Rasch Analysis of the Mental Health Recovery Measure

Yen-Ching Chang, Sarah H. Ailey, Tamar Heller, Ming-De Chen

Consumer-oriented recovery among people with mental illness has been discussed for more than two decades, but few reliable and valid recovery measurements are currently available. This study used Rasch methods to assess the Mental Health Recovery Measure (MHRM). Participants were 156 adults with mental illness who lived in the community. After the Rasch analyses, the MHRM was modified to a 26-item measure with a 4-point Likert scale. Unidimensionality was confirmed for the revised MHRM, and it also showed proper rating scale functioning and high reliability. The revised MHRM is sufficient to assess only those in the initial and middle stages of recovery. More high-recovery-level items are needed to assess people in a high-recovery stage. Occupational therapists can use the revised MHRM in future quantitative studies and program evaluation.


Consumer-oriented recovery has been regarded as a guiding vision of the future of mental health services (Anthony, 1993). The concept of consumer-oriented recovery is new and mainly derived from consumers’ personal experiences. Although recovery is an individual journey and has various meanings for different people, some commonly identified recovery components are hope, empowerment, taking personal responsibility, self-redefinition, and participation in meaningful activities (Davidson, O’Connell, Tondora, Lawless, & Evans, 2005; Ridgway, 2001; Young & Ensing, 1999). By following the recovery concept, people with mental illness are treated as independent individuals rather than dependent patients, and mental health professionals are helpers rather than authorities.

Research on the role of recovery in mental health and clinical program evaluation is hampered by the lack of theoretical models of recovery and reliable and valid theoretically based recovery measures (Bellack, 2006; Jacobson & Curtis, 2000). Young and Ensing (1999) developed a promising model that describes the recovery process of people with mental illness. The model was developed using grounded theory analysis of seven semistructured interviews and two focus groups with 18 people with mental illness. The model covers major recovery concepts found in the literature and includes three phases of recovery.

The initiation of the recovery process—Phase 1—is “overcoming stuckness.” Themes include acknowledging and accepting illness, having the desire and motivation to change, and finding a source of hope and inspiration. For people who have been independent, this phase is difficult because they have to admit that they have limitations and need help. When they overcome this inner stumbling block, however, they can start their recovery journey.

Phase 2 is “regaining what was lost and moving forward.” This phase includes three themes: (1) discovering and fostering self-empowerment, (2) learning and engaging in self-redefinition, and (3) returning to basic functioning. In this phase, people with mental illness take responsibility for their
recovery, illness, and lives. Also, they discover their potential, understand the relationship between self and illness, actively participate in various activities, and have meaningful relationships with others. They are open to more opportunities and growth.

Phase 3 is “improving quality of life.” Two themes are included in this phase: (1) attaining an overall sense of well-being and (2) reaching new potentials of higher functioning. People with mental illness pursue higher life goals in this phase. They have a sense of peacefulness and confidence. Moreover, they strive to take on new challenges or financial responsibilities (e.g., cars and houses) to enjoy better lives. They may also become role models and helpers to encourage other people.

The Mental Health Recovery Measure (MHRM; Young & Bullock, 2003) is a theoretically based recovery measure developed according to Young and Ensing’s (1999) recovery model with the capacity to capture the complete recovery perspective. It has not been published in a peer-reviewed journal but was included in a compendium of published and unpublished recovery-related measurements (Campbell-Orde, Chamberlin, Carpenter, & Leff, 2005).

The MHRM initially had 41 items and then was revised to the current 30-item version (Bullock, 2005). It uses a 5-point Likert scale and includes seven domains organized under the three phases of the recovery process model: overcoming stuckness (Phase 1); self-empowerment, learning and self-redefinition, and basic functioning (Phase 2); and overall well-being, new potentials, and advocacy and enrichment (Phase 3).

Two new concepts added to the 30-item MHRM were (1) advocacy and enrichment and (2) spirituality. The advocacy and enrichment domain expands the measurement to include advocating, participating in work or meaningful activities, coping with stigma, and having fiscal stability to better measure the recovery of people at higher recovery levels. Spirituality includes two items and has not been regarded as one domain. However, religion or spirituality is considered as one source of outside support (Spaniol, Wewiorski, Gagne, & Anthony, 2002). Religion or spirituality can help people with mental illness develop hope and pursue their recovery but may not be part of their inner recovery status.

