The demand for comprehensive outcome assessments applicable across diagnostic groups and post acute care settings has generated considerable interest in the development of instruments to assess functional health concepts. One such instrument, the Activity Measure for Post-Acute Care (AM-PAC) (Haley et al., 2004b), is a patient-reported measure that includes scales assessing physical functioning and movement, personal and instrumental activities of daily living, and applied cognition. Medicare has identified the AM-PAC as a recommended measure to be used in justifying outpatient therapy cap extensions.

Consistent with current practice, the AM-PAC was developed and calibrated using Rasch measurement theory, which makes strong assumptions about, among other things, the dimensionality of the construct being modeled. In particular, the Rasch model used to calibrate each of the three AM-PAC scales assumes unidimensionality, i.e., that all of the items respond to the same underlying construct. This is a particularly important issue with regard to the AM-PAC Applied Cognition scale, because its item content is fairly heterogeneous, including items related to auditory-verbal communication, reading and writing, and more general cognitive functions such as memory and problem-solving. The developers of the AM-PAC used accepted practice, including factor analysis, Rasch-based item fit statistics, and differential item functioning (DIF) tests to support the unidimensionality of their item pool. However, the patient sample providing the data for these analyses was extremely heterogenous and included patients with a variety of neurological, musculoskeletal, and complex medical conditions. Also, they did not carry out analyses targeted at assessing the practical consequences for person measurement of any multidimensionality present in the item pool.

The purpose of the present study was to evaluate the dimensionality of a set of patient-reported items measuring cognitive and communication function in a group of right-hemisphere (RH) and left-hemisphere (LH) stroke survivors. The item pool was drawn from the Burden of Stroke Scale (BOSS) (Doyle et al., 2004; Doyle et al., 2007) and was similar in size and content to the AM-PAC Applied Cognition short form for outpatients (Haley et al., 2004a) (see Tables 1 and 2). Conducting the analyses on relatively homogenous groups of right and left hemisphere stroke survivors permitted the testing of specific hypotheses about the dimensionality of cognitive and communication functioning relative to side of lesion. We predicted that LH stroke survivors would report more difficulty with communication activities and RH stroke survivors would report more difficulty on general cognition items. We also predicted that separating the cognition and communication items would result in more valid measurement of groups and individuals.

METHOD

A total of 459 stroke survivors participated in the initial BOSS field trial (Doyle et al., 2004) or a subsequent longitudinal field trial (Doyle et al., 2007). The present analyses included data from the BOSS Cognition and Communication domain scales on 176 participants with unilateral left hemisphere stroke and 140 with unilateral right hemisphere stroke. The remainder were excluded because of bilateral, posterior fossa, or undetermined site of lesion.

ANALYSIS and RESULTS

Item-level exploratory factor analyses were conducted using Mplus 1.04. Parallel analyses were conducted on the actual data and on an identically-sized data set simulated under the assumption of a single dimension. Results are presented in Table 3. The high proportion of variance accounted for by the first factor (65.5%), a ratio of first-to-second factor variance accounted for > 4, and a high correlation between factors (0.66) are typically taken as evidence of sufficient unidimensionality to permit application of a unidimensional Rasch model. However, the root mean square residual statistic suggested that a 2-factor solution fit data better, as did comparison with a simulated 1-dimensional data set.

The data were fit to a Rasch Partial Credit model, which permits independent estimation of the rating scale structure for each item. Consistent with Haley et al. (2004a), items with standardized information-weighted fit statistics <-2 or >+2 were excluded. One item (CM4) was excluded for an elevated fit value (5.4). Remaining items obtained acceptable fit values.

Next, differential item functioning (DIF) analysis was conducted. DIF analysis tests whether persons from two groups score differently on each item when they are equated on overall score. We predicted that the cognition items would be more difficult for the RH participants and that communication items would be more difficult for LH participants. Results are presented in Figure 1. After correcting for multiple comparisons, five items (CG3-4, CM1, CM5-6) demonstrated significant DIF, each in the expected direction. When Rasch models were estimated separately for the BOSS Cognition and Communication items, no items showed significant DIF. These analyses suggested that patient-reported cognition and communication may represent distinct constructs for unilateral stroke survivors.

