

Routledge
Taylor & Francis Group

Neuropsychological Rehabilitation

An International Journal

ISSN: 0960-2011 (Print) 1464-0694 (Online) Journal homepage: www.tandfonline.com/journals/pnrh20

The measurement structure of the Mayo-Portland Adaptability Inventory-4 (MPAI-4) in individuals with acquired brain injury: Evidence from Rasch analysis and classical test theory in a Japanese cohort

Nozomi Oyama, Fay Crawford , Shinichi Wada , Kaoru Saito & Andrew Bateman

To cite this article: Nozomi Oyama, Fay Crawford , Shinichi Wada , Kaoru Saito & Andrew Bateman (03 Nov 2026): The measurement structure of the Mayo-Portland Adaptability Inventory-4 (MPAI-4) in individuals with acquired brain injury: Evidence from Rasch analysis and classical test theory in a Japanese cohort, Neuropsychological Rehabilitation, DOI: 10.1080/09602011.2026.2758193

To link to this article: https://doi.org/10.1080/09602011.2026.2758193

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 03 Nov 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=pnrh20

NEUROPSYCHOLOGICAL REHABILITATION
https://doi.org/10.1080/09602011.2026.2758193

Routledge
Taylor & Francis Group

CASE REPORT

OPEN ACCESS

Check for updates

The measurement structure of the Mayo-Portland Adaptability Inventory-4 (MPAI-4) in individuals with acquired brain injury: Evidence from Rasch analysis and classical test theory in a Japanese cohort

Nozomi Oyama [a,d], Fay Crawford[a], Shinichi Wada[b], Kaoru Saito[c] and Andrew Bateman[a]

[a]Health and Social Care, University of Essex, Colchester, UK; [b]Department of Rehabilitation Medicine, Sanno Rehabilitation Clinic, Tokyo, Japan; [c]Reinbou Kawasaki, Kanagawa, Japan; [d]Rehabilitation Department, Junwakai Memorial Hospital, Miyazaki, Japan

ABSTRACT

The Mayo-Portland Adaptability Inventory-4 (MPAI-4) is widely used to assess comprehensive outcomes following acquired brain injury (ABI). Although a Japanese version is available, its psychometric properties have not been evaluated. This study examined the measurement properties of the Japanese MPAI-4. A cross-sectional study was conducted with 163 participants receiving post-discharge rehabilitation between June 2024 and June 2025 at rehabilitation facilities in Japan. Rasch analyses examined reliability, unidimensionality, item fit, category functioning, targeting and differential item functioning. Classical Test Theory analyses assessed internal consistency and factor structure. The Japanese MPAI-4 demonstrated a Person Separation Index of 0.89 and high internal consistency (Cronbach's $\alpha = 0.91$). Rasch analysis identified departures from model assumptions, including multidimensionality and disordered category thresholds. Exploratory factor analysis suggested a multifactorial structure, although it did not clearly align with the original three-domain model. These findings indicate that the Japanese MPAI-4 captures multiple aspects of post-ABI functioning, reflecting the complexity of outcomes in community rehabilitation. Although measurement issues should be considered when interpreting scores, the study provides initial psychometric evidence for the Japanese MPAI-4. These findings may support its clinical use for comprehensive assessment of post-ABI functioning in Japan, while contributing to the cross-cultural evidence base for the MPAI-4.

ARTICLE HISTORY
Received 22 June 2026
Accepted 28 October 2026

KEYWORDS
The Mayo-Portland Adaptability Inventory-4 (MPAI-4); Acquired brain injury (ABI); Rehabilitation outcome; Rasch analysis; Classical Test Theory (CTT)

CONTACT Nozomi Oyama no21869@essex.ac.uk University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, United Kingdom Nozomi Oyama @nozomi_oyama

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Introduction

Outcome assessment constitutes a fundamental component of rehabilitation practice and is used to inform clinical judgement, facilitate interprofessional communication and shape evaluation in both practice and research contexts. Given the complexity of rehabilitation outcomes, it is essential to use measures with established reliability and validity and to understand the measurement properties underlying score interpretation. In particular, psychometric evaluation can help determine whether outcome measures function consistently across different populations and contexts and whether their scoring structures adequately reflect the constructs being assessed (Mokkink et al., 2010).

Outcomes are particularly complex in acquired brain injury (ABI), which refers to brain damage occurring after birth and includes both traumatic and non-traumatic causes. ABI can result in a wide range of cognitive, emotional, behavioural, communicative and physical impairments, many of which persist long-term (Goldman et al., 2022). Moreover, the consequences of ABI may extend beyond specific neurological impairments, affecting psychological wellbeing and social functioning. Such consequences can create new challenges in family relationships and social roles or exacerbate pre-existing difficulties (Evans, 2011, 2019).

Given this complexity, rehabilitation for individuals with ABI should extend beyond the recovery of impaired functions to the optimization of functioning across psychological, social and daily life domains (Evans, 2011; Wilson et al., 2009). Accordingly, comprehensive outcome assessment is essential alongside functional assessments, particularly in post-discharge and community-based rehabilitation settings.

Within the Japanese context, rehabilitation services are increasingly required to respond to a rapidly ageing population, and this demand extends to ABI rehabilitation. A recent scoping review (Oyama et al., 2026) identified the Functional Independence Measure (FIM) (Granger & Hamilton, 1994; Keith et al., 1987) as the most commonly used outcome measures in post-discharge ABI rehabilitation in Japan, followed by the Frenchay Activities Index (FAI) (Holbrook & Skilbeck, 1983), the Life-Space Assessment (LSA) (Baker et al., 2003) and the Barthel Index (BI) (Mahoney & Barthel, 1965). These measures primarily assess independence in activities of daily living and mobility, and their use has been promoted in part through policy initiatives aimed at reducing long-term care needs in an ageing society. While such measures are useful for standardized service evaluation, reliance on them alone may be insufficient for capturing the multifaceted consequences of ABI. For example, the FIM has been shown to exhibit ceiling effects in individuals with mild brain injury in outpatient settings (Oyama et al., 2025b). These limitations highlight the need for more comprehensive outcome measures in routine practice.

This study focuses on the psychometric properties of the Japanese version of the Mayo-Portland Adaptability Inventory-4 (MPAI-4). The MPAI-4 is a

comprehensive outcome measure designed to assess the sequelae of ABI, particularly in the post-discharge phase (Bohac et al., 1997; Malec, 2004; Malec et al., 2000, 2003, 2012; Malec & Kean, 2016; Malec & Lezak, 2008; Malec, et al., 2025; Malec & Moessner, 1997; Malec & Thompson, 1994). It evaluates functioning across three domains: Ability, Adjustment and Participation, using a 5-point rating scale, with higher scores indicating greater impairment. The MPAI-4 is characterized by its broad coverage of cognitive, emotional, behavioural and social domains and can be completed by clinicians, individuals with ABI or family members. It has been used to examine different aspects of functioning following ABI, including psychosocial functioning across the Ability, Adjustment and Participation domains (Mann et al., 2023) and participation outcomes following ABI (Thüss et al., 2025). The 3rd edition of the MPAI-4 manual was published in 2024, with revised guidance on administration and rating and updated reference data for Rasch-derived T-scores (Malec, 2024).

The original version of the MPAI-4 has demonstrated robust reliability and validity and has been translated into multiple languages, including French, Danish, Spanish, German, Italian, Portuguese, Swedish, Dutch and Korean (Cattelani et al., 2009; Guerrette & McKerral, 2020). Although a Japanese version was recently developed (Wada et al., 2022a), its psychometric properties have not yet been evaluated, and its use remains limited. To date, only a small number of studies have reported findings using the MPAI-4 (Oyama et al., 2025a, 2025b; Wada, 2024). This limited psychometric evidence may restrict its wider adoption in clinical practice in Japan.

More broadly, evaluation of translated versions of the MPAI-4 is important not only for local clinical implementation but also for understanding whether the instrument functions consistently across different linguistic, cultural and rehabilitation contexts. Evidence from several language versions has generally supported the utility of the MPAI-4; however, a recent systematic review identified considerable variability in the methodological quality of validation studies and highlighted ongoing limitations in the evidence for cross-cultural validity (Ataman et al., 2023). Consequently, further psychometric evaluation in diverse populations remains important for strengthening the international evidence base for the MPAI-4. Although the present study does not directly examine cross-cultural validity, it contributes to this broader literature by providing a comprehensive evaluation of the Japanese version using complementary psychometric approaches.

