Rasch Measurement Analysis of the Mayo-Portland Adaptability Inventory (MPAI-4) in a Community-Based Rehabilitation Sample

Jacob Kean,¹ James F. Malec,¹ Irwin M. Altman,² and Shannon Swick³

Abstract

The precise measurement of patient outcomes depends upon clearly articulated constructs and refined clinical assessment instruments that work equally well for all subgroups within a population. This is a challenging task in those with acquired brain injury (ABI) because of the marked heterogeneity of the disorder and subsequent outcomes. Although essential, the iterative process of instrument refinement is often neglected. This present study was undertaken to examine validity, reliability, dimensionality and item estimate invariance of the Mayo-Portland Adaptability Inventory – 4 (MPAI-4), an outcome measure for persons with ABI. The sampled population included 603 persons with traumatic ABI participating in a home- and community-based rehabilitation program. Results indicated that the MPAI-4 is a valid, reliable measure of outcome following traumatic ABI, which measures a broad but unitary core construct of outcome after ABI. Further, the MPAI-4 is composed of items that are unbiased toward selected subgroups except where differences could be expected [e.g., more chronic traumatic brain injury (TBI) patients are better able to negotiate demands of transportation than more acute TBI patients]. We address the trade-offs between strict unidimensionality and clinical applicability in measuring outcome, and illustrate the advantages and disadvantages of applying single-parameter measurement models to broad constructs.

Key words: outcome measures; rehabilitation; TBI

Introduction

THE MAYO-PORTLAND ADAPTABILITY INVENTORY (MPAI) is a measure of global brain injury outcome that is the product of almost 20 years of research and development. In its current version (MPAI-4), it has become increasingly used to evaluate participants and outcomes in post-hospital brain injury rehabilitation programs throughout the United States and the world. The English language version has been translated into a number of foreign languages. A manual summarizing much of the MPAI-4's development, psychometric qualities, and English and foreign language versions of the inventory itself can be found on the web site sponsored by Center for Outcome Based Measurement in Brain Injury (www.tbims.org/combi/mpai).

The original version was based on the Portland Adaptability Inventory developed by Lezak (Lezak, 1987). Subsequent revisions of the inventory have been conducted using both classical

test theory and Rasch analysis. These analyses have confirmed the construct, concurrent and predictive validity of the measure (Bohac et al., 1997; Malec, 2004; Malec and Degiorgio, 2002; Malec and Lezak, 2008; Malec and Thompson, 1994; Malec et al., 1995; Malec et al., 2000a,b; Malec et al., 2003). Although the psychometric qualities of the measure have been progressively refined through these analyses, each version has appeared to represent a core linear construct describing global outcome after acquired brain injury (ABI). However, this global construct also appears to include subordinate correlated constructs representing physical and cognitive abilities that are often affected by ABI; emotional and interpersonal problems; and involvement in social, community and productive activities. These three subordinate constructs are measured respectively by three subscales: Ability Index, Adjustment Index, and Participation Index in the MPAI-4.

The most recent psychometric analysis of the MPAI was completed on data obtained from 386 individuals with

¹Department of PM&R, Indiana University School of Medicine, Indianapolis, Indiana.

²Gentiva Rehab Without Walls®, Phoenix, Arizona.

³Gentiva Rehab Without Walls®, Marshall, Michigan.

moderate-to-severe traumatic brain injury (TBI) in a range of treatment settings: outpatient, community-based, and residential (Malec et al., 2003). Rasch analyses confirmed satisfactory person reliability (.88) and item reliability (.99); satisfactory internal consistency was also demonstrated by a classical psychometric indicator (Cronbach's $\alpha = .89$) for the entire measure and for each of the three subscales. Factor and cluster analysis of items suggested that subscales developed on a rational clinical basis were generally consonant with empirical relationships among items. The three rational subscales were moderately correlated with each other and appeared to represent progressively more difficult regions of the underlying construct of global outcome after ABI. Items on the Ability Index were most likely to approach normal and difficulties in this area were characteristic of those with most severe global impairment. Items on the Participation Index were generally the least likely to indicate normal functioning and those with difficulties restricted to this area were also typically those with least global impairment. Adjustment items tended to fall in between the two other subscales in their association with the overall level of impairment indicated by the measure. Therefore, previous studies support the broad but unitary nature of the construct of outcome after ABI as well as the relationship and structure of the subscales. These analyses also provided the basis for modifications of the instrument to improve psychometric properties (such as, using a 5-point rather than a 6-point rating scale, or adding or deleting items to obtain optimal item fit to the measure).

