Published in final edited form as:

Aphasiology. 2017 ; 31(8): 861–878. doi:10.1080/02687038.2016.1225274.

Validating the Communicative Participation Item Bank (CPIB) for use with people with aphasia: an analysis of Differential Item Function (DIF)

Carolyn Baylor, PhD, CCC-SLP,

University of Washington, Department of Rehabilitation Medicine, Box 356490, Seattle, WA 98195, 206-221-3563

Megan Oelke, MS, CCC-SLP,

Veterans Affairs Medical Center Puget Sound, University of Washington, Department of Speech and Hearing Sciences; Box 354875, Seattle, WA 98105, 206-685-2140

Alyssa Bamer, MPH,

University of Washington Box 354237, Seattle, WA 98195, 303-953-8085

Eileen Hunsaker, MS, CCC-SLP,

MGH Institute of Health Professions, 36 1st Ave, Charlestown Navy Yard, Boston, MA 02129, 617-724-6847

Catherine Off, PhD, CCC-SLP,

University of Montana, Curry Health Center/CSD/LL, 32 Campus Drive, University of Montana, Missoula, MT 59802, 406-243-2104

Sarah E. Wallace, PhD, CCC-SLP,

Duquesne University, 410 Fisher Hall, Duquesne University, Pittsburgh, PA 15217, 412-396-4129

Suzanne Pennington, MS, CCC-SLP,

MGH Institute of Health Professions, 36 1st Ave, Charlestown Navy Yard, Boston, MA 02129, 617-724-7926

Diane Kendall, PhD, CCC-SLP, and

University of Washington, Veterans Affairs Medical Center Puget Sound, Department of Speech and Hearing Sciences; Box 354875, Seattle, WA 98105, 206-616-0537

Kathryn Yorkston, PhD, BC-NCD

University of Washington, Department of Rehabilitation Medicine, Box 356490, Seattle, WA 98195, 206-543-3345

Abstract

Background: The term ‘communicative participation’ refers to participation in the communication aspects of life roles at home, at work, and in social and leisure situations. Participation in life roles is a key element in biopsychosocial frameworks of health such as the World Health Organization’s (WHO) International Classification of Functioning, Disability and Health (ICF), and the Aphasia Framework for Outcomes Measurement (AFROM). The Communicative Participation Item Bank (CPIB) was developed as a patient-reported measure of communicative participation for adults. Initial validation focused on adults with motor speech or

voice disorders. No prior studies have conducted quantitative validation analyses for the CPIB for persons with aphasia (PWA).

Aims: The primary purpose of this study was to begin validation of the CPIB for PWA by conducting an analysis of differential item functioning (DIF). A DIF analysis was used to identify whether item parameters of the CPIB differed between PWA and the populations used in prior CPIB calibration. Secondary analyses evaluated the level of assistance needed by PWA to complete the CPIB, relationships between the CPIB and a gold-standard patient-reported instrument for PWA - American Speech-Language-Hearing Association Quality of Communication Life Scale (ASHA-QCL), and relationships between PWA and family proxy report on the CPIB.

Methods and procedures: This study included 110 PWA and 90 proxy raters. PWA completed a battery of patient-reported questionnaires in one face-to-face session. Speech-language pathologists (SLPs) provided communication support. Data on aphasia severity from the Western Aphasia Battery - Revised (WAB-R) and demographic data either existed from prior research or were collected during the session. Proxy raters completed a similar battery of self-report questionnaires.

Outcomes and results: Results of the DIF analysis suggested statistically significant DIF on two of the 46 items in the CPIB, but the DIF had essentially no impact on CPIB scores. PWA with WAB-R Aphasia Quotient scores above 80 appeared comfortable reading the CPIB items, although required occasional assistance. Most participants who were unable to complete the CPIB had WAB-R Aphasia Quotient scores lower than 50. Correlation between the CPIB and ASHA-QCL was moderate; and correlation between PWA and proxy scores was low.

Conclusions: Most PWA were able to respond to CPIB items, although most required or requested support. Although these results are preliminary due to a small sample size, the data support that the CPIB may be valid for PWA. Caution is warranted regarding proxy report because of low correlation between PWA and proxy ratings.

Keywords

Aphasia; Communicative participation; Patient-reported outcomes; Quality of life; Proxy

Introduction

Over recent years, biopsychosocial frameworks for health and healthcare have emphasized a holistic approach to assessment and intervention. These frameworks take into consideration not only the impairments associated with a health condition, but also elements of the physical and social environments, as well as personal characteristics that contribute to how each individual experiences the health condition and responds to intervention. The World Health Organization's (WHO) International Classification of Functioning, Disability and Health (ICF) (World Health Organization, 2001) is the biopsychosocial framework most commonly cited, while the Aphasia Framework for Outcomes Measurement (A-FROM) (Kagan et al., 2008) adapts the ICF specifically for people with aphasia (PWA). One key element in these frameworks is the construct of participation, or involvement in life situations. Communicative participation has been defined as taking part in the

communication aspects of life situations. It is the involvement in situations in which information and ideas are exchanged between people (Eadie et al., 2006).

The construct of participation is a promising candidate for outcome measurement because it allows clinicians and researchers to understand how elements of biopsychosocial frameworks - physical, social, and personal variables intersect to influence the targeted endpoint of enhanced participation in life situations (Kagan & Simmons-Mackie, 2007; Mallinson & Hammel, 2010). Furthermore, many of the intervention goals that PWA say are important to them relate to the construct of life participation, thereby increasing the relevance and importance of measuring participation as a clinical outcome (Worrall et al., 2011). Measuring participation from the perspective of PWA is likely to provide valuable insights into consequences of aphasia and efforts to remediate it. This topic has received recent attention in literature reviewing patient-reported participation measures for PWA (Brandenburg, Worrall, Rodriguez, & Bagraith, 2015; Dalemans, de Witte, Lemmens, van dan Heuvel, & Wade, 2008).

