Supplemental Materials 2. Definitions and rationales for DIF grouping variables.

As indicated in the Method section, the purpose of the differential item functioning (DIF) analysis was to identify items whose location estimates may be different for different subgroups of persons with aphasia (PWA). Communication tasks, activities, and situations pervade daily life and occur in a wide variety of cultural and personal contexts. For this reason, self-reports about communicative functioning by PWA are potentially influenced by a large number of clinical and demographic factors that may induce DIF. Definitions for each grouping variable are provided in the table below, and a brief rationale for each follows.

Table S2. Definitions of the covariates used for DIF analyses. Where the covariates referred to function or impairment, the reference group was arbitrarily defined to have worse functioning or more impairment and the focal group was defined to have better functioning or less impairment.

Covariate	Reference group	Focal group
Age	Age ≤ 62, the sample median, *n* = 165	Age > 62, *n* = 164
Aphasia severity	PICA (Porch, 2001) overall raw score ≤ 12.37, the sample median, *n* = 165	PICA (Porch, 2001) overall raw score > 12.37, *n* = 164
Gender	Male, *n* = 214	Female, *n* = 115
Lesion site and number	Single left hemisphere lesion, *n* = 223	Multiple lesions or lesion outside the left hemisphere, *n* = 106
Motor speech diagnosis	Motor speech disorder present, defined by positive diagnosis of dysarthria, apraxia of speech, or neurogenic stuttering, *n* = 158	Motor speech disorder absent; aphasia only, *n* = 171
Self-reported hand function	Impaired, based on self-reported functionality of preferred hand as "non-functional," "limited functionality," or "functional for brief writing tasks (e.g., signature/ address)," *n* = 180	Normal, based on self-reported functionality of preferred hand as "normal function and use for writing," *n* = 149
Self-reported hearing function	*Impaired*, defined by self-reported difficulty hearing in everyday situations or hearing aid use, *n* = 106	*Normal*, defined by self-report of no difficulty hearing in everyday situations and no hearing aid use, *n* = 223
Self-reported vision function	*Impaired*, defined by self-reported difficulty seeing in everyday situations, *n* = 67	*Normal*, defined by self-report of no difficulty seeing in everyday situations, *n* = 262
Time post-onset of aphasia	Months post-onset ≤ 34, the sample median, *n* = 165	Months post-onset > 34, *n* = 164
Veteran status	Non-veteran, *n* = 192	Veteran, *n* = 137

Age: Age is known to affect many aspects of communication (Abrams & Farrell, 2011), including comprehension and production (Fergadiotis & Wright, 2011; Wright, Capilouto, Srinivasan, & Fergadiotis, 2011) and vocabulary and lexical retrieval (Kavé & Nussbaum, 2012; Kavé & Yafé, 2014). It is also reasonable to expect that age may affect individuals' responses to item content having to do with e-mail and computer usage.

Aphasia severity: Although aphasia severity is highly correlated with the general factor measured by the Aphasia Communication Outcome Measure (ACOM), it is possible that individuals with milder and more severe aphasia may interpret item content differently because of differences in how they perceive their own communicative functioning relative to those around them or their own premorbid status. One piece of evidence for this hypothesis comes from our prior work on the ACOM (Doyle et al., 2013) in which we found that persons with more severe aphasia tended to overestimate their functioning relative to responses provided by surrogates, whereas persons with milder aphasia tended to underestimate their functioning relative to surrogates. Also, we note that in their work on the development of the Activity Measure for Post-Acute Care—an item response theory (IRT)–based patient-reported outcome (PRO) measure of physical functioning, activities of daily living, and applied cognition—Haley et al. (2004) and Coster, Haley, Ludlow, Andres, and Ni (2004) found examples of items showing DIF by clinician-rated severity on the modified Rankin Scale.

Gender: Although the existence and nature of gender differences in language processing, language use, and communication remains a matter of debate (Canary & Hause, 1993; Wallentin, 2009), there is are sufficient number of findings of positive differences (e.g., Newman, Groom, Handelman, & Pennebaker, 2008; Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999; Roter, Lipkin, & Korsgaard, 1991) to suggest the possibility of gender-related DIF in the ACOM. We also note that the finding of gender-related DIF is not uncommon in PRO scales targeting health-related constructs (e.g., Haley et al., 2004; Teresi, Ramirez, Lai, & Silver, 2008).

Lesion site and number: We included a test for DIF between participants with aphasia due to single left-hemisphere stroke and participants with multiple and/or right-hemisphere lesions because of the possibility that the responses of the latter group might be influenced by cognitive impairments in addition to aphasia.

Motor speech diagnosis: It is reasonable to hypothesize that concomitant motor speech disorder might cause PWA to report lower levels of functioning on items requiring verbal expression (and possibly correspondingly higher functioning on comprehension items) after conditioning out differences between those with and without motor speech disorder in the underlying latent traits.

