A Peer Support Scale for Adults Treated for Psychoactive Substance–Use Disorder: A Rasch Analysis

Kathryn D. Mazurek and James R. Ciesla
College of Health and Human Sciences, Northern Illinois University, DeKalb, IL, USA.

ABSTRACT: The literature suggests that peer support is protective of relapse for adults treated for substance–use disorder. However, to our knowledge there is no standard measure of peer support. Therefore, the purpose of this research was to use Rasch analysis to assess a 13-item peer support scale used in a group of adults treated for primary psychoactive substance–abuse disorder. The participants (n = 408) are adults who were discharged from an inpatient substance–abuse treatment program from five successive years, 2004–2009. Overall, it is acceptable to surmise that items 1–12 are part of the same dimension for the 13-item scale. Given the prominence of therapeutic communities as a mode of primary treatment and the importance of peer support, it is important to both the academic and treatment communities to have a standard way to measure peer support. The scale presented here can be useful for this purpose.

HIGHLIGHTS:
• We assess a peer support scale for those treated for substance abuse.
• We examine characteristics of the peer support scale.
• We provide an option to measure peer support for those treated for substance abuse.

KEYWORDS: psychoactive substance–use disorder, substance–abuse treatment outcomes, peer/social support, Rasch analysis, item response theory

Introduction
Peer support and social support are prominent aspects of treatment for people with psychoactive substance–use disorder (PSUD).1 For example, therapeutic communities have long been a standard residential program approach for people with primary PSUDs.2 They are characterized by community as method, whereby the community is composed of supportive peers and have an identity, rules, and clear expectations for members’ behavior.3 Peer influence is used to help individuals learn to assimilate social norms and develop more effective social skills.4 Therapeutic communities utilize the strong presence of 12-step programs, such as Alcoholics Anonymous groups, as well as peer support and confrontation, and community governance.5 Indeed, the applied treatment literature is replete with evidence that peer support is an important part of treatment as well as relapse prevention in a number of treatment modalities.6

Peer support as a construct is placed within, and is part of, the broader construct of social support.7 Social support is the perception that an individual is situated in a supportive network of relationships.8 It is a multifaceted construct which includes the perception that one is cared for by others as well as the realistic expectation that one will receive tangible assistance from other people when needed.8 Although social support is conceptualized and operationalized in a number of ways, there is strong evidence that health maintenance and recovery from illness can be influenced significantly by a person’s access to supportive others.9 Social support is linked to self-efficacy and outcome expectancy10 both of which may have a positive impact on recovery.

Social support and peer support have been separated into many categories. Three kinds of peer support have been described by Schaefer et al: emotional, tangible, and informational support.11 Schaefer et al differentiate between different aspects of peer support and perceived peer support wherein a social network is defined by the amount or number of social relationships one has. Peer support is defined as a person’s impression of the benefits of those social relationships.11 More recently, Salzer and Shear delineated four kinds of peer support: emotional, instrumental, informational, and affiliational support.12 In spite of its intuitive appeal and its prominence in treatment programs, there is little evidence that deliberate and systematic research has been undertaken to understand the theoretical underpinnings of peer support and its role in treatment and recovery from PSUD. It appears that peer support variables are a better indication of health and well-being than numbers of social networks.4 However, given the current body of substance–abuse treatment literature, there are no known norms for analyzing peer support including measuring and operationalizing the construct. Relationships between
peer support and theoretical models in broad terms are not known. Therefore, the goal of this research is to use item response theory (IRT) to evaluate a scale designed to measure the peer support for adults treated for SUD and to suggest a possible tie-in of the construct of peer support to the literature on the topic.

Materials and Methods

Purpose. The purpose of this research is to use Rasch analysis to assess a 13-item peer support scale used in a group of adults treated for primary PSUD. The scale is designed to measure the relationship between adults treated for SUDs and their peers. One benefit of providing a measure of peer support to treatment communities is that it will be available for practitioners and researchers to design and evaluate treatment programs.

Participants. Study participants were adults who were treated and discharged from a primary in-patient substance-abuse treatment program located in the Midwest United States. Survey data were collected by independent university-based researchers and included participants who were discharged from treatment from the years 2004–2009. The researchers obtained a list of all the adults who completed treatment successively during the previous year and who provided consent to be part of this posttreatment outcomes study. The sampling frame included only successful treatment discharges in the previous year. The criteria for treatment success comport with the ASAM Level I-A (primary inpatient treatment) short-term inpatient care.