The Communicative Participation Item Bank (CPIB) is a patient-reported outcomes instrument that targets the construct of communicative participation (Baylor et al., 2013). The CPIB was developed with the intent of being appropriate for community-dwelling adults in the variety of conversational situations they frequently engage in as part of life roles at home, at work, and in social and leisure situations. All items in the CPIB start with the stem, 'Does your condition interfere with....' followed by various conversational situations such as 'Making a phone call to get information,' or 'Having a conversation in a noisy place.' Respondents choose from four response categories to rate the level of interference they experience in that situation, ranging from 'Not at all,' to 'Very much.' The item bank consists of 46 items, and a paper-and-pencil 10-item disorder-generic short form is available. The CPIB scores are reported as T-scores where 50 = the mean of the calibration sample and standard deviation = 10.

One of the primary goals in developing the CPIB was to have an instrument that could be appropriate for adults with different types of communication disorders. This goal was motivated by prior qualitative research demonstrating that despite very different underlying impairments and communication disorder characteristics, adults across different communication disorders often experience similar restrictions in participating in the communication aspects of life situations such as difficulties communicating in noisy situations, problems keeping up with fast-moving conversations, and challenges communicating on the phone and in other non-face to face communication situations (Baylor, Burns, Eadie, Britton, & Yorkston, 2011; Garcia, Laroche, & Barrette, 2002). Prior research has shown that PWA experience considerable restrictions participating in life situations that involve communication and interaction as demonstrated by evidence of fewer social contacts, fewer social activities, and fewer friendships (Cruice, Worrall, & Hickson, 2006; Dalemans, de Witte, Wade, & van dan Heuvel, 2010; Dalemans, De Witte, Wade, & Van den Heuvel, 2008; Davidson, Howe, Worrall, Hickson, & Togher, 2008). In addition to the language impairment, there are many environmental barriers to communicative participation for PWA including behaviors of some communication partners, physical barriers, and societal barriers, among others (Howe, Worrall, & Hickson, 2008). Prior

research with people with acquired motor speech disorders, such as those associated with Parkinson's disease and multiple sclerosis, as well as people with acquired voice disorders, has shown restrictions in communicative participation similar to those described by PWA. The barriers to participation described by people with motor speech and voice disorders are also similar to those for PWA, including some behaviors of communication partners, physical barriers, and societal attitudes (Baylor, Yorkston, & Eadie, 2005; Miller, Noble, Jones, & Burn, 2006; Walshe & Miller, 2011; Yorkston, Klasner, & Swanson, 2001). Thus, while the overt symptoms of communication disorders, such as word-finding impairments versus imprecise articulation may vary across different communication disorders, growing research evidence, particularly evidence from qualitative research, suggests that the impact of these disorders on communicative participation is very similar across different conditions. Having patient-reported instruments validated to assess communicative participation across different communication disorders would provide tools for quantitative investigations of these observations. Patient-reported instruments that are validated across different communication disorders might facilitate both clinical and research endeavors aimed at developing and implementing interventions that could improve participation outcomes, particularly interventions focused on environmental modifications that benefit people across different communication disorders.

Patient-reported outcomes for PWA have been well-documented in the literature. Examples of existing instruments provide evidence that PWA can and do participate in self-report of their experiences with stroke and aphasia: the Stroke and Aphasia Quality of Life Scale (Hilari, Byng, Lamping, & Smith, 2003), the Burden of Stroke Scale (Doyle, McNeil, Hula, & Mikolic, 2003), the Aphasia Communication Outcome Measure (Doyle et al., 2013), and the Assessment for Living with Aphasia (Simmons-Mackie et al., 2014). However, these instruments have been developed to specifically focus on stroke and aphasia, and therefore do not serve the need for an instrument that can be used across different disorders. The American Speech-Language-Hearing Association Quality of Communication Life Scale (ASHA-QCL; (Paul et al., 2005) could be applicable to different disorders, but the instrument targets quality of life, a broader and more diverse construct than participation in life situations. This focus on quality of life is reflected in items on the ASHA-QCL that pertain more to personal feelings and self-perceptions, in addition to items about engaging in life situations. The CPIB addresses these gaps in existing instruments by targeting the need for an instrument that focuses specifically on the impact of the communication disorder on participation in life situations and is appropriate for use across different communication disorder populations.

Much of the initial work developing the CPIB was conducted with individuals with neurologic motor speech disorders, speech disorders due to head and neck cancer, or voice disorders (Baylor et al., 2014; Baylor, et al., 2013; Baylor, Yorkston, Eadie, Miller, & Amtmann, 2009; Yorkston et al., 2008). Prior research with the CPIB showed the absence of clinically meaningful bias across these different motor speech and voice disorders (Baylor, et al., 2013), as well as lack of bias across different countries for people with Parkinson's disease (Baylor, et al., 2014). PWA were included in cognitive interviews to provide qualitative feedback and guidance during item development (Baylor, et al., 2011). A current gap in the literature, however, is that quantitative validation studies of the CPIB have not yet

been completed with PWA. The purpose of this study was to contribute to the validation of the CPIB for PWA. The CPIB was developed using the statistical methods of Item Response Theory (IRT) (Embretson & Reise, 2000; Reeve, 2004), and the primary aim of this study was to examine whether or not there is bias between PWA and the populations included in prior development of the CPIB through an IRT analysis of differential item functioning (DIF). Secondary analyses further contributed to validation of the CPIB by examining the association between the CPIB and the ASHA-QCL (Paul, et al., 2005). Finally, relationships were explored between the CPIB scores of PWA to those of family proxy participants to provide guidance as to whether or not the CPIB should be used for proxy report by family if the PWA is unable to complete the questionnaire due to severity of language impairment. The specific research questions of this study were:

1. Is there significant and clinically meaningful DIF on the CPIB between PWA and the prior calibration populations?
2. Is the CPIB accessible for PWA and what levels of assistance are needed for PWA to complete the CPIB?
3. What is the correlation between the CPIB and a different instrument validated for PWA – the ASHA-QCL (Paul, et al., 2005)?
4. What is the correlation between CPIB scores for PWA and family / proxy report?