The impact of multidimensionality of item content on person measurement was examined (following Smith, 2002) by generating a Cognition score and a Communication score for each person and cross-plotting them, as presented in Figure 2. After correcting for differences in the local origin between the two scales, independent t-tests were conducted for each score pair (α=0.05), using the ability-dependent model standard errors. If the two scales respond to the same underlying construct, then the scores should be equivalent within measurement error and the number of significant t-tests should be <5%. In fact, a substantial number of t-tests (12.3%, 95%CI =9.2%, 16.4%) was significant, indicating that the multidimensionality in item content among the two scales does disturb the measurement of individuals. Of those individuals demonstrating significantly higher Communication than Cognition scores, 75% (95%CI=51%, 90%) were RH stroke survivors. Of those demonstrating higher Cognition than Communication scores, 92% (95%CI=74%, 98%) were LH stroke survivors.

We also conducted analyses of relative precision for detecting group differences. Of the RH and LH 150 participants who were tested at 3MPO in the BOSS longitudinal field trial, 81 were identified by a speech-language pathologist as having communication impairment, including aphasia, apraxia of speech, and/or dysarthria. We conducted 3 ANOVAs, one each with the Cognition, Communication, Composite (combined) scales as the dependent variable, and with group (communication impairment vs. no communication impairment) as the independent variable. A significant difference was found on each scale, all $p < 0.001$. Indices of relative precision (RP) for each pairwise comparison were calculated by taking the ratios of the relevant F-values. To the extent that the RP ratio for a given comparison exceeds 1, the numerator measure can be said to be more responsive than the denominator measure. The obtained RP-ratios and their bootstrapped 95% confidence intervals are displayed in Table 4. The Communication and Composite scales each demonstrated significantly better responsiveness than the Cognition scale.

A second, overlapping group of 73 participants was identified as having cognitive impairment. We conducted identical ANOVAs and RP analyses using this grouping as the dependent variable. Again, significant differences between groups were found for each scale. As shown in Table 5, all three RP comparisons were significant.

DISCUSSION

While patient-reported cognition and communication may be sufficiently inter-related to permit their measurement as a single construct for certain purposes and populations, they appear to represent distinct constructs among RH and LH stroke survivors. Outcome evaluation in targeted populations with known communication and cognitive disorders such as stroke survivors should be based upon independent unidimensional scales of cognitive and communicative functioning or employ multidimensional scaling methods.

References

Doyle, P. J., McNeil, M., Mikolic, J. M., Prieto, L. R., Hula, W. D., Lustig, A. P. et al. (2004). The Burden of Stroke Scale (BOSS) provided valid and reliable score estimates of functioning and well-being in stroke survivors with and without communication disorders. *Journal of Clinical Epidemiology, 57,* 997-1007.

Doyle, P. J., McNeil, M. R., Bost, J. E., Ross, K. B., Wambaugh, J. L., Hula, W. D. et al. (2007). The Burden of Stroke Scale (BOSS) provided valid, reliable, and responsive score estimates of functioning and well-being during the first year of recovery from stroke. *Qual.Life Res., 16,* 1389-1398.

Haley, S. M., Andres, P. L., Coster, W. J., Kosinski, M., Ni, P., & Jette, A. M. (2004a). Short-form activity measure for post-acute care. *Archives of Physical Medicine and Rehabilitation, 85,* 649-660.

Haley, S. M., Coster, W. J., Andres, P. L., Ludlow, L. H., Ni, P., Bond, T. L. et al. (2004b). Activity outcome measurement for postacute care. *Med.Care, 42,* I49-I61.

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

Table 1. Activity Measure for Post-Acute Care (AM-PAC) Applied Cognition short form for community-dwelling individuals (Haley et al., 2004a). Patients respond to items 1-9 on a 5 or 6-point difficulty scale, and to item 10 on a 6-point assistance scare.

