

The unidimensionality of the five Brain Injury Rehabilitation Trust Personality Questionnaires (BIRT-PQs) may be improved: preliminary evidence from classical psychometrics

Benedetta Basagni, Daniele Piscitelli, Antonio De Tanti, Leonardo Pellicciari, Lorella Algeri, Serena Caselli, Rita Formisano, Jessica Conforti, Anna Estraneo, Pasquale Moretta, Maria Grazia Gambini, Maria Grazia Inzaghi, Gianfranco Lamberti, Mauro Mancuso, Cristina Quinquinio, Matteo Sozzi, Laura Abbruzzese, Marina Zettin & Fabio La Porta

To cite this article: Benedetta Basagni, Daniele Piscitelli, Antonio De Tanti, Leonardo Pellicciari, Lorella Algeri, Serena Caselli, Rita Formisano, Jessica Conforti, Anna Estraneo, Pasquale Moretta, Maria Grazia Gambini, Maria Grazia Inzaghi, Gianfranco Lamberti, Mauro Mancuso, Cristina Quinquinio, Matteo Sozzi, Laura Abbruzzese, Marina Zettin & Fabio La Porta (2020) The unidimensionality of the five Brain Injury Rehabilitation Trust Personality Questionnaires (BIRT-PQs) may be improved: preliminary evidence from classical psychometrics, *Brain Injury*, 34:5, 673-684, DOI: [10.1080/02699052.2020.1723700](https://doi.org/10.1080/02699052.2020.1723700)

To link to this article: <https://doi.org/10.1080/02699052.2020.1723700>



Published online: 04 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 113



View related articles [↗](#)
















View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)



The unidimensionality of the five Brain Injury Rehabilitation Trust Personality Questionnaires (BIRT-PQs) may be improved: preliminary evidence from classical psychometrics

Benedetta Basagni ^a, Daniele Piscitelli ^b, Antonio De Tanti ^a, Leonardo Pellicciari ^c, Lorella Algeri^d, Serena Caselli ^{e,f}, Rita Formisano ^g, Jessica Conforti^a, Anna Estraneo ^h, Pasquale Morettaⁱ, Maria Grazia Gambini ^j, Maria Grazia Inzaghi^k, Gianfranco Lamberti ^l, Mauro Mancuso ^{m,n}, Cristina Quinquinio^o, Matteo Sozzi ^p, Laura Abbruzzese ⁿ, Marina Zettin^{q,r}, and Fabio La Porta ^s

^aCentro Cardinali Ferrari, Santo Stefano Riabilitazione, Fontanellato, Italy; ^bSchool of Physical and Occupational Therapy, McGill University, Montreal, Canada; ^cNeurorehabilitation Research Laboratory, IRCCS San Raffaele Pisana, Rome, Italy; ^dUOC Psicologia, ASST Papa Giovanni XXIII, Bergamo, Italy; ^eUnità Operativa di Medicina Riabilitativa, Dipartimento di Neuroscienze, Azienda Ospedaliero-Universitaria di Modena, Modena, Italy; ^fScuola di Dottorato in Sanità Pubblica, Università degli Studi di Milano-Bicocca, Milano, Italy; ^gUnità Post-Coma Ospedale di Riabilitazione Fondazione Santa Lucia, Rome, Italy; ^hIRCCS; Fondazione Don Carlo Gnocchi, Severe Brain Injury Department, Florence, Italy; ⁱMaugeri Scientific and Clinical Institutes, IRCCS Department of Neurorehabilitation, Institute of Telesse Terme, Telesse Terme, Italy; ^jDipartimento di riabilitazione, Ospedale Sacro cuore- Don Calabria, Negrar, Verona; ^kLaboratorio di Neuropsicologia, Istituto Clinico Quarenghi, Bergamo, Italy; ^lUnità Spinale e Medicina Riabilitativa Intensiva, AUSL Piacenza, Piacenza, Italy; ^mDepartment of Rehabilitation, National Health Service South-Est Tuscany, Grosseto, Italy; ⁿResearch Centre, Tuscany Rehabilitation Clinic, Monteverchi, Italy; ^oSanto Stefano Riabilitazione, Istituto di Riabilitazione Santo Stefano, Porto Potenza Picena, Italy; ^pDipartimento di Scienze Neuroriabilitative, Casa di Cura del Policlinico, Milan, Italy; ^qCentro Puzzle, Turin, Italy; ^rDipartimento di Psicologia, Università di Torino, Turin, Italy; ^sNeurorehabilitation Unit, IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy

ABSTRACT

Objective: To assess the internal construct validity (ICV) of the five Brain Injury Rehabilitation Trust Personality Questionnaires (BIRT-PQ) with Classical Test Theory methods.

Methods: Multicenter cross-sectional study involving 11 Italian rehabilitation centers. BIRT-PQs were administered to patients with severe Acquired Brain Injury and their respective caregivers. ICV was assessed by the mean of an internal consistency analysis (ICA) and a Confirmatory Factor Analysis (CFA).

Results: Data from 154 patients and their respective caregivers were pooled, giving a total sample of 308 subjects. Despite good overall values (alphas ranging from 0.811 to 0.937), the ICA revealed that several items within each scale did not contribute as expected to the total score. This result was confirmed by the CFA, which showed the misfit of the data to a unidimensional model (RMSEA ranging from 0.077 to 0.097). However, after accounting for local dependency found within the data, fitness to a unidimensional model improved significantly (RMSEA ranging from 0.050 to 0.062).

Conclusion: Despite some limitations, our analyses demonstrated the lack of ICV for the BIRT-PQ total scores. It is envisaged that a more comprehensive ICV analysis will be performed with Rasch analysis, aiming to improve both the measurement properties and the administrative burden of each BIRT-PQ.

ARTICLE HISTORY

Received 31 May 2019
Revised 2 October 2019
Accepted 11 December 2019

KEYWORDS

Brain Injuries; Behavioral Symptoms; Personality Assessment; Outcome Assessment (Health care); Psychometrics; Rehabilitation

Background

Neurobehavioral changes, often defined as ‘personality changes’ (1), are frequent long-term sequelae of traumatic brain injury (TBI), occurring from 30% to 60% of cases (2,3). Frequently reported changes include reduced motivation, emotional regulation difficulties (explosive anger, irritability, aggression, and labile mood), difficulties maintaining social relationships, disinhibition, and impulsivity (4–7). Some authors have termed these changes as non-cognitive neurobehavioral sequelae (8), or as neurobehavioral disability (5). The latter plays a central role in determining the overall outcome of patients with TBI, even more than cognitive impairments (9–12). Furthermore, it has been proven that these non-cognitive neurobehavioral impairments are good predictors of caregivers’ burden, patients’ quality of life in terms of independent living, intimate relationships’

development, employment’s acquisition or maintenance, independently of the severity of the brain injury (13–15). Neurobehavioral disability has also been described in non-traumatic acquired brain injuries (nt-ABI). Indeed, several behavioral changes have been described among these latter patients (16), including disorders of social perceptiveness, self-control and/or emotional regulation, decreased ability to learn from social experience, difficulties with impulse control, development of sociopathic or borderline traits, loss of self-sense, childish behavior, disinhibited behavior, and poor social judgment (17). Thus, non-cognitive behavioral impairments may be considered a landmark of ABI regardless of its etiology.

Given the clinical relevance of neurobehavioral disability, an adequate assessment of non-cognitive neurobehavioral impairments plays a central role in rehabilitation programs for patients with ABI, and such aspects always need to be

carefully investigated. Nevertheless, despite the recognition of its importance, few valid and reliable measures are available for the assessment of non-cognitive neurobehavioral sequelae (18,19). Indeed, at least three criteria seem to be necessary for the effectiveness of such kinds of instruments as they should: (a) be specific for personality changes following ABI and not created for other types of pathologies, (b) cover a broad domain of functioning, and (c) present a robust internal validity. Several papers underline the limits of the measures commonly used in neurorehabilitation (20).

In a series of studies, Cattran and colleagues (8,21–23) introduced a specific set of scales designed to assess non-cognitive neurobehavioral impairments for patients with ABI, recently translated and cross-culturally adapted for the Italian population (24). The five Brain Injury Rehabilitation Trust personality questionnaires (henceforth BIRT-PQs) consist of five separated scales (totaling 150 items) assessing the following areas: motivation, emotional regulation, social cognition, disinhibition, and impulsivity. As each questionnaire is available in a patient's self-administered and caregiver-rated forms, the comparison of these two parallel forms may offer an indirect measure of the problem's awareness of a patient. Selectivity for the ABI population, adequate comprehensiveness of the neuro-behavioral impairments spectrum, and the availability of two different points of view (self and relative) of the same problem are clinical advantages that make the BIRT-PQs attractive for clinicians.

