JSLHR

Research Note

A Differential Item Functioning (DIF) Analysis of the Communicative Participation Item Bank (CPIB): Comparing Individuals With Parkinson's Disease From the United States and New Zealand

Carolyn Baylor,[a] Megan J. McAuliffe,[b] Louise E. Hughes,[b] Kathryn Yorkston,[a] Tim Anderson,[c] Jiseon Kim,[a] and Dagmar Amtmann[a]

Purpose: To examine the cross-cultural applicability of the Communicative Participation Item Bank (CPIB) through a comparison of respondents with Parkinson's disease (PD) from the United States and New Zealand.
Method: A total of 428 respondents—218 from the United States and 210 from New Zealand—completed the self-report CPIB and a series of demographic questions. Differential item functioning (DIF) analyses were conducted to examine whether response bias was present across the 2 groups.
Results: No items were identified as having statistically significant DIF across the U.S. and N.Z. cohorts.
Conclusion: The current CPIB items and scoring parameters are also suitable for use with respondents from New Zealand.

Key Words: communicative participation, Parkinson's disease, dysarthria, outcomes measurement, item response theory, cultural and linguistic diversity

Since the adoption of the World Health Organization's International Classification of Functioning, Disability, and Health (ICF; World Health Organization, 2001) in 2001, the concept of participation has become a central focus of health assessment and intervention. The ICF defines participation as "involvement in life situations," which pertains to fulfillment of required and desired roles in the context of real-life situations. In the field of communication disorders, the term *communicative participation* has been used to refer to the communication aspects of involvement in life situations. Specifically, communicative participation has been defined as "taking part in life situations where knowledge, information, ideas or feelings are exchanged" (Eadie et al., 2006, p. 309). Clinical practice for speech-language pathologists (SLPs) has become increasingly focused on how to integrate communicative participation into assessment and intervention programs that help people with communication disorders fulfill their required and desired communication needs. Improved participation is an important indicator of treatment outcomes by demonstrating that intervention has made a meaningful impact on the ability of individuals to do the things they want and need to do. Adopting a participation-focused approach to intervention, however, requires that SLPs be equipped with the assessment and intervention tools needed to deliver high-quality evidence-based practice.

The Communicative Participation Item Bank (CPIB) was developed as a self-report outcome measure of communicative participation (Baylor et al., 2013). It is intended for community-dwelling adults across a range of communication disorders and life situations, although the majority of the development work to date has focused on voice and motor speech disorders. The items ask about the extent to which the respondent's condition interferes with a range of typical speaking situations. Some examples of item content include talking with people the respondent does not know, communicating in a small group of people, or making a phone call to get information. The CPIB was designed as a self-report instrument out of recognition that only the individual living with the health

[a]University of Washington, Seattle
[b]University of Canterbury, Christchurch, New Zealand
[c]New Zealand Brain Research Institute, Christchurch
Correspondence to Carolyn Baylor: cbaylor@u.washington.edu

Editor: Jody Kreiman
Associate Editor: Julie Liss
Received December 21, 2012
Revision received May 2, 2013
Accepted June 13, 2013
DOI: 10.1044/1092-4388(2013/12-0414)

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

condition experiences the unique combination of physical, environmental, and personal influences that shape participation outcomes. Although participation could be assessed from different perspectives, at some point it must include the perspective of the person with the condition (Brown et al., 2004; Law, 2002; Perenboom & Chorus, 2003; Whiteneck, 1994).