The 30-item MHRM shows appropriate reliability (Cronbach’s α = .93; test–retest reliability = .92) and convergent validity (positive correlations with empowerment, r = .67, and resilience, r = .75). Nevertheless, because the MHRM has not been published in a peer-reviewed journal, detailed information about the validation process is unavailable (e.g., the process and appropriateness of adding two new concepts). In addition, because the MHRM is theoretically based, it is unclear whether the measure represents the three phases of the recovery process model.

Currently, the Rasch model has been widely used in occupational therapy–related measurement evaluation (Classen et al., 2012; Hancock, Bundy, Honey, James, & Tamsett, 2011). Compared with classical test theory (CTT), the Rasch model has many advantages. First, the Rasch model can transform raw scores into an interval scale, whereas CTT analyzes raw scores directly (Smith, Conrad, Chang, & Piazza, 2002). Second, the Rasch model has the characteristic of invariance. The person’s level in the construct (in this case, recovery condition) is not influenced by the tested item (i.e., item-free), and the item’s reflection of the underlying construct (in this case, recovery) is not influenced by the person attempting the item (i.e., sample-free; Wright & Linacre, 1989). On the contrary, CTT is item dependent and sample dependent (Green, 1991; Smith et al., 2002).

Finally, the Rasch model offers more functions and information for evaluating measurements than CTT (Conrad et al., 2004). For instance, person’s level of construct and item’s reflection of the construct are located on the same map in the Rasch model, which provides abundant information to improve the quality of a measurement. Also, the map displays an item hierarchy that can help researchers understand the measured construct and verify the relationship between the observed and the estimated construct (Green, 1996). Overall, the Rasch model can help researchers effectively appraise the construct validity of measures.

Because the MHRM was developed and evaluated by CTT, the purpose of this study was to reevaluate the psychometric characteristics of the MHRM using Rasch psychometric analysis. To improve its construct validity, we sought to detect confusing items and assess whether the MHRM represents the progression toward recovery described in Young and Ensing’s (1999) model.

Method

Research Design

We undertook a psychometric study to evaluate the MHRM using Rasch methods and a convenience sample of people with mental illness living in the Chicago metropolitan area. We obtained institutional review board approval from the University of Illinois at Chicago and the research site. All participants completed the MHRM after providing informed consent.
Participants

Study participants were recruited from a large recovery-oriented community mental health agency located in Chicago. Because inpatient populations are relatively unstable and the consumer-oriented recovery model may be inapplicable to their experience (Frese, Stanley, Kress, & Vogel-Scibilia, 2001), we considered only people living in the community, receipt of services from the collaborating agency, and ability to fill out the study survey independently.

Data Collection and Analysis

With the assistance of program staff, the first author (Chang) convened meetings in several of the agency’s community programs and explained the study’s purpose and procedures to potential participants. Eligible volunteers completed the self-reported, anonymous survey. Data collection occurred from June 2010 through August 2010.

We used the Rasch model to evaluate the collected data. Various facets of construct validity can be addressed by Rasch methods, such as fit statistics and item hierarchy (Smith, 2001). This study included investigation of rating scale functioning, dimensionality, targeting and item hierarchy, and reliability. Data were analyzed using the Winsteps Version 3.70 software program (Linacre, 2010).

Rating Scale Functioning

Because the MHRM uses a 5-point Likert scale, we applied the Rasch rating scale model. A proper rating scale needs to meet several criteria. First, category frequencies should be >10 for each category. Second, average measures should increase when the category value increases. Third, step or threshold calibration should increase monotonically as well. In addition, distances between thresholds should be >1.4 logits to show the differentiation between categories. Finally, the category outfit mean square (MnSq) should be <2.0. Rasch rating scale diagnostics can examine the differentiation in levels of responses (the 5-point scale) and help researchers improve the existing rating scale functioning (Bond & Fox, 2007).