The objective of this study was to examine the reliability and structural validity of the Japanese version of the MPAI-4. To our knowledge, this is the first study to provide a comprehensive psychometric evaluation of the Japanese version. Establishing these properties will support the use of the Japanese MPAI-4 in clinical practice and research and contribute to the broader international evidence base regarding the assessment of long-term functioning, adjustment and participation following ABI.

Methods

Methodological approach

The evaluation of the measurement properties of the Japanese version of the MPAI-4 was guided by the recommendations of the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN), which provide an internationally recognized framework for evaluating health measurement instruments (COSMIN, no date; Mokkink et al., 2010; Prinsen et al., 2018). Within the COSMIN framework, content validity is considered the most important measurement property, followed by the evaluation of internal structure where appropriate for the underlying measurement model.

As described in the Introduction, the MPAI was developed and refined through a series of studies involving theoretical development, classical test theory and Rasch analysis, with the development of the MPAI-4 subsequently documented by Malec and Lezak (2008) and Malec (2024). Importantly, the MPAI subscales were rationally derived to represent distinct domains of outcome following ABI, and their structure was subsequently supported by empirical analyses, strengthening the conceptual basis of the subscale structure. In addition, Lexell et al. (2012) demonstrated that all 30 scoreable items of the MPAI-4 could be mapped to components and categories of the International Classification of Functioning, Disability and Health (ICF). Although the original development study did not report a formal expert-panel procedure, these findings provide indirect support for the conceptual relevance and coverage of the MPAI-4 content.

Following the assessment of content validity, structural validity can be evaluated to determine whether the scores of a measure adequately reflect the dimensionality of the construct being measured (COSMIN, n.d.). The relevance of this evaluation depends on the measurement model underlying the instrument. Specifically, structural validity is relevant when a measure is based on a reflective model. In psychometric theory, relationships between latent constructs and observed indicators are commonly conceptualized using either reflective or formative measurement models (Bollen & Lennox, 1991).

In a reflective model, the latent construct is assumed to influence responses to the individual items; therefore, changes in the underlying construct would be expected to result in systematic changes across related items, which would generally be expected to correlate with one another. In contrast, a formative model assumes that the individual items represent different aspects of the construct and collectively contribute to its overall meaning. Consequently, the items do not necessarily need to be correlated, and removing an item may change the meaning of the overall measure (Bollen & Lennox, 1991; Coltman et al., 2008; Diamantopoulos & Winklhofer, 2001). The choice between reflective and formative models should be guided primarily by the theoretical relationship between

the construct and its items rather than by statistical results alone (Bollen & Lennox, 1991).

With regard to the MPAI-4, the appropriate measurement model remains debatable. Ataman et al. (2023) considered the MPAI-4 to be a formative measure and therefore regarded structural validity and internal consistency as uninterpretable. Nevertheless, earlier development and psychometric studies have provided empirical support for the three-subscale structure of the MPAI-4. Kean et al. (2011) described the instrument as measuring a broad but unitary construct of global outcome after ABI, with three subordinate and correlated constructs represented by the Ability, Adjustment and Participation subscales. Their analyses supported the subscale structure, and each subscale also performed adequately as an individual measure in Rasch analysis. These findings do not establish a reflective measurement model, but they provide a basis for examining the dimensionality and measurement properties of the subscales separately.

In the present study, the structural properties of the Japanese MPAI-4 were therefore examined at both the subscale and whole-scale levels, consistent with previous psychometric investigations of the instrument. At the subscale level, this approach was intended to examine whether the items within each established domain showed sufficient unidimensionality to support interpretation of the corresponding subscale. The findings are therefore interpreted as evidence regarding the internal structure of the MPAI-4, in light of the ongoing uncertainty regarding its underlying measurement model.

Study design

This study employed a cross-sectional design to evaluate the psychometric properties of the Japanese version of the MPAI-4. Rasch analysis was used as the primary analytical framework to examine item functioning, dimensionality. Complementary analyses based on Classical Test Theory (CTT) were conducted to provide additional evidence regarding reliability and structural validity.

Participants

Inclusion and exclusion criteria

Participants were adults aged 18 years or older with ABI, including conditions such as stroke, traumatic brain injury, encephalopathy, encephalitis, brain tumour and subdural haematoma. Individuals with progressive neurological diseases, such as multiple sclerosis, Parkinson's disease or Alzheimer's disease, were excluded. Participants were required to be receiving post-discharge rehabilitation between June 2024 and June 2025, with no restriction regarding time since injury or onset. Eligible settings included outpatient rehabilitation, day-care rehabilitation, home-visit rehabilitation, independent living training

services and employment transition support services. Individuals receiving inpatient hospital-based rehabilitation were excluded. Only individuals able to provide informed consent were included.

Sample size considerations

Sample size requirements in Rasch analysis depend on the intended precision of item parameter estimates, with larger samples generally providing more precise and stable estimates. Previous methodological studies suggest that samples of approximately 100 participants are generally sufficient for stable item calibration, whereas larger samples (e.g., 200–500) may be required when higher levels of precision are desired (Linacre, 1999; Tennant & Conaghan, 2007; Van de Winckel et al., 2022). Based on these recommendations, a target sample of at least 100 participants was considered sufficient for the primary Rasch analyses conducted in this study.

Settings and recruitment

Data were collected from seven rehabilitation facilities across Miyazaki, Tokyo and Kanagawa, Japan. These included one rehabilitation hospital (outpatient, home-visit and day-care rehabilitation departments), two rehabilitation clinics, a community-based group rehabilitation service for individuals with ABI, and three municipal rehabilitation centres.

Potential participants were identified by trained rehabilitation staff at each participating site. Research collaborators at each site received the study protocol and standardized instructions regarding participant recruitment and administration of the MPAI-4. Prior to data collection, orientation sessions were conducted either in person or online to standardize participant recruitment, MPAI-4 administration and scoring.

Ethical approval

Ethical approval was obtained from the Research Ethics Committee of Junwakai Memorial Hospital, Miyazaki, Japan (Approval No. 240422-1, approved on 22 April 2024) and the University of Essex, UK (Approval No. ETH2223-1359, approved on 10 May 2024).

Data collection procedures

Following informed consent, demographic and clinical information was extracted from medical records, including age, gender, diagnosis, date of onset and type of rehabilitation received.

The MPAI-4 was completed by research collaborators in accordance with the Japanese MPAI-4 manual (Wada et al., 2022b). Ratings were primarily based on routine clinical observations and information recorded from medical records.

When necessary, a brief structured interview with the participant was conducted to obtain additional information required for scoring.

All data were stored securely using the University of Essex Box cloud storage system. Participants were assigned unique identification codes, and no personally identifiable information was retained within the dataset.

Statistical analysis

Analyses were performed using RUMM2030 and IBM SPSS Statistics (version 29). Descriptive statistics were first calculated to summarize participants' demographic and clinical characteristics. All analyses were conducted for the full scale, followed by separate analyses for the three domains of the MPAI-4 (Ability, Adjustment and Participation).

Rasch analysis

The structural validity of the original MPAI-4 has primarily been examined using the Rasch measurement model (Malec, 2004; Malec et al., 2003; Malec & Lezak, 2008). In the present study, consistent with the methodological rationale described above, Rasch analysis was used as the primary analytical approach to evaluate structural validity and related measurement properties of the Japanese version of the MPAI-4.

The Rasch model estimates person ability and item difficulty on a common latent continuum and evaluates the extent to which observed responses conform to model expectations (Rasch, 1960/1980; Tennant & Conaghan, 2007). When model assumptions are satisfied, ordinal raw scores can be transformed into interval-level measures, allowing more robust analysis and interpretation of scale scores (Tennant et al., 2004). As summarized in several review studies on Rasch analysis (e.g., Aryadoust et al., 2019; Lombard et al., 2022; Stolt et al., 2022), its use for scale development and validation has become increasingly widespread. In rehabilitation research, its application is now well established (Tennant & Küçükdeveci, 2023), and it is increasingly used for scale development and validation in the field of brain injury rehabilitation (e.g., Bateman et al., 2009; Caracuel et al., 2011; Pretz et al., 2016; Simblett & Bateman, 2011; Turner-Stokes et al., 2019; Wang et al., 2022).

Analyses were based on the Partial Credit Model, which is appropriate for polytomous response categories. Extreme person responses were excluded from the estimation of person locations in accordance with RUMM2030 procedures. The analytical procedures were as follows:

(1) Internal consistency reliability was evaluated using the Person Separation Index (PSI), which indicates the ability of the measure to distinguish between individuals with different levels of the latent trait. Values of 0.70 or higher were considered acceptable for group comparisons, while

values above 0.90 were considered indicative of the ability to distinguish three or more levels of latent trait (Fisher, 1992).