Data analysis. IRT has many applications one of which is the Rasch model. The Rasch model can provide information for understanding if the total score of an instrument is accurate enough to characterize an individual test taker. In addition, Rasch analysis can be a dynamic form of analysis that provides evidence for why certain responses to an item may be invalid. The data were analyzed using Winsteps software with a one-parameter IRT Rasch model. The Rasch model used for this analysis is the polytomous Partial Credit Scale form that uses the equation.

\[
\log \left( \frac{P_{nj}}{P_{n(j-1)}} \right) = B_n - D_i - F_j
\]

where,

- \( P_{nj} \) is the probability that person \( n \) encountering item \( i \) is observed in category \( j \),
- \( B_n \) is the ability or rater-severity measure of person \( n \),
- \( D_i \) is the difficulty-to-endorse measure of item \( i \), and
- \( F_j \) is the calibration measure of category \( j \) relative to category \( P_{n(j-1)} = B_n D_i F_j \) (Eq. 1).

Rasch models are successfully employed in well-defined groups of individuals who are responding to assessment items of a latent trait or characteristic. In these instances, the items are scored by integers in ordered categories with an increase in the level of magnitude of the latent trait or characteristic. The beginning point for estimating the latent trait in Rasch models is the sum of the items.

Results

The descriptive statistics are provided in Table 1. There were 408 participants with a mean age of 36.3 (standard deviation 10.88). The sample was mainly male (65.0%) with 35.0% females. The majority of sample was White (83.1%) while 14.2% were African-American and 1.6% were Latino. Alcohol abuse was reported by 59.1% of the sample. Cocaine was reported by 21.0% of the sample, with opiates and marijuana less frequently reported with 9.4% and 8.9%, respectively.

Table 2 shows the peer support scale, response categories, and response frequencies. Several items were reverse coded in order to identify participants who may not have carefully read and answered the questions. All missing values were excluded from the analysis. Rasch analysis skips responses with missing values and eliminates them from the analysis.

Person and item reliability and separation examination. Table 3 shows reliability testing with separation and reliability coefficients. The person reliability is 0.50 and the Cronbach’s alpha person raw score reliability is 0.92. The person reliability score is low while the Cronbach’s alpha person raw score reliability...
raw score is high. The probability of a correct response by an individual to items that are scaled by difficulty is person reliability. The low person reliability score could indicate the need to lengthen the test.

There were 408 measured extreme and nonextreme people in the model. The 37 extreme people were removed and the reliability changed very little (0.50 and 0.51, respectively), ruling out ceiling and floor effects. The item reliability of 0.98 is high, which demonstrates model reliability. The real separation is 6.69, indicating a fair degree of item discrimination; strongly indicating items are placed reliably on the Rasch ruler of about six levels of importance. The real separation statistic shows that the level of the instrument response choices have equal discernment (ie, distance between 1 and 2 is similar to the distance between 3 and 4, etc).

**Item difficulty statistics.** Table 4 shows all the questions in the 13-item scale. The questions were scored with successive ordinal integers 1 through 5. Table 4 gives item difficulty statistics in misfit order. The Item Number column shows the arrangement of the question in the data. The column headed Total Score is the sum of the scored responses to the corresponding item by survey participants and the column Count gives the completeness for each item that ranges from 387 to 355.

### Table 1. Demographic characteristics.

| Characteristic       | %    | Number |
|----------------------|------|--------|
| Gender               |      |        |
| Male                 | 65.0 | 265    |
| Female               | 35.0 | 143    |
| Race                 |      |        |
| African-American     | 14.2 | 58     |
| Caucasian            | 83.1 | 339    |
| Hispanic             | 1.6  | 7      |
| Asian/PI             | 0.5  | 2      |
| Other                | 0.5  | 2      |
| Drug                 |      |        |
| Alcohol              | 59.1 | 219    |
| Marijuana            | 8.9  | 33     |
| Cocaine              | 21.0 | 78     |
| Opiates              | 9.4  | 35     |
| Other drug           | 1.6  | 6      |

Notes: Valid n = 408. Mean age = 36.3, standard deviation 10.88.

