Authors:

Pamela Gallagher, PhD
F. Franchignoni, MD
A. Giordano, PhD
M. MacLachlan, PhD

Affiliations:

From the Faculty of Science and Health (PG), School of Nursing, Dublin City University, Ireland; Department Occupational Rehabilitation and Ergonomics (FF) and Unit of Bioengineering (AG), Rehabilitation Institute of Veruno (NO), "Salvatore Maugeri" Foundation, Clinica del Lavoro e della Riabilitazione, IRCCS, Italy; and School of Psychology (MM), University of Dublin, Trinity College, Ireland.

Correspondence:

All correspondence and requests for reprints should be addressed to Pamela Gallagher, PhD, Faculty of Science and Health, School of Nursing, Dublin City University, Dublin 9, Ireland.

Disclosures:

Financial disclosure statements have been obtained, and no conflicts of interest have been reported by the authors or by any individuals in control of the content of this article. P. Gallagher and M. MacLachlan are the members of the Dublin Psychoprosthetics Group (available www.tcd.ie/psychoprosthetics).

0894-9115/10/8906-0487/0
American Journal of Physical Medicine & Rehabilitation

DOI: 10.1097/PHM.0b013e3181dd8cf1

ORIGINAL RESEARCH ARTICLE

Trinity Amputation and Prosthesis Experience Scales

A Psychometric Assessment Using Classical Test Theory and Rasch Analysis

ABSTRACT

Gallagher P, Franchignoni F, Giordano A, MacLachlan M: Trinity Amputation and Prosthesis Experience Scales (TAPES): A psychometric assessment using classical test theory and Rasch analysis. Am J Phys Med Rehabil 2010;89: 487–496.

Objective: To perform a detailed psychometric analysis using both classical test theory and Rasch analysis of the three main scales of the Trinity Amputation and Prosthesis Experience Scales (TAPES) in people with a lower-limb amputation.

Design: A sample of 498 persons who were prosthesis users with a lower-limb amputation was retrospectively studied, pooled from a number of studies undertaken across the United Kingdom and Ireland in the past decade in which the TAPES had been completed as part of a postal survey. Both factor analysis techniques and Rasch analysis were performed on TAPES data. Dimensionality, item fit to the model, response category performance, and internal construct validity were assessed. Category collapsing and item removal were considered to improve the questionnaire.

Results: The analyses suggested to restructure the TAPES as follows: (a) three psychosocial adjustment subscales with a four-point rating scale (and a reworded item); (b) an activity restriction scale based on ten items with their original three-point rating scale; and (c) two satisfaction with the prosthesis subscales using a three-point rating scale. All scales and subscales showed acceptable internal consistency and ability to define a distinct hierarchy of persons along the measured construct.

Conclusions: This study empirically identified a revised version of the TAPES (TAPES-R) with a simplified general structure and psychometrically suitable for assessing the complex experience of amputation and adjustment to a lower-limb prosthesis. Additional studies are needed to confirm and further explore its measurement properties in other samples, thereby adding clinical validity to the instrument.

Key Words: Amputation, Psychometrics, Rasch Analysis, Outcome Assessment, Psychological Adjustment

The ability to appropriately measure outcomes is acknowledged as an integral part of evidence-based practice. Consequently, there is an increasing tendency for clinicians and practitioners to use and seek out outcome measurement instruments. In terms of prosthetic rehabilitation, outcomes are predominantly concerned with the physical, functional, and psychosocial factors related to prosthetic use. An individual must adjust to many new personal realities after an amputation and the acquiring of a prosthetic limb. However, the way in which persons negotiate their environment and activities may differ considerably.[1] Furthermore, they must also become accustomed to the prosthesis, which is a new part of their internal and external image.[2,3] The variability in people's adjustment to limb loss and subsequent prosthesis usage has prompted the development of several scales to investigate the success of adjustment to a prosthesis and the factors related to prosthetic use.[4] Among them, the Trinity Amputation and Prosthesis Experience Scales (TAPES) were theoretically and empirically derived to enable an examination of the psychosocial processes involved in adjusting to amputation and the specific demands of wearing a prosthesis.[5] The TAPES has been widely used in both clinical and research contexts, has been translated into different languages (see www.tcd.ie/psychoprosthetics), and has demonstrated good psychometric properties according to classical test theory.[5,6] But, more recently there has also been a deliberate move toward basing the psychometric evaluation of instruments on item response theory and Rasch analysis.[7,8] Rasch analysis provides a wealth of psychometric information that is not obtainable through classical test theory[9,10] on (a) the functioning of rating scale categories; (b) the validity of a measure by evaluating the fit of individual items to the latent trait; and (c) the consistency of item difficulties with the expectations of the construct (and hence a description of the range and hierarchical relationship of the variable). Indeed, Rasch analysis has been recommended as a method for assessing scale properties in addition to classical psychometric criteria for reviewing and assessing surveys and questionnaires for disability outcomes research.[7]

The purpose of this study, therefore, was to perform a detailed psychometric analysis using both classical test theory and Rasch analysis of the three main scales of the TAPES (adjustment, activity restriction, and satisfaction) in people with a lower-limb amputation and to thoroughly investigate a spectrum of psychometric characteristics. This approach aims to provide the rationale for revising and improving the measurement qualities of this outcome measure.