Dimensionality

Unidimensionality is the main requirement of the Rasch model; all items of the instrument must measure a single construct. It can be assessed by item fit statistics and principal components analysis (PCA) of residuals (Smith, 2002). Item fit statistics include two types, outfit and infit. Outfit is sensitive to unexpected responses far from an item’s measure; infit is sensitive to unexpected responses close to an item’s measure. Also, both outfit and infit have two forms, MnSq and standardized mean square (Zstd). MnSq value >1.4 or Zstd value >2.0 indicate misfit, which means that the item performance does not match the expectation of the Rasch model (Bond & Fox, 2007; Wright & Linacre, 1994).

PCA can reveal potential dimensions that may cause multidimensionality. No rule of thumb exists for the unidimensionality judgment. This study used two criteria: >40% of variance explained by the measurement and <15% of residual variance explained by the first principal component of residuals to indicate unidimensionality (Conrad, Iris, Ridings, Langley, & Wilber, 2010; Conrad et al., 2006). In addition, Winsteps can generate unidimensional simulated data. Researchers can compare the PCA of the simulated data with the real data. If the real data show a significantly larger residual variance explained by the first principal component of residuals than the simulated data, the scale may be multidimensional.

Targeting and Item Hierarchy

The logit, or natural logarithm of odds, is the unit of measurement in Rasch analysis. Both item and person data are transformed into the logit metric. Item measures (scores) indicate the degree to which items represent the construct. Person measures indicate the degree to which people possess the construct. In this study, items representing a higher recovery level and participants with a higher recovery condition have higher logit values. The item–person map shows items and people concurrently. Ideally, the distribution of items is expected to cover the distribution of participants. By investigating the map, researchers can determine whether items target people appropriately or whether a ceiling or floor effect exists.

Moreover, the item hierarchy on the map is useful to examine the measurement construct. Because the MHRM is based on a recovery process model, we can compare the observed item hierarchy with the theoretical phases to evaluate the construct validity. Items belonging to Phase 3 of recovery are expected to show higher item measures and to be located on a higher part of the map than items representing Phases 1 and 2.

Reliability

The Rasch model evaluates both person reliability and item reliability. It is important that a measure clearly separate participants in levels of a construct and clearly separate items as reflecting differing levels of a construct.
An alternative way to estimate the spread of people and items in levels of the construct is the separation index \( G_p \) provided in the Rasch model. A higher \( G_p \) indicates that the people or items are more spread out on the measured construct. Also, strata are useful for determining the number of statistically distinct groups of people or items and are calculated by the formula \( (4G_p + 1) / 3 \) (Smith, 2001).

**Results**

A total of 159 participants filled out the survey. We excluded three surveys with inattentive response sets (i.e., the person responded to the whole survey with a specific answer or a pattern), so 156 surveys were included in the analysis. This sample size is adequate for this study according to the suggestion of Linacre (1994). Participants included 107 men and 49 women; 72% were single, 46% were African-American, and 37% were White. Regarding education level, 54% of participants had a high school degree or lower, and 46% reported some college or higher. The four diagnoses reported most often were bipolar disorder (36%), schizophrenia (25%), major depression (18%), and schizoaffective disorder (16%). Mean age was 47 yr (standard deviation \[ SD \] = 11.10).

Initial Rasch analyses found problematic rating scale functioning and misfitting items. We modified the MHRM with the assistance of Rasch methods to improve its performance.

**Rating Scale Functioning**

Rating scale functioning was assessed using several criteria. The MHRM (Table 1) met some requirements but not the step calibration and the distances between thresholds. The second threshold \((-0.26)\) was higher than the third threshold \((-0.35)\); that is, step calibration did not increase monotonically. The distances between the first and second thresholds and the second and third thresholds were 1.19 and 0.09 logits, smaller than the 1.4 logit cutoff. We combined the response categories of 1 (disagree) and 2 (not sure) to improve the rating scale functioning. The revised rating scale had a 4-point scale and met all requirements (Table 2).