### Table 2. Response frequency by category of the 13-item social support scale.

| ITEM | CATEGORIES (%) | NONE | A FEW | SOME | MANY | ALMOST ALL | NA |
|------|----------------|------|-------|------|------|------------|----|
| 1    | How many of your friends support your recovery? | 4.0  | 8.9   | 22.1 | 16.4 | 44.7       | 4.0|
| 2    | How many of your friends work regularly? | 2.5  | 1.3   | 6.1  | 11.9 | 71.1       | 7.1|
| 3    | How many of your friends seem optimistic? | 0.8  | 2.0   | 8.8  | 21.7 | 59.8       | 6.8|
| 4    | How many of your friends get into arguments/fights? | 2.8  | 2.3   | 15.9 | 33.8 | 37.6       | 7.6|
| 5    | How many of your friends spend time with their families? | 1.3  | 3.3   | 13.2 | 22.6 | 51.0       | 8.6|
| 6    | How many of your friends like being with their families? | 2.0  | 2.3   | 10.4 | 21.6 | 53.8       | 9.9|
| 7    | How many of your friends drink too much alcohol? | 47.1 | 19.7  | 16.5 | 4.6  | 4.1        | 8.1|
| 8    | How many of your friends use drugs? | 71.2 | 7.8   | 7.8  | 1.5  | 2.3        | 9.3|
| 9    | How many of your friends trade, sell, deal drugs? | 82.3 | 3.5   | 2.8  | 0.8  | 0.8        | 9.8|
| 10   | How many of your friends break the law? | 76.6 | 8.8   | 3.5  | 0.3  | 1.0        | 9.8|
| 11   | How many of your friends hang out with gangs? | 87.6 | 2.0   | 1.5  | 0.8  | 0.3        | 7.8|
| 12   | How many of your friends go to jail or prison? | 80.0 | 6.8   | 3.5  | 0.5  | 0.8        | 8.4|
| 13   | How many of your friends go to substance abuse treatment? | 60.4 | 11.9  | 10.7 | 3.3  | 5.6        | 8.1|
### Table 3. Person and item reliability.

| RAW SCORE | MEASURE | MODEL ERROR | INFIT MNSQ | Z-STD | OUTFIT MNSQ | Z-STD | ADJ. RMSE | S.D. | SEP. | REL. |
|-----------|---------|-------------|------------|-------|------------|-------|----------|------|------|------|
| **408 Measured Non-Extreme Persons (S.E. of person mean = 0.04)** |
| Mean | 54.9 | 12.4 | 1.59 | 0.45 | 1.03 | 0.1 | 0.89 | 0.0 |
| S.D. | 11.6 | 2.2 | 0.76 | 0.19 | 0.63 | 1.1 | 0.68 | 1.0 |
| Max. | 64.0 | 13.0 | 3.15 | 0.95 | 4.01 | 4.6 | 4.86 | 5.1 |
| Min. | 2.0 | 1.0 | -1.03 | 0.25 | 0.00 | -3.0 | 0.00 | -2.4 |
| Real | | | | 0.53 | | 0.54 | 1.01 | 0.50 |
| Model | | | | 0.49 | | 0.58 | 1.17 | 0.58 |
| **408 Measured Extreme and non-mean persons** | (S.E. of person mean = 0.05) |
| Mean | 54.1 | 12.1 | 1.73 | 0.58 | - | - | - | - |
| S.D. | 14.3 | 2.8 | 1.05 | 0.42 | - | - | - | - |
| Max. | 65.0 | 13.0 | 4.27 | 1.79 | - | - | - | - |
| Min. | 1.0 | 1.0 | -1.03 | 0.25 | - | - | - | - |
| Real | | | | 0.74 | | 0.75 | 1.02 | 0.51 |
| Model | | | | 0.71 | | 0.78 | 1.09 | 0.54 |
| **13 Measured Non-Extreme Items** | (S.E. of person mean = 0.20) |
| Mean | 1,622.8 | 363.7 | 0.00 | 0.08 | 1.12 | 0.7 | 0.92 | -0.5 |
| S.D. | 97.3 | 7.8 | 0.68 | 0.03 | 0.35 | 3.0 | 0.45 | 3.0 |
| Max. | 1,787.0 | 387.0 | 0.80 | 0.17 | 2.15 | 9.8 | 2.33 | 8.7 |
| Min. | 1,489.0 | 355.0 | -1.42 | 0.05 | 0.8 | -2.4 | 0.54 | -2.8 |
| Real | | | | 0.10 | | 0.67 | 6.69 | 0.98 |
| Model | | | | 0.09 | | 0.67 | 7.44 | 0.98 |

Notes: *Person raw score-to-measure correlation = 0.65; Cronbach alpha person raw score reliability = 0.92. Item raw score-to-measure correlation = −0.96; Log-likelihood Chi-square: 7,136.17 with 4,011 d.f., P = 0000.