Methods

All methods for the study were initially approved through the Institutional Review Board at the University of Washington. Ethics approval was also obtained through institutions that collaborated for data collection: the University of Montana, Duquesne University, and the MGH Institute of Health Professions. To facilitate understanding by PWA of the consent process, communication supports were used such as repetition, clarification, and aphasia-friendly forms.

Participants

Inclusion criteria for PWA included that they be adults age 18 years or older, proficient in English prior to their stroke, and have a history of aphasia due to left cerebral hemisphere stroke. Time post-onset was six months or more to ensure that the PWA had lived with aphasia long enough to have experienced the impact of aphasia in a variety of life situations. All participants had been diagnosed with aphasia by a certified speech-language pathologist (SLP). No restrictions were placed on type or severity of aphasia, nor on treatment history of aphasia. The goal for this study was to include a wide range of aphasia severities and types to improve representativeness of the sample and therefore applicability of the results for use with the broader population of PWA. The majority of participants (n=63 of the total 110) were recruited and seen through the University of Washington (UW) Aphasia Research Laboratory's Aphasia Registry and Repository. The UW Aphasia Repository contains demographic and aphasia testing data for the participants, including data from Western Aphasia Battery - Revised (WAB-R), the Raven's Progressive Coloured Matrices (Raven, 1998) to rule out cognitive impairment, and screenings to rule out visual neglect and hearing

impairment. Participants at the other data collection sites were also recruited through similar university clinic and research programs with prior assessments in place to confirm diagnosis of aphasia and to rule out cognitive impairment. PWA were paid \$25 for their participation as well as reimbursed for parking, bus fare, or other local travel to attend the research session.

Persons with aphasia were invited to nominate a family member or other unpaid caregiver to participate in the study as a proxy rater. Involvement of a proxy rater was not required for a PWA to enroll in the study. Inclusion criteria for proxy raters were that they were adults age 18 or older with no self-reported history of cognitive impairment, and that they were in contact with the PWA at least twice per week (either in person or via phone) so that they would have an appreciation of the impact of aphasia on the daily life of the PWA. The targeted proxy raters were family members, but participants were allowed to identify their own proxy raters to include others such as close friends with whom they might have more contact than biological family members.

Data collection

Data collection occurred in a single face-to-face session lasting on average 90 minutes with breaks provided to reduce fatigue. The session was held at the university clinics or at participants' homes, except for one session conducted via telemedicine in Montana. All data collection sessions were conducted by licensed, certified SLPs who had clinical or research experience working with PWA, or by master's students in speech-language pathology under the direct supervision of the licensed and certified SLPs. Data were collected through administration of the instruments described in the next sections (except for the Montana location in which only the CPIB, the WAR, and demographic data were collected):

Communicative Participation Item Bank (CPIB):

The CPIB is described in the introduction. All 46 items in the item bank were administered for validation purposes. The CPIB was always administered first. Although this introduced the risk of order effects, this was done because the CPIB was the primary measure for this study, and because of the possibility of fatigue limiting completion of the entire protocol. However, it should be noted that while the full item bank of 46 items needed to be administered for calibration purposes in this study, typically the full item set would not be administered in clinical settings. This is because adaptive formats or short forms such as the existing 10-item generic short form (Baylor, et al., 2013) are preferable and sufficiently accurate. Communication supports were provided to the PWA for the CPIB and other patient-reported outcomes used. These supports included reading items to participants, rephrasing or explaining items, and using multi-modal communication such as gestures, drawings, or writing key words. The SLPs were cautious not to change the meaning of the items through any communication supports; but instead to ensure that the inclusion of writing key words, drawings, or other clarifications conveyed the meaning of the item. When providing communication support, the SLPs discussed the items with the PWA and/or asked for example situations from the PWA to verify that the PWA understood the intended meaning of the items.

Rating level of assistance needed to complete CPIB:

When using patient-reported outcomes measures, the accessibility of the instrument for individuals with language impairments is always of concern. The goal of the assessment is to understand the point of view of the PWA regarding his or her experiences living with the condition, not whether or not they can complete the questionnaire independently. Therefore, language support was provided as needed including reading items to the PWA and providing clarification of the items through explanations, gestures, and/or drawings (Tucker, Connor, Kirchner, Baum, & Edwards, 2008). For this study, the rating scale provided by Tucker et al. (2012) (Tucker, Edwards, Mathews, Baum, & Connor, 2012) was used, with some modification, to categorize the level of assistance needed by the PWA to complete the CPIB. This rating scale is a seven-level scale with 7 = completing the questionnaire without any help to 1 = unable to do even with extensive help. Each step on the scale from “7” to “1” represents an increase in the amount or frequency of assistance required with guidelines provided. For example, 6 = minimal assistance, 2–3 repetitions or explanations needed; 5 = minimal to moderate assistance, 5–6 repetitions or explanations needed, and so forth. One limitation of the scale by Tucker et al. (2012) is that it does not specify how reading the items to participants is accounted for in the levels of assistance. Furthermore, early in this study the researchers noted that some PWA who had the language ability to read the items to themselves requested that the items be read to them as an energy-conservation strategy. To describe the level of assistance needed on the CPIB for this study, the researchers first divided participants into two groups – those who read the items to themselves versus those who required or requested that the items be read to them (this latter group had the items in front of them so that they could follow along while the SLP read the items to them). Then within those categories, the researchers rated each participant on the 1–7 scale in terms of how much other assistance was needed in terms of repeating or clarifying items through speaking, gestures, writing, or drawing. The SLPs at the different data collection sites were trained and calibrated for using this rating scale consistently by completing a series of practice scenarios assigning ratings to different levels of help described in the scenarios. These practice scenarios were developed and scored by the two SLPs responsible for study implementation and data collection at the UW, and the scores of the other participating SLPs were compared to these ratings for calibration. Any discrepancies in ratings were discussed until agreement was reached among the SLPs regarding how to rate different levels of assistance.

