Rasch Analysis, Dimensionality, and Scoring of the Neuropsychiatric Inventory (NPI) Irritability and Aggression Subscales in Individuals with Traumatic Brain Injury

James F. Malec, PhD, ABPP-Cn, Rp (Corresponding Author)
Senior Research Professor Emeritus,
Department of Physical Medicine and Rehabilitation, Indiana University School of Medicine
Emeritus Professor of Psychology, Mayo Clinic
11050 Brentwood Ave.
Zionsville, IN 46077
Tel: 507-202-5103
jfmalec@iupui.edu

Timothy E. Stump, MA
Biostatistician III
Department of Biostatistics and Department of Public Health
Indiana University School of Medicine
410 W. Tenth St., Suite 3000
Tel: 317-274-2661
Fax: 317-274-2678
tstump@iu.edu

Patrick O. Monahan, PhD
Associate Professor
Department of Biostatistics and Department of Public Health
Indiana University School of Medicine
410 W. Tenth St., Suite 3000
Tel: 317-274-2661
Fax: 317-274-2678
pmonahan@iu.edu

Jacob Kean, PhD
Research Speech-Language Pathologist
Salt Lake City VA Health Care System
Associate Professor
Health System Innovation and Research
Departments of Population Health Sciences and Communication Sciences and Disorders
University of Utah
295 Chipeta Way
Salt Lake City, UT 84108
Tel: 801-213-3751
Fax: 801-581-3623
jacob.kean@gmail.com

Dawn Neumann, PhD
Assistant Professor,

This is the author's manuscript of the article published in final edited form as:

Malec, J. F., Stump, T. E., Monahan, P. O., Kean, J., Neumann, D., & Hammond, F. M. (2017). Rasch Analysis, Dimensionality, and Scoring of the Neuropsychiatric Inventory (NPI) Irritability and Aggression Subscales in Individuals with Traumatic Brain Injury. Archives of Physical Medicine and Rehabilitation. https://doi.org/10.1016/j.apmr.2017.07.020

Department of Physical Medicine and Rehabilitation, Indiana University School of Medicine
Scientist, Rehabilitation Hospital of Indiana
4141 Shore Drive
Indianapolis, IN 46254
Tel: 317-329-2000
Fax: 317-329-2600
dmneuman@iupui.edu

Flora M. Hammond, MD
Professor and Chair,
Department of Physical Medicine and Rehabilitation, Indiana University School of Medicine
Chief Medical Officer, Rehabilitation Hospital of Indiana
Goodman Hall 4700
355 W. 16th St.
Indianapolis, IN 46202
Tel: 317-963-7077
Fax: 317-963-7075
Flora.hammond@rhin.com

Dr. Hammond has a financial relationship with Avanir Steering Committee outside the submitted work. The other authors have nothing to disclose.

The contents of this article were developed under grants from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant numbers H133A20016, H133A080035, and 90DP0001-01-01). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this article do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

Running Head: Rasch NPI Irritability Aggression Subscales

Rasch Analysis, Dimensionality, and Scoring of the Neuropsychiatric Inventory (NPI) Irritability and Aggression Subscales in Individuals with Traumatic Brain Injury

Abstract

Objective: To develop, for versions completed by individuals with traumatic brain injury (TBI) and an observer, a more precise metric for the Neuropsychiatric Inventory (NPI) Irritability and Aggression scales using all behavioral item ratings for use with individuals with TBI and address the dimensionality of the represented behavioral domains. **Design:** Rasch and confirmatory factor analyses of retrospective baseline NPI data from three treatment studies. **Setting:** Postacute rehabilitation clinic. **Participants:** 287 cases with observer ratings; 238 cases with self-ratings by participants with complicated mild, moderate or severe TBI at least 6 months post-injury. **Main Outcome Measure:** Frequency and severity ratings from NPI Irritability/Lability and Agitation/Aggression subscales. **Results:** Confirmatory factor analyses of both observer and participant ratings showed good fit for either a one-factor or two-factor solution. Consistent with this, the Rasch model also fit the data well with aggression items indicating the more severe end of the construct and irritability items populating the milder end. **Conclusions:** Irritability and aggression appear to represent different levels of severity of a single construct. The derived Rasch metric offers a measure of this construct based on responses to all specific items that is appropriate for parametric statistical analysis and may be useful in research and clinical assessments of individuals with TBI.