METHOD

Instrument

Trinity Amputation and Prosthesis Experience Scales

In addition to requesting demographic and disability-related data (regarding sex, age, cause and type of amputation, length of time living with the prosthesis, and degree of prosthetic use), the TAPES consists of three scales (psychosocial adjustment, activity restriction, and satisfaction with the prosthesis), each one containing three subscales.[5]

The three psychosocial adjustment subscales are as follows: general adjustment (e.g., "I have adjusted to having an artificial limb"), social adjustment (e.g., "I don't mind people asking about my artificial limb"), and adjustment to limitation (e.g., "Being an amputee means that I can't do what I want to do"). Each of these subscales contains five items, which are measured along a five-point rating scale (strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree). Items 9 and 11–15 have reversed scores. Higher scores always indicate greater levels of adjustment.

The three activity restriction subscales are as follows: Functional activity restriction (e.g., walking 100 yards), Social activity restriction (e.g., visiting friends), and Athletic activity restriction (e.g., sport and recreation). Each of these activity restriction subscales contains four items, which are measured along a three-point scale (not at all limited, limited a little, and limited a lot). Higher scores are indicative of greater activity restriction.

Three further subscales assess satisfaction with the prosthesis, measured along a five-point scale (very dissatisfied, dissatisfied, neither dissatisfied nor satisfied, satisfied, and very satisfied). The functional satisfaction subscale contains five items (e.g., reliability). There are four items in the aesthetic satisfaction subscale (e.g., color), and there is a single weight satisfaction item. Higher scores in each of the satisfaction subscales are indicative of greater satisfaction with the prosthesis.

Participants

A sample of 498 persons who were prosthesis users with a lower-limb amputation was retrospectively studied, pooled from a number of studies undertaken across the United Kingdom and Ireland in the past decade in which subjects had completed the TAPES as part of a postal survey (see Table 1). For each of the studies from which the data are pooled, ethical approval had been obtained from the respective institutional ethics committee.

TABLE 1 Sample characteristics

	N	%
Gender		
Males	359	72.1
Females	138	27.7
Missing	1	0.2
Type of amputation		
Transfemoral	168	33.7
Transtibial	286	57.4
Through knee	16	3.2
Bilateral	22	4.4
Missing	6	1.2
Cause		
Cancer	52	10.4
Accident	170	34.1
PVD/diabetes	177	35.5
Other	93	18.7
Missing	6	1.2

	Mean	SD	Range
Age (yrs)	54.8	18.6	18–89
Length of time with a prosthesis (mos)	106.5	101.8	2–876

PVD, peripheral vascular disease.

Statistical Analysis

Analysis of Homogeneity and Dimensionality

Internal consistency of the TAPES subscales was analyzed by calculating (a) Cronbach's coefficient α. α Values >0.70 are recommended for group level comparison, whereas a minimum of 0.85–0.90 is desirable for individual judgments[11,12]; and (b) the item-total correlation. Each item should correlate with the total score (omitting that item) with $r > 0.20$.[11] The correlations were calculated as Spearman's ρ (r_s), corrected for ties.

Dimensionality (the structure of underlying latent factors that can be seen as actually causing the observed score on the test items) must be investigated before proceeding with further analyses: ideally, one sole latent factor should be present for the application of Rasch analysis to the entire item set. If unidimensionality is not sustainable, the number of underlying factors and their relationship with each item can be used to separate the scale into subscales. Classical factor analysis was used to evaluate the dimensional structure of each of the three TAPES scales.[13]

First, a confirmatory factor analysis for categorical data (CFA; LISREL version 8.80 software; Scientific Software International, Inc., Lincolnwood, IL) was performed to evaluate the fit of the data to the previously published models[5] (related to the domains psychosocial adjustment, activity restriction, and satisfaction with the prosthesis). The extent to which the model can be used to reproduce the sample data was determined by examining the following indexes:

a. Comparative fit index (CFI): The CFI ranges from 0 (poor fit) to 1 (good fit) with values >0.90 being indicative of a good fit.[14]
b. Nonnormed fit index (NNFI, aka Tucker-Lewis index): Hu and Bentler[15] have recently suggested NNFI ≤ 0.95 as a cutoff for a good model fit and this is widely accepted.
c. Root mean square error of approximation (RMSEA): a RMSEA value <0.1 is considered to reflect adequate fit[14] and <0.05 is usually taken as an indication of good fit.
d. Standardized root mean square residual (SRMR): a value <0.05 is widely considered a good fit, and a value $<0.08-0.1$ is an adequate fit.[16]

If CFA showed poor fit, the parallel analysis by Horn[17] was used to estimate the number of factors to retain in each scale: the size of the eigenvalues obtained from principal component analysis is compared with those obtained from a randomly generated data set of the same size and number of variables, and only factors with eigenvalues exceeding the values obtained from the corresponding random dataset are retained for further investigation. Parallel analysis was conducted using ViSta[18] Parallel Analysis plugin (http://www.mdp.edu.ar/psicologia/vista/). The data were then randomly divided into half. An explanatory factor analysis was performed on the first half ($n = 249$) with a principal component analysis using the number of factors suggested by the parallel analysis. After varimax rotation, the relationships between items and retained factors were used to build a new multifactorial model, verifying at least three items per factor; common conceptual meaning of the items loading on each factor; and different constructs between factors. A CFA was then performed on the other half of the data to confirm the dimensionality of the scales as defined by explanatory factor analysis.