1.	Phone use
2.	Explaining how to do something
3.	Count out correct money
4.	Look up phone number or address
5.	Problem solving of complex tasks
6.	Plan and manage daily routine
7.	Check finances
8.	Read and follow complex instructions
9.	Fill out long written form
10.	Meal preparation

Table 2. Burden of Stroke Scale (BOSS) Cognition and Communication Domain scales.

	Since your stroke, how difficult is it for you to…
Cognition Scale	
	CG1. Concentrate?
	CG 2. Solve a problem?
	CG 3. Remember everyday tasks?
	CG 4. Learn new things?
	CG 5. Remember what people say?
Communication Scale	
	CM1. Talk?
	CM2. Understand what people say to you?
	CM3. Understand what you read?
	CM4. Write a letter?
	CM5. Talk with a group of people?
	CM6. Be understood by others?
	CM7. Find the words you want to say?
Response Scale	
	1. Not at all
	2. A little
	3. Moderately
	4. Very
	5. Cannot do

Table 3. Results of item-level exploratory factor analysis on the 11 BOSS Cognition and Communication items.

	Actual Data	Simulated Data
Variance Accounted for by 1st Factor	65.5%	71.2%
Variance Accounted for by 2nd Factor	10.8%	4.4%
Variance Accounted for by 3rd Factor	4.2%	3.8%
Variance Accounted for by 4th Factor	3.9%	3.6%
Ratio of 1st-to-2nd Factor Variance	6.07	16.2
1st Factor-2nd Factor Correlation in Promax-Rotated 2-factor solution	0.66	0.72
Root Mean Square Residual (RMSR) (< 0.08 indicates acceptable fit)		
1-Factor solution	0.0842	0.0253
2-Factor solution	0.0243	0.0194
3-Factor solution	0.0161	0.0151

Table 4. Relative precision ratios for detecting differences between stroke survivors with and without communication disorders. Ratios reliably greater than 1 ($\alpha=0.05$) are marked with an asterisk (*).

Comparison	Relative Precision (95%CI)
Communication Scale / Cognition Scale	2.02 (1.25, 4.18)*
Communication Scale / Composite Scale	1.17 (0.92, 1.58)
Composite Scale / Cognition Scale	1.72 (1.26, 2.76)*

Table 5. Relative precision ratios for detecting differences between stroke survivors with and without cognitive disorders. Ratios reliably greater than 1 ($\alpha=0.05$) are marked with an asterisk (*).

Comparison	Relative Precision (95%CI)
Cognition Scale / Communication Scale	2.78 (1.29, 10.75)*
Cognition Scale / Composite Scale	1.44 (1.00, 2.38)*
Composite Scale / Communication Scale	1.93 (1.22, 4.67)*