Western Aphasia Battery-Revised (WAB-R) Aphasia Quotient:

The WAB-R Aphasia Quotient score, a measure of overall severity of aphasia, was used to compare the level of assistance needed to complete the CPIB to aphasia severity. For most participants, the WAB-R data already existed from prior research or clinical involvement in the university clinics. Existing data were utilized given that participants were in the chronic phase of aphasia, although their registry records were reviewed to ensure that they had not had any major change in medical status since administration of the WAB-R that might signal a significant change in language function. The portions of the WAB-R needed to calculate the Aphasia Quotient were administered as part of the protocol for this study to participants who did not have scores available from prior clinical or research records.

ASHA Quality of Communication Life (ASHA-QCL):

The researchers selected the ASHA-QCL as a ‘gold-standard’ instrument for comparison to the CPIB for several reasons. The ASHA-QCL was developed based on principles of the WHO ICF and was designed to focus on the impact of living with a communication disorder (Bose, McHugh, Schollenberger, & Buchanan, 2009; Paul, et al., 2005; Threats, 2006), as was the CPIB. In addition, like the CPIB, the ASHA-QCL was not developed for any single communication disorder, but refers more generally to communication experiences as opposed to specific speech or language symptoms. Finally, the ASHA-QCL was developed to be accessible for individuals with language impairments in that the format of the instrument consists of briefly-worded item stems with a visual-analog rating scale and pictures for the anchors to assist with interpretation (Paul, et al., 2005). The ASHA-QCL consists of 17 items that ask about different aspects of living with communication disorders as well as a single item at the end, typically scored separately, that asks for a rating of overall quality of life. Although the rating form presents a visual-analog scale, the scoring instructions convert that to a categorical scoring format.

Demographic data:

Demographic data including age, gender, time post-stroke, education level, living situation, marital status, employment status, and race / ethnicity were collected either from existing registry data or directly from the participants.

Proxy raters:

Family members or an unpaid caregiver completed a questionnaire battery. Questionnaires from that battery included in this paper are the 10-item disorder-generic short form of the CPIB and demographic data. Proxy raters completed their questionnaires independently of the PWA either in-person while the PWA was in the data collection session or at home. Proxy raters were not in the session with the PWA so they did not know how the PWA answered the items. When answering the questions, proxy raters were instructed to answer the items according to their perspective on the PWA’s communicative participation, meaning that the proxy raters were not asked to try to guess how they think the PWA would answer the questions, but instead what their view was of the PWA’s participation.

Data analyses

Descriptive statistics were generated for participant demographic data and the levels of assistance required to complete the CPIB. In addition to these descriptive elements, the following analyses were conducted to answer the research questions:

Differential item function (DIF):

In item response theory, DIF analyses are utilized to examine the presence and extent of bias across different groups (Reeve et al., 2007). DIF analyses examine whether item parameters, such as item difficulty and item discrimination, differ between the groups being considered. For this study, the new data collected for PWA were compared to prior data used in calibrating the CPIB (Baylor, et al., 2013). In that prior study, DIF analyses across the populations of Parkinson’s disease, multiple sclerosis, and head and neck cancer revealed

that statistically significant DIF was present on some items across those three populations. However, the impact of DIF on scoring was considered negligible because of the very strong correlation between scores adjusted for DIF and scores not adjusted for DIF. For that reason, the presence of DIF was regarded as likely to have no impact on clinical scoring and therefore could be disregarded (Baylor, et al., 2013). For this study, the data from the groups in the prior study were pooled to be the reference group against which the data from the PWA were compared, based on the finding that there was no meaningful DIF on the construct of interest across those three groups. Furthermore, the item parameters previously published for scoring and interpreting the CPIB are those of the combined group, and it is this set of item parameters that would be used for scoring and interpreting the CPIB for PWA if no meaningful DIF for PWA is found. Summary demographic data for the reference group are presented in Table 1.

Methods used for this study were patterned after those used by Cook et al., (2011) and Cook et al., (2012)(Cook, Bamer, Amtmann, Molton, & Jensen, 2012), (Cook et al., 2011)and in prior investigations of the CPIB (Baylor, et al., 2014; Baylor, et al., 2013). DIF analyses were conducted using the R software package Lordif (Choi, Gibbons, & Crane, 2011). Different criteria can be utilized for assessing the presence of statistically significant DIF, each with varying levels of sensitivity (Choi, et al., 2011). The chi-square values were not considered for this study given prior suggestion that this criterion is overly sensitive in identifying statistical DIF which may not lead to clinically meaningful differences (Choi, et al., 2011). The results for the criteria of change in pseudo- $R^2 < .13$ representing negligible DIF, a 10% change in beta (i.e., standardized regression coefficient), and a 5% change in beta representing significant DIF are reported.

In addition to identifying statistically significant DIF on items, two additional analyses were conducted to examine the extent to which any DIF would lead to meaningful changes in scores on the CPIB. First, again following methods used in prior studies (Cook, et al., 2011) two scores were generated for participants – one score accounting for any statistically significant DIF on items and a second score ignoring any DIF. The correlation between these two scores was examined using Pearson correlations. High correlations between the DIF-adjusted and non-adjusted scores would suggest that adjusting for DIF may not yield any meaningful changes in scores, and thus the presence of DIF might be disregarded. Low correlations would suggest DIF does make a meaningful difference in scoring, and that adjustments might need to be made such as population-specific scoring guides or reconsideration of appropriateness of items for the populations under consideration. Second, the mean difference between DIF-adjusted and non-adjusted scores was calculated for the aphasia participants to evaluate how much scores changed on average due to DIF.

Correlations between the CPIB and ASHA-QCL:

Pearson correlation was utilized to examine the association between the CPIB and the composite score across the first 17 items on the ASHA-QCL. Spearman correlation was utilized to examine the association between the CPIB and the single overall quality of life rating on the ASHA-QCL.

Correlations between PWA and proxy ratings on the CPIB:

Descriptive statistics were utilized to compare the PWA to proxy scores on the CPIB, and Pearson correlation was utilized to examine the association between the two groups of raters.