(2) Unidimensionality was assessed using principal component analysis of residuals followed by paired *t*-tests according to the procedure proposed by Smith (2002). Fewer than 5% significant *t*-tests were interpreted as evidence supporting unidimensionality.

(3) Local dependency was assessed using residual correlations based on Yen's Q3 statistic (Yen, 1984). Following Christensen et al. (2017), values exceeding 0.20 above the average residual correlation were considered indicative of dependency.

(4) Item fit was evaluated using item-person interaction and item-trait interaction chi-square statistics. Fit residuals within the range of ±2.5 were considered acceptable (Andrich & Marais, 2019) and non-significant chi-square statistics indicated adequate model fit.

(5) Response category functioning was examined by assessing the ordering of category thresholds.

(6) Targeting was assessed through visual inspection of person-item threshold distribution maps to evaluate the alignment between participant ability and item difficulty along the latent trait continuum.

(7) Differential item functioning (DIF) was examined using two-way analysis of variance (ANOVA) of standardized residuals across class intervals. Both uniform and non-uniform DIF were evaluated. Bonferroni adjustments were applied to control for multiple comparisons.

(8) Where substantial item misfit or other notable departures from the Rasch model were identified, exploratory post hoc analyses were conducted to examine the potential impact of rescoring and item removal on model fit and unidimensionality. Where disordered thresholds were identified, adjacent response categories were collapsed as appropriate. Item removal was also explored iteratively based on the overall pattern of item functioning, with items showing substantial misfit and local dependency considered for removal. Following item removal, unidimensionality was re-evaluated and changes in the PSI were examined.

Rasch analyses were reported in accordance with the Rasch Reporting Guideline for Rehabilitation Research (RULER) (Mallinson et al., 2022; Van de Winckel et al., 2022).

Classical Test Theory (CTT) analysis

Complementary analyses based on CTT were conducted to contextualize the Rasch findings and to provide additional information regarding internal consistency and factor structure. Integrating both CTT and Rasch-based approaches is consistent with contemporary recommendations in psychometric research (Petrillo et al., 2015; Prieto et al., 2003). The procedures were as follows:

(1) Internal consistency reliability was assessed using Cronbach's α coefficients (Cronbach, 1951). Values of 0.70 or higher were considered acceptable for research purposes (Fisher, 1992; Nunnally & Bernstein, 1994).
(2) Corrected item–total correlations were calculated to assess the relationship between each item and the remaining items in the scale. Correlation coefficients above 0.30 were considered acceptable (Nunnally & Bernstein, 1994).
(3) Exploratory factor analysis (EFA) was conducted to examine the underlying factor structure. Given the ordinal nature of the data and the expectation that constructs were correlated, principal axis factoring with Promax rotation was used. Sampling adequacy was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity. The number of factors retained was determined based on eigenvalues greater than 1.0, inspection of the scree plot and theoretical interpretability with reference to the established MPAI-4 domains (Cattell, 1966; Kaiser, 1960). Although factor loadings of 0.30 are often considered the minimum threshold for interpretation, loadings of 0.40 or greater were regarded as salient in the present study, consistent with common psychometric practice (Nunnally & Bernstein, 1994).

Results

Participant characteristics

Data from 163 participants (105 males and 58 females) who received post-discharge rehabilitation services between June 2024 and June 2025 were analyzed. Participant characteristics are summarized in Table 1. The median age was 63 years (interquartile range [IQR]: 52–76), and the median time since onset or injury was 24.5 months (IQR: 8–52.8). The sample was predominantly composed of individuals with stroke, particularly cerebral infarction and cerebral haemorrhage, while a range of other ABI diagnoses were also represented.

Table 1. Demographic and clinical characteristics of participants.

	N = 163
Sex (male: female)	105: 58
Age, median (IQR), years	63 (52–76)
Time since onset/injury, median (IQR), months	24.5 (8–52.8)
Diagnosis (%) Cerebral infarction	81 (49.7)
Cerebral haemorrhage	49 (30.1)
Traumatic brain injury	10 (6.1)
Subarachnoid haemorrhage	9 (5.5)
Brain tumour	5 (3.1)
Hypoxic ischaemic encephalopathy	4 (2.5)
Encephalitis	3 (1.8)
Other	2 (1.2)

*IQR: Interquartile range.

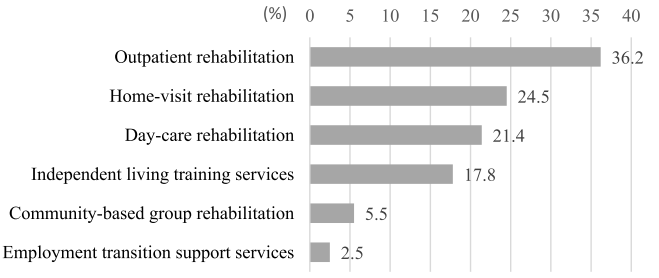


Figure 1. Proportion of post-discharge rehabilitation services utilized by participants.

Participants accessed a range of post-discharge rehabilitation services (Figure 1). Outpatient rehabilitation was the most commonly utilized service, followed by home-visit and day-care rehabilitation. Some participants accessed multiple services.

Rasch analysis

Overall model performance

Of the 163 participants included in the study, four were classified as having extreme scores and were excluded from the estimation of person locations; category frequencies were therefore based on 159 participants.

Rasch analysis indicated that the full scale did not adequately fit the Rasch model and demonstrated evidence of multidimensionality. At the subscale level, model fit and dimensionality showed some improvement, although the Rasch model assumptions were not fully met for any subscale. A summary of the Rasch model fit and measurement properties is presented in Table 2.

Multidimensionality and local dependency

Evidence of multidimensionality was observed for the full scale. Although the proportion of significant paired *t*-tests was lower at the subscale level, none of the subscales satisfied the criterion for strict unidimensionality (Table 2).

Local dependency was identified across all subscales, particularly between conceptually related items (Table 3). The strongest dependencies were observed between "Social contact" and "Leisure and recreational activities," "Anxiety" and "Depression," and "Memory" and "Visuospatial abilities."

Table 2. Summary of Rasch model fit and measurement properties.

	Item–trait interaction χ^2 (df)	*p*-value	PSI	% significant *t*-tests	No. of misfitting items	No. of items with ordered thresholds
Full scale	155.53 (60)	<.001	0.89	26.5%	8	5
Ability	120.89 (29)	<.001	0.69	15.3%	6	3
Adjustment	42.62 (24)	.01	0.76	8.5%	1	2
Participation	29.50 (16)	.02	0.82	9.6%	1	0

Table 3. Residual correlations between item pairs indicating potential local dependency across subscales.

Subscale	Item pair	Residual correlation (*r*)
Ability	Memory – Visuospatial abilities	0.37
	Mobility – Use of hands	0.35
Adjustment	Social contact – Leisure and recreational activities	0.67
	Anxiety – Depression	0.46
Participation	Social contact – Leisure and recreational activities	0.38
	Self-care – Residence	0.33

*Only item pairs exceeding the threshold for local dependency (adjusted Q3 > 0.20) are presented.

Item functioning

Significant item–trait interaction chi-square statistics indicated misfit to the Rasch model for the full scale and all subscales (all $p < .05$) (Table 2). At the item level, eight items demonstrated misfit, predominantly within the Ability subscale (Table 4).

Disordered thresholds were frequently observed, with only five items showing ordered thresholds. This issue was most pronounced in the Participation subscale, in which no items demonstrated ordered thresholds. Response frequencies for the original five response categories varied across items, with many items showing a concentration of responses in the lower categories. Detailed response frequencies for each item are provided in Supplementary Table S1.

Item fit residuals were within acceptable ranges across all subscales (Ability: –1.603–0.748; Adjustment: –0.689–1.270; Participation: –1.428–1.142), indicating no evidence of extreme item misfit.

Reliability and targeting

The full scale demonstrated good reliability (PSI = 0.89). At the subscale level, reliability was acceptable for Adjustment (PSI = 0.76) and Participation (PSI = 0.82), but slightly lower for Ability (PSI = 0.69).

Targeting was generally acceptable, with mean item difficulties centred close to 0 logits across subscales. However, mean person ability exceeded item difficulty in the Ability (1.19 logits) and Adjustment (1.31 logits) subscales, whereas closer alignment was observed in the Participation subscale (0.16 logits). Person–item threshold distributions are shown in Figure 2.

Table 4. Misfitting items identified by item–trait interaction statistics.