The next column displays the statistic for the measure for each item and this is the Rasch estimate of item difficulty or model parameters expressed as logits. Item difficulty shows how items are organized on the scale by how possible it is for the items to be chosen.[13]

**Unidimensionality and fit statistics.** Table 4 also provides the infit and outfit statistics. The infit and outfit statistics are representative of how well the data conform to the IRT model.[17] The outfit statistics are based on "conventional sum of squared standardized residuals and infit statistics are based on information-weighted sum squared standardized residuals".[17] The outfit and infit statistics provide information used to identify "(1) items that are not part of the same dimension; (2) the item is subject to misunderstanding; and (3) the likelihood that a response is guesswork or a person possesses special familiarity or expertise".[17] The mean-square (MNSQ) and the z-score standardized t-tests comprise statistics for both the infit and outfit measures. With the exception of question 13 (see Table 4), the mean-square statistics of the infit and outfit values were between 0.5 and 1.5. Since they indicate item randomness and therefore unidimensionality, it is clear that item 13 falls outside of the prediction of the IRT model. In addition, Table 4 gives the model average infit MNSQ and outfit MNSQ, which are 1.12 and 0.92, respectively, and have relatively small and similar standard deviations (0.35 and 0.45). These numbers indicate a near-ideal fit. Overall, it is acceptable to surmise that items 1–12 are part of the same dimension and can be considered logically independent of each other.

**Item difficulty statistics.** The last pair of statistics measuring randomness of items in the scale in terms of the entire model is under the heading Exact Match in Table 4. The OBS% is the percentage of data points that are within 0.5 score points of their expected values. The EXP% is the percentage of data points that are anticipated to be within 0.5 score points of their expected values. If OBS% is less than the EXP%, then data are more random than what the model predicts and if OBS% is greater than EXP%, then the data are more predictable than what the model predicts.[16] The OBS% statistics are greater than the EXP% statistics with the exception of the following items 1, 7, and 13. These items show a small degree of more randomness than the model predicts. Given the magnitude of the infit and outfit statistics in the preceding columns, the finding is appropriate. Additionally, deviation from the predicted randomness is slight except for item 13. The model average OBS% is less than the model average EXP% (63.0, 60.7) and the standard deviations of the averages are proportionally
small (19.5 and 17.8, respectively). This indicates that the model comports to the level of predictability inherent in the model.

Table 5 provides the observed average statistic. These statistics are the average of the measures produced in each category by the model and the average measure is anticipated to increase with category value. The following items 1, 3, 4, 5, 6, 7, 8, 9, and 12 are ordered categories because the observed average of items increases with the category value. Items 2 and 13 have one or two categories that are disordered. Ordered items imply that respondents found the response categories to be meaningful, explicit, and understandable. It should be noted that the average ability scores for items 2, 3, and 4 tend to be close.
Table 5. Observed average by category.

| ITEM | OBSERVED AVERAGE (CATEGORIES) | 1 | 2 | 3 | 4 | 5 |
|------|--------------------------------|---|---|---|---|---|
| 1    |                                | -0.10 | 0.82 | 1.36 | 1.71 | 2.27 |
| 2    |                                | 0.67 | *0.09 | 1.01 | 1.34 | 2.04 |
| 3    |                                | -0.02 | 0.41 | 1.03 | 1.46 | 2.14 |
| 4    |                                | 0.39 | 1.15 | 1.31 | 1.62 | 2.37 |
| 5    |                                | 0.46 | 0.59 | 1.14 | 1.42 | 2.29 |
| 6    |                                | 0.70 | 0.73 | 0.98 | 1.41 | 2.26 |
| 7    |                                | 0.54 | 0.91 | 1.32 | 1.67 | 2.26 |
| 8    |                                | 1.01 | 0.82 | 0.93 | 1.20 | 2.08 |
| 9    |                                | -0.57 | 0.15 | 0.59 | 0.89 | 1.96 |
| 10   |                                | -0.34 | *-0.98 | 0.65 | 1.06 | 2.02 |
| 11   |                                | -1.03 | 0.51 | *0.15 | 0.71 | 1.90 |
| 12   |                                | -0.02 | 0.33 | 0.46 | 1.11 | 1.98 |
| 13   |                                | 1.50 | *1.22 | 1.51 | *1.35 | 2.04 |

Note: *Equals disordered category.

Discussion

Given the prominence of therapeutic communities as a mode of primary treatment, it is important to both the academic and treatment communities to have a standard way to measure peer support. The scale presented here can be useful for this purpose.