Abbreviations

AIMS	Amantadine Irritability Multi-site Study
F+S	Frequency plus severity
FXS	Frequency times severity
NPI	Neuropsychiatric Inventory

Rasch NPI Irritability Aggression Subscales

PROMIS Patient Reported Outcomes Measurement Information System

SASNOS St. Andrews-Swansea Neurobehavioral Outcome Scale

STAXI State-Trait Anger Expression Inventory

TBI Traumatic brain injury

For many survivors of traumatic brain injury (TBI) and their families, the pervasive aftermath of emotional and behavioral impairments are the most troublesome and challenging consequences.[1-4] Irritability and aggression after TBI can be particularly concerning; these deficits have been associated with a variety of negative outcomes in home life, family and caregiver burden, relationships, social interactions, work, and general community integration.[4-11] Studies indicate the incidence of chronic (≥6 months) post-TBI irritability ranges from 15% and 74%[12-17] and aggression from, 12% to 41%.[18-22] Beyond the heterogeneity of the samples, the variety of different tests used to evaluate irritability and aggression across studies likely contributes to the marked variation in prevalence estimates.

Despite the number of measures available, there are no well-accepted operational definitions[23] or assessment tools[24] for irritability and aggression after TBI, which complicates the evaluation of these behaviors. . This has been a long-standing and commonly acknowledged problem, with little progress made in the last several decades. In 1992, Prigatano remarked, "irritability and angry outbursts are poorly understood. There is a clear need for a classification system and for behavioral based definitions and measurements to enhance research in this area."[23, p. 363] Primarily because of a continued reliance on theory without empirical support, we are no closer to a consensus on universal definitions of irritability and aggression than we were in the early nineties. At a fundamental level, there is no research to-date that addresses whether these constructs are conceptually distinct after a TBI, or if they represent different degrees of emotional and behavioral dysfunction along a unified continuum.

Developing an empirically-based conceptual understanding of irritability and aggression is important for establishing meaningful operational definitions, a more accurate evaluation and understanding of the problem, and ultimately being able to identify effective treatments. Essential to such empirical study is determination of sound measures that reflect the constructs of interest. Measures commonly used in TBI research include the Anger scale in the Traumatic Brain Injury Quality of Life (TBI-QOL) suite of measures,[25] the State-Trait Anger Expression Inventory (STAXI),[26] the physical and verbal aggression and anger subscales of the Buss Perry Aggression Questionnaire,[27] the Aggression domain and Irritability subdomain of the St. Andrews-Swansea Neurobehavioral Outcome Scale (SASNOS),[28] the Aggression subscale of the Neurobehavioral Functioning Inventory (NFI)[29], and the Irritability/Lability and Agitation/Aggression subscales of the Neuropsychiatric Inventory (NPI).[30] However, none have captured general consensus as the measure of choice. Furthermore, a measure that can be completed both by those with TBI and an observer would be a value in research and practice. Both individuals with TBI and their close others may have biases, limited awareness, or imperfect memory in assessing irritability and aggression. However, distinct and important information regarding dysfunctional behavior after TBI can be gained from separate reports provided by individuals with TBI and observers and address biases and imperfect perception or recall by assessing the behavior from multiple perspectives. Most of the measures listed above were designed to be completed by the person with TBI or an observer, but not both. The exception is the NFI; however, studies of the NFI have been critical of the psychometric properties of this measure.[24,31]

The NPI is an extended inventory of neuropsychiatric symptoms divided into a number of subscales that indicate specific neuropsychiatric symptom complexes or syndromes. In our prior

research we have used two subscales, administered independently both to participants with TBI and their observers: *Irritability/Lability* (subsequently referred to as the Irritability subscale) and *Agitation/Aggression* (subsequently referred to as the Aggression subscale). This assessment involves asking an observer or the participant to indicate whether the symptom is present, and if so, its frequency, and its severity as well as the level of distress experienced due to the symptom. In standard administration, the respondent is then asked to identify the symptom that is "most problematic" and the frequency score multiplied by the severity score for that item indicates the score for the entire subscale. However, identification of the "most problematic" item can be controversial. Should this the item be the one that the respondent indicates is "most problematic" when asked that specific question? Or should the "most problematic" item be the item with the highest frequency times severity score, i.e., the *worst* item? Mirroring the controversy about the nature of irritability and aggression, it has also been unclear whether the NPI Irritability and NPI Aggression subscales indicate two distinct dimensions or two extremes of the same dimension with symptoms of irritability representing the milder end and symptoms of aggression, the more severe.

Because of these issues, we believed that further psychometric evaluation of this measure within the TBI population would advance empirical study in this area. In our prior research, we have always asked respondents to rate all items for frequency, severity, and distress in addition to identifying which behavior is "most problematic." Distress about a symptom is considered to be a different construct from the ratings of symptom frequency and severity. Nonetheless, frequency and severity ratings for all items may provide useful information to evaluate irritability and aggression in contrast to basing the score for a subscale only on a single item (either *most problematic* or *worst*). A version of the NPI with these characteristics would also

give clinicians a tool for assessing irritability and aggression after TBI from the perspectives of both the individual with TBI and a close other and for assessing change in response to treatment. A more straightforward approach to administration and scoring would reduce burden on both interviewer and respondent and open the possibility of self-administration.