Subscale	Item	*p*-value
Ability	Use of hands	<.001
	Dizziness	<.001
	Attention	<.001
	Nonverbal communication	<.001
	Novel problem-solving	<.001
	Audition	.02
Adjustment	Pain and headache	.001
Participation	Residence	.02

*Items with statistically significant item–trait interaction chi-square probabilities ($p < .05$) are presented.

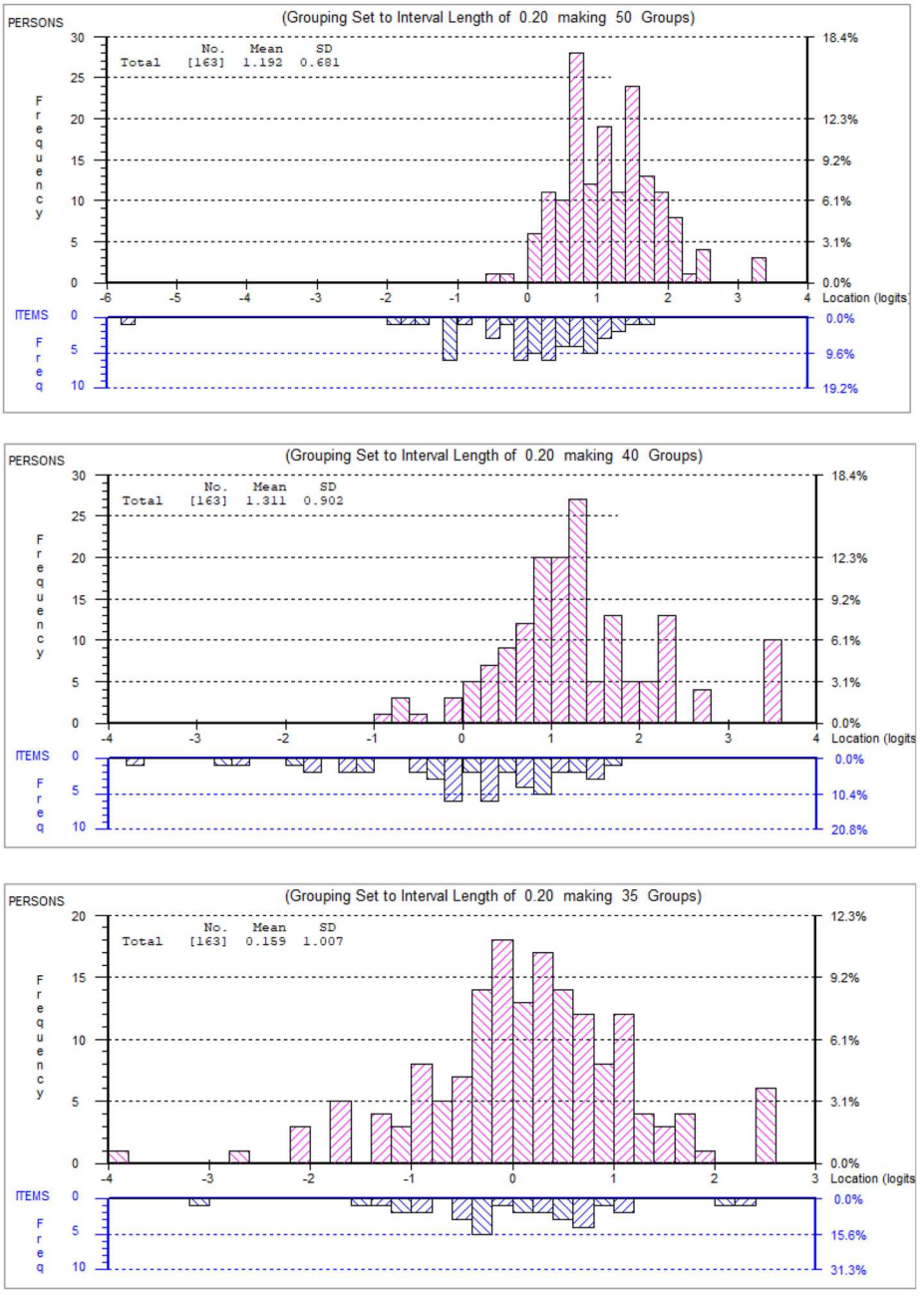


Figure 2. Person–item threshold distributions for each subscale.

Distributions of person ability and item thresholds for the Ability, Adjustment and Participation subscales (top to bottom). Overall, item thresholds were centred near 0 logits. Person ability and item thresholds were less closely alignment in the Ability and Adjustment subscales, whereas closer alignment was observed in the Participation subscale.

Differential item functioning (DIF)

DIF analyses based on ANOVA of standardized residuals identified significant uniform DIF for "Nonverbal communication" within the Ability subscale by gender ($p < .001$). By age, significant uniform DIF was observed for "Mobility" ($p < .001$) and "Nonverbal communication" ($p < .001$) within the Ability subscale, and for "Residence" within the Participation subscale ($p = .001$). No evidence of non-uniform DIF was identified. No DIF was detected in the Adjustment subscale.

Model modifications

As none of the subscales satisfied the assumption of unidimensionality, post hoc analyses including modifications were conducted at the subscale level.

Category rescoring was first undertaken to address disordered thresholds, with rescoring decisions guided by the category probability curves and category frequencies for individual items rather than by applying a uniform rescoring scheme, while avoiding unnecessary further category collapsing. Excluding the five items that had ordered thresholds with the original five response categories, the remaining items were rescored. After rescoring, the number of participants with extreme scores increased from four to nine, resulting in 154 participants being included in the analysis. The rescored items had ordered thresholds with one item with four categories, 22 items with three categories and four items with dichotomous formats. Thus, although rescoring improved threshold ordering for several items, the PSI was no longer estimable for the Ability and Adjustment subscales and decreased substantially for the Participation subscale. Detailed rescoring results and threshold distributions are provided in Supplementary Table S2 and Figure S1. Given these findings, item removal was subsequently explored to examine whether removing items with evidence of misfit or local dependency could improve the measurement properties of the subscales.

For the Ability subscale, removal of three items ("Mobility," "Use of hands" and "Nonverbal communication") reduced the proportion of significant *t*-tests from 15.28% to 4.91%; however, the PSI decreased from 0.69 to 0.63. In the Adjustment subscale, removal of "Social contact" and "Leisure and recreational activities" reduced the proportion of significant *t*-tests from 8.20% to 3.07%, while the PSI decreased from 0.76 to 0.70. In the Participation subscale, removal of "Residence" reduced the proportion of significant *t*-tests from 9.62% to 4.91%, with the PSI decreasing from 0.82 to 0.79. Detailed results are presented in Table 5.

Classical Test Theory (CTT)

Missing data were minimal across items, with item-level missing rates below 5%. Therefore, no imputation procedures were applied. CTT analyses were conducted using data from all 163 participants.

Table 5. Effects of item removal on unidimensionality and reliability of each subscale.

Subscale	Items removed	% significant paired *t*-tests (<5%)		PSI	
		Before	After	Before	After
Ability	Mobility Use of hands Nonverbal communication	15.28	4.91	0.69	0.63
Adjustment	Social contact Leisure and recreational activities	8.20	3.07	0.76	0.70
Participation	Residence	9.62	4.91	0.82	0.79

Internal consistency

The total scale demonstrated excellent internal consistency (Cronbach's $\alpha = 0.91$). Subscale reliabilities were acceptable to good (Ability: $\alpha = 0.79$; Adjustment: $\alpha = 0.84$; Participation: $\alpha = 0.86$).

Item–total correlations

Corrected item–total correlations were generally positive across subscales, although several items in the full scale and the Ability subscale demonstrated weak or negative correlations.

Exploratory factor analysis (EFA)

Full scale. EFA of the full scale demonstrated good sampling adequacy (KMO = 0.86; Bartlett's test $p < .001$). Although eight factors had eigenvalues greater than 1.0, inspection of the scree plot suggested a more parsimonious four-factor solution (Table 6).

The extracted factors were broadly interpretable as cognitive–behavioural functioning (factor 1), physical independence and daily activities (factor 2), emotional and somatic symptoms (factor 3), and visual–perceptual abilities (factor 4) (Table 6). Most items loaded meaningfully onto one factor, although several items demonstrated cross-loadings, and the overall structure did not yield a clearly interpretable simple structure. A three-factor solution corresponding to the original MPAI-4 domains was also examined; however, substantial cross-loadings and inconsistent item groupings limited its interpretability.