In general terms, the measurement of a latent construct by a scale or questionnaire (i.e. the personality changes as measured by the five BIRT-PQs) is entirely based on the assumption of unidimensionality, by which all items in a scale should contribute to measuring a single underlying construct. Unidimensionality contributes to the so-called internal construct validity (ICV) of a scale (25). Only upon demonstration of adequate ICV, it would be legitimate to sum together the item scores to generate the total score of the scale.

As the ICV of the five BIRT-PQs has not been assessed so far, this study aims to evaluate the internal consistency, the unidimensionality, and the respondent burden of the BIRT-PQs under the Classical Test Theory (CTT) framework.

Methods

Study design and setting

This was a multicenter cross-sectional study carried out from April 2016 to December 2017 in eleven Italian rehabilitation centers¹ with expertise in care and management of patients with ABI.

Subjects

Participants were patients and their respective caregivers, who were consecutively enrolled within the participating centers. The following inclusion criteria were employed for enrollment: a diagnosis of severe ABI (s-ABI), characterized at the onset by lack of consciousness, as defined by an initial Glasgow Coma Scale ≤ 8 (26), lasting more than 24 h; age between 18 and 70 years; a Level of Cognitive Functioning

(LCF) (27) score ≥ 7 at the time of enrollment; a premorbid Modified Barthel Index (MBI) (28) score equal to 100; attending the outpatient clinic after 6 months but not later than 6 years since inpatient discharge from the same center; availability of a caregiver willing to participate. Exclusion criteria were: aphasia severe enough to impair the ability to read or to understand the spoken language; previous history of any neurological and psychiatric disorders; being domiciled in long-term care, nursing home, or other cared residential facility; inability to give informed written consent.

Local Ethical Committees approved the study of the participating centers, which was carried out under the principles outlined in the Helsinki declaration (29). Participants and their respective caregivers gave their written informed consent to take part in the study.

Outcome measures

The primary outcome measures were the BIRT-PQs (21–23), which include five independent scales assessing several dimensions of personality that might be altered after an ABI. The scales are motivation (BMQ, 34 items), regulation of emotions (BREQ, 32 items), social cognition (BSCQ, 28 items), disinhibition (BDQ, 24 items), and impulsivity (BIQ, 32 items). Two versions of the questionnaire are available: a patient version, where the respondent is asked to answer questions regarding his/her aspects of personality, and a caregiver-rated version, which is identical to the former, but it is self-administered by a caregiver, and the questions regards the aspects of personality of his/her cared one affected by ABI. Each item is scored using a Likert format with 4-response options ranging from 'always' to 'never' (item score range: 1 to 4). Most of the items are worded as such that the higher the score, the more frequent is the related aspect of personality disturbance. However, to reduce acquiescence, some items are reversed in meaning and scoring, so that the higher the score, the less frequent is the personality change. The total score of each scale is obtained adding up the scores of the pertaining items, so that the higher the total score, the greater the degree of personality disturbance. As each scale has a different number of item, the total score varies by questionnaire: BMQ ranges from 34 to 136 points, BREQ and BIQ from 32 to 128, BSCQ from 28 to 112, and BDQ from 24 to 96. Within this study, the Italian version of the BIRT-PQs was administered (24).

Also, the following other scales and questionnaires were administered for sample description purposes:

- Disability Rating Scale (DRS) (30). This scale tracks the recovery of functioning after severe head trauma from 'coma to community'. It includes physiologic measures from the Glasgow Coma Scale (GCS), measures of cognitive ability to perform self-care activities, measures of dependence on others, and measures of employability.
- Satisfaction Profile (SAT-P). It assesses the patient's subjective satisfaction within different aspects of functioning (31), including psychological functioning (10 items), physical functioning (9 items), work (5 items), basic needs and free time (5 items), and social

functioning (3 items). Respondents are asked to mark a sign over a 10 cm horizontal line with the extremes semantically defined (left completely unsatisfied, right completely satisfied). As the total score of each sub-scale is computed by calculating the means of each item score, it ranges from 0 to 100 (indicating, respectively, the worst and the best satisfaction).

- Frontal Behavioral Inventory (FBI). It quantifies the personality and behavior changes in persons with dementia from the perspective of caregivers (32). Within the questionnaire, both negative (12 items) and positive behaviors (12 items) are assessed. The total scores of both subscales range from 0 to 36, where higher scores are indicative of a higher level of personality and behavioral change. In this study, the Italian version was employed (33).
- Caregiver Burden Inventory (CBI). It assesses the burden of care for caregivers of dementia patients (34), by exploring five main dimensions: time-dependence, developmental, physical, social, and emotional burden. Each dimension is assessed by five items, except for the physical burden, which has only four items. Each item is scored between 0 (not at all descriptive) and 4 (very descriptive), where higher scores indicate a more significant caregiver burden. Therefore, all total scores range from 0 to 20, whereas the total score for physical burden ranges from 0 to 16.

Procedures

The demographic and clinical characteristics of the patients and their respective caregivers were collected. All subjects (patients and caregivers) were requested to complete the five BIRT-PQs independently. Also, patients were asked to complete the SAT-P, whereas caregivers were required to complete the FBI and CBI. All administration procedures were shared and uniformed between the participating centers.

Statistical analysis

Descriptive sample and scale statistics

Descriptive statistics were used to report all the collected variables. Mean \pm standard deviation (SD), median with 10th and 90th percentile, and absolute frequency with percentage, were calculated for interval, ordinal, and nominal variables, respectively (35). Ceiling and floor effects were defined as the occurrence of the highest or the lowest possible score for each scale, respectively, for more than 15% of the subjects in the sample (36).

Internal consistency analysis

The internal consistency of each five BIRT-PQs was analyzed on a pooled sample including data from the patients and their respective caregivers. Particularly, it was assessed by calculating:

- the Cronbach's alpha (37) for each BIRT-PQs total score, where values ranging between 0.70 and 0.95 are recommended (38);
- the average inter-item correlations, expressed by the Spearman's correlation coefficient (39), is the mean of the inter-item correlations between each pair of items; values ≥ 0.2 were considered acceptable (40);
- Cronbach's Alpha with deleted variable, where the alpha was calculated after removing each item in turn; values below the total Cronbach's alpha are expected (41);
- the item-to-total correlations, which are the non-parametric correlations (based on Spearman's ρ) between each item and its restscores (i.e. the total score minus the item score); values ≥ 0.40 were considered satisfactory (40).

Analysis of unidimensionality: confirmatory factor analysis

A confirmatory factor analysis (CFA), here based on polychoric correlations to fit ordinal data, was undertaken to assess the unidimensionality of the BIRT-PQs. CFA has been described in detail elsewhere (35,42,43).

Within this analysis, the assessment of model fit was performed using the following indicators (44):

- The model chi-square (χ^2), an overall indicator of model fit, which is a measure of the discrepancy between the covariance matrices of the model and of the sample (42). For a perfectly fitting model, the χ^2 probability values are not significant, although these statistics are sensitive to a larger sample size.
- The Root Mean Square Error of Approximation (RMSEA), which is a measure of the discrepancy between the covariance matrix predicted by the model and the population covariance matrix, if it were available. In other words, the RMSEA is an estimate of how well the model fits the observed data. It is considered one of the most informative index of model fit (42) and, unlike the χ^2 , it is less influenced by sample size. Values ≤ 0.06 are considered indicative of a good fit.
- The Standardized Root Mean Square Residual (SRMR), which is the average value across all residual values derived from the comparison between the predicted and the observed variance-covariance matrix (42). In practice, SRMR is the amount of error by which the model explains the correlations (42). Values ≤ 0.08 are typical of well-fitting models (45).
- The Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI). They both measure the proportionate improvement in the model fit by comparing the hypothesized model and the null model. However, while the CFI is normed (i.e. its values range between 0 and 1), the TLI is not (i.e. its values can extend beyond the range [0, 1]). Furthermore, the TLI includes a penalty for overly complex models. For both indicators, values ≥ 0.95 are considered indicative of a well-fitting model (45).

Table 1. Demographic characteristics of the sample.

