Work-related measures of Physical and Behavioral Health Function: Test-Retest Reliability

Molly Elizabeth Marino, MPH[1], **Mark Meterko, PhD**[1,2], **Elizabeth E. Marfeo, PhD, OT, MPH**[1], **Christine M. McDonough, PT, PhD**[1], **Alan M. Jette, PT, PhD**[1], **Pengsheng Ni, MD, MPH**[1], **Kara Bogusz, BA**[1], **Elizabeth K. Rasch, PT, PhD**[3], **Diane E. Brandt, PT, MS, PhD**[3], and **Leighton Chan, MD, MPH**[3]

[1]Boston University School of Public Health, Health & Disability Research Institute, Boston, MA

[2]VA Boston Healthcare System, Center for Healthcare Organization and Implementation Research, Boston, MA

[3]National Institutes of Health, Rehabilitation Medicine Department, Mark O. Hatfield Clinical Research Center, Bethesda, MD

Abstract

Background—The Work Disability Functional Assessment Battery (WD-FAB), developed for potential use by the US Social Security Administration to assess work-related function, currently consists of five multi-item scales assessing physical function and four multi-item scales assessing behavioral health function; the WD-FAB scales are administered as Computerized Adaptive Tests (CATs).

Objective—The goal of this study was to evaluate the test-retest reliability of the WD-FAB Physical Function and Behavioral Health CATs.

Methods—We administered the WD-FAB scales twice, 7–10 days apart, to a sample of 376 working age adults and 316 adults with work-disability. Intraclass correlation coefficients were calculated to measure the consistency of the scores between the two administrations. Standard error of measurement (SEM) and minimal detectable change (MDC_{90}) were also calculated to measure the scales precision and sensitivity.

Results—For the Physical Function CAT scales, the ICCs ranged from 0.76–0.89 in the working age adult sample, and 0.77–0.86 in the sample of adults with work-disability. ICCs for the Behavioral Health CAT scales ranged from 0.66–0.70 in the working age adult sample, and 0.77–0.80 in the adults with work-disability. The SEM ranged from 3.25–4.55 for the Physical Function scales and 5.27–6.97 for the Behavioral Health function scales. For all scales in both samples, the MDC_{90} ranged from 7.58–16.27.

Corresponding Author: Molly Marino, MPH, Health & Disability Research Institute, Boston University School of Public Health, 715 Albany Street, T5W, Boston MA, 02118, Phone: 617-638-1990, memarino@bu.edu.

Conflict of Interest: None

Conclusion—Both the Physical Function and Behavioral Health CATs of the WD-FAB demonstrated good test-retest reliability in adults with work-disability and general adult samples, a critical requirement for assessing work related functioning in disability applicants and in other contexts.

Keywords

Disability evaluation; Psychometrics; United States Social Security Administration; computer adaptive test; reliability

Introduction

The United States Social Security Administration's (SSA) disability programs provide financial support to over 10.1 million workers with disability and their families through the SSDI program, and 8.3 million additional individuals through the SSI program.[1] In determining eligibility for work disability benefits, the SSA uses a definition of disability based on a medical model, focusing on symptoms and diagnoses. This definition does not comprehensively address a person's ability to perform tasks and activities required in work environments.[2,3] In conjunction with SSA, we developed new instruments to measure physical and behavioral health function relevant for work that may improve the disability determination process.[4–8]

The Work Disability Functional Assessment Battery (WD-FAB) consists of five Physical Function (PF scales which measure changing and maintaining body positions, upper body function, upper extremity fine motor, whole body mobility, and wheelchair mobility; as well as four Behavioral Health (BH) scales which measure self- efficacy, social interactions, behavioral control, and mood and emotions. PF questions use a five-point response scale from "unable to do" to "no difficulty." Examples of items include: "Are you able to get in and out of bed," "Are you able to write for 20 minutes," and "Are you able to stand for one hour." BH items include two response scales: a four-point response scale from "Never" to "Always," and a five-point scale from "Strongly Agree" to "Strongly Disagree." Examples items include: "I am willing to accept help from others," "When I'm stressed, I can't figure out what to do," and "I can't stop myself from doing the same thing over and over."