Rasch Analysis

The matrix of single raw scores in each unidimensional scale/subscale underwent Rasch analysis using the WINSTEPS software (WINSTEPS Rasch Measurement, version 3.58.1). As a first step, we investigated whether each rating scale was being used in the expected manner. We evaluated the response categories using criteria suggested by Linacre[10]: (a) at least ten cases per category; (b) even distribution of category use; (c) monotonic increase in both average measures of persons with a given score/category and thresholds [thresholds—or step calibrations—are the ability levels at which the response to either of two adjacent categories is equally likely]. Additional criteria were (d) category outfit mean square values <2 (see below) and (e) threshold differences >1.4 log-odd units and <5.

When necessary, categories were collapsed following specific guidelines, and different patterns of categorization were compared, looking not only at the above indicators of category diagnostics but also at the solution maximizing the person separation and reliability indexes.

After this phase, we assessed the validity of each scale (the three psychosocial adjustment subscales, the activity restriction scale, and the two satisfaction with the prosthesis subscales; see Results) by evaluating the fit of individual items to the latent trait as per the Rasch model and examining whether the pattern of item difficulties was consistent with the model expectancies. Depending on the string of ordinal raw scores, the Rasch model estimates goodness-of-fit (or simply "fit") of the observed data to the model-expected data. If the differences between observed and expected scores are not too large, it is said that "the data fit the model" (see below), and this is seen as equivalent proving the theoretical construct validity and adequacy of the scale.[8,19] Information-weighted (infit) and outlier-sensitive (outfit) mean-square statistics (MnSq) for each item were calculated to test whether there were items which did not fit with the model expectations. Both of these fit statistics are expected to approach 1. In accordance with the literature,[19,20] we considered MnSq >0.6 and <1.4 as an indicator of acceptable fit: items outside this range were considered underfitting (MnSq $\geq$ 1.4, suggesting the presence of unexpectedly high variability) or overfitting (MnSq $\leq$ 0.6, indicating a too predictable pattern). Rasch analysis provides estimates of the level of difficulty achieved by each item ("item difficulty") and of the location of each individual subject along the continuum ("subject ability," representing the global amount of trait in the individual). Item difficulty and subject ability are expressed—on a common interval scale—in logit units, a logit being the natural logarithm of the ratio (odds) of mutually exclusive alternatives (e.g., pass *vs.* fail or higher response *vs.* lower response).[19,21] Logit-transformed measures represent linear measures (i.e., the intended amount of the trait). Conventionally, 0 logit is ascribed to the mean item difficulty. For Rasch analysis, it is reported that a sample size of >250 persons will estimate item difficulty with an α of 0.01 to within ±0.5 logits.[22]

Reliability was evaluated in terms of "separation" (G), defined as the ratio of the true spread of the measures with their measurement error.[19,21] The item separation index gives an estimate (in standard error units) of the spread or "separation" of items along the measurement construct; the person separation index gives an estimate of the spread or separation of persons along the measurement construct. A separation index of 1.5 is considered an acceptable level of separation, and an index of 2.0 indicates a good level of separation. A related index is the reliability of these separation indexes, providing the degree of confidence that can be placed in the consistency of the estimates (range, 0–1; coefficients >0.80 are considered as good and >0.90 excellent).[19]

A principal component analysis of the standardized residuals was also performed using the WINSTEPS software to further confirm unidimensionality: it is expected that after the removal of the trait that the scale intends to measure (the so-called Rasch dimension), the residuals for pairs of items should be uncorrelated and normally distributed (i.e., there are no principal components). The following criteria have been recommended to determine whether additional factors are likely to be present in the residuals: (a) a cutoff of 60% of the variance explained by the Rasch factor; (b) eigenvalue of the first residual factor <3; and (c) percentage variance explained by the first contrast of <5%.[23]

RESULTS

Analysis of Homogeneity and Dimensionality

The Cronbach's coefficient α was as follows: 0.89 (psychosocial adjustment), 0.89 (activity restriction), and 0.95 (satisfaction with the prosthesis). The item-total correlation coefficients (r_s) ranged from 0.42–0.77 in all subscales, except for item 9 of the psychosocial adjustment scale ($r_s = 0.30$).

The initial CFA showed a satisfactory fit of the data to the original three-factor model only for the psychosocial adjustment scale (CFI = 0.99; NNFI = 0.98; RMSEA = 0.057; and SRMR = 0.059), with the three factors loading respectively on items 1–5, 6–10, and 11–15.