Results

Participants

Consistent with prior analyses to remove participants with high levels of missing data (Baylor, et al., 2013), data from seven PWA were removed from the data set because they had difficulty responding to the items on the CPIB and were missing responses to more than 10 items. Five of these participants had WAB-R Aphasia Quotient scores lower than 50 suggesting aphasia severity may have impacted their ability to complete the CPIB (although 19 of the 110 participants who remained in the final data set also had WAB-R Aphasia Quotient scores lower than 50). WAB-R data for one participant removed from the analysis was missing, and the final participant had a higher WAB-R Aphasia Quotient score but participated in the session via telemedicine, which might have impacted the ability to provide the support needed to complete the CPIB. These seven participants were from three different data collection sites. After removal of these seven participants, data from 110 PWA were available for the primary analysis of DIF for this study. The number of participants from each data collection location was as follows: University of Washington – 70; Duquesne University – 12; MGH – 24, and University of Montana – 4. Demographic data are presented in Table 2.

Differential Item Function—Using the two criteria of change in pseudo- $R^2 < .13$ and 10% change in beta, no items on the CPIB were identified as having statistically significant DIF between the aphasia and reference group. Using the criterion of a 5% change in beta, two items were identified as having statistically significant DIF. The two items were, ‘... saying something to get someone’s attention,’ and ‘...having a conversation in a noisy place.’ These two items are included in the full set of 46 items, but are not included on the disorder-generic short form for the CPIB. The Pearson correlation between scores on the full item set that were adjusted for DIF and not adjusted for DIF was $r = 1.0$ indicating that adjusting scoring procedures to account for DIF had essentially no impact on the final scores. On average, the CPIB scores that were not adjusted for DIF were 0.7 points (i.e. less than 1 point on the T-score metric) lower than scores adjusted for DIF. Because the impact of DIF on CPIB scoring for PWA could be disregarded due to negligible impact, the remainder of the analyses for this paper were conducted using the original item parameters generated in the CPIB calibration and short form development (Baylor, et al., 2013). For this sample the mean CPIB T-score was 47.3 (SD 8.4) with a range of 27.1 to 66.6. The Pearson correlation between scores based on the full item set and the disorder-generic short form was $r = 0.932$.

Levels of Assistance to Complete the CPIB—Figure 1 presents the distribution of participants across the different levels of assistance required for them to complete the CPIB. Figure 2 presents a graph of the mean and standard deviation of WAB-R Aphasia Quotient scores for participants in each of the different levels of assistance. Two patterns are

suggested in Figure 2. First, it appears the mean WAB-R Aphasia Quotient score was above 80 for participants who read the CPIB items to themselves, regardless of how much support they required on the 7-level support rating scale (although the standard deviation reveals that some PWA with WAB-R Aphasia Quotient scores lower than 80 did read the items to themselves). Second, for the participants who required or requested that items be read to them, there is a clear pattern of increasing levels of independence as mean WAB-R Aphasia Quotient scores increase, although there is considerable overlap across the groups. Two possible conclusions to be drawn from this data are that PWA with WAB-R Aphasia Quotient scores of 80 or higher can likely complete the CPIB with a relatively high level of independence, although some assistance may still be required. In addition, PWA with WAB-R Aphasia Quotient scores below 50 may have significant difficulty completing the CPIB.

Correlation between the CPIB and ASHA-QCL—The Pearson correlation between the CPIB and the ASHA-QCL score that averages responses across the first 17 items was $r = 0.647$ indicating a moderate association between the measure of communicative participation and the broader communication quality of life measure. The Spearman correlation between the CPIB and the single overall quality of life rating on the ASHA-QCL was $\rho = 0.286$ indicating a low correlation between communicative participation and overall quality of life.

Correlation between PWA and proxy ratings on the CPIB—Proxy data were available for 90 PWA. Sixty (54.5%) of the proxy raters were the spouse or partner of the PWA. The second most common category of proxy raters was parents of PWA who represented 13 (11.8%) of the proxy raters. Children of the PWA represented five (4.5%) of the proxy raters, and the same number of proxies were siblings of the PWA. Finally, there were three (2.7%) proxy raters in each of the categories of extended family and close friends. One proxy rater identified as “other,” but did not state their relationship to the PWA. Mean age of proxy raters was 62.1 (SD 12.9) years with a range of 23 to 87 years. The mean CPIB short form T-score for proxy raters was 41.8 (SD 10.2), while the mean CPIB short form T-score for PWA was 47.1 (SD 7.6), revealing that in general PWA rated their own communicative participation as better than the proxy raters did. This difference was statistically significant using a paired t-test ($p < .001$). The Pearson correlation for CPIB scores between PWA and proxy raters was $r = 0.255$. Pearson correlation between PWA and proxy scores on the ASHA-QCL average score was $r = 0.200$. Spearman correlation between PWA and proxy scores on the single ASHA-QCL quality of life item was $\rho = 0.196$. As a group, these results suggest overall low correlations between PWA and proxy raters on measures of both communicative participation and communication-related quality of life.

Discussion

The purpose of this study was to begin validation of the CPIB for PWA by conducting a DIF analysis to assess whether the IRT item parameters and scoring guides developed for the original calibration sample could be used with PWA. The results of this study revealed statistically significant DIF between the original calibration sample and this aphasia sample on only two items in the CPIB (‘...saying something to get someone’s attention,’ and ‘...having a conversation in a noisy place’). However, adjustment in the items or scoring

parameters for PWA does not appear warranted because the presence of DIF made essentially no difference in the scores generated for participants. Therefore, this study supports growing evidence for the applicability of the CPIB and the existing disorder-generic short form across different speech and language disorders (Baylor, et al., 2013). While most participants were able to respond to the CPIB items, participants required a range of assistance levels to complete the CPIB. In general, participants with WAB-R Aphasia Quotient scores higher than 80 appeared able to read the items to themselves with minimal assistance. Gradually increasing amounts of assistance were required with increasing levels of aphasia severity. Seven participants who initially enrolled in the study were unable to complete the task, and five of these had WAB-R Aphasia Quotient scores less than 50. One observation to note is that the amount of assistance required for this study might be more than would be needed in typical uses of the CPIB. For this study, participants had to complete all 46 items in the item bank to have the data for validating the instrument. When using the CPIB in typical clinical settings, however, the full item set is not administered. Either the short form or a (future) computerized adaptive testing format would involve much smaller item sets, thus placing far less burden on the PWA and reducing the risk of fatigue associated with longer questionnaires.

