; however, use of the categories “To a Great Degree” and “All the Time” were not endorsed as frequently as the others. Revising the Likert category structure may be needed.
Limitations/Future Directions

- Sample Size
  - Although the sample size was adequate ($N = 116$) to maintain item calibration stability within $\pm 1/2$ logits (i.e., 95% CI - 100 people, more (and diverse) participants could be recruited in future research (Linacre, 1994).

- Dimensionality
  - The three-factor structure established by Sullivan and colleagues (1995) should be further examined for the purposes of using the PCS to measure workplace bullying.
    - Future studies might include Confirmatory Factor Analysis (CFA) to confirm the three-factor structure of Helplessness, Magnification, and Rumination.
Selected References

Contact Information

Aryn C. Karpinski, Ph.D.

College of Education, Health, and Human Services
Foundations, Leadership, and Administration
Evaluation and Measurement

316 D White Hall
Kent, OH 44240
E-mail: akarpins@kent.edu
Phone: (330) 672-2012