The initial instrument development process involved an extensive literature review, the identification of hypothesized key dimensions of functioning, the development of comprehensive item pools refined by expert panel review, and cognitive interviewing with potential respondents.[4] The item pools were then administered to samples of SSA claimants and US adults from the general population. Factor analyses were used to assess the structure of the item pools. Analyses using Item Response Theory (IRT) were conducted to calibrate the items and create quantitative measures for both Physical and Behavioral Health Function.[5–8]

The use of IRT methods to develop functional assessment instruments allows for their administration using computer-adaptive testing (CAT) methods. CAT algorithms tailor administration by selecting the most appropriate subset of items from the larger item bank based on the respondent's answers to previous items. For each FAB scale, a minimum of 5

items are be administered up to a maximum of 8 for PF scales and 10 for BH scales to achieve a reliability score of 0.85. Each CAT generates a score, reported as a T-score on a scale with a mean of 50 and a standard deviation of 10. Higher scores represent higher functioning. The development process provided evidence of measurement validity at several stages, including the positive goodness-of-fit assessment of the confirmatory factor analysis, and finding differences in score distributions in the expected direction between SSA work disability claimants and a general sample of adults living in the US.[5–8]

The WD-FAB scales were subsequently evaluated regarding efficiency of CAT administration, measurement accuracy as tested by person fit, and construct validity in a separate sample of adults unable to work due to a physical (n=497) or mental (n=476) disability.[9] Data quality, CAT efficiency, person fit and concurrent validity (convergent and discriminant) were well supported and suggest that the WD-FAB could be used to assess physical and behavioral health functioning related to work disability.

The purpose of this study was to continue the psychometric evaluation of the WD-FAB scales by examining the test re-test reliability of the Physical Function and Behavioral Health CATs in a sample of working age adults from the general U.S. population and a sample of adults with work-disability. Reliability in this context speaks to the measure's ability to produce consistent scores for a respondent at different points in time. Assessing the reliability of these measures represents another important step toward establishing the feasibility of using the WD-FAB in SSA disability determination and for work disability assessment in other contexts.

Methods

Sample/Participants

There were two samples for the study: 1) working age adults (the "working-age sample", and 2) adults with work-disability (the "work-disability sample".) The samples were provided by a survey research organization that maintains a large (>1 million) panel of voluntary Internet survey participants. Inclusion criteria for both samples included age 21–66 years and, for the adults with work-disability, self-reported inability to work due to a permanent disability. The working age adult sample was matched to a national sampling frame on gender, age, race, and education.[10] The sampling frame was constructed by stratified sampling from the full 2010 American Community Survey sample. Eligible members were identified based upon background information forms completed when opting into the panel, then invited to participate via email. Each email contained a participant specific link to the survey. A second email was sent to those who completed the Time 1 survey one week after completion.

We estimated the sample size required for the study assuming an ICC of .70, and determined that a sample of 238 would yield a 95% confidence interval of sufficient precision (±0.07).[11] This target was rounded up to 300 to allow for possible post-data collection disqualifications for various reasons, including failure to meet age criteria (21–66 years old), failure to complete the second survey within target time frame (7–10 days), self-report of substantial change in health status during the testing period, and uncertainty regarding the response rate

and losses between the first and second administration. Based on previous experience with longitudinal studies using the panel, it was anticipated that 75% would complete both administrations. Thus to ensure 300 completed pairs of administrations, a target of 400 completed surveys was set for the initial (Time 1) administration. The study procedures were the same for the working age and work disability samples.

Procedure

To assess the test-retest reliability of the WD-FAB, we administered two surveys to the same individuals 7 days apart, requiring that respondents complete the second administration within 72 hours of notification. The survey had three components: demographics, health and functional status assessment, and the WD-FAB scales. Demographics included age, sex, racial and ethnic background, zip code, education and relationship status. The adults with work-disability were also asked when they became unable to work due to their disability, and whether they were receiving any type of disability benefits. Participants were also asked whether or not they use a wheelchair, and if they did, whether they used the wheelchair exclusively to get around. As a result, there were respondents who completed both the wheelchair and whole body mobility scales and individuals who completed only the appropriate scale.