Variable	Patients (n = 154)	Caregivers (n = 154)	Variable	Patients (n = 154)
Age in years	41.9 ± 14.4*	52.1 ± 12.3*	Housing conditions	
	41.5 [23.1, 62.9]†	53.0 [35.7, 66.4]†	Own family (with spouse, partner and/or children)	50.0%
Gender			Family of origin (parents and/or brothers/sisters)	33.8%
Males	68.8%	31.2%	Other cohabitation (i.e. friends)	1.9%
Females	31.2%	57.8%	Sheltered house with educators	1.3%
Missing	0.0%	11.0%	Lives alone	7.1%
			Missing	5.9%
Education			Previous occupation	
Primary	2.6%	5.8%	Retired	5.2%
Secondary	31.2%	30.5%	Full time worker	65.6%
High	48.7%	37.7%	Part time worker	3.9%
Degree	16.9%	13.6%	Seasonal worker	0.6%
Post-degree	0.0%	1.3%	Housewife	3.2%
Missing	0.6%	11.1%	Student	12.3%
			Unemployed (none of the previous categories)	1.3%
Marital Status			Unknown	2.6%
Unmarried	47.4%	-	Missing	5.3%
Married	39.6%	-	Present occupation:	
Separated	2.6%	-	Retired	20.1%
Divorced	3.2%	-	Full time worker	20.8%
Widowed	1.9%	-	Part time worker	14.3%
Missing	5.3%	-	Seasonal worker	0.6%
			Housewife	1.9%
Caregiver			Student	5.8%
Parents	38.3%	-	Unemployed (none of the previous categories)	29.2%
Brother/sister	7.8%	-	Unknown	1.3%
			Missing	6.0%
Spouse/partner	40.9%	-	Occupational capacity:	
			Not able to resume any occupation in any way	14.3%
Son/daughter	5.2%	-	Possible occupation with close supervision or non-competitive work	28.6%
Other relative (not include the previous ones)	0.6%	-	Possible modified occupation independently or selected competitive jobs	31.2%
Friend	0.6%	-	No apparent participation restrictions in employment/activity	19.5%
Other (i.e. educator)	1.3%	-	Missing	6.4%
Missing	5.3%	-		

Note: All data are reported as percentage, except *, which indicates mean ± SD, and †, which indicates median [10P, 90P]

The first analytical step involved testing for multigroup invariance by respondents (patients vs. caregivers), following the approach outlined in Byrne (46). Particularly, an initial separate model was established for each group. Should the initial baseline model fail to fit, we would assess the modification indices (MI) on pairs of items (42,47), which are indicators of model misspecification. Particularly, high MIs suggest the presence of residual covariance between items (42). In other words, MIs indicate local dependency between items, where the response to one of the items within the pair is influenced by the response to the other item (48–50). In these instances, it is possible to re-specify the model after accounting for local dependency by allowing correlation of the error terms of the items in the pair (42,47). Following this, the model fit is reassessed again. Once the two group-specific final models were established (one baseline model for patients and another one for caregivers), we would build a further multigroup baseline model called the ‘configural model.’ The latter model allowed to test

simultaneously for both groups the same parameters that were estimated separately in the group-specific baseline models without imposing any invariance constraint. Following this, increasing levels of constraints would be imposed on the parameters of the configural model (involving, in order, factor loadings, observed variables intercepts, residual variances, and, finally, factor variances) to test a series of increasingly restrictive hypotheses about multi-group invariance. Should the latter approach fail, we would proceed to pool the data from patients and caregivers to collect evidence about the ICV of the BIRT-PQs using MI extensively to adjust for any model misspecifications.

Analysis of respondent burden

The respondent burden was estimated separately for each BIRT-PQs by recording the time needed to administer each questionnaire (51). Individual administration times (for both patients and caregivers) were recorded separately.

Statistical, software, and sample size notes

All descriptive statistics and internal consistency statistics were calculated using SPSS software (version 21 for Windows; SPSS Inc., Chicago, IL; 2004). The non-parametric CFA was performed with the Mplus software (version 6.0. Muthen & Muthen, Los Angeles, CA; 1998–2010; www.statmodel.com). For CFA, we estimated that a sample size of 154 cases would guarantee a ratio subjects-to-parameters to be calculated for the initial analyses between 6.4:1 and 4.5:1, which are values somewhat below the recommended ratio of 10:1 (52). On the other hand, for a 308 subjects sample, the same proportion would be between 12.8:1 and 9.1:1, which would be adequate for these analyses (52).

Results

Sample characteristics

One-hundred and fifty-four patients and their respective caregivers were enrolled in this study. Thus, a pooled sample of 308 subjects was available for analysis. Detailed demographic and clinical characteristics of the sample are reported in Tables 1 and 2, respectively.

The enrolled patients were, on average, 42 years old, and more than two-thirds of them (68.8%) were men. Almost 50% of them had been awarded a high school diploma, whereas only 16.9% of them had achieved a degree. Nearly half of them were unmarried (47.4%), whereas about two-fifths of them were married. For about 40% of them, the primary caregiver was the spouse/partner, whereas slightly less than two-fifth of them were cared for by their parents. One-half of the patients lived with their own family (50%), whereas about one-third of them lived with the family of origin. About two-thirds of the patients were full-time workers before the brain injury, whereas only one in five of them had returned to full-time work at the time of assessment, although only 42.9% of them had a severe or complete restriction of their occupational capacity. On the other hand, the caregivers were predominantly women (57.8%), they were on average 10 years older (52.1 years), and they had achieved, on average, a lower educational level (37.7% high school diploma), although 11% of the education data were missing.

Almost two-thirds of the patients had suffered a TBI (61% of cases), whereas cerebral hemorrhage (either intracerebral or subarachnoid) accounted for about one-fifth of the etiologies. The brain injury at onset was severe (median GCS: = 5 points) with a prolonged disorder of consciousness (almost 3 weeks). The median hospital stay was 4.6 months, although an average length of stay of 6.3 months suggests that some hospital stays had been unusually prolonged. Despite the severity of the initial ABI, patients at discharge had achieved good cognitive and motor functioning (median LCF and MBI were, respectively, 8 and 100), although a Disability Rating Score of 3 suggested a 'partial' level of disability (30).

On average, the assessments took place more than 2 years after hospital discharge (range 9–61 months). No floor or ceiling effect for each BIRT-PQs was detected. Missing item data for each BIRT-PQs were minimal, ranging from 0.05% for BSCQ to 0.09% for BMQ and BIQ; no systematic missing item data pattern could be identified across each scale. The median scores of the

BIRT-PQs were reported separately for patients and caregivers in Table 2. The score distributions of all five scales were left-skewed, with all median scores falling below the median of their respective score distributions. Notably, the lowest levels of personality change were reported for the BREQ and the BSCQ, with median scores lying for both questionnaires in the first quartile of the score distribution for both patients (BREQ: 50; BSCQ: 46) and caregivers (BREQ: 51; BSCQ: 45). Median higher scores lying in the second quartile of the score distribution were reported instead for the BMQ, the BDQ and the BIQ, both by patients (BMQ: 63; BDQ: 42; BIQ: 56) and by caregivers (BMQ: 68; BDQ: 43; BIQ: 59). Caregivers rated, on average, a higher degree of personality change than the patients for the BMQ (5 points), BIQ (3 points), BREQ (1 point), and BDQ (1 point), whereas patients self-reported a higher degree of disturbance than their caregivers within the social cognition scale (1 point).

The levels of satisfaction with life reported by the patients, measured by the SAT-P (Table 2), were in the 3rd quartile for all domains, with the lowest level of satisfaction regarding the physical functioning (median SAT-P = 53.2) and the highest level reported within the social functioning (median SAT-P = 71.0) domains. The lower limit of the score distribution was in the 2nd quartile for the basic needs and free time (42.8), the psychological functioning (40.9), the social functioning (40.0), and the physical functioning domains (29.0), whereas it was in the first quartile only for the work domain (21.0).

According to their caregivers, the total scores of the negative and positive behavior subscales, as measured by the FBI, lied on average in the 2nd and the 1st quartile, respectively (negative behaviors: 10; positive behaviors: 4). The FBI total score, indicating the overall amount of personality disturbance, laid in the first quartile of the score distribution (14). All the total scores of the CBI subscales lied in the first quartile of the score distributions, near the upper limit of the quartile for the time dependence (4), social (4) and physical (3.5) subscales, whereas for the remaining two scales, the total scores lied toward the quartile's lower limit (developmental = 2; emotional = 1). The overall caregiver burden, measured by the CBI total score, laid near the upper limit of the 1st quartile of the score distribution (19).

Internal consistency analysis

All Cronbach's α values for the patients and caregivers data pooled together were >0.9 for all subscales (Table 3), except for BDQ ($\alpha = 0.811$). Similar results were yielded by the analysis of the average inter-item correlations, which were ≥ 0.2 , except for BDQ (0.177).