The moderate correlation between the CPIB and the ASHA-QCL overall score is expected given the overlapping but not identical constructs targeted in the two instruments. As mentioned earlier in this paper, the ASHA-QCL was selected as a comparison measure because it focuses on the impact of the communication disorder on life experiences. Many of the items in the ASHA-QCL reflect the construct of participation such as, '... staying in touch with family and friends,' and being, '...included in conversations.' To that extent, the ASHA-QCL and CPIB were expected to overlap. However, quality of life is typically regarded as a multi-dimensional construct encompassing physical, social, and emotional aspects of well-being (Cella et al., 2015). The ASHA-QCL scale functions as a communication quality of life scale in that in addition to participation it includes other dimensions of living with a communication disorder. Questions such as, 'I see the funny things in life,' and 'I like myself,' in the ASHA-QCL address more coping and emotional issues as opposed to life participation. In contrast, the CPIB focuses on a more circumscribed aspect of living with the communication disorder – the impact on participating in life activities. Participation in life roles is one construct that would be expected to contribute to quality of life, hence the expectation of at least a moderate correlation between the two instruments. However, the more specific focus on life participation in the CPIB as opposed to the inclusion of other quality of life domains in the ASHA-QCL would prevent complete overlap of the two measures. The lower correlation between the CPIB and the single ASHA-QCL item that specifically asks for an overall quality of life rating suggests that the constructs of communicative participation and quality of life overlap less when communication is not specifically represented in the quality of life questions. This raises caution against using the constructs of communicative participation and quality of life interchangeably.

The low correlation between the PWA and proxy ratings raises concern about using the CPIB in a proxy format. In related research using other patient-reported measures for PWA, moderate correlations were found between PWA and surrogate ratings using the Aphasia

Communication Outcomes Measurement (ACOM) (Doyle, et al., 2013). Given the associations were moderate at best, Doyle et al., (2013) cautioned against use of proxy ratings for the ACOM. A comparison of PWA and proxy scores on the Stroke and Aphasia Quality of Life (SAQOL) instrument found that differences in scores between proxy and PWA were statistically significant, but with small to moderate effect sizes (Hilari, Owen, & Farrelly, 2007). In that SAQOL study, proxy participants rated quality of life worse than that rated by PWA (Hilari, et al., 2007). Cruice et al., (2005)(Cruice, Worrall, Hickson, & Murison, 2005) found differing relationships between PWA and proxy ratings depending on the construct being measured, but when differences were present the proxy participants rated lived experiences more negatively than did the PWA. In general, it appears that the relationships between PWA and proxy ratings were weaker on the CPIB than have been shown in other studies with other instruments. This might be due to several factors including the nature of the instruments, the particular sample, or how the proxy participants were instructed to respond to the items (i.e., respond according to how you as the proxy perceive the PWA's participation but do not try to guess how the PWA would answer the questions). However, the results of this study were in agreement with prior studies in finding that proxy participants often rate lived experience constructs more negatively than do PWA; and that caution should be utilized in using proxy ratings in place of self-report by PWA.

The primary limitation of this study is the sample size which is lower than recommended for DIF analyses. Ideally, sample sizes might be 200 or more, although 100 is sometimes considered the lower bound of acceptability for DIF (Lai, Teresi, & Gershon, 2005). Given the small sample size, these results should be considered preliminary, although we suggest they are still highly informative for two key reasons. First, as evidenced by the results of this study, the sample size was sufficient to detect DIF on two items. This suggests that the items with the largest DIF were likely identified, and while DIF may be present on other items in the item bank, the magnitude of DIF is likely smaller than that on the two items identified. Second, for practical purposes of utilizing the CPIB for clinic and research purposes, the final consideration is not whether or not there is statistically significant DIF present, but whether or not any DIF would make a meaningful difference in the scoring of the instrument such that adjustments in items or scoring guidelines would be needed. This study demonstrated that the DIF on the two items that were identified would not make any meaningful difference in scoring the CPIB and thus could be disregarded. Similar observations were noted in prior studies (with larger sample sizes) in the development of the CPIB showing that any statistically significant DIF did not lead to meaningful changes in final CPIB scores (Baylor, et al., 2013).

A second limitation of this study is that when administering patient-reported questionnaires to PWA, there is always the risk that comprehension impairments (auditory or reading) would lead PWA to not fully understand the items, or that assistance provided by the examiners might alter the meaning of the items. In this study, the face-to-face interviews allowed the researchers to discuss the responses of PWA with them, which allowed for opportunities to confirm that the PWA understood the items, or to clarify for them if there appeared to be any misunderstandings. This is a risk that has to be balanced with the concerns for patient autonomy and respect for the patient's perspective that can only be captured by including some element of patient-report in clinical assessment and outcomes

measurement. While clinicians and researchers must be mindful of the potential impact of aphasia on patients' abilities to participate in patient-reported outcomes, PWA should not be excluded from patient-reported outcomes because of these risks. Systematic, careful, and reflective administration of patient-reported outcomes by skilled clinicians and researchers should promote the balance of including PWA in patient-reported outcomes to the extent they are able, while ensuring that patient responses reflect the authentic viewpoint of the PWA based on sufficient understanding of the questions.

In conclusion, this study suggests the CPIB, including the disorder-generic short form as developed initially with people with motor speech and voice disorders, is accessible for PWA. SLPs should be prepared to provide support for PWA across severity levels to complete the CPIB, with increasing levels of support needed as aphasia severity increases. While the statistical analyses should be considered preliminary due to sample size, the absence of clinically meaningful DIF in this study provides early evidence of validity of the CPIB for PWA. Caution should be exercised in using the CPIB as a proxy rating by family members given the low correlation between PWA and proxy ratings in this study. Future research should further explore details related to the use of the CPIB with PWA including relationships between communicative participation, other patient-reported constructs, and aphasia characteristics. Future research should also explore the impact of treatment on self-reported participation. Finally, further investigation should address the variables that may influence the differing perspectives of PWA and family members on the experience of living with aphasia.