Self-reported hand function: We expected that peripheral upper limb motor impairments might affect responses to writing-related items after conditioning out the intended underlying latent traits.

Self-reported vison and hearing function: We expected that peripheral sensory impairments might affect responses to reading and listening comprehension-related items after conditioning out the underlying latent traits.

Time post-onset: As individuals with aphasia recover from and adapt to their communication impairments, their appraisal of their own functioning and their interpretation of item content may change over time. This recovery and adaptation is likely most dynamic in the initial months post-onset, but changes in specific areas of communicative functioning can be induced by treatment in the chronic phases, and relatively little is known about psychological and social adaptations that may occur in the chronic phase (Hersh, 1998). These concerns are related to the concept of response shift as it is discussed in the PRO literature (e.g., Schwartz, 2010), which describes changes in health state, internal standards,

values, or meaning of the target construct that may influence patient reports in ways that compromise psychometric assumptions such as measurement invariance.

Veteran status: We examined DIF by veteran status because we believed that cultural differences between veterans and nonveterans might influence their responses to or interpretations of some item content. For example, veterans presenting to Veterans Health Association (VHA) facilities are routinely asked to identify themselves by their last name and the final four digits of their social security number, so we hypothesized that content related to saying, writing, or spelling names or social security numbers might demonstrate DIF between veterans and nonveterans. Also, we note that lack of measurement invariance between veterans and nonveterans has been demonstrated for two health-related quality of life PRO measures in patients receiving hemodialysis (Saban, Bryant, Reda, Stroupe, & Hynes, 2010). Finally, because this was VA-supported research, it was incumbent upon us to (1) ensure that the tool we are developing functions well for the veteran population, and (2) actively consider veteran-specific issues and concerns in carrying out our research.

References

Abrams, L., & Farrell, M. T. (2011). Language processing in normal aging. In J. Guendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders* (pp. 49–74). New York, NY: Psychology Press.

Canary, D. J., & Hause, K. S. (1993). Is there any reason to research sex differences in communication? *Communication Quarterly*, *41*(2), 129–144.

Coster, W. J., Haley, S. M., Ludlow, L. H., Andres, P. L., & Ni, P. S. (2004). Development of an applied cognition scale to measure rehabilitation outcomes. *Archives of Physical Medicine and Rehabilitation*, *85*(12), 2030–2035.

Doyle, P. J., Hula, W. D., Hula, S. N. A., Stone, C. A., Wambaugh, J. L., Ross, K. B., & Schumacher, J. G. (2013). Self-and surrogate-reported communication functioning in aphasia. *Quality of Life Research*, *22*(5), 957–967.

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*(11), 1414–1430.

Haley, S. M., Coster, W. J., Andres, P. L., Ludlow, L. H., Ni, P., Bond, T. L., . . . Jette, A. M. (2004). Activity outcome measurement for postacute care. *Medical Care*, *42*(1 Suppl.), I49–I61.

Hersh, D. (1998). Beyond the "plateau": Discharge dilemmas in chronic aphasia. *Aphasiology*, *12*(3), 207–218.

Kavé, G., & Nussbaum, S. (2012). Characteristics of noun retrieval in picture descriptions across the adult lifespan. *Aphasiology*, *26*(10), 1238–1249.

Kavé, G., & Yafé, R. (2014). Performance of younger and older adults on tests of word knowledge and word retrieval: Independence or interdependence of skills? *American Journal of Speech-Language Pathology*, *23*(1), 36–45.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, *45*(3), 211–236.

Porch, B. (2001). *Porch Index of Communicative Ability*. Albuquerque, NM: PICA Programs.

Randolph, C., Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Hermann, B. P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology*, *14*(6), 489–496.

Roter, D., Lipkin, M., Jr., & Korsgaard, A. (1991). Sex differences in patients' and physicians' communication during primary care medical visits. *Medical Care*, 1083–1093.

Saban, K. L., Bryant, F. B., Reda, D. J., Stroupe, K. T., & Hynes, D. M. (2010). Measurement invariance of the kidney disease and quality of life instrument (KDQOL-SF) across veterans and non-veterans. *Health and Quality of Life Outcomes*, *8*(1), 120.

Schwartz, C. E. (2010). Applications of response shift theory and methods to participation measurement: A brief history of a young field. *Archives of Physical Medicine and Rehabilitation*, *91*(9), S38–S43.

Teresi, J. A., Ramirez, M., Lai, J.-S., & Silver, S. (2008). Occurrences and sources of differential item functioning (DIF) in patient-reported outcome measures: Description of DIF methods, and review of measures of depression, quality of life and general health. *Psychology Science Quarterly*, *50*(4), 538.

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language*, *108*(3), 175–183.

Wright, H. H., Capilouto, G. J., Srinivasan, C., & Fergadiotis, G. (2011). Story processing ability in cognitively healthy younger and older adults. *Journal of Speech, Language, and Hearing Research*, *54*(3), 900–917.