The goals of the psychometric studies reported here were to develop, using the information for all behaviors rated on the NPI Irritability and Aggression subscales, a more precise measure of irritability and aggression with a standard approach to administration and scoring and to address the issue of dimensionality in the behavioral items contained in these two subscales.

Method

Participants

Analyses reported here were conducted on de-identified baseline NPI data from three separate studies conducted in outpatient rehabilitation settings for observer data: (1) a study of the effects of carbamazepine on irritability and aggression,[32] (2) a single site study,[33] and (3) the Amantadine Irritability Multi-site Study (AIMS)[34] of the effects of amantadine on irritability and aggression. Observers were persons who had regular contact with the participant with TBI enrolled in the study. Participant self-ratings were available for two of these studies: the carbamazepine and the multi-site AIMS trials. All participants with TBI included in these trials had a history ranging from complicated mild to severe TBI as indicated by post-resuscitation Glasgow Coma Scale (GCS) score 13 or lower or GCS Motor < 6 off paralytics; loss of consciousness, unresponsiveness or coma attributable to TBI; disorientation attributable to TBI and persisting ≥ 24 hours; post-traumatic amnesia lasting ≥ 24 hours; neuroimaging consistent with TBI; or other evidence of TBI-related focal neurological findings indicating significant

injury to the brain sustained at least 6 months prior to enrollment. 287 unique cases with observer NPI ratings and 238 cases with participant self-ratings were available. Table 1 provides basic demographic and injury-related information about these aggregated samples. Additional information about participants and studies is available in the original reports cited previously. Since all data were de-identified, this research was classified as exempt by the Indiana University IRB.

Procedure

As mentioned in the introduction, the NPI was administered in a nonstandard format in English in all 3 studies which were conducted in the United States. In all studies, both participants with TBI and observers were administered the NPI at baseline prior to the initiation of the clinical trial. They were asked to indicate whether each item on the NPI Irritability and Aggression subscales was present during the preceding month, identify the most problematic item, and rate its severity (mild, moderate, marked), frequency (occasionally, often, frequently, very frequently), and the distress it caused. After rating the most problematic item, the respondent then rated the frequency, severity, and distress of the other items. Severity ratings were coded from 1-3 indicating increasing severity; frequency ratings were coded 1-4 representing increasing frequency. Items that were reported as nonproblematic were coded as zero for both frequency and severity.

Statistical analyses

Analyses were conducted separately for observer and for participant NPI ratings. Rasch and principal components analyses of residuals (PCA) were conducted using Winsteps Version 3.91.2. Desirable item fit was set at 1±.4 although a degree of variance was tolerated when only one of the fit indices or only one of the severity-frequency item pairs for an item failed to meet

this criterion. Confirmatory factor analyses were conducted with Mplus Version 7.4 using the mean and variance-adjusted weighted least squares estimator (WLSMV). Items were treated as categorical indicators. Both frequency and severity items were included simultaneously with a correlated error term for each severity-frequency item pair. Two models were considered: (1) a single factor model including all irritability and aggression items and (2) a 2-factor model separating irritability and aggression items and estimating a correlation between factors. Criteria of good overall CFA model fit included the following: comparative fit index (CFI) > .95,[35] , 1999), root mean square error of approximation (RMSEA) < .06,[35] and weighted root mean square residual (WRMR) < 1.00.[36] The general irritability and aggression items (i.e., Does the patient show any other signs of irritability? Does the patient have any other aggressive or agitated behaviors?) were not included in any analysis since they were nonspecific. Summary demographic statistics were computed with SPSS version 24. Missing item data were rare for observer ratings (0.24%); four observers were missing 2 items and two observers were missing 4 items. By default, Mplus includes cases with partial item-level data in the models. There were no missing data for participant self-ratings; consequently, no attempt was made to impute missing data.

Results

NPI Observer Ratings

Rasch analyses: FrequencyXSeverity (FXS Model)

Rasch analyses were first conducted on frequency and severity ratings separately for the 6 specific items on the Irritability subscale and 7 specific items on the Aggression subscale.