The current study was conducted to reassess psychometric qualities and structure of the MPAI-4 in a large and more homogenous group of patients involved in community-based brain injury rehabilitation. The application of Rasch analyses to rehabilitation measures has developed in parallel with the development of the MPAI-4. In recent years, guidelines have been offered for key elements of Rasch analysis and reporting (Smith et al., 2003; Tennant and Conaghan, 2007). More detailed analyses (e.g., differential item functioning, local dependency) that have not been previously completed on the MPAI-4 have become routine. The current study follows these guidelines and includes recommended analyses to provide a systematic reevaluation of specific psychometric qualities of the MPAI-4.

Methods

Subjects

The sample consisted of 603 consecutive admissions with TBI who were evaluated prior to participating in a home- and community-based rehabilitation program provided by Gentiva® Rehab Without Walls. Table 1 shows demographic and injury-related variables for the sample. Injury severity was determined based on available data as follows: "severe" if the patient had a Glasgow Coma Scale (GCS) score < 9 , loss of consciousness (LOC) for > 24 h, or post-traumatic amnesia (PTA) lasting > 1 week; "moderate" if the patient had a GCS score between 9 and 12, LOC between 30 min and 24 h, or PTA that lasted between 24 h and 1 week; or "mild" if the patient had a GCS score > 12 , LOC < 30 min, and PTA lasting < 24 h. If more than one indicator was present and differed in level of severity, the most severe level was assigned. PTA, GCS, and LOC data were available for 212, 216, and 195 patients, respectively. Ten percent of the sample had insufficient data for classification.

TABLE 1. DEMOGRAPHIC CHARACTERISTICS AND INJURY SEVERITY MEASURES

<i>Gender</i>		
Male	436	72%
Female	167	28%
<i>Age</i>		
Mean	40.2	
SD	14.8	
Median	39.0	
Range	14-86	
<i>Chronicity</i>		
Median	81.0	
SD	1418.9	
<i>Severity</i>		
Unknown	58	9.6%
Mild	103	17.1%
Moderate	115	19.1%
Severe	327	54.2%

Measure

The instrument evaluated in this study was the MPAI-4. The MPAI-4 consists of 30 items designed to assess commonly occurring limitations after ABI. It is divided into three subscales: Ability Index, Adjustment Index, and Participation Index. Most items are rated on a 5-point scale ranging from "normal" function to "severe limitations." Some items, particularly on the Participation Index, are anchored to more discrete behaviors and activities. A brief history of the development of the MPAI-4 was provided in the introduction to this article.

Procedures

All study participants were actively involved in a post-acute home- and community-based rehabilitation program. When participants were admitted to the program, clinical evaluations were conducted by rehabilitation professionals who would be treating them, and, following completion of the assessments, the MPAI-4 was completed by the treatment team based on professional consensus. The treatment team rating the MPAI-4 was provided with written materials regarding the process of data collection and instructed to refer to the MPAI-4 Manual for detailed rating instructions for each item.

Statistical analysis

All Rasch analyses were conducted with WINSTEPS version 3.69.1 using a partial credit model, and all other analyses were conducted with SPSS, version 17.0 (SPSS, 2008). Briefly, Rasch analysis creates a linear model of the latent variable—a yardstick of the construct being measured—based on the log odds of a specific response on an item by a person in the measurement sample. Crucial measurement properties derived from the analysis include: 1) the reliability with which the measurement items sort persons into different perfor-

mance strata (i.e., person separation); 2) the validity with which the measurement items correspond to the latent linear variable (i.e., fit statistics); and 3) the correspondence of increases on the rating scale response structure to increases in ability of persons responding at the given levels (i.e., point-biserial correlation and average measure). The sample of 603 independent patients is large enough for definitive item calibration (Linacre, 1994).

Good reliability and validity are critical and suggest fit between the data and linear latent variable modeled in Rasch analysis but can also be insensitive to violations of unidimensionality (i.e., measuring only the target construct). For example, a scale may include items that are biased to subgroups (e.g., males or females) within the sample but exhibit good reliability and validity across the entire sample. Accordingly, differential item functioning (DIF) analysis was conducted to determine if subgroups within the sample exhibited systematic difference in their performance on MPAI-4 items. Next, item residual correlations were examined to ensure that the items were locally independent, indicating that the variance in item responses was accounted for by person ability (i.e., the Rasch dimension) rather than by other dimensions (Marais and Andrich, 2008). Finally, principal components analysis (PCA) of the item residuals was conducted to determine if additional factors in the data suggested by strong item residual correlations were indicative of the subscale structure established in previous psychometric study of the MPAI-4.