The item-by-item internal consistency analysis showed that all items had an item-to-total correlation coefficient higher than 0.40 for all item in each subscale (Table 3), except for five items in BMQ (BMQ03, BMQ06, BMQ13, BMQ18, BMQ32), three items in BREQ (BREQ06, BREQ08, BREQ28), two items in BSCQ (BSCQ19, BSCQ26), ten items in BDQ (BDQ07–BDQ9, BDQ11–BDQ14, BDQ18, BDQ20, and BDQ21), and seven items in BIQ (BIQ01, BIQ02, BIQ04, BIQ18, BIQ20, BIQ22, and BIQ23). Across all five scales, all correlations were positive, except for BDQ and BIQ, which both had two items with negative correlation coefficients (≤ -0.2).

The alpha with item deleted suggested that the deletion of most items with item-to-total correlation coefficients <0.4 would lead to

Table 2. Clinical characteristics of the sample.

Variable	Patients (n=154)	Variable	Patients (n=154)	Caregivers (n=154)
Aetiology*				
Traumatic brain injury	61.0%	BIRT-PQ^d		
Intra-cerebral hemorrhage	13.0%	Motivation (range 34-136)	63.0 [49.0, 84.0]	68.00 [48.0, 98.1]
Subarachnoid hemorrhage	9.7%	Emotional regulation (range 32-138)	50.0 [38.0, 72.8]	51.00 [40.0, 78.0]
Ischemic stroke	6.5%	Social cognition (range 28-112)	46.0 [38.0, 62.8]	45.00 [34.0, 69.0]
Anoxia	4.5%	Disinhibition (range 24-96)	42.0 [34.0, 52.7]	43.00 [34.0, 57.0]
Other	5.3%	Impulsivity (range 32-128)	56.0 [45.0, 73.8]	59.00 [44.0, 78.2]
Glasgow Coma Scale at onset (range 3-15) ^a	5.0 [3.0, 8.0]	SAT-P scale^e		
Duration of disorder of consciousness (days)		Psychological Functioning (range 0-100)	63.4 [40.9, 90.2]	-
Mean ± SD	20.6 ± 16.3	Physical Functioning (range 0-100)	53.2 [29.0, 84.0]	-
Median [10-90P]	16.5 [5.0, 40.0]	Work (range 0-100)	64.1 [21.0, 90.7]	-
Length of hospital stay as inpatient (months)		Basic needs and free time (range 0-100)	68.0 [42.8, 92.6]	-
Mean ± SD	6.3 ± 5.7	Social Functioning (range 0-100)	71.0 [40.0, 98.3]	-
Median [10-90P]	4.6 [1.7, 13.1]	FBI scale^f		
Level of Cognitive Functioning on discharge (range 1-10) ^b	8.0 [7.0, 9.0]	Negative Behaviour (range 0-36)	-	10.0 [1.0, 22.0]
Barthel Index on discharge (range 0-100) ^c	100.0 [82.4, 100.0]	Disinhibition (range 0-36)	-	4.0 [0.0, 17.0]
Disability Rating Scale on discharge (range 1-10) ^d	3.0 [0.0, 6.9]	Total Score (range 0-72)	-	14.0 [1.6, 35.0]
Interval between hospital discharge and enrolment		CBI scale^g		
Mean ± SD	20.5 ± 18.8	Time dependence (range 0-20)	-	4.0 [0.0, 15.0]
Median [10-90P]	12.2 [6.1, 55.4]	Social (range 0-20)	-	4.0 [0.0, 14.2]
Time since lesion (months)		Physical (range 0-16)	-	3.0 [0.0, 11.0]
Mean ± SD	26.3 ± 19.6	Developmental (0-20)	-	2.0 [0.0, 11.0]
Median [10-90P]	19.0 [9.0, 61.1]	Emotional (range 0-20)	-	1.0 [0.0, 8.0]
		Total Score (range 0-96)	-	19.0 [1.6, 52.4]

Abbreviations: BIRT-PQ, Brain Injury Rehabilitation Trust; SD, standard deviation; CBI, Caregiver Burden Inventory; FBI, Frontal Behavioral Inventory; N, number; SAT-P, Satisfaction Profile

Note: All statistics are reported as median [10P, 90P], unless specified otherwise (*indicates percentage).

^ahighest values indicate a higher level of consciousness.

^bhighest values indicate a higher level of disability.

^chighest values indicate a higher level of cognitive function.

^dlower values indicate a higher level of disability.

^ehigher values indicate a greater level of behavioral disturbance.

^fhighest values indicate a higher level of subjective satisfaction.

^ghighest values indicate a higher level of personality and behaviour changes.

^hhighest values indicate a higher level of charge for caregivers.

an increase of the Cronbach's alpha for the remaining item set (Table 3).

Confirmatory factor analysis

The summary results of the CFA are reported in Table 4. The initial CFA failed to support the ICV of all five BIRT-PQs, both for the patient and the caregiver versions (RMSEA ranging from 0.073 to 0.097, and from 0.078 to 0.095, respectively). Except for the BSCQ, the RMSEA values were lower for the patient version than for the caregiver version. High values of MIs were detected for all scales, both for patients and caregivers. After allowing correlation of the error terms of the locally dependent pairs of items within a scale, it was possible to achieve a better fit to a one-factor model, although the RMSEAs were lower than the recommended cutoff of 0.06 only for the patient's versions of BREQ and BDQ. For the patient's version of BIQ, the final model was un-identified. Furthermore, it was not possible to proceed further with the analysis of multigroup invariance because for all BIRT-PQs (both patient and caregiver versions), the configural models were un-identified too. As a consequence, it was decided to pool the patient and caregiver data together and to proceed with the analysis.

The pooled CFA base analyses failed to support the ICV of all five BIRT-PQs (RMSEA ranging from 0.077 to 0.097; SRMR ranging from 0.081 to 0.105). High values of MIs were detected for all scales, indicating local dependence

between a substantial number of item pairs (23, 17, 26, 13, and 31 locally dependent pairs of items for BMQ, BREQ, BSCQ, BDQ, and BIQ, respectively). After accounting for all locally dependent pairs by allowing correlations of their error terms, fit to one-factor model improved for all scales, with RMSEA values at or below (for BREQ, BDQ, and BIQ), and just above (for BMQ and BSCQ) the cutoff value of 0.06, and SRMR ≤ 0.08 for all scales. CFI and TLI were adequate only for BREQ (0.962 and 0.958, respectively).

Analysis of respondent burden

The average administration time for each BIRT-PQ questionnaire ranged from 5.7 to 7.5 min for the patient version and 4.1 to 6.2 min for the caregiver version (Table 5). The average overall administration times of the five BIRT-PQs were 32 min (range: 23.5–40.5) and 23.8 min (range: 16.4–31.2) for the patient and caregiver versions, respectively.

Discussion

Within this paper, we undertook a CTT analysis of the internal construct validity of the five BIRT-PQs, here based on internal consistency and CFAs. Our results suggested that the total scores of the five BIRT-PQs lacked adequate internal construct validity. Particularly, despite good overall internal consistency values for each scale, several items within each

Table 3. Results of internal consistency analysis.