However, the Rasch model did not fit these data well. Subsequently, the frequency X severity (FXS) score was evaluated for fit with the Rasch model. Since it is a product, the FXS score has an accelerating distribution. To develop a more linear distribution, we combined adjacent levels of the original FXS score with the following objectives: (a) minimize disordered response levels, (b) extreme scores remain extreme (i.e., 0→0 and 12→4), (c) the middle level (2) had the highest proportion (~25-40%) and (d) levels 1 and 3 at lower proportions (~5-20%). The conversion below best approached these objectives and resulted in adequate separation between rating levels for each item with optimal person fit for the overall measure.

FXS score:	0	1	2	3	4	6	8	9	12
Converted item score:	0	1	1	1	2	2	3	3	4

The 13 Irritability and Aggression items were submitted to Rasch analysis using the converted item score. Initial analyses indicated that three items were significantly misfitting. When these items were eliminated, Mean Square Infit and Outfit ranged from .74 to 1.28 for the remaining items. One case with abnormal response patterns (i.e., Person Infit or Outfit > 3.0) was then eliminated. This final 10-item model had Person reliability/separation=.84/2.29; Item reliability/separation=.98/8.02 with a Cronbach's alpha=.85. The difference between the means of the measure and population was -.34, indicating better targeting of the more aggressive and irritable respondents.

Rasch Analysis: Frequency+Severity (F+S model)

We recognized that frequency and severity ratings for a specific item were not highly correlated in most cases and consequently might function as separate items in Rasch analysis. In order to improve on Person fit, we subsequently conducted Rasch analyses using both the frequency and the severity scores for each of the 6 items on the Irritability subscale and 7 items

on the Aggression subscale. These 26 items were submitted to Rasch analysis using a partial-credit model because the number of rating levels differed between frequency and severity items. Initial analysis revealed no item markedly misfitting items. However, 19 cases with abnormal response patterns (i.e., Person Infit or Outfit > 3.0) were identified. After eliminating these cases, the final 26-item model had Person reliability/separation=.89/2.88; Item reliability/separation=.99/10.24 with Cronbach's alpha=.90. Infit ranged from .84 to 1.30; Outfit from .83 to 1.50.Outfit for only two frequency items exceeded 1.40; these items were retained. In addition to better item fit statistics than the FXS model, the mean for measure of -.19, indicating improved targeting of the sample. On the Person-Item map (Figure 1), most of the Aggression items populated the more severe end of the spectrum with the Irritability items at the milder end. One item showed minimally disordered response categories. Dimensionality was difficult to interpret. A PCA of residuals found eigenvalues greater than 2 for the first four contrasts; however, these factors each explained only 4-5% of the variance. The factors themselves were not clearly interpretable.

Confirmatory factor analyses

Because the PCA of residuals raised concern regarding dimensionality, we further examined these data using confirmatory factor analysis. Both 1- and 2-factor models fit the data well. The 1-factor model yielded a chi-square of 677.79 (286 df, p<.0001), RMSEA=.069 (.062-.076), CFI=.973, WRMR (weighted root mean square residual)=1.306. With the exception of five items in the .4 range, factor loadings were all in the .5 and .6 range with a low of .42 (behaviors hard to handle--frequency) and high of .66 (slam doors, kick furniture--frequency).

The 2-factor model resulted in a chi-square value of 625.30 (285 df, p<.0001), RMSEA=0.065 (0.058-0.071), CFI=0.977 and WRMR=1.237. The correlation between aggression and irritability factors in the 2-factor model was estimated to be 0.83. Factor loadings on the aggression factor ranged from .44 (behaviors hard to handle--frequency) to .69 (slam doors, kick furniture--frequency). Irritability factor loadings ranged from .44 (impatient, trouble coping with delays--frequency) to .68 (bad temper, flying off the handle--frequency). The chi-square difference test showed that the 2-factor model provided statistically better fit compared to the 1-factor model (chi-square value=23.4, 1 df, p<.0001).

Rasch analyses of Irritability and Aggression subscales

Since the factor analyses suggested that the Irritability and Aggression subscales may be separable factors, we attempted to fit a Rasch model to items contained in each of these subscales. These analyses included both frequency and severity items. Rasch analysis of the Irritability subscale showed acceptable Person reliability/separation (.83/2.24) but inadequate Item reliability/separation (.88/2.71). Examination of the Person-Item Map (Figure 1) showed that Irritability subscale items were tightly clustered and thus provided coverage of only a small portion of the distribution. Rasch analysis revealed only marginally acceptable Person reliability/separation for the Aggression subscale (.79/1.96) but good Item reliability/separation (.99/9.04).

NPI Participant Ratings

NPI data from the 238 cases with participant self-ratings at baseline (before treatment) were used in these analyses. As for the observer ratings, we evaluated the fit of frequency, severity, and frequencyXseverity scores to the Rasch model. None of these models fit as well as the frequency+severity (F+S) model which we describe in more detail below.