Subscale analyses. Sampling adequacy was satisfactory for all subscales (see notes to Tables 7–9). The number of factors retained for each subscale was determined with consideration of eigenvalues, scree plot inspection, interpretability and relevant conceptual considerations.

In the Ability subscale, EFA indicated a multidimensional structure, with items loading across cognitive, motor, sensory and visual domains (Table 7). The Adjustment subscale also demonstrated a heterogeneous structure, with items distributed across emotional, behavioural, social and somatic domains (Table 8). For the Participation subscale, a two-factor structure representing independent and social participation was identified (Table 9).

Table 6. Pattern matrix from EFA of the full scale (four-factor solution) Pattern matrix obtained from EFA of the full scale using principal axis factoring with Promax rotation.

Item	Factor 1	Factor 2	Factor 3	Factor 4
Impaired awareness	.94			
Novel problem-solving	.85			
Nonverbal communication	.84			
Attention	.75			
Irritability	.74			
Fund of information	.70			
Verbal communication	.69			
Initiation	.63			
Family relationships	.59			
Memory	.58			
Managing money	.45	(.44)		
Mobility		.85		
Self-care		.84		
Transportation		.83		
Residence		.81		
Use of hands		.59		
Leisure and recreational activities		.47		
Employment		.46		
Pain and headache			.72	
Depression			.65	(.41)
Sensitivity to mild symptoms			.64	(.44)
Anxiety			.61	
Fatigue			.46	
Dizziness			.43	
Vision				.72
Visuospatial abilities				.58

Note: Factor loadings ≥.40 are presented. Cross-loadings are shown in parentheses. Items with loadings <.40 on all factors are not displayed.

Table 7. Pattern matrix from EFA of the Ability subscale (four-factor solution) Pattern matrix obtained from EFA of the Ability subscale using principal axis factoring with Promax rotation.

Item	Cognitive	Motor	Sensory	Visual
Memory	.90			
Fund of information	.87			
Novel problem-solving	.86			
Attention	.75			
Nonverbal communication	.56			
Use of hands		.88		
Mobility		.83		
Motor speech		.60		
Audition			.86	
Verbal communication	.40		(−.51)	
Dizziness			.49	
Vision				.87
Visuospatial abilities				.67

Note: Primary factor loadings (≥.40) are presented. Cross-loadings are shown in parentheses. Loadings <.40 are suppressed unless conceptually relevant.
KMO = 0.80; Bartlett's test $p < .001$.

Discussion

Summary of main findings

This study evaluated the structural validity and selected measurement properties of the Japanese version of the MPAI-4 using Rasch analysis as the primary

Table 8. Pattern matrix from EFA of the Adjustment subscale (four-factor solution) Pattern matrix from EFA of the Adjustment subscale items using principal axis factoring with Promax rotation.

Item	Social participation	Behavioural	Emotional	Somatic
Social contact	.94			
Leisure and recreational activities	.93			
Initiation	.59			
Inappropriate behaviour		.92		
Irritability		.81		
Impaired awareness		.78		
Anxiety			.90	
Sensitivity to mild symptoms			.79	
Depression			.78	
Pain/headache				.82
Fatigue				.80
Family relationships				(.41)

Note**.** Primary factor loadings (≥.40) are presented. Cross-loadings are shown in parentheses. Loadings <.40 are suppressed unless conceptually relevant.
KMO = 0.80; Bartlett's test $p < .001$.

Table 9. Pattern matrix from EFA of the Participation subscale (two-factor solution) Pattern matrix from EFA of the Participation subscale using principal axis factoring with Promax rotation.

Item	Independent Participation	Social Participation
Self-care	.89	
Residence	.88	
Transportation	.85	
Managing money	.68	
Employment	.56	
Social contact		.99
Leisure and recreational activities		.79
Initiation		.65

Note: Primary factor loadings (≥.40) are presented. Loadings <.40 are suppressed unless conceptually relevant.
KMO = 0.82; Bartlett's test $p < .001$.

analytical framework, supported by complementary CTT approaches. Rasch analysis indicated that neither the full scale nor subscales functioned as strictly unidimensional measures. Complementary CTT analyses similarly suggested that the full scale encompasses multiple related dimensions. At the subscale level, these structures also showed evidence of internal complexity. Overall findings support the interpretation of the Japanese MPAI-4 as a multidomain measure. The full-scale score may provide an overall representation of post-ABI functioning, with the three subscale scores providing additional information about specific domains.

Interpretation of Rasch analysis findings

Dimensionality and local dependency

The pattern of findings suggests that the multidimensionality observed in the full scale was not simply a consequence of combining the three conceptual

subscales, as indications of departures from strict unidimensionality were also observed within the individual subscales.

Local dependency was observed among conceptually related items, including "Memory" and "Visuospatial abilities," "Mobility" and "Use of hands," and "Social contact" and "Leisure and recreational activities." Such dependencies can arise when items contain overlapping content or when the response to one item is logically related to the response to another, and may lead to overestimation of reliability and affect the independence of item responses assumed by the Rasch model (Christensen et al., 2017). In the present study, the occurrence of local dependency may therefore have affected the independence of item responses and potentially influenced estimates of measurement properties. However, local dependency alone does not establish that an item is unsuitable for the scale. In rehabilitation measures, items may provide clinically relevant information despite sharing some variance with related items, and the decision to remove an item should therefore consider its other psychometric and conceptual characteristics (Tennant & Küçükdeveci, 2023).

Scale modification analyses further supported this interpretation. Across the subscales, item removal reduced the indications of multidimensionality but also reduced the PSI, indicating that modifications did not consistently improve person separation. This suggests that some misfitting items may still capture clinically meaningful aspects of functioning. These findings indicate a trade-off between reducing multidimensionality and retaining clinically meaningful information (Tennant & Conaghan, 2007).

Item-level measurement properties

Most items demonstrated generally acceptable fit, with limited DIF effects of small magnitude. Age-related DIF in "Mobility" may reflect genuine age-related differences in functional capacity rather than clear evidence of measurement bias. DIF for "Nonverbal communication" across age and gender may be associated with reported age- and gender-related differences in the recognition of nonverbal emotional cues (Hall et al., 2025; Ruffman et al., 2008; Sze et al., 2012; Thompson & Voyer, 2014). Within the Participation subscale, age-related DIF for "Residence" likely reflects contextual variation (Tyler et al., 2023; Vaughan et al., 2016).

However, disordered thresholds were observed in several items, particularly in Participation, suggesting that the five-category response format did not function optimally for many items. Post hoc rescoring analyses showed that threshold ordering was generally achieved by collapsing adjacent response categories, with many items rescored to three categories. This suggests that respondents may have had difficulty distinguishing between adjacent response categories, particularly at intermediate levels of severity. Although collapsing categories improved model fit, it also reduced the number of distinguishable response levels and may have reduced the ability of the response scale to

distinguish between different levels of severity. Therefore, these findings do not indicate that a three-category response format is necessarily more appropriate for the Japanese MPAI-4.

Despite threshold disordering, the Participation subscale demonstrated relatively stable measurement characteristics compared with the other subscales. Participation-related outcomes are inherently complex and influenced by multiple interacting factors, including impairments, emotional adjustment, environmental context and social expectations (Geyh et al., 2012; World Health Organization, 2001). This complexity may be relevant to the observed threshold disordering, as distinctions between adjacent response categories may be less clear for multidimensional aspects of participation. Nevertheless, the subscale may capture clinically meaningful aspects of everyday functioning.

The reliability findings provided further information on the measurement performance of the scales. The total score showed good reliability (PSI = 0.89), with subscale PSI values ranging from 0.69 to 0.82. The PSI for the full scale was similar to the person reliability reported by Malec et al. (2012) in individuals with cerebrovascular accident (0.91). Approximately 80% of participants in the present study had cerebral infarction or cerebral haemorrhage, providing a broadly comparable clinical context. Although the full scale demonstrated multidimensionality, the relatively high PSI suggests that the full scale can still distinguish between individuals with different levels of functioning across the domains represented by the scale. Thus, the total score may provide a broad representation of functioning across several related functional domains.

The person–item threshold distributions also provided information on targeting. Participation showed closer alignment between person ability and item thresholds than Ability and Adjustment, suggesting that this subscale was better targeted to the functioning range represented in the present sample. The relatively lower item locations observed for Ability and Adjustment may indicate that difficulties in these domains were represented across lower levels of the measured trait, whereas Participation items tended to represent higher levels of the trait. This pattern is broadly consistent with the item hierarchy reported by Kean et al. (2011), in which Ability items were generally easier than Adjustment items, while Participation items were generally the most difficult. Malec (2025) noted that item locations may also provide information relevant to rehabilitation planning by indicating the levels of functioning at which different aspects of functioning are represented.