Results

Crucial measurement properties

Initial calibration. Analysis of crucial Rasch measurement properties revealed a person reliability score of .90 and a person separation score of 3.03, which exceeded the critical threshold of 2.0 and allows the identification of four performance strata. Item separation in this initial calibration was 14.50, and item reliability approached 1.0. The analysis of item fit in this first calibration indicated three items with infit statistical values that were outside the relatively strict criterion of $1.0 \pm .3$: audition, productive activity, and dizziness. Items with outfit statistic values outside acceptable cutoffs were: audition, productive activity, transportation, dizziness, pain, use of hands, and anger.

Second calibration: elimination of outlying persons. Outfit is sensitive to unexpected observations, therefore the most extreme outlying persons (5%, $n = 30$) were excluded from the analysis, and fit statistics were plotted again. The demographic and injury-related characteristics of these outlying cases were very similar to the entire sample: 70% were male, mean age was 43.4 (14.0), and more than half (53.3%) had severe injuries. The median injury chronicity for excluded cases (265.5) was, however, different than that for the entire group (81.0). This second calibration ($n = 573$), conducted after the removal of these statistically outlying persons, improved person separation (3.09) and reliability (.91) from the original calibration and remedied excessive outfit for all items except audition, productive activity, and dizziness. Excessive infit remained for audition and dizziness.

Third calibration: rescaling select items. Previous psychometric investigation of the MPAI-4 suggested that the 0–4

rating scale structure for audition, pain, transportation, and productivity activity items produced infrequently observed intermediate categories (Malec et al., 2003), therefore, the responses to these items were revalued according to the previous report and the measure was calibrated a third time. This final recalibration resulted in further improvement in person separation (3.15) but no change to person reliability (.91). Despite the revalued response, the productive activity item exhibited disordering of average measures and point biserial correlation of 0.22, which did not exceed the critical threshold ($\geq .30$). Furthermore, the dizziness item exceeded both infit and outfit cutoffs, whereas use of hands, vision, and anger items exceeded acceptable outfit cutoffs. Finally, the audition item demonstrated acceptable fit statistics but failed to demonstrate point biserial correlation (.24) that exceeded the critical cutoff. Table 2 includes the items of the MPAI-4, their location on the linear dimension created by the Rasch analysis (i.e., measure), and associated outfit statistics and point biserial correlations. In the final calibration, the operational range of item targeting accounted for the abilities of 559 of the 573 persons retained for the analysis (97.5%).

Figure 1 shows item outfit plotted against measure. Note that the item numbering in Figure 1 does not correspond with numbering on the MPAI-4. The y-axis marks item difficulty; items in the upper area of the figure are more difficult than those in the lower area of the figure. The hierarchy of item difficulties observed in the current study was similar to that observed in a previous investigation (Malec et al., 2003): items comprising the Ability Index were generally easier than items in the Adjustment Index, whereas items in the Participation Index were generally the most difficult. The size of the circle corresponding to the item indicates the size of the standard error, or the measurement precision of the item, with smaller circles representing more precise items. Figure 1 shows that the standard error (i.e., imprecision) of the audition item (0.09) was significantly greater than the other items ($\chi = 0.04$). The gray areas on the left- and rightmost portions of the x-axis mark areas which are outside the range ($1.0 \pm .3$) of acceptable fit, indicating these items (i.e., dizziness, use of hands, vision, anger) have subtle deviance from the linear latent variable. As illustrated in Figure 1, the concordance of hierarchical ordering with previous investigations and generally acceptable point biserial correlations and fit statistics indicates the items of the MPAI-4 demonstrate fit with the linear model created by Rasch analysis.

Dimensionality

DIF. The recalibrated measure was used to investigate dimensionality through analysis of item bias. Item bias was examined for injury chronicity and patient age and gender. A sample size of 200 per group is sufficient for the detection of moderate uniform DIF in a scale the length of the MPAI-4 (Scott et al., 2009). Age subgroups exceeded 200; however females ($n = 159$) in the gender contrast and time since injury > 6 months ($n = 174$) in the chronicity contrast only approached this threshold. Items were considered to be biased if they exhibited both substantive (i.e., DIF contrast size > 0.5 logits) and statistical ($p < .05$) significance (Zwick et al., 1999). Three items showed moderate-to-large bias: audition, independent living, and transportation. Table 3 shows the contrasts, biased items and size of the bias.