Rasch Analysis: Frequency+Severity (F+S model)

Initial Rasch analysis found no markedly misfitting items. However, 6 cases with an abnormal response pattern (i.e., Person Infit or Outfit > 3.0) were eliminated. This final 26-item model had Person reliability/separation=.85/2.37; Item reliability/separation=.98/7.83; Cronbach's alpha=.91. Mean for measure= -.60, suggesting limited coverage of the lower end of the distribution (See also Figure 2). The Person-Item map (Figure 2) showed most of the Aggression items defining the more severe end of the spectrum with the Irritability items at the milder end. Two items showed minimally disordered response categories. Dimensionality was unclear. A PCA of residuals indicated that the eigenvalue for the first five contrasts were greater than 2; however, each of these factors accounted for only between 4.3% and 6% of the variance. As in the observer data, factors were difficult to interpret.

Confirmatory factor analyses

Both 1- and 2-factor models fit the data well. The 1-factor model yielded a chi-square of 416.70 (286 df, p<.0001), RMSEA=.044 (.034-.053), CFI=.991, WRMR (weighted root mean square residual)=.95. With the exception of two items in the .2 range (gets upset--frequency and severity) and one item in the .3 range (hurt or hit others--severity), factor loadings were all in the .4 and .8 range with a low of .40 (hurt or hit others--frequency) and high of .81 (bad temper, flying off the handle--severity). The 2-factor model resulted in a chi-square value of 410.87 (285 df, p<.0001), RMSEA=.043 (.033-.052), CFI=.991 and WRMR=.93. The correlation between aggression and irritability factors in the 2-factor model was.91. Factor loadings on the aggression factor ranged from .24 (gets upset--severity) to .71 (shout or curse angrily--severity). Irritability factor loadings ranged from .50 (impatient, trouble coping with delays--frequency) to .82 (bad temper, "flying off the handle"--severity). The chi-square difference test showed that the 2-

factor model provided slightly statistically better fit compared to the 1-factor model (chi-square value=4.22, 1 df, p=.0415).

Rasch analyses of Irritability and Aggression subscales

Rasch analysis of the frequency and severity items on the Irritability subscale showed acceptable Person Fit/Separation (.84/2.29) but marginal Item Fit/Separation (.93/3.59). Rasch analysis of the Aggression subscale revealed inadequate Person Fit/Separation (.70/1.54) with acceptable Item Fit/Separation (.98/7.25).

Discussion

Taken together, Rasch and factor analysis of data from NPI Irritability and Aggression subscales indicate that these behavioral domains represent a single construct composed of two ordinally-related factors: irritability (e.g. impatience, bad temper) in its milder form and aggression (e.g., slamming or kicking things, hurting others) in its more severe manifestation. The good fit of the data to both one factor and two factor models supports this conclusion since it indicates that behaviors describing both irritability and aggression can be accounted for on a single dimension and that irritability and aggression can also be described as separate factors. While these factors are separable, they have an ordinal relationship, that is, aggression items represent greater symptom severity than irritability items. The Rasch model and associated Person-Item maps illustrate more clearly that the aggression factor tends to represent the more severe form of this behavioral domain and the irritability factor, the milder form. Although the fit indices of the 2-factor model were slightly better compared to the 1-factor model, the sample sizes were large enough for chi-square difference tests to detect small deviations of good fit.

The method of administration used to obtain data in this study was nonstandard, that is, both observers and participants were asked to rate all items on the NPI Irritability and

Aggression scales for both frequency and severity. This method of administration coupled with Rasch analysis and scoring provides a means to integrate ratings for all items included in the NPI Irritability and Aggression subscales. For individuals with TBI, this may provide a more representative assessment of Irritability and Aggression than the standard approach estimating these variables based on a single item identified as most problematic.

From a measurement perspective, the fit of a Rasch model to both observer and participant ratings indicates that these data can be translated into a metric appropriate for use in parametric data analyses. Separate subscales for irritability and aggression were not sufficiently reliable to be acceptable for clinical and research use. Items contributing to each of these subscales cover a relatively small proportion of the distribution; whereas, a metric based on items from both subscales covers the entire distribution relatively well. However, since the aggression items generally are associated with higher scores (see Figures 1 and 2), examination of the score for the overall Rasch NPI Irritability and Aggression Scale reveals whether the behavior of the person rated is characterized primarily by irritability (i.e., scores below the mean) or by both irritability and aggression (scores above the mean). Tables are available as supplemental material to convert raw scores for either observer or participant ratings to a Rasch metric on a 0-100 scale with a mean of approximately 46.