BMQ	Alpha		TotCorr		Alpha		TotCorr		Alpha		TotCorr		Alpha		TotCorr	
	Alpha	TotCorr	BREQ	BREQ	Alpha	TotCorr	BSCQ	BSCQ	Alpha	TotCorr	BDQ	BDQ	Alpha	TotCorr	BIQ	BIQ
BMQ01	0.930	0.284	BREQ	0.937	0.329	BSCQ	0.909	0.267	0.811	BDQ	0.177	0.904	0.904	0.239	BIQ01	0.239
BMQ02	0.929	0.415	BREQ01	Alpha-R	0.719	BSCQ01	Alpha-R	0.572	0.798	BDQ01	TotCorr-R	Alpha-R	Alpha-R	TotCorr-R	BIQ02	TotCorr-R
BMQ03	0.928	0.534	BREQ02	0.933	0.581	BSCQ02	0.904	0.566	0.799	BDQ02	0.546	0.907	0.907	0.546	BIQ03	0.165
BMQ04	0.930	0.335	BREQ03	0.935	0.691	BSCQ03	0.905	0.535	0.800	BDQ03	0.546	0.898	0.898	0.546	BIQ04	0.391
BMQ05	0.928	0.537	BREQ04	0.934	0.581	BSCQ04	0.906	0.472	0.795	BDQ04	0.612	0.902	0.902	0.612	BIQ05	0.391
BMQ06	0.928	0.557	BREQ05	0.933	0.691	BSCQ05	0.904	0.573	0.800	BDQ05	0.567	0.898	0.898	0.567	BIQ06	0.624
BMQ07	0.930	0.385	BREQ06	0.939	0.232	BSCQ06	0.906	0.479	0.802	BDQ06	0.459	0.899	0.899	0.459	BIQ07	0.578
BMQ08	0.929	0.489	BREQ07	0.935	0.580	BSCQ07	0.906	0.479	0.811	BDQ07	0.228	0.899	0.899	0.228	BIQ08	0.566
BMQ09	0.928	0.552	BREQ08	0.938	0.280	BSCQ08	0.907	0.427	0.807	BDQ08	0.293	0.900	0.900	0.293	BIQ09	0.532
BMQ10	0.927	0.606	BREQ09	0.934	0.652	BSCQ09	0.905	0.520	0.807	BDQ09	0.297	0.900	0.900	0.297	BIQ10	0.492
BMQ11	0.927	0.632	BREQ10	0.935	0.629	BSCQ10	0.905	0.550	0.794	BDQ10	0.641	0.900	0.900	0.641	BIQ11	0.563
BMQ12	0.927	0.597	BREQ11	0.936	0.519	BSCQ11	0.906	0.496	0.842	BDQ11	-0.257	0.900	0.900	-0.257	BIQ12	0.483
BMQ13	0.931	0.271	BREQ12	0.937	0.428	BSCQ12	0.907	0.431	0.816	BDQ12	0.167	0.901	0.901	0.167	BIQ13	0.445
BMQ14	0.928	0.510	BREQ13	0.936	0.478	BSCQ13	0.906	0.475	0.807	BDQ13	0.306	0.897	0.897	0.306	BIQ14	0.692
BMQ15	0.928	0.557	BREQ14	0.936	0.421	BSCQ14	0.905	0.521	0.809	BDQ14	0.224	0.901	0.901	0.224	BIQ15	0.440
BMQ16	0.929	0.446	BREQ15	0.935	0.589	BSCQ15	0.907	0.414	0.798	BDQ15	0.486	0.899	0.899	0.486	BIQ16	0.548
BMQ17	0.928	0.511	BREQ16	0.934	0.695	BSCQ16	0.903	0.629	0.795	BDQ16	0.534	0.901	0.901	0.534	BIQ17	0.435
BMQ18	0.930	0.338	BREQ17	0.935	0.594	BSCQ17	0.903	0.649	0.801	BDQ17	0.417	0.902	0.902	0.417	BIQ18	0.420
BMQ19	0.927	0.651	BREQ18	0.934	0.662	BSCQ18	0.904	0.625	0.806	BDQ18	0.313	0.914	0.914	0.313	BIQ19	-0.271
BMQ20	0.927	0.600	BREQ19	0.935	0.665	BSCQ19	0.913	0.144	0.800	BDQ19	0.475	0.901	0.901	0.475	BIQ20	0.349
BMQ21	0.929	0.503	BREQ20	0.936	0.553	BSCQ20	0.905	0.513	0.807	BDQ20	0.302	0.903	0.903	0.302	BIQ21	0.439
BMQ22	0.928	0.573	BREQ21	0.936	0.523	BSCQ21	0.905	0.514	0.810	BDQ21	0.249	0.901	0.901	0.249	BIQ22	0.439
BMQ23	0.928	0.543	BREQ22	0.936	0.420	BSCQ22	0.904	0.596	0.800	BDQ22	0.430	0.902	0.902	0.430	BIQ23	0.384
BMQ24	0.927	0.613	BREQ23	0.937	0.400	BSCQ23	0.904	0.563	0.798	BDQ23	0.477	0.902	0.902	0.477	BIQ24	0.387
BMQ25	0.929	0.436	BREQ24	0.936	0.461	BSCQ24	0.907	0.419	0.801	BDQ24	0.421	0.899	0.899	0.421	BIQ25	0.544
BMQ26	0.927	0.620	BREQ25	0.936	0.461	BSCQ25	0.906	0.451	-	-	-	0.899	0.899	-	BIQ26	0.544
BMQ27	0.928	0.522	BREQ26	0.934	0.700	BSCQ26	0.910	0.168	-	-	-	0.897	0.897	-	BIQ27	0.682
BMQ28	0.928	0.540	BREQ27	0.934	0.673	BSCQ27	0.907	0.426	-	-	-	0.901	0.901	-	BIQ28	0.436
BMQ29	0.929	0.485	BREQ28	0.940	0.222	BSCQ28	0.904	0.570	-	-	-	0.901	0.901	-	BIQ29	0.438
BMQ30	0.927	0.643	BREQ29	0.936	0.533	-	-	-	-	-	-	0.898	0.898	-	BIQ30	0.651
BMQ31	0.928	0.506	BREQ30	0.934	0.644	-	-	-	-	-	-	0.899	0.899	-	BIQ31	0.597
BMQ32	0.930	0.373	BREQ31	0.934	0.650	-	-	-	-	-	-	0.899	0.899	-	BIQ32	0.542
BMQ33	0.929	0.493	BREQ32	0.935	0.560	-	-	-	-	-	-	0.901	0.901	-	BIQ33	0.442
BMQ34	0.928	0.544	BREQ33	-	-	-	-	-	-	-	-	-	-	-	-	-
Recommended values→	>Alpha	≥0.40	>Alpha	>Alpha	≥0.40	>Alpha	>Alpha	≥0.40	>Alpha	>Alpha	≥0.40	>Alpha	>Alpha	>Alpha	>Alpha	≥0.40

Abbreviations: BIRT, Brain Injury Rehabilitation Trust; BMQ, BIRT Motivation Questionnaire; BREQ, BIRT Regulation of Emotions Questionnaire; BSCQ, BIRT Social Cognition Questionnaire; BDQ, BIRT Disinhibition Questionnaire; BIQ, BIRT Impulsivity Questionnaire; TotCorr: Average inter-item correlations; Alpha: Cronbach's alpha; TotCorr-R: item-to-total correlation; Alpha-R: alpha if an item is deleted. Statistics values beyond the recommended cutoff are in bold.

Table 4. Summary results of confirmatory factorial analysis for Brain Injury Rehabilitation Trust Personality Questionnaires.

BIRT-PQ	N	Initial analysis						Final analysis					
		FP	Chi-sq	RMSEA	SRMR	CFI	TLI	FP	Chi-sq	RMSEA	SRMR	CFI	TLI
Motivation (BMQ)													
Patient	154	34	978.4₅₂₇*	0.075	0.106	0.822	0.811	42	855.3₅₁₉*	0.065	0.098	0.868	0.857
Caregiver	154	34	1023.4₅₂₇*	0.078	0.096	0.900	0.893	36	991.4₅₁₉*	0.076	0.094	0.906	0.899
Configural	154			Un-identified model									
Overall	308	34	1496.9₅₆₁*	0.077	0.087	0.863	0.854	59	1113.4₅₀₂*	0.063	0.073	0.917	0.908
Emotional regulation (BREQ)													
Caregiver	154	32	939.5₄₆₄*	0.082	0.091	0.925	0.920	43	728.2₄₅₃*	0.063	0.079	0.957	0.953
Configural	154			Un-identified model									
Overall	308	32	1338.4₄₆₄*	0.078	0.081	0.903	0.896	51	783.9₄₄₅*	0.050	0.063	0.962	0.958
Social cognition (BSCQ)													
Patient	154	28	860.7₃₅₀*	0.097	0.138	0.717	0.695	42	615.9₃₃₆*	0.074	0.115	0.845	0.826
Caregiver	154	28	772.6₃₅₀*	0.089	0.101	0.894	0.885	39	580.7₃₃₉*	0.068	0.087	0.939	0.932
Configural	154			Un-identified model									
Overall	308	28	1365.5₃₅₀*	0.097	0.105	0.812	0.797	62	672.3₃₁₆*	0.061	0.073	0.933	0.920
Disinhibition (BDQ)													
Patient	154	24	521.1₂₅₂*	0.083	0.116	0.769	0.747	29	378.7₂₄₇*	0.059	0.102	0.887	0.873
Caregiver	154	24	584.4₂₅₂*	0.093	0.107	0.855	0.842	31	397.8₂₄₅*	0.089	0.064	0.934	0.925
Configural	154			Un-identified model									
Overall	308	24	863.3₂₅₂*	0.089	0.098	0.794	0.774	39	422.9₂₃₇*	0.050	0.072	0.937	0.927
Impulsivity (BIQ)													
Patient	154	32	998.2₄₆₄*	0.086	0.120	0.762	0.746			Un-identified model			
Caregiver	154	32	1106.5₄₆₄*	0.095	0.109	0.861	0.851	46	775.0₄₅₀*	0.068	0.089	0.929	0.922
Configural	154			Un-identified model									
Overall	308	32	1818.0₄₆₄*	0.097	0.104	0.772	0.756	63	905.8₄₃₃*	0.060	0.074	0.920	0.909
Recommended values			<i>n.s.</i>	≤ 0.06	≤ 0.08	≥ 0.95	≥ 0.95		<i>n.s.</i>	≤ 0.06	≤ 0.08	≥ 0.95	≥ 0.95

Abbreviations: BIRT-PQs, Brain Injury Rehabilitation Trust Questionnaires. N, sample size; FP, number of free parameters; Chi-sq, Chi-square; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; CFI, Comparative Fit Index; TLI, Tucker Lewis Index; *n.s.*, not significant. For chi-squares, the degrees of freedom were reported. * indicates significant chi-squares *p*-values. Statistics values outside the recommended cutoff are in bold.