TABLE 2. MPAI-4 ITEM STATISTICS IN MEASURE ORDER

<i>Total score</i>	<i>Measure</i>	<i>Model S.E.</i>	<i>Outfit mnsq</i>	<i>Ptbis</i>	<i>Item</i>
172	1.94	.09	1.17	.22	Audition
599	.98	.04	1.43	.32	Dizziness
533	.95	.04	1.06	.50	Inappropriate Behavior
559	.72	.04	.97	.55	Motor Speech
739	.72	.04	1.30	.38	Sensitivity
766	.67	.04	1.32	.37	Anger
806	.52	.04	1.36	.40	Vision
782	.47	.04	1.02	.57	Information
891	.44	.04	1.24	.39	Depression
827	.40	.04	.85	.63	Visuospatial Abilities
910	.34	.04	.83	.63	Nonverbal Communication
926	.30	.04	.89	.60	Verbal Communication
989	.25	.04	1.30	.40	Anxiety
918	.21	.04	.94	.61	Self-Care
987	.14	.04	1.38	.41	Use of Hands
1083	.04	.04	.79	.66	Initiation
1157	-.05	.04	.92	.58	Family/SO Relationships
1175	-.06	.04	.99	.55	Self-Awareness
1205	-.18	.04	1.09	.49	Mobility
1326	-.35	.04	.73	.67	Problem-Solving
1399	-.49	.04	.80	.61	Social Contact
689	-.59	.07	1.05	.33	Pain
1439	-.63	.05	.81	.61	Memory
1414	-.68	.05	.77	.64	Attention
1586	-.77	.04	.85	.58	Money Management
1708	-.94	.04	.81	.54	Independent Living
1314	-1.01	.04	1.16	.36	Transportation
1782	-1.22	.05	.85	.52	Leisure
1547	-1.49	.06	1.23	.24	Productivity

Total Score, the total of all partial credit scores; Measure, the calibrated value; Model S.E., the error of measurement; Outfit MnSq, the misfit in terms of the mean square outfit; Ptbis, the point biserial correlation; Item, A brief description of item content.

Subscales and item residuals. The multidimensionality reflected in the subscale structure established in previous analysis of the MPAI-4 was examined by calibrating each subscale separately and also through PCA of the entire set of item residuals. The subscales performed adequately as individual measures, a finding concordant with previous investigations, with person separation and reliability values approaching or exceeding acceptable cutoffs of 2.0 and 0.8, respectively. These results are reported in Table 4.

Analysis of item residual correlations indicated 21 pairs of items were correlated at ≥ 0.3 , indicating significant local dependence. The 21 locally dependent items are listed in Table 5.

Principal components analysis tests the possibility that variance unexplained by the latent dimension created in Rasch analysis (first component), and suggested by item residual correlations, could be modeled as another dimension(s) (nth component). Subordinate dimensions in the data suggest that the target construct is best measured using different scales or subscales. Analysis indicated that the first component, the latent linear dimension created by the Rasch model, accounted for 28.0 eigenvalue units (48.3%) of the variance in the data. The second, third, and fourth components constructed by PCA met or exceeded the 2.0 eigenvalue units, i.e., accounted for more than chance variance. These components accounted for 4.0 (6.8%), 3.7 (6.4%), and 2.0 (3.5%) eigenvalue units of variance, respectively.

In order to further evaluate dimensionality, items loading positively and negatively on the components derived from

PCA were used to make separate person estimates and then compared using independent *t*-tests as suggested by Smith (2002). If $>5\%$ of the *t*-tests exceed the range of ± 1.96 , the measure is considered multidimensional. For the second, third, and fourth components constructed by PCA, the percentage of *t*-tests $> \pm 1.96$ were 30.4%, 30.2%, and 12.0%, respectively. The sets of items that loaded on the subordinate dimensions are listed in Table 6, and items in bold represent items of previously defined MPAI-4 subscales.

Discussion

The primary purpose of the current study was the replication and extension of psychometric analyses of the MPAI-4. The results of the study support the item selection and rating scale adjustments made in previous analyses. The MPAI-4 demonstrates excellent coverage of the range of person abilities and activities among individuals participating in an outpatient rehabilitation program. It reliably places and meaningfully stratifies persons along a latent variable, and assesses domains that are important to outcome following ABI. Improvements shown in crucial Rasch measures through the iterative calibration process supported the revaluing of certain responses based on previous investigations. Additionally, the subscales demonstrate adequate core psychometric properties and can be used independently.