Interpretation of CTT findings

CTT analyses provided complementary information on the internal structure of the Japanese version of the MPAI-4, as CTT and Rasch analysis examine different aspects of measurement properties (Bond & Fox, 2015; Floyd & Widaman, 1995). EFA of the full scale indicated a multifactorial structure that did not clearly align

with the original three-domain model. These findings suggest that, in the Japanese version, the relationships among items may be more complex than implied by the original structure, although further cross-cultural research is required to determine whether this reflects translation-related, cultural or sample-specific influences.

At the subscale level, however, the factor structures appeared more distinct than that of the full scale. The Ability and Adjustment subscales demonstrated heterogeneous item clustering, suggesting that these domains encompass multiple related but distinct functional components. This may reflect the multidimensional nature of these domains, which incorporate different aspects of cognitive, emotional and behavioural functioning following ABI (Lezak et al., 2012). In contrast, the Participation subscale showed a comparatively clearer factor structure than the other subscales. Although the items represent different areas of everyday participation, their relationships may be more readily explained by a common underlying construct of participation, resulting in a clearer empirical structure.

Integrated interpretation

The combined use of Rasch analysis and CTT provided complementary insights into the measurement properties of the Japanese MPAI-4. Rasch analysis identified item-level issues, including local dependency and threshold disordering, while CTT provided information on the empirical factor structure. Despite their different analytical perspectives, both approaches indicated that the full scale does not conform to a strictly unidimensional structure. At the subscale level, the structures appeared somewhat more distinct, but they also showed evidence of multidimensionality or heterogeneity. These findings indicate that the Japanese MPAI-4 should not be interpreted as a measure of a single homogeneous latent trait.

However, the findings do not necessarily indicate that the measure lacks value as an outcome measure. In the present study, the full scale demonstrated relatively high person separation, and most items showed generally acceptable fit with generally small DIF effects. The Participation subscale also showed comparatively clearer structural characteristics and better targeting than the Ability and Adjustment subscales. These findings indicate that the Japanese MPAI-4 shows some favourable measurement characteristics alongside the identified limitations. Thus, the scale may still provide useful information about functioning across the domains represented by the instrument.

The MPAI-4 was developed as a measure of global outcome following ABI while capturing multiple domains of functioning (Malec, 2024; Malec & Lezak, 2008). This broad conceptualization is consistent with the multidimensional nature of post-ABI outcomes, in which functioning, activity and participation may involve related but distinct dimensions (Hobart & Cano, 2009; Streiner &

Norman, 2008; World Health Organization, 2001). Previous Rasch analysis has described the MPAI-4 as representing a "broad but unitary core construct" (Kean et al., 2011). The present findings therefore broadly support the multidomain scope of the MPAI-4, while indicating that its empirical structure cannot be regarded as strictly unidimensional in the present sample.

An important implication of these findings concerns the interpretation of post hoc modifications. Although rescoring and item removal improved some statistical indicators, these changes were not consistently associated with improved measurement properties and could reduce the representation of conceptually important content. From a COSMIN perspective, structural validity findings should therefore be considered alongside the conceptual coverage of the construct, rather than used in isolation to justify modification of a measure (Mokkink et al., 2010; COSMIN, n.d.). The present findings therefore do not provide sufficient evidence to recommend a revised or shortened version of the Japanese MPAI-4. Rather, they indicate that the measure has limitations that should be recognized when interpreting its total and subscale scores.

Clinical implications

From a clinical perspective, the present findings support the potential clinical use of the Japanese version of the MPAI-4 for obtaining a broad profile of functioning following ABI, while recognizing the measurement limitations identified in this study. Given the multidimensionality observed in both the full scale and the subscales, the total score may be useful for obtaining an overall indication of functioning following ABI, while the subscale and item-level scores can provide more specific information about different aspects of functioning. Clinical interpretation should therefore consider the pattern of scores across the domains rather than relying on the total score alone.

In practice, the domain-specific information provided by the MPAI-4 may help clinicians identify relative strengths and difficulties across different aspects of functioning and inform rehabilitation planning. The Participation subscale may be particularly relevant to everyday functioning and community participation, including leisure activities, transportation and employment-related activities, which may not be fully captured by measures of basic functional independence. Such information may complement other outcome measures and contribute to goal setting and monitoring of changes in community-based functioning.

Within the Japanese context, where relatively few instruments comprehensively assess multiple domains of outcome following ABI, the Japanese MPAI-4 may therefore provide a useful framework for considering these domains within a single measure and for facilitating communication about rehabilitation goals and outcomes across professional disciplines.

Limitations and future directions

Several limitations should be acknowledged. First, the study excluded individuals unable to provide informed consent, which may have resulted in underrepresentation of individuals with severe cognitive or communicative impairments.

Second, the MPAI-4 manual recommends administration following a thorough evaluation of the individual and based on consensus among an interdisciplinary team. In the present study, assessments were conducted by research personnel, which may have introduced variability in ratings and influenced response category functioning. Future studies should examine multidisciplinary scoring procedures, including consensus among raters and involvement of family members where appropriate, to improve rating consistency.

Third, although the sample was considered adequate for a preliminary Rasch analysis based on commonly cited recommendations, a larger sample would have provided greater precision in parameter estimates and more stable estimates of item functioning. The sample size may also have limited the stability and interpretability of the EFA. Future studies should therefore use larger samples to further investigate the factor structure of the Japanese MPAI-4.

Fourth, the sample was not designed for normative standardization, and normative scores were not derived in the present study. The original MPAI-4 manual recommends the use of standardized scores derived from a US normative sample. However, the direct application of these norms to the Japanese version may not be appropriate, as standardized scores are dependent on the characteristics of the reference population and may be influenced by cultural and healthcare system differences. Future studies should therefore aim to establish Japanese normative data using larger and more representative samples (de Vet et al., 2011; Terwee et al., 2018).

Finally, participants were recruited from community-based rehabilitation settings in selected regions of Japan. Although these settings included different geographical and service contexts, the restricted geographic coverage may limit generalisability. Future research should prioritize larger multi-site studies and further examine potential cultural and contextual influences on item functioning, as well as the performance of the response categories in larger and more diverse samples.

Conclusion

This study examined the psychometric properties of the Japanese version of the MPAI-4 using both Rasch analysis and CTT approaches. The findings indicate that the instrument demonstrates acceptable reliability, although strict unidimensionality was not supported. The results suggest that the Japanese MPAI-4 is more appropriately interpreted as a multidomain measure rather than as a measure of a single homogeneous construct. The total score may

provide a broad indication of overall functioning, but should not be interpreted in isolation. The Ability, Adjustment and Participation domains may provide complementary information about different aspects of functioning.

Within Japan, where few instruments comprehensively assess multidomain outcomes following ABI, the present findings provide evidence supporting the potential clinical use of the MPAI-4. More broadly, the findings contribute to the international literature on the measurement properties of the MPAI-4 across different cultural and linguistic contexts.

Acknowledgements

The authors gratefully acknowledge the clinicians at all participating facilities for their assistance with data collection. The authors also thank all participants for their valuable contributions to this research.

AB benefits from support of NIHR HealthTech Research Centre for Brain Injury (NIHR205305) and NIHR Brain Injury Social Care Network Incubator (NIHR305277). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

ChatGPT (OpenAI) was used for English language editing during manuscript preparation. All AI-assisted outputs were reviewed and approved by the authors.

Author contributions

CRediT: **Nozomi Oyama:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Visualization, Writing – original draft, Writing – review & editing; **Fay Crawford:** Supervision, Writing – review & editing; **Shinichi Wada:** Resources, Writing – review & editing; **Kaoru Saito:** Resources, Writing – review & editing; **Andrew Bateman:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Data availability statement

The data that support the findings of this study are not publicly available due to privacy restrictions but are available from the corresponding author upon reasonable request.

ORCID

Nozomi Oyama http://orcid.org/0000-0003-4355-7709

References

Andrich, D., & Marais, I. (2019). *A course in Rasch measurement theory: Measuring in the educational, social and health sciences*. Springer.