Table 5. Summary results of the analysis of respondent burden (estimated as administration time) for Brain Injury Rehabilitation Trust Personality Questionnaires.

	Patient						Caregiver					
	BMQ	BREQ	BSCQ	BDQ	BIQ	Overall	BMQ	BREQ	BSCQ	BDQ	BIQ	Overall
Mean (minutes)	6.7	7.5	6.0	5.7	6.1	32.0	6.2	4.7	4.2	4.1	4.5	23.8
SD (minutes)	2.3	3.0	2.0	1.8	1.7	10.8	3.5	1.6	1.9	1.8	1.7	10.5
Median (minutes)	7.5	8.0	8.0	6.0	6.0	35.5	5.0	6.0	7.0	6.0	5.0	29.0
Min (minutes)	4.9	5.5	4.3	4.2	4.5	23.5	4.5	3.3	2.7	2.8	3.0	16.4
Max (minutes)	8.5	9.4	7.7	7.3	7.6	40.5	7.9	6.2	5.6	5.5	6.0	31.2

Abbreviations: BIRT-PQs, Brain Injury Rehabilitation Trust Questionnaire; BMQ: BIRT Motivation Questionnaire; BREQ: BIRT Regulation of Emotions Questionnaire; BSCQ: BIRT Social Cognition Questionnaire; BDQ: BIRT Disinhibition Questionnaire; BIQ: BIRT Impulsivity Questionnaire; Overall: global administration time of the five BIRT-PQs. SD, standard deviation; Min, minimum; Max, maximum.

scale were problematic as they contributed much less than expected to the total score. This issue was also confirmed by the CFA, which showed the misfit of the data to a one-factor model, suggesting a lack of unidimensionality for the total scores. Furthermore, the baseline CFAs indicated that a source of misfit could be represented by the presence of local dependency within the data. Indeed, after accounting for this local dependency, the fit to a one-factor model improved significantly, although some un-modeled sources of misfit remained concealed within the data. It cannot be excluded that violations of multigroup invariance contributed to these unknown sources of misfit, as the latter could not be investigated because of the larger sample size needed for this kind of analysis. Finally, the study of respondent burden suggested that the BIRT-PQs are demanding instruments in terms of administration time, both for patients and caregivers.

As we employed the same inclusion criteria and etiology characteristics of the previous BIRT-PQs studies (21,22), patients with mild or moderate brain injuries could not be included in the sample. Despite this, in comparison to the sample enrolled in the initial validation studies of the BIRT-PQs (21,22), our sample was

twice as large (154 subjects vs. 72 subjects). Furthermore, as pooling patients and caregiver together further doubled the sample size, we were able to reach a total sample size of >300, which was sufficient to guarantee a ratio subjects-to-item between 8.8:1 and 12.5:1, values which were close or above the recommended ratio 10:1 for CFA (52). Overall, as our patients' characteristics are coherent with epidemiological data available on s-ABI regarding age, gender, and etiology (53), our sample is not only numerically appropriate, but it can also be considered representative of the population of patients with s-ABI.

Before this study, the psychometric performance of the BIRT-PQs had been assessed only in terms of classic reliability (test-retest reliability and internal consistency) and external construct validity (concurrent and predictive validity) (8,21,22). The evidence provided in these studies about the validity of the BIRT-PQs is limited. Firstly, the psychometric methods employed were based on parametric statistics (e.g. multiple regression), whereas the BIRT-PQs total scores, being ordinal, would have indeed required the use of psychometric methods based on non-parametric statistics (25,54,55). Secondly, these methods compare the psychometric performance of the instrument's total score with

an external indicator, which is supposedly considered a valid indicator of the same or similar construct (25,56). The main shortcoming of these methods is that they assume the validity of the total score as a unidimensional indicator (i.e. they assume that all items within a scale contribute to measuring a single underlying construct) so that their scores can be summed together to generate a total score. Unfortunately, this assumption has never been tested in detail so far.

On the other hand, the internal construct validity methods employed in this paper aim to establish 'internally' the validity of the total score by comparing its psychometric performance to that predicted by a measurement model (25,44). The only previously published analyses which could provide some information about the internal construct validity of the BIRT-PQs were those related to internal consistency, as the latter is considered a necessary, although not sufficient, pre-requisite for unidimensionality (40). These analyses reported high values for all BIRT-P scales, thus suggesting a high internal consistency for all five scales. Particularly, in a sample of 72 subjects with ABI, Cronbach's alphas of 0.940 and 0.950 were reported for the self- and relative-rated BMQ, respectively (21). Cattran et al. (22) found a Cronbach's alpha of 0.960 for both self- and relative-rated BREQ in a sample of 74 patients with ABI. Finally, Cattran et al. (8) found internal consistency values (respectively, for the self- and relative-rated version) of 0.915 and 0.949 for BSCQ, of 0.899 and 0.942 for BDQ, and 0.924 and 0.952 for BIQ on a sample of 72 participants.

The internal consistency reported in this paper showed Cronbach's alpha values systematically lower than those reported in the literature (BMQ: 0.930; BREQ: 0.937; BSCQ: 0.909; BDQ: 0.811; BIQ: 0.904). This finding can be explained considering that, as the alphas were calculated on a bigger sample size ($n = 308$), larger measurement errors could be expected. Notwithstanding this limitation, all of Cronbach's alpha values suggested high internal consistency. However, as Cronbach's alpha values are known to increase with the number of items in the scale, the supplementary analyses here performed cast some further light on the internal consistency of the five scales. Particularly, the average inter-item correlation analysis suggested that the internal consistency of BDQ was less satisfactory than as indicated by the Cronbach's alpha, and that BREQ had a higher internal consistency (0.329) than BMQ (0.284), although based on a slightly lesser number of items. This finding can be explained considering that the average inter-item correlation is an internal consistency indicator which, unlike Cronbach's alpha, is not influenced by the number of items in the scale (40). Furthermore, despite the overall consistency reliability values, our item-by-item analyses suggested that in each scale, there were some problematic items. This aspect was especially evident for BDQ and BIQ, where some low negative item-to-total correlations were detected, thus suggesting that these items did not contribute at all to the total scores. Furthermore, all items correlation <0.4 to the restscore led to an increase of the Cronbach's alpha should that item be deleted, thus suggesting that these items were less consistent than the majority of other items and, therefore, less contributing to the total score.

As internal consistency analyses cannot establish whether a scale is unidimensional (40), we performed a CFA to fit

a one-factor model, to test specifically the assumption that the items in each scale were unidimensional. Therefore, we performed a non-parametric CFA based on polychoric correlations to fit ordinal data, as the assumptions of normality and linearity of item scores, which are required by parametric CFA (47), are unlikely to be met by ordinal data (25). Within CFA, the first analyses were performed on each scale separately for patients and caregivers to test the ICV assumption of invariance. Indeed, for valid measurements, it is necessary to demonstrate that the various components of the measurement models are equivalent (i.e. invariant) across particular sample subgroups, such as patients and caregivers (46). Unfortunately, we were unable to test this important ICV assumption, as the standard errors of the models' parameters could not be computed, leading to un-identified configural models. This issue could be explained considering that with a sample size of 154 subjects, the sample size\ parameters ratio for the final analyses fell well below the recommended ratio of 10:1 (actually well below 5:1 for most scales, except for BDQ) (52). Thus, it is highly likely that the additional constraints imposed by the configural model did reduce even further this ratio, thus leading to un-identified models.

The CFA on pooled data rejected the unidimensionality assumption. This finding was somehow expected, given the results of the item-by-item internal consistency analyses. However, the base analyses showed the presence of high values of MIs within each scale (42,47), which are indicators of local dependency (48–50,57). Indeed, local dependence frequently occurs in health outcome scales (56), although it is frequently unreported or inadequately addressed (56,58,59). Local dependency is often a source of model misfit, as it may be linked either to response dependency (the response to an item can be predicted on the basis of the response to a similar item) (48–50), or to multidimensionality, where some of the response variances are accounted for by an additional latent variable (59,60). Indeed, after accounting for local dependency by correlating the error terms of the item pairs with high MIs (42,47), a substantial improvement of the fitness to the one-factor CFA model was observed for all the scales, thus suggesting a remarkable improvement of the unidimensionality of each scale. Given the impossibility to test the assumption of multigroup invariance by the respondent, we could not exclude that the suboptimal fit to a unidimensional model for most of the BIRT-PQs for the pooled sample data could be related to the presence of un-modeled violations of this requirement within the data.