The current study revealed subtle differences from previous investigations. Visual inspection of the location of items

FIG. 1. Bubble plot of MPAI-4 items. The radius of the bubble corresponds to the standard error of the item. x-axis = outfit mean square values for MPAI-4 items; y-axis = item difficulty estimates on a logit (log odds unit) scale, with positive values indicating more severe impairment levels.

on the latent variable suggests that productivity, leisure, transportation, and fatigue items were “more difficult” than reported in previous investigations; whereas verbal and nonverbal communication, information, inappropriate behavior, and pain items were “easier.” These differences may be explained by the fact that the current sample was taken at an earlier point in the recovery continuum than had been previous samples, making participation items such as productivity and leisure more difficult because the persons had less time to adapt to new circumstances of life. One could speculate that the differences in the pain item could be explained by better access to care in the current, more acute sample when compared with the more chronic cases sampled in previous investigations. The current sample was composed

of individuals who were able to participate in home- and community-based rehabilitation and therefore were likely to be less impaired overall than some of the individuals in prior samples (such as those requiring residential care). This may explain the relative ease (i.e., absence of impairment) demonstrated in this sample for items measuring communication, fund of information, and inappropriate behavior.

Differences in item fit between the current and previous investigations appear largely attributable to the choice in the current investigation to focus on outfit statistics rather than infit statistics. Infit is a statistic that downweights outliers, whereas outfit does not. Consequently, infit statistical values outside the chosen range of acceptability are a greater threat to the measurement. Because the most outlying persons were

TABLE 3. DIFFERENTIAL ITEM FUNCTIONING (DIF) CONTRASTS

<i>DIF contrast</i>	<i>Biased items</i>	<i>Favored group</i>	<i>Size of bias (in logits)</i>
Male - Female	Audition	Female	0.85
Age to 45 years - Age >45 years	None	N/A	N/A
Chronicity to 6 months - Chronicity >6 months	Independent Living, Transportation	Chronicity >6 months Chronicity >6 months	0.51 0.75

TABLE 4. RASCH RELIABILITY AND SEPARATION STATISTICS FOR MPAI-4 AND SUBSCALES

	<i>Person reliability</i>	<i>Person separation</i>	<i>Item reliability</i>	<i>Item separation</i>
Abilities Subscale	.83	2.21	1.00	14.29
Adjustment Subscale	.80	1.98	.99	12.34
Participation Subscale	.79	1.91	1.00	14.77
Full MPAI-4	.91	13.15	1.00	15.16

removed from the analysis in the iterative calibration process, we felt it most appropriate to focus on the more stringent outfit statistic. It should be noted that in the final calibration, no items demonstrated infit statistics exceeding acceptable cutoffs, in accordance with previous investigations. Further, the cutoff used in the current investigation for both infit and outfit statistics ($1.0 \pm .3$) was a stringent one, as Linacre and Wright recommend different cutoff values for different types of data being modeled (e.g., multiple-choice questions, rating scales, clinical observations) and suggest of $1.0 \pm .4$ for rating scales (Linacre and Wright, 1994). The latter is commonly used but ignores differences in the sample size among studies. Given the relatively large sample in the current study and the previous psychometric refinement of the MPAI, we felt that the more stringent criterion was most appropriate. All misfitting items were slightly misfitting (> 1.3) because of excess unmodeled noise (i.e., unpredictable success and failure on items) and were restricted to Ability and Adjustment items in this sample. This perhaps reflects the heterogeneity common in ABI rehabilitation populations: limitations and the need for rehabilitation in the domain of Participation are common among these patients despite a myriad of different injuries and sequelae. In light of the choice to focus on outfit and the stringent criterion choice, these misfitting items may best be viewed as less reliable but yet meaningful contributors to

measurement of outcomes following ABI (Linacre and Wright, 1994).

The assessment of dimensionality revealed what appear to be both chronicity- and gender-based item bias and general support for the subscale structure developed in previous analysis of the MPAI-4. Bias in favor of patients with more chronic ABI on independent living and transportation items is logical; those with more experience dealing with the effects of the injury are better able to negotiate the demands of independent living and transportation. Bias in favor of females on the audition item was marked; the size of the bias was $\sim 25\%$ of the range of the scale and roughly equivalent to the logit difference between items assessing self-awareness and independent living. This difference may be explained by longitudinal studies of hearing that suggest that sensitivity declines more than twice as fast in men as in women at most ages and frequencies and that hearing levels in both sexes are highly variable (Pearson et al., 1995), which was reflected in the imprecision of the audition item noted previously. Replicating these DIF findings in another sample may warrant rescaling of these items to account for the differences in item performance according to sample characteristics. However, given that the items were biased against subpopulations with expected difficulties on those items, it appears that the items are detecting differences in ability, although perhaps premorbid ability in the case of the audition item. A more direct solution may be to amend the administration manual to call attention to issues in rating participation items for patients with relatively acute ABI and to possible sex differences in audition.