The CPIB was developed using item response theory (IRT), a set of statistical methods for instrument development and measurement that has been widely used in other disciplines, such as in educational testing, and is now being incorporated into several speech-language pathology instruments (Baylor et al., 2011; Embretson, 1996; Embretson & Reise, 2000; Reeve et al., 2007). IRT uses mathematical models to explain the relationship among the characteristics or parameters of items in an instrument (e.g., item difficulty and item discrimination), how an individual responds to each item, and that individual's underlying latent trait (i.e., a variable such as ability or attitude that cannot be measured directly). In IRT, the relationship among item parameters, person response, and latent trait is modeled separately for each item as opposed to for the test as a whole. Because this relationship is modeled on an item-by-item basis, individual items can be extracted from the full item bank and presented in smaller subsets through adaptive testing to meet the goals of different assessment situations. Adaptive testing tailors a test for each individual by taking into account the individual's response to a previous item and then choosing a next item that is appropriate to the individual's currently estimated latent trait level. Adaptive testing with IRT item banks can facilitate "measurement efficiency" in that measurement is as precise, if not more precise, than traditional instruments but with far fewer items needed in most assessment situations (Cook, O'Malley, & Roddey, 2005). Further details regarding the background and development of the CPIB are available elsewhere (Baylor et al., 2013; Baylor, Yorkston, Eadie, Miller, & Amtmann, 2009; Yorkston et al., 2008). The full set of item bank items as well as a general short form with scoring instructions are available in Baylor et al. (2013).

During development of any new instrument, one important consideration is that the instrument be free of bias. Respondents who have the same level of the latent trait (e.g., communicative participation) would be expected to respond to items in the same way on the basis of that latent trait (they have the same response probability given a latent trait level; Embretson & Reise, 2000). If bias is present, the relationship between the item and the latent trait varies across different groups. Said another way, participants will answer items differently because of the influence of a variable other than the latent trait, such as membership in a certain group (age, gender, or culture). Presence of bias confounds measurement of the trait of interest.

In considering possible bias in the CPIB, one area of concern is the risk of bias across different communication disorder populations. A second possible source of bias is different cultural backgrounds. As stated by Battle (1993),

> since the roots of communication are embedded in culture, it is logical to assume that one cannot study communication or communication disorders without reference to the cultural, historical, or societal basis of the communication style of the language used by members of the culture. (p. xix)

There is a growing recognition of the need for assessment tools that are applicable across cultures, and thus far, the CPIB has been developed and tested solely on respondents in North America. Evaluating whether response bias occurs cross-culturally is a key step in the development of the CPIB. To investigate this question, we selected for analysis an English-speaking country that differed culturally from the United States. Hence, this study compares data from two groups—respondents from the United States and Aotearoa (i.e., New Zealand),[1] respectively.

New Zealand is a nation of just over four million people located in the southwestern Pacific Ocean. It differs from the United States in both societal composition and mode of delivery of health services. New Zealand is a bicultural society, with the Treaty of Waitangi, or Te Tiriti O Waitangi—which enshrines the principles of protection, partnership, and participation—as its founding document. Approximately 15% of the population identify as Māori, 7% as Pacific peoples, and 9% as Asian, with the remaining 68% identifying as N.Z. European and a small percentage as "other" (Statistics New Zealand, 2006). Given the unique cultural composition of New Zealand, assessment tools that have been validated on this population are sought after. The United States and New Zealand also differ considerably with regards to the provision of health care services. In New Zealand, health care is provided via government funding, and individuals have the option of private funding of health care. SLP services for adults are primarily available through regionally based District Health Boards. These services provide inpatient, outpatient, and/or community-based care (McLellan, McCann, & Worrall, 2011). The differences in societal composition and health care provision across the two countries ensured that a comparative study of the United States and New Zealand was an appropriate starting point for investigation of potential cultural bias in the CPIB.

In IRT, bias is evaluated by using an analysis of differential item functioning (DIF; Reeve et al., 2007). We will briefly summarize the use of DIF in IRT to orient readers as to what the DIF analysis involves and how bias, if detected, can be managed. Before conducting a DIF analysis, researchers need to assess, based on theory or prior research, whether there is reason to be concerned about possible bias (Reeve et al., 2007). In this study, concerns about bias were cultural, but bias could arise for other reasons, such as age, gender, and disorder characteristics, or from many other sources. Once researchers determine qualitatively that there is concern for possible bias, statistical analyses are conducted to identify any items in the item bank that contain statistically significant DIF. There are multiple statistical approaches for identifying DIF that are described elsewhere (Choi, Gibbons, & Crane, 2011; Cook et al., 2011; Crane et al.,

[1] *Aotearoa* is the Māori word for New Zealand.