From a theoretical perspective, our results suggest that irritability and aggression, as measured by the NPI, are not different behavioral domains but represent two ends of a continuum. The measurement procedures used in this study are a step toward better operationalization of this construct and have implications for future research and practice. For example, much like the distinction between “major” and “minor” depression, evaluation of irritability/aggression along the continuum described by the Rasch scale may support future

research to determine what types of treatments are most effective for individuals evidencing the milder elements of this problematic behavioral domain and which treatments are most effective for those at the more severe end.

Limitations. This was a retrospective, secondary analysis of convenience data. Although the sample used was relatively large and data was gathered from three different research studies, these data may not be representative of all individuals with TBI in the postacute phase or of individuals with brain injury more generally.

Conclusions. Psychometric analysis of data from the NPI Irritability and Aggression scales indicates that behaviors identified by items in these scales describe a single behavioral domain representing irritability alone in its milder expression and including aggressive behaviors in in its more severe form. These analyses contribute to establishing the validity of this construct. The Rasch metric developed from these analyses may provide a more representative assessment of irritability/aggression since it is based on ratings of the entire array of behaviors described by items in the NPI Irritability and Aggression scales. Such a metric may be useful in practice to assess the severity of disordered behavior in this domain and to monitor response to treatment. In research, the Rasch metric proposed here meets criteria for use in parametric statistical analyses.

References

1. Anderson MI, Parmenter TR, Mok M. The relationship between neurobehavioural problems of severe traumatic brain injury (TBI), family functioning and the psychological well-being of the spouse/caregiver: path model analysis. Brain Inj 2002;16(9):743-757.
2. Gainotti G. Emotional and psychosocial problems after brain injury. Neuropsychol Rehabil 1993;3(3):259-277.

3. Morton MV, Wehman P. Psychosocial and emotional sequelae of individuals with traumatic brain injury: A literature review and recommendations. Brain Inj 1995;9:81-92.
4. Lezak MD, O'Brien KP. Longitudinal study of emotional, social, and physical changes after traumatic brain injury. J Learn Dis 1988;21(8):456.
5. Demark J, Gemeinhardt M. Anger and its management for survivors of acquired brain injury. Brain Inj 2002;16(2):91-108.
6. Felmingham KL, Baguley IJ, Crooks J. A comparison of acute and postdischarge predictors of employment 2 years after traumatic brain injury. Arch Phys Med Rehabil 2001;82(4):435-439.
7. Khan F, Baguley IJ, Cameron ID. 4: Rehabilitation after traumatic brain injury. Med J Australia 2003;178(6):290-297.
8. Lezak MD. Psychological implications of traumatic brain damage for the patient's family. Rehabil Psychol 1987;31(4):241-250.
9. Marsh N, Kersel D, Havill J, Sleigh J. Caregiver burden at 6 months following severe traumatic brain injury. Brain Inj 1998;12(3):225-238.
10. Oddy M, Coughlan T, Tyerman A, Jenkins D. Social-adjustment after closed head-injury - a further follow-up 7 years after injury. J Neurol Neurosurg Psychiat 1985;48(6):564-568.
11. Winkler D, Unsworth C, Sloan S. Factors that lead to successful community integration following severe traumatic brain injury. J Head Trauma Rehabil 2006;21(1):8-21.
12. Kim SH, Manes F, Kosier T, Baruah S, Robinson R. Irritability following traumatic brain injury. J Nerv Ment Dis1999;187(6):327-335.
13. Hammond F, Knotts A, Hirsch M, Norton J, Demakis G, Cook J, Kilmer R. Posttraumatic irritability and related factors. J Head Trauma Rehabil 2006;21(5):412.

348 14. McKinlay WW, Brooks DN, Bond MR, Martinage DP, Marshall MM. The short-term
349 outcome of severe blunt head injury as reported by relatives of the injured persons. J
350 Neurol Neurosurg Psychiat 1981;48:527-533.
351 15. Yang C-C, Hua M-S, Lin W-C, Tsai Y-H, Huang S-J. Irritability following traumatic brain
352 injury: divergent manifestations of annoyance and verbal aggression. Brain Inj
353 2012;26(10):1185-1191.4
354 16. Kilmer RP, Demakis GJ, Hammond FM, Grattan KE, Cook JR, Kornev AA. Use of the
355 neuropsychiatric inventory in traumatic brain injury: A pilot investigation. Rehabil Psychol
356 2006;51(3):232-238.
357 17. Brooks N, Campsie L, Symington C, Beattie A, McKinlay W. The effects of severe head
358 injury on patient and relative within seven years of injury. J Head Trauma Rehabil
359 1987;2(3):1-13.
360 18. Draper K, Ponsford J, Schönberger M. Psychosocial and emotional outcomes 10 years
361 following traumatic brain injury. J Head Trauma Rehabil 2007;22(5):278-287.
362 19. Dyer KFW, Bell R, McCann J, Rauch R. Aggression after traumatic brain injury:
363 Analysing socially desirable responses and the nature of aggressive traits. Brain Inj
364 2006;20(11):1163-1173.
365 20. Johansson SH, Jamora CW, Ruff RM, Pack NM. A biopsychosocial perspective of
366 aggression in the context of traumatic brain injury. Brain Inj 2008;22(13-14):999-1006.
367 21. Kilmer RP, Demakis GJ, Hammond FM, Grattan KE, Cook JR, Kornev AA. Use of the
368 neuropsychiatric inventory in traumatic brain injury: A pilot investigation. Rehabil Psychol
369 2006;51(3):232-238.