**Dimensionality**

Using the 4-point-scale data, five items were misfitting: Items 30, 4, 25, 26, and 16. Rasch analysis detects problematic items but does not provide reasons why they do not fit the Rasch model. We talked with mental health experts to find the most appropriate explanation for the misfit of the items. After consulting three experts (two mental health professionals and one consumer), we kept Item 30 and deleted the other four items.

The item with the greatest misfit was 30—“I have enough money to spend on extra things or activities that enrich my life.” This item is intended to represent a higher level of recovery (Bullock, 2005). Being able to afford certain possessions and leisure activities was under the subtheme “ability to improve standard of living” in Young and Ensing’s model. Many participants in this study (84%) were unemployed and lived on Social Security benefits. Having money did not necessarily indicate a better recovery condition; people in the initial recovery stage who received benefits might respond positively to this high-recovery-level item, so it needed to be reworded to better represent the spirit of the subtheme. Despite the need for revision, because it had the highest item measure, we decided to keep it to better measure this sample.

The second misfitting item was 4—“I take risks to move forward with my recovery.” On the basis of the recovery process model, this item indicates that the person has the courage to try something new or face a challenge. Nevertheless, taking risks implies negative behavior for many people, and respondents may not have caught the positive implied meaning while reading this item. In addition, the item was located in the middle of the item map and had a measure similar to some other recovery items. We decided to delete this item.

Items 25—“When I am feeling low, my religious faith or spirituality helps me feel better”—and 26—“My religious faith or spirituality supports my recovery”—were also misfitting. Young and Ensing (1999) found that spirituality plays an important role in many people’s recovery. Nevertheless, spirituality is one kind of support that people may receive, and some people may prefer other kinds of support (e.g., from friends or family). Knowing the degree to which a specific support helps the respondent’s recovery is not the purpose of the recovery measure. The misfit directly proved this argument, and the two items were deleted.

The final misfitting item was 16—“I am comfortable with my use of prescribed medications.” Mental health consumers have argued that medication is one method, but not the only one, for controlling symptoms. Many people struggle with the side effects of psychiatric medication (Mead & Copeland, 2000). This item may have been inappropriate for people who used other strategies to cope with their symptoms. Even for people who do take medication, this item may not adequately represent their recovery condition. Medication can be considered as
one kind of support, like spirituality. It can help recovery but is not the focus of the recovery measure. Referring to the recovery process model, this item might aim to measure the subtheme “being able to take care of self.” This item needed to be reworded and had a similar measure to some other items. Therefore, we decided to delete this item.

After deleting 4 items, the 26-item MHRM still had 4 items with misfit (see Table 3). Besides Item 30, the 3 items with misfit were 13—“I eat nutritious meals every day,” 8—“I socialize and make friends,” and 3—“I ask for help when I am not feeling well.” Item 13 showed relatively high misfit in addition to Item 30. Some people may be unable to endorse this item because of a limited budget or reliance on food stamps rather than a lack of ability to take care of self. However, both outfit and infit MnSq values were <1.4. Also, Item 13 presented a unique concept in recovery and showed a relatively high item measure. Hence, it was decided to keep this item. Items 8 and 3 had marginal misfit and did not show obvious wording problems, so we kept them as well. We kept all 26 items for the subsequent analyses. Moreover, after 4 items were deleted, the functioning of the rating scale was still appropriate.

Table 1. Items of the Mental Health Recovery Measure

<table>
<thead>
<tr>
<th>Item</th>
<th>Item Measure</th>
<th>Phase of Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.96</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>−0.32</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>−0.07</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.22</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>−0.58</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0.30</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0.16</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>−0.07</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>−0.60</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>−0.42</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>−0.67</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>−0.27</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>0.77</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>0.36</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>0.12</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>0.04</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>−0.14</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>0.05</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>0.77</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>0.25</td>
<td>3</td>
</tr>
<tr>
<td>21</td>
<td>0.46</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>−0.30</td>
<td>3</td>
</tr>
<tr>
<td>23</td>
<td>−0.10</td>
<td>3</td>
</tr>
<tr>
<td>24</td>
<td>−0.10</td>
<td>3</td>
</tr>
<tr>
<td>25</td>
<td>−0.37</td>
<td>NA</td>
</tr>
<tr>
<td>26</td>
<td>−0.23</td>
<td>NA</td>
</tr>
<tr>
<td>27</td>
<td>−0.04</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>0.16</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>0.36</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>1.25</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. NA = not applicable.