2007; Reeve et al., 2007). The overarching question in the DIF analyses is whether item parameters, such as item difficulty and item discrimination, are consistent across the groups being evaluated (e.g., across the U.S. and N.Z. samples). DIF is examined on an item-by-item basis, with the results revealing the specific items that demonstrate DIF. If no items are identified as demonstrating statistically significant DIF, the item bank can be used with the groups tested without concern for bias. If items with statistically significant DIF are identified, this suggests that a different set of item parameters are needed for those items for each of the groups tested. If that were to occur in this particular study, each item with DIF would have particular item difficulty and item discrimination values for the U.S. sample and different item difficulty and item discrimination values for the N.Z. sample. Several approaches can be considered to address DIF when it is present (Reeve et al., 2007). First, the researchers may explore whether the DIF makes a meaningful change in practical applications of the item bank. Researchers can compare participants' scores by using a common set of item parameters (not adjusted for DIF) with scores generated using the different group-specific item parameters (scores adjusted for DIF). If the adjusted and nonadjusted scores are highly correlated, it is likely that DIF would have little clinical significance, and it can be ignored. If there is concern about a practical or clinical significance of DIF, there are several options to consider: These include developing separate scoring guides for the different groups so that scoring can be conducted using the group-specific item parameters; developing group-specific short forms or item banks; removing items with DIF from the item bank entirely; or revising and retesting the items with DIF. This last option would be the most difficult in terms of time and labor, requiring extensive new data collection to recalibrate the item bank.

During recent item calibration of the CPIB, a DIF analysis was conducted to compare three diagnostic groups: multiple sclerosis, Parkinson's disease (PD), and head and neck cancer (Baylor et al., 2013). These analyses examined the possibility of bias across disorder groups. Although the presence of statistically significant DIF was identified on several items, a comparison of DIF-adjusted scores with the original unadjusted scores (which did not account for DIF) showed such a high correlation (r = .995) that the DIF was regarded as negligible. The DIF was highly unlikely to produce any clinically meaningful change in scoring and was therefore ignored. These findings supported the use of the CPIB as a "cross-disorder" instrument appropriate for at least those three populations.

The present study continues this line of research to examine the possibility of bias across individuals from different cultures. To minimize potential confounds, we selected a single diagnostic group to compare the CPIB responses of individuals from the United States and New Zealand. Parkinson's disease (PD) was chosen for a number of reasons. First, PD is one of the most common degenerative neurological disorders encountered by SLPs in clinical practice. It affects approximately 1% of people over the age of 60 years, rising to as high as 4% as age approaches 80 years (de Lau & Breteler, 2006). Of those with PD, it is estimated that 50%–89% will develop the speech disorder hypokinetic dysarthria (Hartelius & Svensson, 1994; Johnson & Pring, 1990). Second, there exists a distinct lack of outcome measures that adequately address communicative participation in people with acquired neurological disorders (Eadie et al., 2006). Although there is growing evidence to support behavioral intervention for speech disorders associated with PD, the effect of treatment on communicative participation is not clear. This is particularly important given that individuals with PD are known to be concerned by the effect of speech changes on their ability to communicate and on how they view themselves (Miller, Noble, Jones, & Burn, 2006). Finally, measures designed to assess communicative participation are crucial to the development and evaluation of intervention approaches as well as to the education of those with acquired communication disorders and their families (Yorkston, 2007). Given these issues, PD was selected as an appropriate starting point for a comparison of CPIB responses by those from the United States and New Zealand.

The purpose of this study was to conduct a DIF analysis of the 46-item CPIB item bank in a comparison of PD samples from the United States and New Zealand. The absence of meaningful DIF would support use of the CPIB across these countries without concern for bias. Presence of meaningful DIF would suggest the possible presence of cultural bias in the CPIB and would lead to recommendations for either using a different item set in the two countries or using scoring adjusted for DIF.