22. Neumann, D., Malec, J. F., & Hammond, F. M. (2016). The relations of self-reported aggression to alexithymia, depression, and anxiety after traumatic brain injury. J Head Trauma Rehabil Epub ahead of print: 10.1097/HTR.0000000000000261.

23. Prigatano GP. Personality disturbances associated with traumatic brain injury. J Consult Clin Psychol 1992;60(3):360-368.

24. Cusimano MD, Holmes SA, Sawicki C, Topolovec-Vranic J. Assessing aggression following traumatic brain injury: a systematic review of validated aggression scales. J Head Trauma Rehabil 2014;29(2):172-184.

25. Tulsky DS, Kinsella PA, Victorson D, Carlozzi N, Bushnik T, Sherer M, Choi SW, Heinemann AW, Chiaravalloti N, Sander AM, Englander J, Hanks R, Kolakowsky-Hayner S, Roth E,Gershon R, Rosenthal M, Cella D. TBI-QOL: Development and Calibration of Item Banks to Measure Patient Reported Outcomes Following Traumatic Brain Injury. J Head Trauma Rehabil. 2016;31(1):40–51.

26. Spielberger, C. D. (1999). Manual for the State-trait Anger Expression Inventory-2 (STAXI-2). Odessa, Fl: Psychological Assessment Resources.

27. Buss AH, Perry M. The aggression questionnaire. J Pers Soc Psychol 1992;63(3):452-459.

28. Alderman N, Wood RL, Williams C. The development of the St Andrew's-Swansea Neurobehavioural Outcome Scale: validity and reliability of a new measure of neurobehavioural disability and social handicap. Brain Inj 2011;25:83-100.

29. Kreutzer JS, Seel RT, Marwitz JH. *Neurobehavioral Functioning Inventory: NFI*. San Antonio, TX: Psychological Corporation; 1999.

30. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurol 1994;44(12):2308-2314.

31. Czuba KJ, Kersten P, Kayes NM, Smith GA, Barker-Collo S, Taylor WJ, McPherson KM. Measuring neurobehavioral functioning in people with traumatic brain injury: Rasch analysis of the Neurobehavioral Functioning Inventory. J Head Trauma Rehabil. 2016; 31(4): E59–E68.32.

32. Hammond FM. Carbamazepine for the treatment of chronic post-traumatic brain injury irritability and aggression. https://clinicaltrials.gov/ct2/show/NCT00621751. Accessed 2/24/17.

33. Hammond FM, Bickett AK, Norton JH, Pershad R. Effectiveness of amantadine hydrochloride in the reduction of chronic traumatic brain injury irritability and aggression. J Head Trauma Rehabil. 2014 Sep-Oct;29(5):391-9.

34. Hammond FM, Sherer M, Malec JF, Zafonte RD, Whitney M, Bell K, Dikmen S, Bogner J, Mysiw J, Pershad R, Amantadine Irritability Multisite Study Group. Amantadine effect on perceptions of irritability after traumatic brain injury: Results of the Amantadine Irritability Multisite Study (AIMS). J Neurotrauma 2015;32(16):1230-8.

35. Hu L-T, Bentler PM. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling 1999;6:1-55.

36. Yu C-Y. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Doctoral Dissertation, University of California, Los Angeles. 1999. Retrieved from http://www.statmodel.com/download/Yudissertation.pdf.