Aryadoust, V., Tan, H. A. H., & Ng, L. Y. (2019). A scientometric review of Rasch measurement: The rise and progress of a specialty. *Frontiers in Psychology*, *10*, 0219. https://doi.org/10.3389/fpsyg.2019.02197

Ataman, R., Thomas, A., Roberge-Dao, J., McKerral, M., Auger, C., Wittich, W., Talla, P. K., Boychunck, A., & Ahmed, S. (2023). Measurement properties of the Mayo-Portland Adaptability Inventory (MPAI-4) and related measures: A systematic review. *Archives of Physical Medicine and Rehabilitation*, *104*(8), 1200–1313. https://doi.org/10.1016/j.apmr.2022.12.196

Baker, P. S., Bonder, E. V., & Allman, R. M. (2003). Measuring life-space mobility in community-dwelling older adults. *Journal of the American Geriatrics Society*, *51*(11), 1610–1614. https://doi.org/10.1046/j.1532-5415.2003.51512.x

Bateman, A., Teasdale, T. W., & Willmes, K. (2009). Assessing construct validity of the self-rating version of the European Brain Injury Questionnaire (EBIQ) using Rasch analysis. *Neuropsychological Rehabilitation*, *19*(6), 941–954. https://doi.org/10.1080/09602010903021170

Bohac, D. L., Malec, J. F., & Moessner, A. M. (1997). Factor analysis of the Mayo-Portland Adaptability Inventory: Structure and validity. *Brain Injury*, *11*(7), 469–482. https://doi.org/10.1080/bij.11.7.469.482

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305–314. https://doi.org/10.1037/0033-2909.110.2.305

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

Caracuel, A., Bateman, A., Teasdale, T. W., Verdejo-Garcia, A., & Perez-Garcia, M. (2011). Spanish, French, and British cross-cultural validation of the European Brain Injury Questionnaire. *The Journal of Head Trauma Rehabilitation*, *26*(6), 478–488. https://doi.org/10.1097/HTR.0b013e3181fc042c

Cattelani, R., Corsini, D., Posteraro, L., Agosti, M., & Saccavini, M. (2009). The Italian version of the Mayo-Portland Adaptability Inventory-4. A new measure of brain injury outcome. *European Journal of Physical and Rehabilitation Medicine*, *45*(4), 513–519.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical value for Yen's Q3: Identification of local dependency in the Rasch model using residual correlations. *Applied Psychological Measurement*, *41*(3), 178–194. https://doi.org/10.1177/0146621616677520

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of informative measurement. *Journal of Business Research*, *61*(12), 1250–1262. https://doi.org/10.1016/j.jbusres.2008.01.013

COSMIN. (no date). *COSMIN taxonomy of measurement properties*. Retreived October 31, 2025, from https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, *38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Evans, J. J. (2011). Positive psychology and brain injury rehabilitation. *Brain Impairment*, *12*(2), 117–127. https://doi.org/10.1375/brim.12.2.117

Evans, J. J. (2019). Neuropsychological rehabilitation: The origins of contemporary practice. *Higher Brain Function Research*, *39*(3), 273–282. https://doi.org/10.2496/hbfr.39.273

Fisher, W. P., Jr. (1992). Reliability statistics. *Rasch Measurement Transactions*, *6*(3), 238. Retrieved December 9, 2025, from https://www.rasch.org/rmt/rmt63i.htm

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Geyh, S., Nick, E., Stirnimann, D., Ehrat, S., Peter, C., & Lude, P. (2012). Self-efficacy and self-esteem as predictors of participation in spinal cord injury – An ICF-based study. *Spinal Cord*, *50*(9), 699–706. https://doi.org/10.1038/sc.2012.18

Goldman, L., Siddiqui, E. M., Khan, A., Jahan, S., Rehman, M. U., Mehan, S., Sharma, R., Budkin, S., Kumar, S. N., Sahu, A., Kumar, M., & Vaibhav, K. (2022). Understanding acquired brain injury: A review. *Biomedicines*, *10*(9), 2167. https://doi.org/10.3390/biomedicines10092167

Granger, C. V., & Hamilton, B. B. (1994). The uniform date system for medical rehabilitation report of first admissions for 1992. *American Journal of Physical & Rehabilitation*, *73*(1), 51–55. https://doi.org/10.1097/00002060-199473010-00012

Guerrette, M. C., & McKerral, M. (2020). Validation of the Mayo-Portland Adaptability Inventory-4 (MPAI-4) and reference norms in a French-Canadian population with traumatic brain injury receiving rehabilitation. *Disability and Rehabilitation*, *44*(18), 5250–5256. https://doi.org/10.1080/09638288.2021.1924882

Hall, J. A., Gunnery, S. D., & Schlegel, K. (2025). Gender and accuracy in decoding affect cues: A meta-analysis. *Journal of Intelligence*, *13*(3), 38. https://doi.org/10.3390/jintelligence13030038

Hobart, J., & Cano, S. (2009). Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. *Health Technology Assessment*, *13*(12), 1–177. https://doi.org/10.3310/hta13120

Holbrook, M., & Skilbeck, C. E. (1983). An activities index for use with stroke patients. *Age and Aging*, *12*(2), 166–170. https://doi.org/10.1093/ageing/12.2.166

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kean, J., Malec, J. F., Altman, I. M., & Swick, S. (2011). Rasch measurement analysis of the Mayo–Portland Adaptability Inventory (MPAI-4) in a community-based rehabilitation sample. *Journal of Neurotrauma*, *28*(5), 745–753. https://doi.org/10.1089/neu.2010.1573

Keith, R. A., Granger, C. V., Hamilton, B. B., & Sherwin, F. S. (1987). The functional independence measure: A new tool for rehabilitation. *Advances in Clinical Rehabilitation*, *1*, 6–18.

Lexell, J., Malec, J. F., & Jacobsson, L. J. (2012). Mapping the Mayo–Portland adaptability inventory to the international classification of functioning, disability and health. *Journal of Rehabilitation Medicine*, *44*(1), 65–72. https://doi.org/10.2340/16501977-0897

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.

Linacre, J. M. (1999). Understanding Rasch measurement: Estimation methods for Rasch measures. *Journal of Outcome Measurement*, *3*(4), 382–405.

Lombard, K., Nolan, C., & Heron, E. (2022). A scoping review of the use of Rasch analysis methodology to strengthen self-report occupational therapy mental health measures.

Occupational Therapy in Mental Health, *38*(2), 124–150. https://doi.org/10.1080/0164212X.2021.2005736

Mahoney, F. I., & Barthel, D. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, *14*, 56–61.

Malec, J. F. (2004). Comparability of Mayo-Portland Adaptability Inventory ratings by staff, significant others and people with acquired brain injury. *Brain Injury*, *18*(6), 563–575. https://doi.org/10.1080/02699050310001646134

Malec, J. F. (2024). *Manual for the Mayo-Portland Adaptability Inventory (MPAI-4) for adults, adolescents and children* (3rd ed.). Foundation to Advance Brain Rehabilitation.

Malec, J. F. (2025). Research letter: Rehabilitation planning using Rasch analysis of the Mayo-Portland Adaptability Inventory, 4th edition. *Journal of Head Trauma Rehabilitation*, *40*(5), 387–392. https://doi.org/10.1097/HTR.0000000000001042

Malec, J. F., & Kean, J. (2016). Post-inpatient brain injury rehabilitation outcomes: Report from the national outcome info database. *Journal of Neurotrauma*, *33*(14), 1371–1379. https://doi.org/10.1089/neu.2015.4080

Malec, J. F., Kean, J., Altman, I. M., & Swick, S. (2012). Mayo-Portland adaptability inventory: Comparing psychometrics in cerebrovascular accident to traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *93*(12), 2271–2275. https://doi.org/10.1016/j.apmr.2012.06.013

Malec, J. F., Kragness, M., Evans, R. W., Finlay, K. L., Kent, A., & Lezak, M. D. (2003). Further psychometric evaluation and revision of the Mayo-Portland Adaptability Inventory in a national sample. *The Journal of Head of Trauma Rehabilitation*, *18*(6), 479–492. https://doi.org/10.1097/00001199-200311000-00002

Malec, J. F., & Lezak, M. D. (2008). *The Mayo-Portland Adaptability Inventory (MPAI-4) for adults, children and adolescents. First edition April 2003, Revised with adaptations for paediatric version added January 2008*, Retrieved February 24, 2025, from http://www.tbims.org/combi/mpai

Malec, J. F., Logan, D. M., McGrath, C, Parrott, D, & Walters, G. J. (2025). Rasch analysis of the Mayo-Portland Adaptability Inventory for posthospital brain injury rehabilitation. *Journal of Head Trauma Rehabilitation*, *40*(1), 8–16. https://doi.org/10.1097/HTR.0000000000000935