The psychometric properties of the instrument showed that the person reliability score is low (0.50); however, increasing the length of the test could potentially enhance the person reliability score. Another noteworthy finding is that it is clear from the infit and outfit coefficients that item 13 exhibits a high degree of variation from the model prediction than is commonly accepted. Similarly, the exact match measures indicate a high degree of randomness. Given the questionable face validity of the item to treatment success and failure, it would be appropriate to recode item 13 or potentially remove it all together. Doing so will almost certainly improve the reliability statistics for the extreme and nonextreme measures. This finding caused us to look back in the literature and it is unclear if having friends that go to substance-abuse treatment is either protective or not for relapse. Therefore, the results show that it is hard to tell how to code item 13 and it may need to be coded differently.

Another improvement that could be made to the scale would be to collapse categories 2 and 3 or to collapse categories 3 and 4. This would eliminate three of the four disordered categories and potentially enhance the usability of the scale.

Additionally, participants had some difficulty differentiating between the choices rarely and sometimes. In the future, the scale should have clearer definitions for response categories and should include specific periods for the items (ie, the number of times in a month or within a certain period of time). By giving better anchors to the descriptions of the response categories, the observed average statistic orderings could be enhanced.

Conclusion

Despite these improvements, our results give treatment providers and researchers an opportunity to use a psychometrically sound instrument to determine treatment success in terms of peer support. Treatment providers may be better able to create or enhance treatment programs and they may be better able to assess posttreatment abstinence based on peer support. Measures of peer support will also enable treatment provided to better design treatment approaches because providers will be better able to screen, track, and measure outcomes of care. Some advantages to this instrument is that it is short, easy to administer, and can easily be used as a screening tool. The questions are based on the actual experiences of the patients and the results of the scale can be useful to treatment providers by creating a feedback loop. Because we know that the scale is psychometrically valid, it makes for a good research instrument as well. For example, from a research perspective, the case can be made for cardinality. The data gathered can be used in an array of quantitative research.

Author Contributions

Principle investigator of the study: JC. Designed the study and helped draft the manuscript: JC. Helped draft the manuscript: KM. Both authors approved of the final manuscript.

REFERENCES

1. Center for Substance Abuse Treatment. *What are Peer Recovery Support Services?* Rockville, MD: Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services; 2009. [HHS Publication No. (SMA) 09–4454].
2. De Leon G. Residential therapeutic communities in the mainstream: diversity and issues. *J Psychoactive Drugs,* 1995;27:1–15.
3. Brunette MF, Mueser KT, Drake RT. A review of research on residential programs for people with severe mental illness and co-occurring substance use disorders. *Drug Alcohol Rev,* 2004;23:471–81.
4. Smith LA, Gates S, Foxcroft D. Therapeutic Communities for Substance Related Disorder. The Cochrane Collaboration. New York, NY: John Wiley & Sons; 2008.
5. Polkin D. Rethinking confrontation in alcohol and drug treatment: consideration of the clinical context. *Subst Use Misuse,* 2003;38(165–1):84.
6. White WL. Peer-Based Addiction Recovery Support: History, Theory, Practice, and Scientific Evaluation. Chicago IL: Great Lakes Addiction Technology Transfer Center; 2009.
7. Lakey B. Social support: basic research and new strategies for intervention. In: Maddux JE, Taingney JP, eds. *Social Psychological Foundations of Clinical Psychology.* New York: Guildford; 2010:177–94.
8. Sarason IG, Sarason BR, Shearin EN, Pierce GR. A brief measure of social support: practical and theoretical implications. *J Soc Pers Relat,* 1997;4:497–510.
9. Uchino B. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J Behav Med,* 2006;29:377–87.
10. Tardy C. Social support measurement. *Am J Community Psychol,* 1985;13:187–202.
11. Schaefer C, Coyne J, Lazarus R. The health related functions of social support. *J Behav Med,* 1981;4:381–406.
12. Saltzer MS, Liptrap Shear S. Identifying consumer-provider benefits in evaluating outcomes of care. *J Soc Pers Relat,* 2006;29:377–87.
13. Linacre JM. *Winsteps Rasch measurement* (Version 3.64.2). Available at: http://www.winsteps.com. Accessed March 1, 2014.
14. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* Mahwah, NJ: LEA Publishers; 2001.
15. Alagumalai S, Curtis DD, Hungi N. *Applied Rasch Measurement: A Book of Exemplars.* New York: Springer-Kluwer; 2005.
16. Wright BD, Masters GN. *Rating Scale Analysis.* Chicago, IL: MESA Press; 1982.
17. Olenen P. An introduction to using multidimensional item response theory to assess latent factor structures. *J Soc Work Res,* 2010;1:66–82.