Acknowledgements

This work was supported by the American Speech-Language-Hearing Foundation under a Clinical Research grant, and by the National Institute for Deafness and Other Communication Disorders under grant 1R03DC010044; PI – Baylor. We would like to thank the students who contributed to the project through data entry. Finally, we express deep appreciation to the individuals with aphasia and their family members and friends who participated in this study.

References

- Baylor C, Burns M, Eadie T, Britton D, & Yorkston K (2011). A qualitative study of interference with communicative participation across communication disorders in adults. *American Journal of Speech-Language Pathology*, 20, 269–287. [PubMed: 21813820]
- Baylor C, McAuliffe M, Hughes L, Yorkston K, Anderson T, Kim J, & Amtmann D (2014). 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. *Journal of Speech Language and Hearing Research*, 57, 90–95.
- Baylor C, Yorkston K, & Eadie T (2005). The consequences of spasmodic dysphonia on communication-related quality of life: A qualitative study of the insider's experiences. *Journal of Communication Disorders*, 38, 395–419. [PubMed: 15963338]
- 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 RM, & 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.
- Bose A, McHugh T, Schollenberger H, & Buchanan L (2009). Measuring quality of life in aphasia: results from two scales. *Aphasiology*, 23(7–8), 797–808.

- Brandenburg C, Worrall L, Rodriguez A, & Bagraith K (2015). Crosswalk of participation self-report measures for aphasia to the ICF: what content is being measured? *Disability and Rehabilitation*, 37(13), 1113–1124. [PubMed: 25163830]
- Cella D, Hahn E, Jensen S, Butt Z, Nowinshi C, Rothrock N, & Lohr K (Eds.). (2015). *Patient-reported outcomes in performance measurement*. Research Triangle Park, NC: RTI Press Publication.
- Choi SW, Gibbons L, & Crane PK (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, Bamer A, Amtmann D, Molton I, & Jensen M (2012). Six patient-reported outcome measurement information system short form measures have negligible age-or diagnosis-related differential item functioning in individuals with disabilities. *Archives of Physical Medicine and Rehabilitation*, 93, 1289–1291. [PubMed: 22386213]
- Cook K, Bombardier C, Bamer A, Choi SW, Kroenke K, & Fann JR (2011). Do somatic and cognitive symptoms of traumatic brain injury confound depression screening? *Archives of Physical Medicine and Rehabilitation*, 92, 818–823. [PubMed: 21530731]
- Cruice M, Worrall L, & Hickson L (2006). Quantifying aphasic people's social lives in the context of non-aphasic peers. *Aphasiology*, 20(12), 1210–1225.
- Cruice M, Worrall L, Hickson L, & Murison R (2005). Measuring quality of life: Comparing family members' and friends' ratings with those of their aphasic partners. *Aphasiology*, 19(2), 111–129.
- Dalemans R, de Witte L, Wade DT, & van dan Heuvel WJA (2010). Social participation through the eyes of people with aphasia. *International Journal of Language and Communication Disorders*, 45(5), 537–550. [PubMed: 19839875]
- Dalemans R, de Witte LP, Lemmens J, van dan Heuvel WJA, & Wade DT (2008). Measures for rating social participation in people with aphasia: a systematic review. *Clinical Rehabilitation*, 22, 542–555. [PubMed: 18511534]
- Dalemans R, De Witte LP, Wade DT, & Van den Heuvel WJA (2008). A description of social participation in working-age persons with aphasia: A review of the literature. *Aphasiology*, 22(10), 1071–1091.
- Davidson B, Howe T, Worrall L, Hickson L, & Togher L (2008). Social participation for older people with aphasia: The impact of communication disability on friendships. *Topics in Stroke Rehabilitation*, 15(4), 325–340. [PubMed: 18782736]
- Doyle P, Hula W, Austermann Hula SN, Stone C, Wambaugh J, Ross K, & Schumacher J (2013). Self and surrogate-reported communication functioning in aphasia. *Quality of Life Research*, 22, 957–967. [PubMed: 22729711]
- Doyle P, McNeil MR, Hula WD, & Mikolic JM (2003). The Burden of Stroke Scale (BOSS): validating patient-reported communication difficulty and associated psychological distress in stroke survivors. *Aphasiology*, 17(3), 291–304.
- Eadie T, Yorkston K, Klasner ER, Dudgeon BJ, 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. [PubMed: 17102143]
- Embretson S, & Reise SP (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Garcia LJ, Laroche C, & Barrette J (2002). Work integration issues go beyond the nature of the communication disorder. *Journal of Communication Disorders*, 35, 187–211. [PubMed: 12036151]
- Hilari K, Byng S, Lamping D, & Smith S (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of Acceptability, Reliability, and Validity. *Stroke*, 34, 1955–1950.
- Hilari K, Owen S, & Farrelly J (2007). Proxy and self-report agreement on the Stroke and Aphasia Quality of Life Scale - 39. *Journal of Neurology Neurosurgery and Psychiatry*, 78, 1072–1075.
- Howe T, Worrall LE, & Hickson LMH (2008). Interviews with people with aphasia: Environmental factors that influence their community participation. *Aphasiology*, 22(10), 1–29.
- Kagan A, & Simmons-Mackie N (2007). Beginning with the end: Outcome-driven assessment and intervention with life participation in mind. *Topics in Language Disorders*, 27(4), 309–317.