Method

The study was approved by the institutional review board at the University of Washington (United States) and the Multiregion Ethics Committee of the Ministry of Health (New Zealand).

Participant Recruitment

Recruitment goals were for 200 participants from each country. Participants met the inclusion criteria if they had been diagnosed with PD at least 3 months prior to participation in the study, they felt that their ability to communicate had been affected by PD, but they still used natural speech for at least some communication. Individuals relying solely on augmentative and alternative communication were not included out of concern that the relevance and acceptability of the CPIB items has not yet been tested in that population. There were no restrictions on treatment history either for the PD in general or speech in particular. Participants were community dwelling. Residents of skilled nursing facilities or other medical facilities were not included because the items in the CPIB target the types of communication situations experienced by community-dwelling adults. The items have not been tested for relevance and acceptability to residents in medically based facilities. Participants needed to be able to provide their own responses and were asked whether they used help either reading the items or marking

their responses. However, participants were not excluded for using help.

For the U.S. sample, the data from the PD participants used in the item calibration of the CPIB were used in this analysis. Detailed information about recruitment and data collection is reported elsewhere (Baylor et al., 2013). Participants were recruited through postings in speech pathology clinics, through dissemination of information about the study to support groups, and through the Washington Parkinson's Disease Registry affiliated with the University of Washington. The N.Z. sample was recruited via Parkinson's New Zealand, postings in speech pathology clinics, and a database kept by the New Zealand Brain Research Institute.

Data Collection

Data collection methods used in New Zealand were identical to those reported by Baylor et al. (2013). All data were collected through self-report questionnaires. Questionnaires were administered either online using the Assessment Center website, available through the NIH Patient-Reported Outcomes Measurement Information System (www.assessmentcenter.net via www.nihpromis.org), or on paper forms that were mailed to participants. Participants were asked to complete a battery of questionnaires covering a range of topics related to self-report of communication disorder symptoms and severity, communicative participation and other psychosocial aspects of living with a communication disorder, global health-related quality of life, PD-specific symptoms, and demographic information. This article focuses on the CPIB DIF analysis with a summary of demographic information.

Before administering the CPIB, we conducted a content review to ensure that the wording was relevant and appropriate to New Zealand. Wording changes were made on three items and are presented in Table 1. A formal copy of the N.Z. version can be obtained from Megan J. McAuliffe at http://www.cmds.canterbury.ac.nz/research/motorspeechdisorders.shtml.

Data Analysis

Descriptive analyses were completed for the demographic data using Microsoft Excel and SPSS. The entire set of 46 CPIB items included in the full item bank were included in the DIF analysis. DIF was analyzed using the Lordif software package in R (Choi et al., 2011). Lordif uses an ordinal logistic regression framework combined with Samejima's (1969) graded response model to examine DIF (Crane, Gibbons, Jolley, & Van Belle, 2006). As mentioned above, different statistics and interpretation criteria can be used for identifying DIF. For this study, two criteria were used to detect meaningful DIF: 5% and 10% changes in the regression coefficient beta (Choi et al., 2011; Crane et al., 2007; Crane, Van Belle, & Larson, 2004). Two other indices that have been cited in the literature were not considered: Chi-square values have been suggested to be too sensitive and to detect DIF that does not lead to meaningful changes in scores, whereas the pseudo-R^2 has been cited as not being adequately sensitive for detecting meaningful DIF (Choi et al., 2011). The two criteria of 5% and 10% changes in the regression coefficient beta appear to be the more moderate indicators, balancing adequate sensitivity with the ability to detect meaningful DIF (Choi et al., 2011).

When statistically significant DIF is identified in any items, the next step is to evaluate whether the DIF would create any clinically meaningful differences (bias) in respondents' scores. The presence of clinically meaningful DIF can be evaluated by recalibrating items to the graded response model by using the DIF-adjusted item parameters and then rescoring the respondents by using the DIF-adjusted item parameters. Then the original scores (that do not account for DIF) can be compared with the DIF-adjusted scores by using a Pearson correlation to examine the association between the two sets of scores (Cook et al., 2011). Strong correlations between the original and DIF-adjusted scores would suggest that adjusting the scores for DIF would make no meaningful clinical change in scores, and DIF could likely be ignored. Weaker correlations between the two sets of scores would suggest that the DIF adjustments do make a meaningful change and should be addressed by removing items or developing scoring that accounts for DIF.