The parallel analysis by Horn suggested retaining two factors for the activity restriction (items 1–8, 9–12) scale and two for the satisfaction with the prosthesis scale (items 1–4, 5–9 with item 10, "overall satisfaction," preliminarily excluded from the analysis, being considered a separate global measure of satisfaction). After randomly splitting the data into half, an explanatory factor analysis with varimax rotation of the first half of the data of each scale was used to determine the structure of the relationships between items and the factors.

The successive CFA on the second half of the data provided an acceptable two-factor solution for the satisfaction with the prosthesis scale (CFI = 0.98; NNFI = 0.97; RMSEA = 0.089; and SRMR = 0.057).

The two-factor solution for the activity restriction scale showed only a moderate fit according to conventional standards (CFI = 0.97; NNFI = 0.98; RMSEA = 0.10; and SRMR = 0.10), very similar to that obtained with the one-factor solution (CFI = 0.98; NNFI = 0.97; RMSEA = 0.10; and SRMR = 0.12). Thus, for parsimony, we decided to provisionally proceed, as the activity restriction scale was one unidimensional item pool, and leave the task of further verification to the subsequent Rasch analysis.

Rasch Analysis

The rating scale diagnostics showed that some rating categories did not comply with the preset criteria for category functioning (average measures, thresholds, etc.). In all the three psychosocial adjustment subscales, the "neutral" mid-point category 2 (neither agree nor disagree) was very far below the model level (see Fig. 1), whereas the other categories met all the established criteria. Thus, it was decided to eliminate the middle category, obtaining a new four-level rating scale (0–3 points). Moreover, the five categories of the satisfaction with the prosthesis scale (0, very dissatisfied; 1, dissatisfied; 2, neither dissatisfied nor satisfied; 3, satisfied; and 4, very satisfied) showed some disordered thresholds (see Fig. 2*A*). The best model was the one that collapsed category 0, 1, and 2 into a unique category (representing answers from persons not clearly satisfied). Figure 2*B* shows the probability curves of the three resulting categories after the collapsing procedure. The analysis of the activity restriction rating-scale functioning confirmed that the responses in the three categories behaved in a valid manner.

After the rating-scale modifications, Rasch analysis showed that all items of the three psycho-

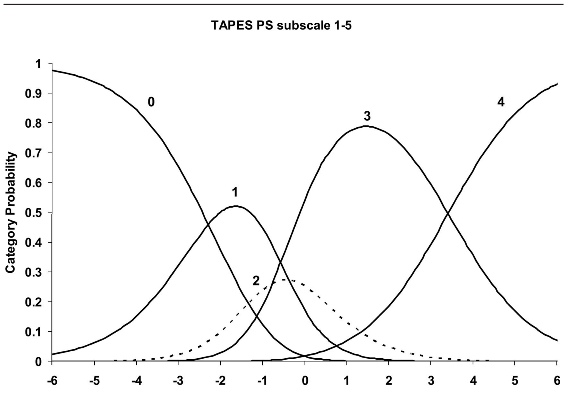

FIGURE 1 *Category probability curves of psychosocial adjustment subscale 1 (five categories: 0 = strongly disagree; 1 = disagree; 2 = neither agree nor disagree; 3 = agree; and 4 = strongly agree). The y-axis represents the probability (0–1) of responding to one of the rating categories, and the x-axis represents the different performance values (difference between subject ability and the item difficulty) in logits. The "0" curve declines as the subject's ability increases; the crossing point (where 0 and 1 are equally probable) is the first "threshold." The same applies for the other curves. The ideal plot should look like a range of hills, with an "emerging" crest for each category. It can be seen from the figure that the probability of using category 2 (dashed line) is never higher than that of adjacent ratings.*