Malec, J. F., & Moessner, A. M. (1997). Factor analysis of the Mayo–Portland Adaptability Inventory: Structure and validity. *Brain Injury*, *11*(6), 469–482. https://doi.org/10.1080/bij.11.7.469.482

Malec, J. F., Moessner, A. M., Kragness, M., & Lezak, M. D. (2000). Refining a measure of brain injury sequelae to predict postacute rehabilitation outcome: Rating scale analysis of the Mayo-Portland Adaptability Inventory (MPAI). *Journal of Head Trauma Rehabilitation*, *15*(1), 670–682. https://doi.org/10.1097/HTR.0000000000000935

Malec, J. F., & Thompson, J. M. (1994). Relationship of the Mayo-Portland Adaptability Inventory to functional outcome and cognitive performance measures. *Journal of Head Trauma Rehabilitation*, *9*(4), 1–15. https://doi.org/10.1097/00001199-199412000-00003

Mallinson, T., Kozlowski, A. J., Johnston, M. V., Weaver, J., Terhorst, L., Grampurohit, N., Juengst, S., Ehrlich-Jones, L., Heinemann, A. W., Melvin, J., Sood, P., & Van de Winckel, A. (2022). Rasch reporting guideline for rehabilitation research (RULER): the RULER statement. *Archives of Physical Medicine and Rehabilitation*, *103*(7), 1477–1486. https://doi.org/10.1016/j.apmr.2022.03.013

Mann, G., Troeung, L., Singh, K. A., Reddell, C., & Martini, A. (2023). Psychosocial functioning mediates change in motor and cognitive function throughout neurorehabilitation for adults with acquired brain injury (ABI-RESTaRT). *Neurological Sciences*, *44*(7), 1235–1248. https://doi.org/10.1007/s10072-023-06645-8

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stantford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Oyama, N., Aoki, S., Williams–Macklin, T., & Bateman, A. (2026). Post-discharge outcome measurement tools in occupational therapy for people with acquired brain injury in Japan: A scoping review. *Peer J*, *14*, e20765. https://doi.org/10.7717/peerj.20765

Oyama, N., Nakatake, J., Ando, R., & Arakawa, H. (2025a). Developing community support through public–private collaboration: A comprehensive group day programme for individuals with acquired brain injury. *Cognitive Rehabilitation*, *30*, 76–103. https://doi.org/10.50970/cogrehab.2025.004

Oyama, N., Sakoda, K., & Kawano, H. (2025b). Outcome measures in outpatient rehabilitation for patients with higher brain dysfunction: A comparison of the Mayo-Portland Adaptability Inventory-4 (MPAI-4) and the Functional Independence Measure (FIM) at a single facility. *Sogo Rehabilitation*, *54*(7), 721–728 (in Japanese). https://doi.org/10.11477/mf.038698220530070721

Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Value in Health*, *18*(1), 25–34. https://doi.org/10.1016/j.jval.2014.10.005

Pretz, C. R., Kean, J., Heinemann, A. W., Kozlowski, A. J., Bode, R. K., & Gebhardt, E. (2016). A multidimensional Rasch analysis of the functional independence measure based on the national institute on disability, independent living, and rehabilitation research traumatic brain injury model systems national database. *Journal of Neurotrauma*, *33*(14), 1358–1362. https://doi.org/10.1089/neu.2015.4138

Prieto, L., Alonso, J., & Lamarca, R. (2003). Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes*, *1*(1), 27. https://doi.org/10.1186/1477-7525-1-27

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press. (Original work published 1960)

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). Recognition of facial, auditory, and bodily emotions in older adults. *Journal of Gerontology: Psychological Sciences*, *63B*(2), 75–87. https://doi.org/10.1093/geronb/gbp072

Simblett, S. K., & Bateman, A. (2011). Dimensions of Dysexecutive Questionnaire (DEX) examined using Rasch analysis. *Neuropsychological Rehabilitation*, *21*(1), 1–25. https://doi.org/10.1080/09602011.2010.531216

Smith, E. V., Jr. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, *3*(2), 205–231.

Stolt, M., Kottorp, A., & Suhonen, R. (2022). The use and quality of reporting of Rasch analysis in nursing research: A methodological scoping review. *International Journal of Nursing Studies*, *132*, 104244. https://doi.org/10.1016/j.ijnurstu.2022.104244

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use* (4th ed.). Oxford University Press.

Sze, J. A., Goodkind, M. S., Gyurak, A., & Levenson, R. W. (2012). Aging and emotion recognition: Not just a losing matter. *Psychology and Aging*, *27*(4), 940–950. https://doi.org/10.1037/a0029367

Tennant, A., & 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 & Rheumatism*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108

Tennant, A., & Küçükdeveci, A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: Early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences*, *17*(4), 1209670. https://doi.org/10.3389/fresc.2023.1208670

Tennant, A., McKenna, S., & Hagell, P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. *Value in Health*, *7*(1), S22–S26. https://doi.org/10.1111/j.1524-4733.2004.7s106.x

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouterm, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion*, *28*(7), 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Thüss, N. S., den Hertog, H. M., van Bennekom, C. A. M., Spikman, J. M., & van der Naalt, J. (2025). The post-ABI effect – Long-term impact of aging on participation after acquired brain injury. *Journal of Neurology*, *272*(10), 675. https://doi.org/10.1007/s00415-025-13423-3

Turner-Stokes, L., Medvedev, O. N., & Siegert, R. J. (2019). Rasch analysis of the UK functional assessment measure in a sample of patients with traumatic brain injury from the UK national clinical database. *Journal of Rehabilitation Medicine*, *51*(8), 566–574. https://doi.org/10.2340/16501977-2580

Tyler, C. M., Dini, M. E., & Perrin, P. B. (2023). Group-based patterns of life satisfaction and functional independence over the 10 years after traumatic brain injury in older adults: A model systems study. *International Journal of Environmental Research and Public Health*, *20*(9), 5643. https://doi.org/10.3390/ijerph20095643

Van de Winckel, A., Kozlowski, A. J., Johnston, M. V., Weaver, J., Grampurohit, N., Terhorst, L., Juengst, S., Ehrlich-Jones, L., Heinemann, A. W., Melvin, J., Sood, P., & Mallinson, T. (2022). Reporting guideline for RULER: Rasch reporting guideline for rehabilitation research: Explanation and elaboration. *Archives of Physical Medicine and Rehabilitation*, *103*(7), 1487–1498. https://doi.org/10.1016/j.apmr.2022.03.019

Vaughan, L., Leng, X., Monte, M. L. A., Tindle, H. A., Cochrane, B. B., & Shumaker, S. A. (2016). Functional independence in late-life: Maintaining physical functioning in older adulthood predicts daily life function after age 80. *The Journals of Gerontology – Series A Biological Sciences and Medical Sciences*, *71*(Suppl 1), S79–S86. https://doi.org/10.1093/gerona/glv061

Wada, S. (2024). The usefulness of the Japanese version of the Mayo-Portland Adaptability Inventory-4 for acquired brain injuries. *Cognitive Rehabilitation*, *29*, 37–52. https://doi.org/10.50970/cogrehab.2024.003

Wada, S., Saito, K., Mochizuki, M., Iwaya, T., & Kawate, N. (2022a). *The Japanese version of the Mayo–Portland Adaptability Inventory (MPAI-4) for adults, children and adolescents* (in Japanese). Retrieved January 23, 2026, from https://www.tbims.org/combi/mpai/Japanese/MPAI-4(scoring%20sheet)%20Japanese.pdf

Wada, S., Saito, K., Mochizuki, M., Iwaya, T., & Kawate, N. (2022b). *Manual for the Japanese version of the Mayo-Portland Adaptability Inventory (MPAI-4) for adults, children and*

adolescents (in Japanese). Retrieved January 23, 2026, from http://www.tbims.org/combi/mpai/Japanese/MPAI-4(manual)%20Japanese.pdf

Wang, I., Wang, Y. C., Wu, T. Y., Chou, C. Y., & Hsieh, C. L. (2022). Rasch analysis of the Stroke Impact Scale-16. *American Journal of Occupational Therapy*, *76*(6), 7606205130. https://doi.org/10.5014/ajot.2022.049335

Wilson, B. A., Gracey, F., Evans, J. J., & Bateman, A. (2009). *Neuropsychological rehabilitation: Theory, models, therapy and outcomes*. Cambridge University Press.

World Health Organization. (2001). *International classification of functioning, disability and health (ICF)*. Retrieved March 28, 2025, from https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health

Yen, W. J. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, *8*(2), 125–145. https://doi.org/10.1177/014662168400800201