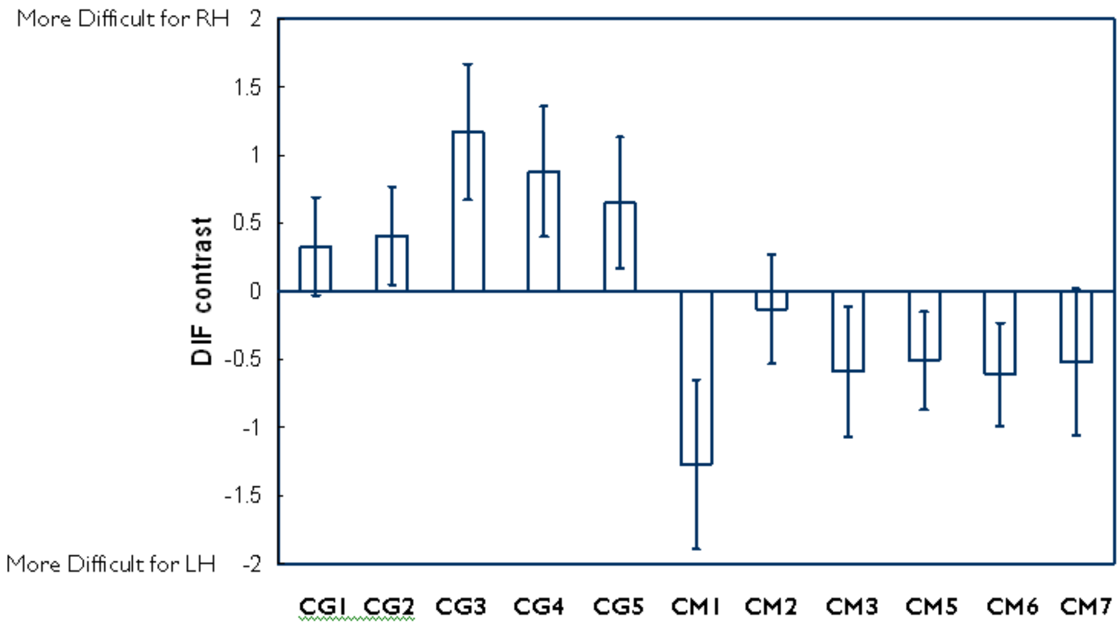

Figure 1. Results of differential item functioning (DIF) analysis. The DIF contrast is the difference in Rasch item difficulty values estimated separately for the two groups. Error bars indicate 95% confidence intervals (uncorrected for multiple comparisons) about the RH-LH differences in item difficulty.

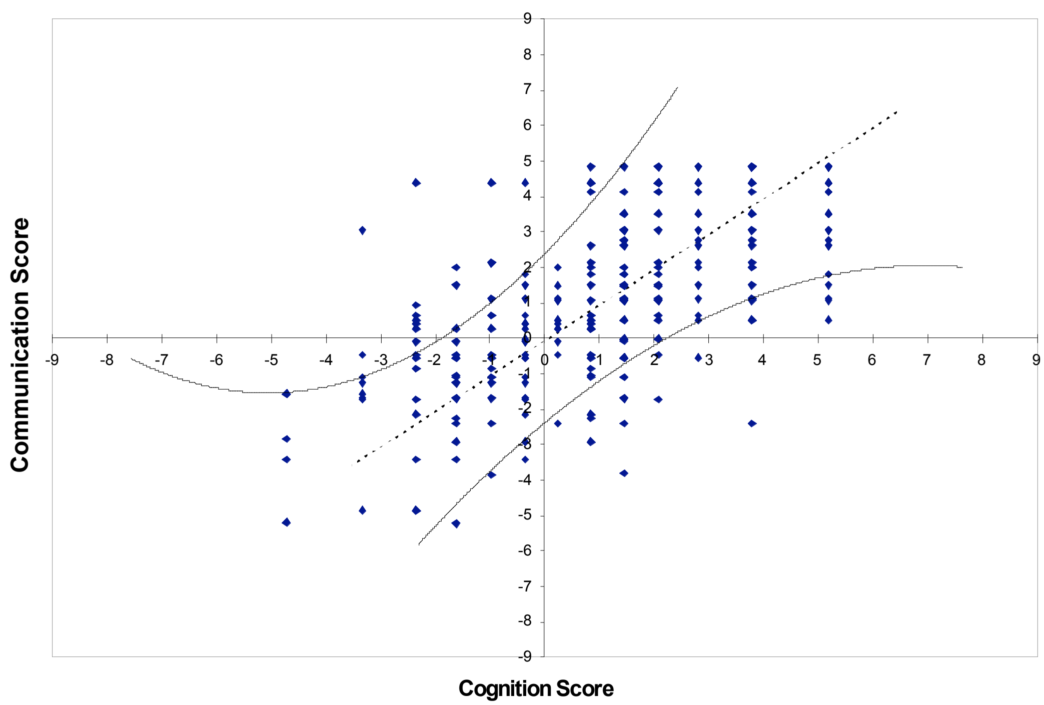

Figure 2. Scatterplot of person scores on the Cognition and Communication scales. The dotted line through the origin represents a perfect match between the two scales. The curved solid lines represent an approximate 95% confidence interval about the identity line, roughly equivalent to the t-tests described in the text. The correlation between the two measures was 0.71, >0.99 when disattentuated for measurement error.