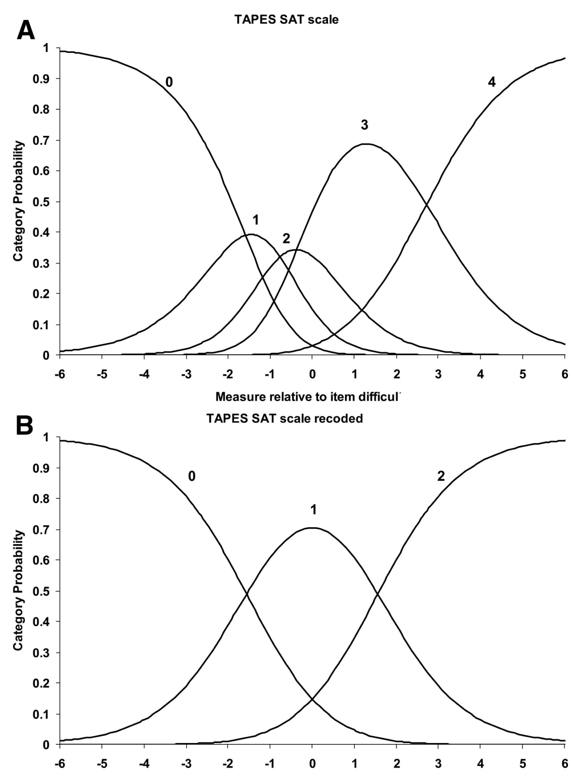

FIGURE 2 *Category probability curves of the satisfaction with prosthesis subscale 2: (A) original five categories (0 = very dissatisfied, 1 = dissatisfied, 2 = neither dissatisfied nor satisfied, 3 = satisfied, and 4 = very satisfied); (B) revised rating scale after collapsing categories 0, 1, 2, and renumbering (new categories: 0 = not satisfied; 1 = satisfied; and 2 = very satisfied). For details, see Figure 1. Again, the plot should look—as in (B)—like a range of hills, with a clear "emerging" crest for each category. A, The probability of using categories 1 and 2 is very slim. Conversely, it can be seen that the probability of selecting each of the three revised rating categories is now an evident function of the level of ability shown by the subject in the x-axis (B).*

social adjustment subscales fitted the respective constructs (MnSq between 0.6 and 1.4), except for item 9 ("I have difficulty in talking about my limb loss in conversation") in the subscale social adjustment, which was misfitting (infit MnSq = 1.94 and outfit MnSq = 1.97). In the activity restriction scale, all items showed good infit and outfit values, except for items 9 ("maintaining friendships"; Outfit MnSq = 3.13) and 10 ("visiting friends"; Outfit MnSq = 1.84). In the two satisfaction with the prosthesis subscales, the only item that did not fit the Rasch model was item 3 "noise" (infit MnSq = 1.60; outfit MnSq = 1.51).

Regarding the hierarchic ordering of items of the activity restriction scale, Figure 3 shows—according to the Rasch model—the distribution map of subject ability and item difficulty. The ability levels and reliability indexes of the subjects for all scales and subscales of the TAPES are reported in Table 2. The average levels of ability for our sample were sometimes quite far from the mean difficulty of the items, set by convention at 0 logits (see Table 2): this finding indicates problems with targeting, i.e., the extent to which items are of appropriate difficulty for the sample. All scales and subscales contained items distributed into many difficulty strata (except for satisfaction with the prosthesis subscale 1), and each one was able to distinguish 2–3 levels of subject ability in this study sample.

The results regarding the principal component analysis of the standardized residuals (analyzing the variance explained by the estimated measures—Rasch factor—and the first residual factor) are shown in Table 2. They indicate acceptable to borderline levels of unidimensionality.

```
<higher              ||<more difficult>
activity restriction>||
    5       #######  ++
                     ||
                 ##  ||
                .## T||
                     ||
                     ||
    4                ++T
                     ||
            .######  ||
                  .  ||
                  .  ||
                .##  ||  1 vigorous activities    3 running for a bus
    3            .#  ++
                  .  ||
            #######  ||
                 ## S||
            #######  ||
               .###  ||
    2                ++S
         ##########  ||  4 sport and recreation
                 .#  ||
              #####  ||
        .##########  ||
                     ||  6 walking more than a mile
    1    ##########  ++
                  .  ||
           ######## M||  2 climbing several flights of stairs
                     ||
         ##########  ||
                     ||  7 walking half a mile
    0    .#########  ++M'
            .######  ||
                ###  ||
              .####  ||  12 going to work
                  #  ||
            .######  ||
   -1             .  ++  5 climbing one flight of stairs
               ####  ||
                    S||  11 working on hobbies
            .######  ||  8 walking 100 yards
                     ||
             .#####  ||
   -2                ++S
                     ||
               ####  ||
                     ||  10 visiting friends
               .###  ||
                     ||
   -3                ++
              .####  ||
                    T||  9 maintaining friendships
                     ||
                     ||
                .##  ||
   -4                ++T
                     ||
                     ||
                     ||
                  .  ||
                     ||
   -5          .###  ++
<lower               ||<less difficult>
activity restriction>||
```

FIGURE 3 *Subject-ability and item-difficulty maps of the activity restriction scale. The vertical line represents the measure of the variable, in linear logit units. The left-hand column locates the individual's ability along the variable: each "#" denotes three persons. The right-hand column locates the 12-item difficulty measures along the variable (for each item, the difficulty estimate represents the mean calibration of the threshold parameters according to the rating scale model). Misfitting items are visiting friends and maintaining friendships. The double line indicates that the two sides have opposite orientations. From bottom to top, measures indicate higher activity restriction for patients and higher difficulty for items, respectively. By convention, the average difficulty of items in the test is set at 0 logits (and indicated with M'). Accordingly, a candidate with average ability is indicated with M.*

DISCUSSION

To validate a tool measuring a multidimensional phenomenon such as adjustment, it is necessary to analyze several datasets to confirm, refine, or modify the conceptual model of measurement as initially defined. First of all, we tested dimensionality using factor analytic methods, because Rasch analysis fit statistics alone can be inadequate for determining unidimensionality.[24] The original three-factor model for the psychosocial adjustment scale was confirmed. Conversely, two factors instead of the original three came out in the satisfaction with the prosthesis scale. In particular, the item relating to satisfaction with the "weight" of the prosthesis (which was originally a single-item factor) was subsumed into the factor "functional satisfaction." On the other hand, we considered the activity restriction scale as sufficiently unidimensional to undergo Rasch analysis as a whole.