There was general support for the subscale structure developed in previous psychometric investigation of the MPAI-4. A number of item residuals were correlated, indicating subsidiary dimensions not accounted for by the latent linear variable created in the measurement calibration process. PCA demonstrated that these item residuals loaded together on three statistically significant factors that were similar to the Ability, Adjustment, and Participation domains suggested by the subscale structure. Notable exceptions to the previously defined subscale structure were the dizziness item, which fell outside acceptable limits for both infit and outfit statistics and may have loaded errantly as a result of unmodeled noise; the self-awareness item, which loaded negatively on all three dimensions; and the significant factor loadings of physical functioning items on the Participation dimension, which may reflect barriers to participation by those with physical impairments.

Although these results affirm the subscale structure, they also violate the unidimensional Rasch model, which assumes that person ability is the single dimension that accounts for variance in item responses (Marais and Andrich, 2008). As a consequence, this multidimensionality can disrupt parameters such as person reliability and separation and can com-

TABLE 5. ITEM RESIDUAL CORRELATIONS

<i>Correlation values</i>	<i>Items</i>
0.56	anxiety, depression
0.50	mobility, self-care
0.45	social contact, leisure
0.44	memory, problem-solving
0.44	attention, memory
0.41	self-care, independent living
0.41	anxiety, sensitivity
0.40	anger, inappropriate behavior
0.36	anxiety, anger
0.36	depression, sensitivity
0.36	mobility, use of hands
0.35	inappropriate behavior, self-awareness
0.35	attention, problem-solving
0.33	independent living, transportation
0.33	vision, visuospatial abilities
0.33	depression, anger
0.33	use of hands, self-care
0.32	dizziness, pain
0.32	memory, information
0.30	motor speech, verbal communication
0.30	verbal communication, nonverbal communication

TABLE 6. PRINCIPAL COMPONENTS ANALYSIS (PCA) FACTOR LOADINGS OF ITEM RESIDUALS

<i>PCA contrast (items worth of variance)</i>	<i>Items with positive load (loading)</i>	<i>Items with negative load (loading)</i>
1st (4.0)	Anxiety (.65) Depression (.64) Pain (.57) Sensitivity (.56) Dizziness (.48) Anger (.42) Fatigue (.29) Family/SO Relationships (.24) Audition (.23) Vision (.12) Leisure (.10) Social Contact (.08) Inappropriate Behavior (.03)	Problem-solving (-.57) Information (-.49) Verbal communication (-.48) Memory (-.46) Money Management (-.41) Nonverbal communication (-.37) Self-Awareness (-.34) Attention (-.30) Independent Living (-.29) Visuospatial Abilities (-.28) Motor Speech (-.23) Transportation (-.20) Initiation (-.19) Self-Care (-.17) Use of Hands (-.04) Mobility (-.02) Productivity (-.02)
2nd (3.7)	Mobility (.72) Self-Care (.66) Independent Living (.52) Use of Hands (.52) Leisure (.41) Transportation (.37) Fatigue (.25) Pain (.21) Vision (.19) Motor Speech (.13) Productivity (.12) Social Contact (.12) Dizziness (.11) Visuospatial Abilities (.08) Money Management (.03) Vision (.59) Visuospatial abilities (.48) Motor Speech (.32) Dizziness (.32) Audition (.31) Verbal communication (.21) Use of hands (.19) Information (.17) Sensitivity (.06) Memory (.05) Productivity (.03) Mobility (.02) Anxiety (.02)	Anger (-.51) Inappropriate behavior (-.50) Attention (-.43) Problem-Solving (-.41) Memory (-.39) Nonverbal communication (-.39) Self-Awareness (-.37) Family/SO Relationships (-.31) Sensitivity (-.29) Anxiety (-.28) Depression (-.28) Verbal communication (-.16) Information (-.13) Initiation (-.07) Audition (-.06) Social contact (-.58) Leisure (-.54) Inappropriate Behavior (-.29) Independent Living (-.27) Initiation (-.27) Self-Care (-.21) Fatigue (-.20) Family/SO Relationships (-.19) Self-Awareness (-.13) Nonverbal communication (-.13) Attention (-.11) Money Management (-.10) Anger (-.05) Pain (-.05) Problem Solving (-.03) Depression (-.02) Transportation (-.01)
3rd (2.0)		