The analysis of respondent burden demonstrated that the BIRT-PQs were demanding in terms of administration time, both for patients and caregivers. This finding was somehow expected, considering a total item set of 150 questions. The substantial length of the questionnaires does not only limit the acceptability and feasibility of the BIRT-PQs in routine clinical settings (61), but it is also a hindrance because of the substantially larger sample sizes needed to perform more advanced and detailed psychometric analyses, as the multi-group invariance CFA attempted in this paper. Considering the item redundancy uncovered by the CFA on the pooled data, we envisage that the next step of analysis will be a Rasch analysis conducted on the pooled data for each scale aimed at

reducing the item set. Within the Rasch analysis, it will be possible to test all ICV requirements (54), including subgroup invariance for a variety of factors, such as gender, age, etiology, and respondent (so-called analysis of differential item functioning or DIF). The advantage of testing group invariance within the context of Rasch analysis is that the DIF analysis must be performed on pooled data, thus overcoming the sample size issues highlighted in this paper.

This study has some limitations that deserve discussion. Firstly, it should be noted that fitness to the one-factor CFA model was entirely adequate only for the BREQ. Indeed, apart from multidimensionality and local dependency, it is likely that other unrecognized sources of misfit could not be modeled within the CFA performed in this study (54), such as the already mentioned violations of multigroup invariance. Another limitation is that, as all psychometric methods here employed are sample-dependent (54,62), the generalizability of our results may not be assumed for different samples and/or people living in other countries.

Conclusion

Despite the above limitations, our analyses demonstrated the lack of internal construct validity for the original BIRT-PQs total scores. Consequently, the external validity analyses already published should be considered somehow biased, as relying on biased total scores. At the same time, the present analyses demonstrated that the unidimensionality of each BIRT-PQ not only can be improved, but it is sufficient to perform a Rasch analysis (44). The latter will be aimed at deleting misfitting and/or redundant items according to a logic that will take into account both the statistical constraints and the clinical knowledge of the construct being measured. Considering also that the BIRT-PQs' administrative burden is substantial, the perspective of creating more valid and shorter forms of the questionnaires is undoubtedly desirable for improving both the measurement properties of each BIRT-PQ, as well as their acceptability and feasibility in routine clinical settings.

Note

1. Centro Cardinal Ferrari, Fontanellato, Parma; Clinica di Riabilitazione Toscana Spa, Montevarchi, Arezzo; Istituto Clinico Quarenghi, San Pellegrino Terme, Bergamo; Centro Puzzle, Turin; Istituto Santo Stefano, Porto Potenza Picena, Ancona; IRCCS Salvatore Maugeri, Telese, Benevento; Ospedale Sacro Cuore Don Calabria, Negrar, Verona; Ospedale Papa Giovanni XXIII, Bergamo; Casa di Cura del Policlinico, Milan; Fondazione Santa Lucia, Rome; Viglietta, Cuneo.

Acknowledgments

The authors want to thank Maria Paola Ciurli, Elisa Scarano, Maria Daniela Lo Sapio for their help in collecting the data.


Declaration of interest

The authors report no conflict of interest.

Funding

LP was (partially) supported by funding of the Italian Ministry of Health (Ricerca Corrente)

ORCID

Benedetta Basagni  <http://orcid.org/0000-0002-6173-584X>
 Daniele Piscitelli  <http://orcid.org/0000-0002-5240-3798>
 Antonio De Tanti  <http://orcid.org/0000-0003-4803-419X>
 Leonardo Pellicciari  <http://orcid.org/0000-0002-9521-3759>
 Serena Caselli  <http://orcid.org/0000-0002-4187-9475>
 Rita Formisano  <http://orcid.org/0000-0002-2516-0540>
 Anna Estraneo  <http://orcid.org/0000-0001-6646-5626>
 Maria Grazia Gambini  <http://orcid.org/0000-0002-6211-2169>
 Gianfranco Lamberti  <http://orcid.org/0000-0002-3288-5275>
 Mauro Mancuso  <http://orcid.org/0000-0002-8451-9277>
 Matteo Sozzi  <http://orcid.org/0000-0002-3969-7330>
 Laura Abbruzzese  <http://orcid.org/0000-0002-5304-3974>
 Fabio La Porta  <http://orcid.org/0000-0002-5031-6658>

Data availability statement

The data that support the findings of this study are available from the corresponding author [LP], upon reasonable request.

References

1. Yeates GN, Gracey F, McGrath JC. A biopsychosocial deconstruction of "Personality change" following acquired brain injury. *Neuropsychol Rehabil.* 2008;18(5-6):566-89. doi:10.1080/09602010802151532.
2. Norup A, Mortensen EL. Prevalence and predictors of personality change after severe brain injury. *Arch Phys Med Rehabil.* 2015;96(1):56-62. doi:10.1016/j.apmr.2014.08.009.
3. Osborne-Crowley K, McDonald S. A review of social disinhibition after traumatic brain injury. *J Neuropsychol.* 2018;12(2):176-99. doi:10.1111/jnp.2018.12.issue-2.
4. Oddy M. He's no longer the same person: how families adjust to personality change after head injury. *Traumatic brain injury rehabilitation.* Boston, MA: Springer; 1995.
5. Wood RL. Understanding neurobehavioural disability. In: McMillan TM, Wood RL, editors. *Neurobehavioural disability and social handicap following traumatic brain injury.* Hove: Psychology Press Ltd; 2001, pp. 1-28.
6. Gouick J, Gentleman D. The emotional and behavioural consequences of traumatic brain injury. *Trauma.* 2004;6(4):285-92. doi:10.1191/1460408604ta323oa.
7. Ciurli P, Formisano R, Bivona U, Cantagallo A, Angelelli P. Neuropsychiatric disorders in persons with severe traumatic brain injury: prevalence, phenomenology, and relationship with demographic, clinical, and functional features. *J Head Trauma Rehabil.* 2011;26(2):116-26. doi:10.1097/HTR.0b013e3181dedd0e.
8. Cattran CJ, Oddy M, Wood RL, Moir JF. Post-injury personality in the prediction of outcome following severe acquired brain injury. *Brain Inj.* 2011;25(11):1035-46. doi:10.3109/02699052.2011.607787.
9. Morton MV, Wehman P. Psychosocial and emotional sequelae of individuals with traumatic brain injury: a literature review and recommendations. *Brain Inj.* 1995;9(1):81-92. doi:10.3109/02699059509004574.
10. Wood RL, Yurdakul LK. Change in relationship status following traumatic brain injury. *Brain Inj.* 1997;11(7):491-501. doi:10.1080/bij.11.7.491.501.
11. Kreutzer JS, Gervasio AH, Camplair PS. Primary caregivers' psychological status and family functioning after traumatic brain injury. *Brain Inj.* 1994;8(3):197-210. doi:10.3109/02699059409150973.