Rasch analysis was then performed on each scale/subscale suggested by factor analysis. In these conditions, Rasch analysis provides a valuable tool for testing whether the properties of a questionnaire comply with a wide spectrum of psychometric requirements for measurement, which are not analyzed by traditional techniques.[8] Overall, the modifications suggested by Rasch analysis improve the measurement properties of an instrument.[19] Indeed, Rasch analysis is a medium to weigh up the strengths and weaknesses of the TAPES scales in a way that is different from classical test theory, and decide whether and how to refine the scales.

According to rating scale diagnostics performed using Rasch analysis, one response category of the psychosocial adjustment subscales and two of the satisfaction with the prosthesis subscales did not comply with the set criteria for category functioning (average measures, thresholds, etc.): this suggests that respondents were unable to appreciably discern between them as indicating different levels of frequency. Our collapsing procedures improved the measurement qualities of the scales (without decreasing their reliability indexes), minimizing irrelevant construct variance and ensuring that each rating category represents a clearly distinct level of ability. These findings are in line with previous observations[25,26] and indicate an inherent problem with the use of nonresponse categories, i.e., categories labeled with an indifferent, neutral or undecided phrase; such as "neither agree nor disagree," "neither dissatisfied nor satisfied," "does not apply," "don't know," "no opinion." Such categories are used as a "dumping ground," and it has been argued that, from a measurement perspective, it would be better for the respondents to leave these items blank, rather than provide error-laden ratings.[27] Our findings also indicate a problem with the "traditional" five levels of satisfaction rating (from "very dissatisfied" to "very satisfied"). This drawback has been already addressed by Waugh and Chapman,[28] who pointed out that these categories are not ordered from low to high and thus respondents tend to not respond to the items in a logical and consistent pattern. Our recommended revision of the TAPES (the TAPES-R) is shown in the Appendix. The key differences with regard to rating scales are as follows: (a) for the psychosocial adjustment subscales, a simpler four-point rating scale with the categories "strongly agree," "agree," "disagree," "strongly disagree," and an additional nonscored category, "not applicable"; and (b) for the satisfaction with the prosthesis subscales, a simpler three-point rating scale with the categories "not satisfied," "satisfied," and "very satisfied."

After collapsing the categories, the data were analyzed again to extract Rasch-modeled parameters of ability and difficulty and to face the validity and reliability issues. Close inspection of the most misfitting response strings suggested that the re-

TABLE 2 Subjects' ability levels and reliability indices for the different scales and subscales of TAPES

	PS Subscale 1 (General Adjustment[a])	PS Subscale 2 (Social Adjustment[a])	PS Subscale 3 (Adjustment to Limitation[a])	Activity Restriction	SAT Subscale 1[b]	SAT Subscale 2[b]
Average subjects' ability levels (range)	1.34 (−6.34 to +6.77)	1.41 (−6.81 to +7.04)	−0.87 (−6.22 to +5.14)	0.68 (−6.06 to +5.93)	−0.83 (−4.82 to +4.82)	−0.83 (−5.20 to +5.16)
Person separation index	2.17	1.92	1.98	2.51	1.61	1.83
Person separation reliability	0.82	0.79	0.80	0.86	0.72	0.77
Cronbach's alpha	0.90	0.89	0.86	0.89	0.85	0.86
Range of item difficulty estimates	−0.70 to +1.17	−0.68 to +0.87	−0.76 to +0.83	−3.23 to +3.39	−0.25 to +0.20	−1.17 to +1.07
Item separation index	6.00	5.43	6.91	18.48	0.91	8.32
Item separation reliability	0.97	0.97	0.98	1.00	0.46	0.99
Variance explained by the Rasch factor, %	65	60	64	95	55	62
Eigenvalue of the first residual factor	1.3	1.4	1.4	2.4	1.4	1.8
Variance explained by the first residual factor, %	9	13	10	1	19	13

[a]Psychosocial Adjustment (PS) subscale 1: items 1–5; PS subscale 2: items 6–8 and 10; PS subscale 3: items 11–15.
[b]Satisfaction with prosthesis (SAT) subscale 1: items 1–4; SAT subscale 2: items 5–9.

versed phrasing of the misfitting item in the psychosocial adjustment subscale "social adjustment" (item 9, "I have difficulty in talking about my limb loss in conversation") may have caused confusion for some readers who did not pay sufficient attention to the wording and answered using a "response set behavior."[27] For these reasons, this item is reworded in the revised TAPES-R to omit the necessity for its reverse scoring; "I find it easy to talk about my limb loss in conversation" (the statement is more generic than in item 7 and incorporates the wider experience of limb loss and not simply the artificial limb itself). Finally, we are aware that the psychosocial adjustment subscale "adjustment to limitation" contains items written in the opposite direction compared with the first two subscales and that this choice could have contributed in factor analysis to show a separate factor.[28,29] Nonetheless, it is important to note that compared with the two other psychosocial adjustment subscales, the "adjustment to limitations" subscale is focused more on behaviors and less on attitudes, and therefore was likely to emerge as a different factor irrespective of the direction of the items.