*aThe two spirituality items do not have an affiliated phase.

Table 2. Summary of the Rating Scale Functioning After Revising the Categories

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Category Frequency</th>
<th>Average Measure</th>
<th>Step Calibration</th>
<th>Outfit MnSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = strongly disagree</td>
<td>126</td>
<td>−1.50</td>
<td>NA</td>
<td>1.25</td>
</tr>
<tr>
<td>1 = disagree</td>
<td>919</td>
<td>0.38</td>
<td>−2.45</td>
<td>1.20</td>
</tr>
<tr>
<td>2 = agree</td>
<td>2,053</td>
<td>1.34</td>
<td>0.08</td>
<td>0.85</td>
</tr>
<tr>
<td>3 = strongly agree</td>
<td>1,533</td>
<td>2.73</td>
<td>2.36</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Note. MnSq = mean square; NA = not applicable.
In addition to item fit, we evaluated the dimensionality of the revised MHRM using PCA. The revised MHRM had 46.9% of variance explained by the measure (i.e., 23.0 out of 49.0 units) and 10.4% of residual variance explained by the first principal component of residuals (i.e., 2.7 out of 26.0 residual units). No significant difference was found between the revised MHRM and the simulated data. The results indicated unidimensionality.

Targeting and Item Hierarchy

After adaptation, the revised MHRM was examined continuously regarding targeting, item hierarchy, and reliability. Figure 1 is the item–person map of the revised MHRM. On the right side, items reflecting a higher recovery level are placed in higher positions and items reflecting a lower recovery level are placed in lower positions. The mean of item measures is represented by $M$ on the right side of the line, and its default value is 0. Similarly, on the left side, people with a higher recovery level are placed in the higher positions and people with a lower recovery level are placed in the lower positions. The mean of person measures is represented by $M$ on the left side of the line.

A ceiling effect was shown on the map. Namely, many people were located in the upper part of the map, and few items were in the corresponding level. The ability of this sample was higher than that reflected in the items. Only one item (30) had the measure close to the mean of person measures. The mean of item measures was about 1 standard deviation lower than the mean of person measures, which indicates that these items were easy for this sample.

Moreover, we investigated the construct validity by comparing the item hierarchy of the revised scale with the recovery process model. Because people have different backgrounds and experiences, it may be awkward to define the hierarchy of item measures, which may not have been the purpose of the recovery model and measurement. However, it seems to be a trend that the later phase represents a higher recovery level than the first two phases. It was expected that items affiliated with the later phase would show higher measures. The results showed that although some items affiliated with Phase 3 worked as expected, 5 of 12 Phase 3 items had measures lower than the mean of item measures (i.e., 0): Items 17, 22, 23, 24, and 27.

![Figure 1. The item–person map of the revised Mental Health Recovery Measure.](image)

**Note.** Each $\# = 2$ participants; Each $\# = 1$ participant; $M$ = mean; $S$ = 1 standard deviation; and $T = 2$ standard deviations.
Reliability

The revised MHRM showed high reliability: Cronbach’s $\alpha = .95$, person reliability = .92, person separation = 3.38, item reliability = .95, and item separation = 4.22. Using the formula for calculating strata, people could be separated into about five groups, and items could be separated into about six groups.

Discussion

Rasch analyses not only allowed us to investigate the psychometrics of the MHRM but also provided sufficient information to assist us in improving the instrument. In our study, Rasch methods revealed some problems with the MHRM. After modification, the 26-item MHRM with a 4-point Likert scale shows appropriate validity and high reliability. It is also unidimensional, indicating that it measures one concept—that is, recovery.