Table 1. Items for which wording was changed between the U.S. and N.Z. versions for relevance and appropriateness for N.Z. populations.

Wording in U.S. version	Wording in N.Z. version
Does your condition interfere with talking to a store clerk who is in a hurry?'	Does your condition interfere with talking to a shop assistant who is in a hurry?
Does your condition interfere with talking with a clerk in a store about a problem with a bill or purchase?	Does your condition interfere with talking with a shop assistant about a problem with a bill or purchase?
Does your condition interfere with having a conversation while riding in a car?	Does your condition interfere with having a conversation while travelling in a car?

Results

Demographic Data

A total of 428 individuals with PD completed the questionnaires: 218 from the United States and 210 from New Zealand. The U.S. sample was more inclined to use the online response format (56.4%), compared with only 2.4% of the N.Z. respondents. Participants from the two countries were well matched on age, time since diagnosis, and the presence of other significant medical conditions. The two samples were also similar on self-ratings of speech severity, with over 60% of each sample indicating that they "sometimes have to repeat words to be understood." The U.S. sample exhibited a relatively even number of men and women, whereas three quarters of the N.Z. respondents were men. Furthermore, a higher proportion of individuals from New Zealand had accessed speech pathology services relative to the U.S. sample. Full details of the demographic data are provided in Table 2.

Table 2. Summary of participant demographics.

Characteristic	United States (*n* = 218)	New Zealand (*n* = 210)
Enrollment and completion		
Total enrolled	318	249
Total completed	218	210
% completion rate	68.6%	84.3%
Survey response format		
Online	123 (56.4%)	5 (2.4%)
Paper	95 (43.6%)	205 (97.6%)
Age (years)		
M (*SD*)	65.9 (10.0)	69.8 (7.8)
Range	43–99	47–89
Gender		
Men	119 (54.6%)	152 (72.4%)
Women	99 (45.4%)	58 (27.6%)
Ethnicity		
Caucasian	207 (95.0%)	
Hispanic	2 (0.9%)	
Black	1 (0.5%)	
N.Z. European or European		203 (96.7%)
Māori		1 (0.5%)
Pacific peoples[a]		2 (1.0%)
Asian	1 (0.5%)	3 (1.9%)
More than one	5 (2.3%)	
Other		1 (0.5%)
What statement best describes your speech?[b]		
Normal	32 (14.7%)	26 (12.4%)
"sounds different but people understand me"	40 (18.3%)	47 (22.4%)
"sometimes have to repeat words to be understood"	140 (64.2%)	131 (62.4%)
"use gesture, writing or drawing to help people understand my speech"	6 (2.8%)	4 (1.9%)
"not understandable"	0 (0.0%)	1 (0.5%)
Time since diagnosis (years)		
M (*SD*)	8.1 (6.2)	8.3 (6.4)
Range	0–45	0–42
Presence of other significant medical conditions	102 (46.8%)	99 (47.1%)
Had received prior speech pathology services	81 (37.2%)	98 (46.7%)

Note. Where totals do not sum to 100%, a participant or participants did not respond to that particular question. Percentages are based on the total number of participants who completed the survey in each country.

[a]Both respondents in the Pacific peoples category reported dual ethnicity: Pacific peoples and Asian and Pacific peoples and N.Z. European, respectively. [b]Self-reported speech severity index is adapted from Cedarbaum et al. (1999). The item was originally developed as part of the ALS Functional Rating Scale.

DIF Results

No items were identified as having statistically significant DIF according to either the 5% or 10% beta change criterion. Because no items were identified as having significant DIF, the additional analyses of comparing original scores to DIF-adjusted scores were not needed.