12. Milders M. Relationship between social cognition and social behaviour following traumatic brain injury. *Brain Inj.* 2018;16:1–7.
13. Nightingale EJ, Soo CA, Tate RLJBI. A systematic review of early prognostic factors for return to work after traumatic brain injury. *Brain Impairment.* 2007;8(2):101–42. doi:10.1375/brim.8.2.101.
14. Moretta P, Masotta O, Crispino E, Castronovo G, Ruvalo S, Montalbano C, Loreto V, Trojano L, Estraneo A. Psychological distress is associated with altered cognitive functioning in family caregivers of patients with disorders of consciousness. *Brain Inj.* 2017;31(8):1088–93. doi:10.1080/02699052.2017.1290278.
15. Moretta P, Estraneo A, De Lucia L, Cardinale V, Loreto V, Trojano L. A study of the psychological distress in family caregivers of patients with prolonged disorders of consciousness during in-hospital rehabilitation. *Clin Rehabil.* 2014;28(7):717–25. doi:10.1177/0269215514521826.
16. Zasler ND, Martelli MF, Nicholson K. Post-traumatic pain disorders: Medical Assessment and Management. In: *Brain Injury Medicine: Principles and Practice*. 2nd edition. Zasler N, Katz D, Zafonte R, editors. New York: Demos Medical; 2013. p. 954–73.
17. Bivona U, Formisano R, Mastrilli L, Zabberoni S, Caltagirone C, Costa A. Theory of mind after severe acquired brain injury: clues for interpretation. *Biomed Res Int.* 2018;2018:5205642. doi:10.1155/2018/5205642.
18. Silver JM, McAllister TW, Arciniegas DB, American Psychiatric Association. *Textbook of traumatic brain injury*. 3rd ed. Washington, DC: American Psychiatric Association Publishing; 2019.
19. Kaufer DI. Neurobehavioral assessment. *Continuum (Minneapolis)*. 2015;21(3Behavioral Neurology and Neuropsychiatry):597–612. doi:10.1212/01.CON.0000466655.51790.2f.
20. Shukla D, Devi BI, Agrawal A. Outcome measures for traumatic brain injury. *Clin Neurol Neurosurg.* 2011;113(6):435–41. doi:10.1016/j.clineuro.2011.02.013.
21. Oddy M, Cattran C, Wood R. The development of a measure of motivational changes following acquired brain injury. *J Clin Exp Neuropsychol.* 2008;30(5):568–75. doi:10.1080/13803390701555598.
22. Cattran C, OddyM, Wood R. The development of a measure of emotional regulation following acquired brain injury. *J Clin Exp Neuropsychol.* 2011;33(6):672–79. doi:10.1080/13803395.2010.550603.
23. Hyde CJ. *The measurement and impact of five personality changes after brain injury*. Swansea: University of Wales; 2006.
24. Basagni B, Navarrete E, Bertoni D, Cattran C, Mapelli D, Oddy M, De Tanti A. The Italian version of the Brain Injury Rehabilitation Trust (BIRT) personality questionnaires: five new measures of personality change after acquired brain injury. *Neurol Sci.* 2015;36(10):1793–98. doi:10.1007/s10072-015-2251-9.
25. Kucukdeveci AA, Tennant A, Grimby G, Franchignoni F. Strategies for assessment and outcome measurement in physical and rehabilitation medicine: an educational review. *J Rehabil Med.* 2011;43(8):661–72. doi:10.2340/16501977-0844.
26. Jones C. Glasgow coma scale. *Am J Nurs.* 1979;79(9):1551–53.
27. Hagen C, Malkmus D, Durham P. Levels of cognitive functioning, rehabilitation of the head injured adult: comprehensive physical management. Dowey, CA: Professional Staff Association of the Rancho Los Amigos Hospital, Inc.; 1979.
28. Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? *Int Disabil Stud.* 1988;10(2):64–67. doi:10.3109/09638288809164105.
29. World Medical Association General A. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects (revised October 7, 2000). *HIV Clin Trials.* 2001;2(1):92–95. doi:10.1310/GTFR-2DRX-M6YE-ELXR.
30. Rappaport M, Hall K, Hopkins K, Belleza T, Cope D. Disability rating scale for severe head trauma: coma to community. *Arch Phys Med Rehabil.* 1982;63(3):118–23.
31. Majani G, Callegari S. SAT-P/Satisfaction Profile. Soddifazione soggettiva e qualità della vita. Trento; Erickson; 1998.
32. Kertesz A, Davidson W, Fox H. Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Can J Neurol Sci.* 1997;24(1):29–36. doi:10.1017/S0317167100021053.
33. Alberici A, Geroldi C, Cotelli M, Adorni A, Calabria M, Rossi G, Borroni B, Padovani A, Zanetti O, Kertesz A. The frontal behavioural inventory (Italian version) differentiates frontotemporal lobar degeneration variants from Alzheimer’s disease. *Neurol Sci.* 2007;28(2):80–86. doi:10.1007/s10072-007-0791-3.
34. Caserta MS, Lund DA, Wright SD. Exploring the Caregiver Burden Inventory (CBI): further evidence for a multidimensional view of burden. *Int J Aging Hum Dev.* 1996;43(1):21–34. doi:10.2190/2DKF-292P-A53W-W0A8.
35. Long JS. *Confirmatory factor analysis: a preface to LISREL*. Newbury, London, New Delhi: Sage Publications; 1983.
36. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res.* 1995;4(4):293–307. doi:10.1007/BF01593882.
37. Bland JM, Altman DG. Cronbach’s alpha. *BMJ.* 1997;314(7080):572. doi:10.1136/bmj.314.7080.572.
38. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42. doi:10.1016/j.jclinepi.2006.03.012.
39. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Qual Life Res.* 1995;4(3):187–206. doi:10.1007/BF02260859.
40. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess.* 1995;7(3):309. doi:10.1037/1040-3590.7.3.309.
41. Yusoff MSB, Rahim AFA, Yaacob MJ. The development and validity of the Medical Student Stressor Questionnaire (MSSQ). *ASEAN J Psychiatry.* 2010;11(1):231–35.
42. Byrne BM. *Structural equation modeling with AMOS: basic concepts, applications, and programming*. 2nd ed. New York: Routledge; 2010.
43. Hoyle RH. *Confirmatory factor analysis. Handbook of applied multivariate statistics and mathematical modeling*. New York: Academic Press; 2000. 465 p.
44. La Porta F, Franceschini M, Caselli S, Cavallini P, Susassi S, Tennant A. Unified balance scale: an activity-based, bed to community, and aetiology-independent measure of balance calibrated with Rasch analysis. *J Rehabil Med.* 2011;43(5):435–44. doi:10.2340/16501977-0797.
45. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model: Multidiscip J.* 1999;6(1):1–55. doi:10.1080/10705519909540118.
46. Byrne BM. *Structural equation modeling with Mplus: basic concepts, applications, and programming*. New York: Routledge; 2013.
47. Muthén LK, Muthén BO. *MPlus user’s guide 2010*. 6th ed. <http://www.statmodel.com/download/usersguide/>.
48. Lundgren-Nilsson A, Jonsdottir IH, Ahlborg G Jr., Tennant A. Construct validity of the Psychological General Well Being Index (PGWBI) in a sample of patients undergoing treatment for stress-related exhaustion: a Rasch analysis. *Health Qual Life Outcomes.* 2013;11:2. doi:10.1186/1477-7525-11-2.
49. Revicki DA, Chen W, Tucker C. Developing item banks for patient-reported health outcomes. In: *Handbook of item response theory modeling: applications to typical performance assessments* Reise SP, Revicki DA (eds.), 334–363. New York; Routledge; 2014.
50. Watt T, Groenvold M, Deng N, Gandek B, Feldt-Rasmussen U, Rasmussen AK, Hegedüs L, Bonnema SJ, Bjorner JB. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. *Health Qual Life Outcomes.* 2014;12:126. doi:10.1186/s12955-014-0126-z.
51. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, Stein RE. Assessing health status and quality-of-life instruments:

- attributes and review criteria. *Qual Life Res.* 2002;11(3):193–205. doi:10.1023/A:1015291021312.
52. Tabachnick BG, Fidell LS. Using multivariate statistics. 4th ed. Boston (MA): Allyn and Bacon; 2001.
53. Chiavaroli F, Derraik JG, Zani G, Lavezzi S, Chiavaroli V, Sherwin E, ViBasaglia N. Epidemiology and clinical outcomes in a multicentre regional cohort of patients with severe acquired brain injury. *Disabil Rehabil.* 2016;38(20):2038–46. doi:10.3109/09638288.2015.1111439.
54. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess.* 2009;13(12):iii, ix-x, 1–177. doi:10.3310/hta13120.
55. Piscitelli D, Pellicciari L. Responsiveness: is it time to move beyond ordinal scores and approach interval measurements? *Clin Rehabil.* 2018;32(10):1426–27. doi:10.1177/0269215518794069.
56. La Porta F, Caselli S, Susassi S, Cavallini P, Tennant A, Franceschini M. Is the Berg balance scale an internally valid and reliable measure of balance across different etiologies in neurorehabilitation? A revisited Rasch analysis study. *Arch Phys Med Rehabil.* 2012;93(7):1209–16. doi:10.1016/j.apmr.2012.02.020.
57. Pellicciari L, Piscitelli D, Caselli S, La Porta F. A Rasch analysis of the Conley Scale in patients admitted to a general hospital. *Disabil Rehabil.* 2019;41(23):2807–16. doi: 10.1080/09638288.2018.1478000.
58. Lundgren Nilsson A, Tennant A. Past and present issues in Rasch analysis: the functional independence measure (FIM) revisited. *J Rehabil Med.* 2011;43(10):884–91. doi:10.2340/16501977-0871.
59. Marais I. Local dependence. London Hoboken, NJ: ISTE, John Wiley & Sons; 2013.
60. Tennant A, Conaghan PG. 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.* 2007;57(8):1358–62. doi:10.1002/(ISSN)1529-0131.
61. Rolstad S, Adler J, Ryden A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value Health.* 2011;14(8):1101–08. doi:10.1016/j.jval.2011.06.003.
62. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health.* 2015;18(1):25–34. doi:10.1016/j.jval.2014.10.005.