As for the activity restriction scale, the high values of the outfit statistics for items 9 ("maintaining friendships") and 10 ("visiting friends") were due to unexpected responses from a few respondents with very low global activity restriction scores who rated these two activities as limited. We chose to delete the two items because they likely belong to a construct different from "general activity restriction," as already hypothesized by previous factor analyses[5] and confirmed by the present ones. As an additional demonstration of the internal construct validity of the activity restriction scale, the general hierarchic arrangement found by Rasch analysis is consistent with clinical expectations (e.g., see the difficulty order of items such as walking 100 yards, climbing one flight of steps, walking half a mile, and climbing several flights of steps).

Regarding satisfaction with the prosthesis, this analysis confirms the presence of two subscales related respectively to aesthetic (items: 1, color; 2, shape; and 4, appearance) and functional aspects (items 5–9, including the item "weight" that previously constituted a separate factor).

The reliability values indicate that all scales and subscales have acceptable internal consistency and ability to define a distinct hierarchy of persons along the measured construct. Furthermore, all but one (satisfaction with the prosthesis subscale 1) have also a high item separation reliability, which indicates that great confidence can be placed in the replicability of item placement across future samples.[19]

The composite data analysis reported in this article has allowed us to empirically identify a revised version of the TAPES (TAPES-R, see Appendix) that has adequate psychometric characteristics for assessing the complex experience of amputation and adjustment to a lower-limb prosthesis.[30] Importantly, this revised version simplifies the general structure of the TAPES and the interpretation of its scores. The TAPES-R consists of (a) the three psychosocial adjustment subscales with a four-point rating scale (strongly agree, agree, disagree, and strongly disagree) and a reworded item; (b) an activity restriction scale based on ten items with their original three-point rating scale (limited a lot, limited a little, and not at all limited); and (c) two satisfaction with the prosthesis subscales using a three-point rating scale (dissatisfied, satisfied, and very satisfied) and a single overall index of satisfaction with the prosthesis. This overall index provides a more detailed numerical rating scale ranging from 0 to 10 and two verbal anchors at its two limbs ("not at all satisfied" and "very satisfied," respectively).

Although these preliminary results suggest the adequacy of the new instrument, further studies are currently ongoing to analyze the actual performance of the new response structures and to confirm its measurement properties in other samples, thereby adding clinical validity to the instrument. Furthermore, a reasonable next step would be to assess whether the TAPES-R items have local independence and stability of hierarchy across subsamples defined according to relevant clinical criteria.

ACKNOWLEDGMENTS

We thank Prof. Trevor G. Bond for his valuable and constructive comments and suggestions regarding specific aspects of our Rasch analyses. We also acknowledge the researchers who allowed us to pool their data in these analyses.

REFERENCES

1. Ephraim PL, MacKenzie EJ, Wegener ST, Dillingham TR, Pezzin LE: Environmental barriers experienced by amputees: The Craig Hospital Inventory of Environmental Factors—short form. *Arch Phys Med Rehabil* 2006;87:328–33
2. Novotny M: Psychosocial issues affecting rehabilitation. *Prosthetics* 1991;2:373–93
3. Rybarczyk B, Edwards R, Behel J: Diversity in adjustment to a leg amputation: Case illustrations of common themes. *Disabil Rehabil* 2004;26:944–53
4. Condie E, Scott H, Treweek S: Lower limb prosthetic outcome measures: A review of the literature 1995 to 2005. *J Prosthet Orthot* 2006;18:13–45
5. Gallagher P, MacLachlan M: Development and psychometric evaluation of the Trinity Amputation and Prosthesis Experience Scales (TAPES). *Rehabil Psychol* 2000;45:130–54
6. Gallagher P, MacLachlan M: The Trinity Amputation

and Prosthesis Experience Scales and quality of life in people with lower-limb amputation. *Arch Phys Med Rehabil* 2004;85:730–6