Discussion

The purpose of this study was to evaluate the potential presence of cultural bias in the CPIB. An IRT DIF analysis compared responses from individuals with PD in the United States and New Zealand. The analysis found no statistically significant DIF on any of the items in the item bank (as modified for the N.Z. version), suggesting both a low risk of response bias across these two countries and that the items and scoring parameters generated for the U.S. sample and reported by Baylor et al. (2013) can also be used for respondents from New Zealand.

The findings support prior research on the CPIB that suggested an absence of clinically meaningful DIF across three different diagnostic groups within the United States: multiple sclerosis, PD, and head and neck cancer (Baylor et al., 2013). Although caution is warranted in that absence of bias across diagnostic groups should not be generalized to absence of bias across cultural groups, converging evidence from both studies provides support for the appropriateness of the CPIB for adults across a wide range of communication disorders, backgrounds, and situations. Two caveats may be particularly relevant to this study. The first relates to ethnicity: New Zealand-based respondents were predominantly Pākehā (non-Māori of European descent), with a considerably lower number of Māori and Pacific peoples responding relative to population averages. Given the central place of communication and storytelling in both cultures and prior research noting that Māori people view certain communication activities as more important than do other participants (Larkins, Worrall, & Hickson, 2004), future use of the CPIB with higher proportion of Māori and Pacific peoples as respondents may highlight differences not identified in the current analysis. The second caveat relates to gender: Although there were commensurate numbers of men and women in the U.S. sample, 72% of the N.Z. sample was male. Because gender is a possible source of bias in measurement in general, this observation should be noted.

In sum, the current study has demonstrated that the CPIB, developed based on responses from participants from North America, is also suitable for use with respondents from New Zealand. The primary limitation of the study is that it was limited to individuals with PD in New Zealand. Future research could strengthen the results by including additional countries in comparative analyses as well as a broader range of people with different communication disorders. Particular attention is also needed toward evaluating possible DIF in populations with language impairments, such as aphasia. DIF may be explored related to other variables as well, such as gender or age. Caution is warranted, therefore, in extrapolating these current results about lack of meaningful DIF to other countries, communication disorders, or demographic characteristics.

Acknowledgments

This study was made possible by funding from the National Institute on Deafness and Other Communication Disorders (Grant 1R03DC010044; principal investigator Carolyn Baylor) and the

New Zealand Institute of Language, Brain and Behaviour. We thank subjects for their time and efforts participating in this study. We would also like to acknowledge student research assistants in the United States who assisted with data collection, Christina Gray and Brittney Skrupky, and the Washington Parkinson's Disease Registry at the University of Washington. Assistance with participant recruitment in New Zealand was provided by Parkinson's New Zealand, the New Zealand Brain Research Institute, and speech therapists from district health boards across New Zealand.

References

Battle, D. E. (1993). *Communication disorders in multicultural populations*. Boston, MA: Andover.

Baylor, C., Hula, W., Donovan, N., Doyle, P., Kendall, D., & Yorkston, K. (2011). An introduction to item response theory and Rasch models for speech-language pathologists. *American Journal of Speech-Language Pathology, 20,* 243–259.

Baylor, C., Yorkston, K., Eadie, T., Kim, J., Chung, H., & Amtmann, D. (2013). The Communicative Participation Item Bank (CPIB): Item bank calibration and development of a disorder-generic short form. *Journal of Speech, Language, and Hearing Research, 56,* 1190–1208.

Baylor, C., Yorkston, K., Eadie, T., Miller, R. M., & Amtmann, D. (2009). Developing the Communicative Participation Item Bank: Rasch analysis results from a spasmodic dysphonia sample. *Journal of Speech, Language, and Hearing Research, 52,* 1302–1320.

Brown, M., Dijkers, M. P. J. M., Gordon, W. A., Ashman, T., Charatz, H., & Cheng, Z. (2004). Participation objective, participation subjective: A measure of participation combining outsider and insider perspectives. *Journal of Head Trauma Rehabilitation, 19,* 459–481.