Figure Legends

Figure 1. Person-item map for observer ratings

Figure 2. Person-item map for participant self-ratings

Rasch NPI Irritability Aggression Subscales

417

Table 1. Demographic and injury-related summary for combined samples		
	Participant Data Sample	Observer Data Sample
Gender (% Female)	38.2%	41.0%
Race (% White)	84.0%	84.3%
Mean Age (SD)	39.02 (12.71)	38.60 yrs (13.10)
Time Since Injury (SD)	6.70 yrs (8.97)	6.26 yrs (8.23)
Duration of Post-traumatic Amnesia		
<24 h	9.2%	NA
1-6 d	13.0%	
7-13 d	5.9%	
14-20 d	10.9%	
21-29 d	9.7%	
30-59 d	18.9%	
>60 d	30.3%	
Missing	2.1%	
Glasgow Coma Scale score		
3-8	26.5%	NA
9-12	2.9%	
13-15	22.3%	
Chemically-paralyzed, chemically-induced coma, or intubated	44.5%	
Missing	3.8%	

Figure 1. Person-item map for observer ratings

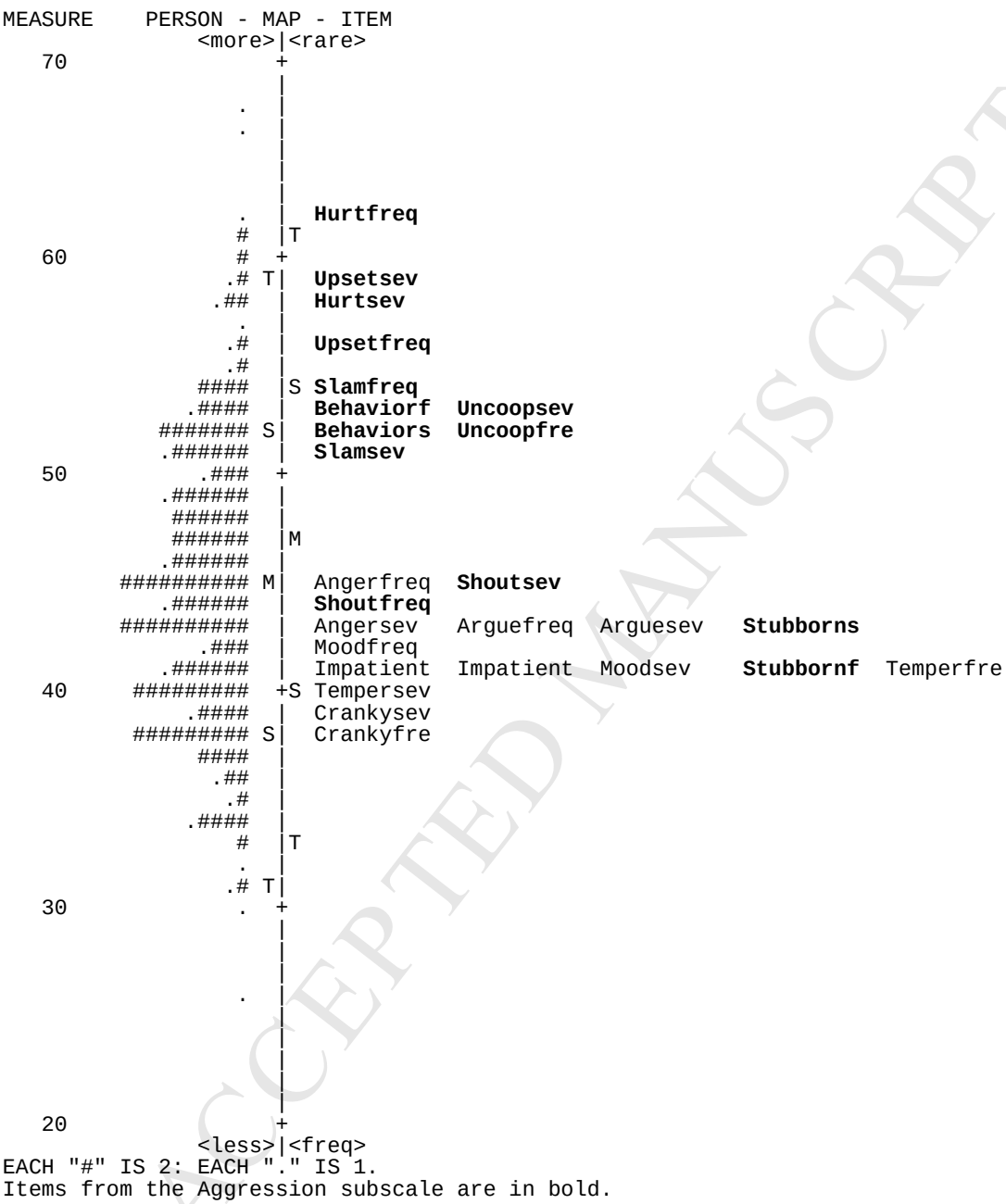

Figure 2. Person-item map for participant self-ratings