7. Andresen EM: Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000;81(suppl 2):S15–20
8. Tesio L: Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation. *J Rehabil Med* 2003;35:105–15
9. Dekker J, Dallmeijer A, Lankhorst G: Clinimetrics in rehabilitation medicine: Current issues in developing and applying measurement instruments 1. *J Rehabil Med* 2005;37:193–201
10. Linacre JM: Investigating rating scale category utility. *J Outcome Meas* 1999;3:103–22
11. Streiner DL, Norman GR (eds): *Health Measurement Scales. A Practical Guide to Their Development and Use*, ed 2. Oxford, Oxford University Press, 1995
12. Bland JM, Altman DG: Cronbach's alpha. *BMJ* 1997; 314:572
13. Coste J, Bouée S, Ecosse E, Leplège A, Pouchot J: Methodological issues in determining the dimensionality of composite health measures using principal component analysis: Case illustration and suggestions for practice. *Qual Life Res* 2005;14:641–54
14. Haley SM, Coster WJ, Andres PL, et al: Activity outcome measurement for postacute care. *Med Care* 2004;42(1 suppl):I49–61
15. Hu LT, Bentler PM: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Eq Model* 1999;6:1–55
16. Schermelleh-Engel K, Moosbrugger H, Muller H: Evaluating the fit of structural equation models: Descriptive goodness-of-fit measures. *Methods Psychol Res Online* 2003;8:23–74
17. Horn JL: A rationale and test for the number of factors in factor analysis. *Psychometrika* 1965;30: 179–85
18. Young, FW. ViSta: The Visual Statistics System, UNC L.L. *Thurstone Psychometric Laboratory Research Memorandum* 1996;94-1
19. Bond TG, Fox CM: *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, Lawrence Erlbaum Associates, 2001
20. Wright BD, Linacre JM: Reasonable mean-square fit values. *Rasch Meas Trans* 1994;8:370
21. Wright BD, Masters GN: *Rating Scale Analysis.* Chicago, Mesa Press, 1982
22. Linacre JM: Sample size and item calibration stability. *Rasch Meas Trans* 1994;7:328
23. Linacre JM: A user's guide to Winsteps. Rasch-Model computer programs. Chicago, IL, 2004. Available at: http://www.winsteps.com/aftp/winsteps.pdf. Accessed May 22, 2009
24. Tennant A, Pallant JF: Unidimensionality matters! (A tale of two Smiths?). *Rasch Meas Trans* 2006;20: 1048–51. Available at: http://www.rasch.org/rmt/rmt201c.htm. Accessed May 22, 2009
25. Lopez W: Communication validity and rating scales. *Rasch Meas Trans* 1996;10:482–3
26. Kyngdon A: Comparing factor analysis and the Rasch model for ordered response categories: An investigation of the scale of gambling choices. *J Appl Meas* 2004;5:398–418
27. Wolfe EW, Smith EV Jr: Instrument development tools and activities for measure validation using Rasch models: Part I—Instrument development tools. *J Appl Meas* 2007;8:97–123
28. Waugh RF, Chapman ES: An analysis of dimensionality using factor analysis (true-score theory) and Rasch measurement: What is the difference? Which method is better? *J Appl Meas* 2005;6:80–99
29. Spector PE, van Katwyk PT, Brannick MT, et al: When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *J Manage* 1997;23:659–77
30. Miller LA, McCay JA: Summary and conclusions from the Academy's sixt state-of-the-science conference on lower limb prosthetic outcome measure. *J Prosth Orthot* 2006;18:2–7

APPENDIX: TAPES-R

In addition to sections on demographic,prosthetic, phantom limb pain, residual limb pain, and other medical problems (see www.tcd.ie/psychoprosthetics), the TAPES-R consists of the following items:

Psychosocial Adjustment	Strongly Disagree	Disagree	Agree	Strongly Agree
a. I have adjusted to having an artificial limb	[]	[]	[]	[]
b. As time goes by, I accept my artificial limb more	[]	[]	[]	[]
c. I feel that I have dealt successfully with this trauma in my life	[]	[]	[]	[]
d. Although I have an artificial limb, my life is full	[]	[]	[]	[]
e. I have gotten used to wearing an artificial limb	[]	[]	[]	[]
f. I don't care if somebody looks at my artificial limb	[]	[]	[]	[]
g. I find it easy to talk about my artificial limb	[]	[]	[]	[]
h. I don't mind people asking about my artificial limb	[]	[]	[]	[]
i. I find it easy to talk about my limb loss in conversation	[]	[]	[]	[]
j. I don't care if somebody notices that I am limping	[]	[]	[]	[]
k. An artificial limb interferes with the ability to do my work	[]	[]	[]	[]
l. Having an artificial limb makes me more dependent on others than I would like to be	[]	[]	[]	[]
m. Having an artificial limb limits the kind of work that I can do	[]	[]	[]	[]
n. Being an amputee means that I can't do what I want to do	[]	[]	[]	[]
o. Having an artificial limb limits the amount of work that I can do	[]	[]	[]	[]

Activity Restriction	Yes, Limited a Lot	Limited a Little	No, Not Limited at All
a. Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	[]	[]	[]
b. Climbing several flights of stairs	[]	[]	[]
c. Running for a bus	[]	[]	[]
d. Sport and recreation	[]	[]	[]
e. Climbing one flight of stairs	[]	[]	[]
f. Walking more than a mile	[]	[]	[]
g. Walking half a mile	[]	[]	[]
h. Walking 100 yards	[]	[]	[]
i. Working on hobbies	[]	[]	[]
j. Going to work	[]	[]	[]

Satisfaction with Prosthesis	Dissatisfied	Satisfied	Very Satisfied
a. Color	[]	[]	[]
b. Shape	[]	[]	[]
c. Appearance	[]	[]	[]
d. Weight	[]	[]	[]
e. Usefulness	[]	[]	[]
f. Reliability	[]	[]	[]
g. Fit	[]	[]	[]
h. Comfort	[]	[]	[]

Please circle the number (0–10) that best describes how satisfied you are with your prosthesis?

0	1	2	3	4	5	6	7	8	9	10
Not at all Satisfied									Very satisfied	