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., & Nakanishi, A. (1999). The ALSFRS-R: A revised ALS functional rating scale that incorporates assessments of respiratory function. *Journal of the Neurological Sciences, 169,* 13–21.

Choi, S. W., Gibbons, L., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software, 39*(8), 1–30.

Cook, K., Bombardier, C., Bamer, A., Choi, S. W., Kroenke, K., & Fann, J. R. (2011). Do somatic and cognitive symptoms of traumatic brain injury confound depression screening? *Archives of Physical Medicine and Rehabilitation, 92,* 818–823.

Cook, K., O'Malley, K. J., & Roddey, T. S. (2005). Dynamic assessment of health outcomes: Time to let the CAT out of the bag? *Health Services Research, 40*(5, Pt. 2), 1694–1711.

Crane, P. K., Gibbons, L., Jolley, L., & Van Belle, G. (2006). Differential item functioning analysis with ordinal logistic regression techniques: DIF detect and difwithpar. *Medical Care, 44*(11, Suppl. 3), S115–S123.

Crane, P. K., Gibbons, L., Ocepek-Weiklson, K., Cook, K., Cella, D., Narasimhalu, K., … Teresi, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research, 16*(Suppl. 1), 69–84.

Crane, P. K., Van Belle, G., & Larson, E. B. (2004). Test bias in a cognitive test: Differential item functioning in the CASI. *Statistics in Medicine, 23,* 241–256.

de Lau, L., & Breteler, M. (2006). Epidemiology of Parkinson's disease. *Lancet Neurology, 5,* 525–535.

Eadie, T., Yorkston, K., Klasner, E. R., Dudgeon, B. J., Deitz, J., Baylor, C., … Amtmann, D. (2006). Measuring communicative participation: A review of self-report instruments in speech-language pathology. *American Journal of Speech-Language Pathology, 15,* 307–320.

Embretson, S. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Embretson, S., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Hartelius, L., & Svensson, P. (1994). Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: A survey. *Folia Phoniatrica et Logopedica, 46,* 9–17.

Johnson, J., & Pring, T. (1990). Speech therapy and Parkinson's disease: A review and further data. *British Journal of Disorders of Communication, 25,* 183–194.

Larkins, B., Worrall, L., & Hickson, L. (2004). Stakeholder opinion of functional communication activities following traumatic brain injury. *Brain Injury, 18,* 691–706.

Law, M. (2002). Participation in the occupations of everyday life. *American Journal of Occupational Therapy, 56,* 640–649.

McLellan, K., McCann, C., & Worrall, L. (2011). Maori with aphasia: A people without a voice? *New Zealand Medical Journal, 124,* 48–57.

Miller, N., Noble, E., Jones, D., & Burn, D. (2006). Life with communication changes in Parkinson's disease. *Age and Ageing, 35,* 235–239.

Perenboom, R. J. M., & Chorus, A. M. J. (2003). Measuring participation according to the International Classification of Functioning, Disability and Health (ICF). *Disability and Rehabilitation, 25,* 577–587.

Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., … PROMIS Cooperative Group. (2007). Psychometric evaluation and calibration of health-related quality of life item banks. *Medical Care, 45*(5, Suppl. 1), S22–S31.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometric Monograph No. 17). Richmond, VA: Psychometric Society.

Statistics New Zealand. (2006). *2006 census of population and dwellings.* Wellington, New Zealand: Author.

Whiteneck, G. G. (1994). Measuring what matters: Key rehabilitation outcomes. *Archives of Physical Medicine and Rehabilitation, 75,* 1073–1076.

World Health Organization. (2001). *International Classification of Functioning, Disability and Health: ICF.* Geneva, Switzerland: Author.

Yorkston, K. (2007). The degenerative dysarthrias: A window into critical clinical and research issues. *Folia Phoniatrica et Logopedica, 59,* 107–117.

Yorkston, K., Baylor, C., Deitz, J., Dudgeon, B. J., Eadie, T., Miller, R. M., & Amtmann, D. (2008). Developing a scale of communicative participation: A cognitive interviewing study. *Disability and Rehabilitation, 30,* 425–433.