The 5-point Likert scale used in the original MHRM showed some inappropriate performances. The improper rating scale functioning was related to the middle category, $2 = \text{not sure}$. It was more favorable to use an even number of rating categories—that is, a 4-point scale. The problem with an odd number of rating categories (e.g., a 5-point scale) is that the middle category tends to be labeled with an indifferent phrase, such as neutral or not sure (Wolfe & Smith, 2007). Participants may choose the middle category for different reasons, such as reading difficulty or inapplicability, yet they are scored the same. The less-defined category presents less clear and more misleading information. After combining the 1 and 2 responses, we found proper rating scale functioning. For future use of the MHRM, the 4-point Likert scale is recommended.

The item–person map showed a much higher mean of person measures than item measures and a ceiling effect for the revised MHRM. The recovery concept has been a main focus of the collaborating mental health agency; recovery was not a new concept for these participants, and they had learned about recovery in various activities and groups. These findings imply that the revised MHRM needs more high-recovery-level items to better measure the recovery of people in higher stages of recovery, such as those in this sample.

In addition, by investigating item hierarchy, we found that five Phase 3 items had measures lower than the mean of item measures. One potential reason was that these items did not adequately represent the later phase. For example, Item 22—“My quality of life will get better in the future”—did not describe the advanced recovery condition. Instead, “I have good quality of life” may better represent the later recovery phase. Items 23—“Every day that I get up, I do something productive,” 17—“I feel good about myself,” and 24—“I am making progress towards my goals” had a similar problem. They were descriptions of general, rather than advanced, recovery, which explained why their scores were lower than other items affiliated with the later phase. As for Item 27—“I advocate for the rights of myself and others with mental health problems”—because many participants had received advocacy services, this item might have been easier to respond to in the affirmative. These problematic Phase 3 items represent important recovery concepts, however, and should be revised to better measure people in higher recovery phases and to improve the targeting of the MHRM.

Implications for Occupational Therapy Practice

Although Hancock et al. (2011) suggested that occupational therapists are well suited to advance the Recovery Assessment Scale (Giffort, Schmook, Woddy, Vollendorf, & Gervain, 1995), a popular recovery measure that has incomplete recovery content, the 26-item MHRM with a 4-point Likert scale is an improvement over previous measures of recovery. Its advantages are as follows:

- The revised MHRM is short and easy to use.
- Because its items were developed on the basis of a comprehensive recovery model, the revised MHRM has better capacity to evaluate people’s recovery status than other existing recovery measurements and recovery-related measurements (e.g., hope or empowerment scales).
- The revised MHRM is useful to evaluate people in the initial and middle recovery stages, who also need relatively more support and interventions. Although some further revisions can improve the scale, it can be a reliable and valid recovery assessment for occupational therapists and in recovery-related studies.

Implications for Occupational Therapy Research

Future research can improve on the scale by adding more high-recovery-level items and revising the wording of certain items. Items with misfit and certain problematic Phase 3 items are good candidates. It would be helpful to conduct focus groups and interviews with people in recovery, whose input can significantly advance item concept and wording. After retesting the psychometrics, the revised scale can be expected to measure people with different levels of recovery, including people with high recovery status. Moreover, further studies can collect data from various mental health agencies and attempt to reach people receiving limited services. People with low literacy may be included by using face-to-face interviews. A bigger
and more diverse sample can provide more accurate information about item hierarchy, and factor analysis can be conducted to investigate the existence of phases. Future research can also explore thresholds between phases, which can indicate clinically relevant differences between people affiliated with each phase. Moreover, a differential item functioning analysis can be conducted to investigate whether people with certain conditions answer the scale differently (Scott et al., 2009).

**Limitations**

Because of limited staffing and financial support, we were able to work with only one recovery-oriented community mental health agency in Chicago. Participants’ demographics and responses to the survey indicate that this sample had relatively stable conditions and high recovery status. This sample was relatively homogeneous, so our findings may not be generalizable to people living in other geographic areas, people with unstable conditions, or people receiving limited or no services. Moreover, because participants were required to fill out the survey independently, this study did not capture the experiences of people with limited literacy.

**Acknowledgments**

The study was supported by the Provost’s Award of the University of Illinois at Chicago. We thank Susan Pickett, Glenn T. Fujiura, and Sarah Parker Harris for their assistance in conceptualizing this research. We also thank the agency members who participated in the study and staff who provided assistance in recruitment.

**References**