promise the validity of measurement when persons are near a clinical cut point or estimates are made on an individual's performance on just a few items, such as in computer adaptive testing (Tennant and Conaghan, 2007). At the same time, every latent construct is multidimensional at some level of analysis, and therefore the tolerance of multidimensionality in a measurement instrument is a relative issue, driven by the measurement goal. Linacre illustrates the trade-off between measuring broad constructs and violating the unidimensional Rasch model: "Mathematics includes arithmetic, algebra, ge-

ometry and word problems. Sometimes we want a 'geometry test'. But, for most purposes, we want a 'math test'" (Linacre, 2008). This illustration is analogous to measuring different facets of outcome after ABI. According to Clark and Watson, a single scale structured into subscales must simultaneously reflect a single, general factor and also demonstrate greater relatedness within subscale items than among subscale items (Clark and Watson, 1995). The coherence of the MPAI-4 items to the Rasch dimension indicates that they define a single, general factor, and the interrelatedness between subscale

items as demonstrated by PCA is greater than the intra-relatedness across subscales. As such, the structure of the MPAI-4 seems to strike a balance between these qualities. There is sufficient coherence of all items to a general linear construct of outcome after ABI to recommend the MPAI-4 total score as a measure of this construct. However, there is also sufficient evidence of multidimensionality to recommend examination of subscale indices, and even individual items, for more detailed clinical characterization of individual cases.

One way to deal with the multidimensionality is to collapse polytomous (e.g., mild, moderate, severe) items into dichotomous (e.g., normal/impaired) items and sum the number of dichotomous items to create a polytomous "superitem" (e.g., 0 items normal, 1 item normal, 2 items normal, 3 items normal) (Wilson, 1989). This polytomous superitem can then be submitted to Rasch analysis with existing polytomous scale items without violating the assumptions of the model. However, the cost of this manipulation is a loss of statistical and diagnostic information (Baghaei, 2008). For example, whereas combining items such as anxiety and depression (or other Adjustment Index items) would accommodate local dependence and thus reduce multidimensionality, it is very unlikely that clinical users would find combining these constructs in any way useful in evaluating their patients.

It is important to acknowledge that we embrace the complexity of outcome after ABI. The choice to use a unidimensional (i.e., Rasch) psychometric model reflects the clinical realities faced by multidisciplinary teams required, at certain points in clinical care, to get together to take a broad view of and document a patient's recovery, and to determine the need for and nature of appropriate clinical services. The MPAI-4 is designed to help them do this. We would argue that no therapists have a unidimensional concept of outcome, but the clinical systems of care require at some points that clinicians treat recovery in a unidimensional way. A physical therapist, for example, understands that range of motion, strength, and balance are essential components of walking. He or she also understands that walking, cognition, and physical endurance are key components of mobility. Considering, for a moment in a team meeting with other rehabilitation professionals, how the dimensions of physical functioning contribute to a patient's overall picture of outcome does not cause the therapist to forget the nested contributions of different domains and subdomains to the broader construct of physical functioning. A total score on the MPAI-4 is not a perfect characterization of outcome, but it is useable and more refined than other similar measurement options. As larger patient samples become available, the degree of compromise to person estimates can be assessed with bifactor models and confirmatory factor analysis approaches.

As with previous analyses of the MPAI, the current study is limited by sample-specific features of the data. Consecutive analyses of the MPAI with different samples have generally shown the measure to represent a broad construct that defines outcome after ABI with high construct validity and internal consistency. This line of research has also identified component interrelated factors with the MPAI that represent domains broadly defined as Ability, Adjustment, and Participation. However, across studies, there have been subtle variations in item fit and hierarchical relationships, which were again apparent in the current more detailed study. Whereas none of these violations of the linear Rasch model

appear sufficiently marked to justify modifications to the instrument at this time, they do support ongoing analyses and evaluation as larger and better-defined, as well as more representative, samples are acquired. Evaluation of alternative measurement models, such as multiple-parameter logistic models, also merits exploration in attempting to capture the hierarchical relationship among subdimensions of the MPAI-4. A primary goal in the development of the MPAI has always been to provide a clinically usable instrument that targets common clinically relevant sequelae and limitations after ABI as briefly as possible. In its current form, the MPAI-4 appears to meet the goals of clinical relevance, usability, and psychometric quality, not perfectly, but well. Nonetheless, ongoing evaluation in progressively larger and more specific samples is desirable to further evaluate the need for modifications of the instrument to improve fit in specific samples and settings and to account for new theoretical developments.