- Kagan A, Simmons-Mackie N, Rowland A, Huijbregts M, Shumway E, McEwen S, . . . Sharp S (2008). Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*, 22(3), 258–280.
- Lai J, Teresi JA, & Gershon R (2005). Procedures for the analysis of differential item functioning (DIF) for small sample sizes. *Evaluation and the Health Professions*, 28(3), 283–294. [PubMed: 16123258]
- Mallinson T, & Hammel J (2010). Measurement of participation: intersecting person, task and environment. *Archives of Physical Medicine and Rehabilitation*, 91(Supplement 1), S29–S33. [PubMed: 20801276]
- Miller N, Noble E, Jones D, & Burn D (2006). Life with communication changes in Parkinson's disease. *Age and Ageing*, 35, 235–239. [PubMed: 16540492]
- Paul DR, Frattali C, Holland AL, Thompson CK, Caperton CJ, & Slater SC (2005). *The American Speech-Language-Hearing Association Quality of Communication Life Scale (QCL) Manual*. Rockville, MD: American Speech-Language-Hearing Association.
- Raven JC (1998). *Raven's Coloured Progressive Matrices (CPM)*: Pearson Publishing.
- Reeve B (2004). Applications of Item Response Theory (IRT) modeling for building and evaluating questionnaires measuring patient-reported outcomes. Retrieved January 9, 2008
- Reeve B, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, . . . Cella D (2007). Psychometric evaluation and calibration of health-related quality of life item banks. *Medical Care* 45(5 Suppl 1), S22–S31. [PubMed: 17443115]
- Simmons-Mackie N, Kagan A, Victor J, Carling-Rowland A, Mok A, Hoch J, . . . Streiner D (2014). The assessment for living with aphasia: reliability and construct validity. *International Journal of Speech-Language Pathology*, 16(1), 82–94. [PubMed: 24160320]
- Threats TT (2006). Towards an international framework for communication disorders: use of the ICF. *Journal of Communication Disorders*, 39, 251–265. [PubMed: 16597447]
- Tucker FM, Connor LT, Kirchner LE, Baum C, & Edwards DF (2008). Inclusion of people with aphasia in self-report outcome measures. Paper presented at the American Speech Language Hearing Association Annual Convention, Chicago, IL.
- Tucker FM, Edwards DF, Mathews L, Baum C, & Connor LT (2012). Modifying health outcome measures for people with aphasia. *American Journal of Occupational Therapy*, 66, 42–50. [PubMed: 22251830]
- Walshe M, & Miller N (2011). Living with acquired dysarthria: the speaker's perspective. *Disability and Rehabilitation*, 33(3), 195–203. [PubMed: 20735274]
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. Geneva: World Health Organization.
- Worrall L, Sherratt S, Rogers P, Howe T, Hersh D, Ferguson A, & Davidson B (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, 25(3), 309–322.
- Yorkston K, Baylor C, Deitz J, Dudgeon BJ, Eadie T, Miller RM, & Amtmann D (2008). Developing a scale of communicative participation: a cognitive interviewing study. *Disability and Rehabilitation*, 30(6), 425–433. [PubMed: 17943516]
- Yorkston K, Klasner ER, & Swanson KM (2001). Communication in context: A qualitative study of the experiences of individuals with multiple sclerosis. *American Journal of Speech-Language Pathology*, 10, 126–137.

Figure 1.

Distribution of participants across the levels of assistance required to complete the CPIB. Participants are divided into those who read the items to themselves (light bars) versus those who required or requested that items be read to them (dark bars). A rating of 7 = independent, and a rating of 1 = unable to do even with assistance. Seven participants received a rating of “1” and were excluded from the study. The percentages in this graph are based on the sample of 110 participants remaining for analysis after removal of the 7 who were unable to complete the task. Note the range of the y-axis has been restricted for presentation. Data for the level of assistance ratings are missing for 7 of the 110 participants.

Figure 2.

This figure shows the mean and standard deviation of the WAB-R Aphasia Quotient scores (y-axis) for PWA in each of the categories for levels of assistance to complete the CPIB. On the x-axis, participants are grouped into those who read the items to themselves independently versus those who required or requested that the items be read to them. The level of assistance scale ranges from 1 to 7 with 1 = unable to complete task even with ongoing assistance, and 7 = independent. In this figure, there are no participants at a level of assistance of “1” because those participants were removed from the analysis due to being unable to complete the CPIB. In the group of participants who read the items to themselves, no participants were rated at the assistance levels of “2” or “4” so those points are not represented on the x-axis.

Table 1.

Demographic data for the reference group for the analysis of differential item functioning (DIF). These three diagnostic groups were combined to form the reference group based on a prior study demonstrating no meaningful DIF across these three groups (Baylor, et al., 2013).

	Reference Group for Differential Item Functioning Analysis		
	Parkinson's Disease	Multiple Sclerosis	Head and Neck Cancer
Sample size	218	216	197
Age in years Mean (SD)	65.9 (10.0)	50.0 (9.6)	61.5 (12.3)
Gender	Male = 54.6% Female = 45.4%	Male = 18.1% Female = 81.5%	Male = 61.4% Female = 38.6%
Years post diagnosis Mean (SD)	8.1 (6.2)	12.2 (10.1)	8.4 (8.1)

Table 2.

Demographic information and WAB-R Aphasia Quotient scores for the aphasia sample. Where totals do not sum to 100%, the remainder is due to missing data.

		Mean (SD) Range
Age in years		60.2 (13.3) 22 – 91
Time post-stroke in years		5.5 (4.9) <1 – 29
Western Aphasia Battery- Revised Aphasia Quotient		70.8 (22.2) 19.0 – 99.6
		Frequency (%) (n=110)
Gender	Male	67 (60.9%)
	Female	43 (39.1%)
Level of education	High school graduate	16 (14.5%)
	Vocational / technical school	9 (8.2%)
	Some college	20 (18.2%)
	College graduate	31 (28.2%)
	Post-graduate (Masters; PhD)	29 (26.4%)
Race / ethnicity	Caucasian	87 (79.1%)
	Black	8 (7.3%)
	Asian	7 (6.4%)
	Hispanic	4 (3.6%)
	American Indian / Alaskan Native	2 (1.8%)
	Other	1 (0.9%)
Work status	Currently working for pay	13 (11.8%)
	Currently not working for pay	91(82.7%)
Marital status	Married / committed relationship	63 (57.3%)
	Single	40 (36.4%)
Living situation	With family / spouse	78 (70.9%)
	Alone	22 (20.0%)
	With friends / roommate	2 (1.8%)
	Assisted living	2 (1.8%)