Acknowledgments

This manuscript was prepared with support from a Small Business Technology Transfer (STTR) grant from the National Institute for Neurological Diseases and Stroke (NINDS).

Author Disclosure Statement

No competing financial interests exist.

References

- Baghaei, P. (2008). The Rasch model as a construct validation tool. *Rasch Measurement Transactions* 22, 1145–1146.
- Bohac, D.L., Malec, J.F., and Moessner, A.M. (1997). Factor analysis of the Mayo-Portland Adaptability Inventory: structure and validity. *Brain Inj.*, 11, 469–482.
- Clark, L.A., and Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychol. Assess.* 7, 309–319.
- Lezak, M. (1987). Relationship between personality disorders, social disturbances, and physical disability following traumatic brain injury. *J. Head Trauma Rehabil.* 2, 57–69.
- Linacre, J.M. (1994). Sample size and item calibration (or person measure) stability. *Rasch Measurement Transactions*, 7, 328.
- Linacre, J.M. (2008). *Winsteps: Rasch-Model Computer Program* (Version 3.32). Chicago, IL: Winsteps.
- Linacre, J.M., and Wright, B.D. (1994). Dichotomous mean-square infit and outfit chi-square fit statistics. *Rasch Measurement Transactions* 8, 360.
- Malec, J.F. (2004). Comparability of Mayo-Portland Adaptability Inventory ratings by staff, significant others and people with acquired brain injury. *Brain Inj.* 18, 563–576.
- Malec, J.F., Buffington, A.L., Moessner, A.M., and Thompson, J. M. (1995). Maximizing vocational outcome after brain injury: integration of medical and vocational hospital-based services. *Mayo Clin. Proc.* 70, 1165–1171.
- Malec, J.F., Buffington, A.L.H., Moessner, A.M., and Degiorgio, L. (2000a). A medical/vocational case coordination system for persons with brain injury: an evaluation of employment outcomes. *Arch. Phys. Med. Rehabil.* 81, 1007–1015.
- Malec, J.F., and Degiorgio, L. (2002). Characteristics of successful and unsuccessful completers of three postacute brain injury rehabilitation pathways. *Arch. Phys. Med. Rehabil.* 83, 1759–1764.

- Malec, J.F., Kragness, M., Evans, R.W., Finlay, K.L., Kent, A., and Lezak, M.D. (2003). Further psychometric evaluation and revision of the Mayo-Portland Adaptability Inventory in a national sample. *J. Head Trauma Rehabil.* 8, 479–492.
- Malec, J.F., and Lezak, M.D. (2008). *Manual for the Mayo-Portland Adaptability Inventory*, 2nd ed. Available at www.tbimimis.org/combi/mpai.
- Malec, J.F., Moessner, A.M., Kragness, M., and Lezak, M.D. (2000b). Refining a measure of brain injury sequelae to predict postacute rehabilitation outcome: rating scale analysis of the Mayo-Portland Adaptability Inventory (MPAI). *J. Head Trauma Rehabil.* 15, 670–682.
- Malec, J.F., and Thompson, J.M. (1994). Relationship of the Mayo-Portland Adaptability Inventory to functional outcome and cognitive performance measures. *J. Head Trauma Rehabil.* 9, 1–15.
- Marais, I., and Andrich, D. (2008). Formalising dimension and response violations of local independence in the unidimensional Rasch model. *J. Appl. Meas.* 9, 200–215.
- Pearson, J.D., Morrell, C.H., Gordon-Salant, S., Brant, L.J., Metter, E.J., Klein, L.L., and Fozard, J.L. (1995). Gender differences in a longitudinal study of age-associated hearing loss. *J. Acoust. Soc. Am.* 97, 1196–1205.
- Scott, N.W., Fayers, P.M., Aaronson, N.K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M.A., and Sprangers, M.A. (2009). A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J. Clin. Epidemiol.* 62, 288–295.
- Smith, E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J. Appl. Meas.* 3, 205–231.
- Smith, R.M., Linacre, J.M., and Smith, E.V., Jr. (2003). Guidelines for manuscripts. *J. Appl. Meas.*, 4, 198–204.
- Tennant, A., and Conaghan, P.G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum.* 57, 1358–1362.
- Wilson, M. (1989). Saltus: a psychometric model of discontinuity in cognitive development. *Psychol. Bull.* 105, 276–289.
- Zwick, R., Thayer, D.T., and Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement*, 36, 1–28.

Address correspondence to:

Jacob Kean, Ph.D.

Department of PM&R

Indiana University School of Medicine

4141 Shore Drive

Indianapolis, IN 46254

E-mail: jakean@indiana.edu