Reliability would not be expected to be as high among those respondents whose health or function had changed substantially between the two survey administrations. Therefore, participants were asked "Has your physical health changed a lot in the past week?" and "Has your mental health changed a lot in the past week?" as an indicator of individuals whose WD-FAB scores had changed between the two survey administrations. Three response options allowed participants to indicate whether their physical or mental health had improved, worsened, or stayed the same over the past week.

Finally, all respondents received either 8 or 9 WD-FAB scales (depending on wheelchair use), administered using CAT methodology. This included the five Physical Health Function scales (changing and maintaining body positions, upper body function, upper extremity fine motor, whole body mobility, and wheelchair) and the four Behavioral Health scales (self-efficacy, social interactions, behavioral control, mood and emotions). All contacts and survey administration occurred over the Internet. Ethics approval was obtained from the Boston University institutional review board.

Data analysis

To assess test-retest reliability of each of the WD-FAB scales, we calculated the intraclass correlation coefficient using a two way mixed model (ICC 3, 1) for all respondents who completed the survey at both time points.[12,13] This was done separately for each sample. As sensitivity analyses, we identified respondents who had reported that their health had changed (either improved or worsened). The physical health "changers" were removed from the sample and the ICCs re-computed for the five Physical Function scales; likewise, the mental health "changers" were removed from the sample and the ICCs for the four Behavioral Health Function scales were re-computed. Folded cumulative distribution curves (called mountain plots) were constructed for each FAB scale in each sample in order to

identify any potential systematic increase or decrease in difference scores from Time 1 to Time 2.[14] To construct a mountain plot, we calculated the difference scores between Time 1 and Time 2 for each scale, sorted the difference scores in ascending order, then computed the percentile rank for each difference score. It should be noted that because we "folded" the empirical cumulative distribution plot at the line y=50%, percentile ranks for scores ranked on the second half were calculated using 100 minus the actual percentile rank. Finally, the mountain plot was generated by plotting the percentile rank against the difference score for each scale.

In addition to test-retest reliability, we also calculated two additional measurement properties that are related to WD-FAB reliability. First, we calculated the standard error of measurement (SEM), which quantifies the precision of individual scores on a scale. SEM was calculated as $S_B * \sqrt{(1-ICC)}$ where S_B is the standard deviation at baseline.[15] We then calculated the minimal detectable change (MDC_{90}), which is the minimal threshold for change in an individual's score that would statistically identify real change.[16] MDC_{90} was calculated as $SEM*1.645*\sqrt{2}$ where 1.645 is derived from the 90% CI of no change.

Results

The final samples included 376 adults from the general US population, and adults with work-disability. The working age adult sample was 45% male and 82% white with an average age of 46.6 years (SD 13.2). The adults with work-disability sample was 45% male and 80% white with an average age of 50.1 (SD 9.36) years. The average duration of work disability was 10.5(SD 8.8) years (Table 1). Figure 1 displays the mountain plots showing the magnitude and direction of the difference scores from Time 1 to Time 2 for each WD-FAB scale in all four samples. As indicated by general symmetry observed around the y-axis, there does not appear to be any systematic bias in the direction of those difference scores.

For the WD-FAB Physical Function scales, ICCs ranged from 0.76 to 0.89 in the working age adult sample, and 0.77–0.86 in the adults with work disability sample, with no significant difference in reliability between the two groups (Table 2a). We did not calculate ICC estimates for the wheelchair scale in the working age adult sample due to the small number of such respondents in that group (n=8). For the WD-FAB Behavioral Health scales, ICCs ranged from 0.66 to 0.70 in the working age adult sample, and 0.77 to 0.80 in the adults with work-disability (Table 2b). The scores were more reliable in the sample of adults with work-disability for the self-efficacy, social interactions, and mood and emotions scales.

At Time 2, in the working age adult sample, 17.6% (n=66) self-reported that their physical health had changed "a lot" in the past week, either for the better or the worse, and 21.3% (n=80) reported that their mental health had changed "a lot" in the past week. In the adults with work-disability sample, the corresponding figures were 31.0% (n=98) and 28.8% (n=91) for physical and mental health change respectively. No significant differences between full-sample ICCs and restricted-sample ICCs were observed in either the physical or mental health domain when the "changers" in each domain were removed from the sample as a sensitivity analysis. (Data not shown).

The SEM ranged from 3.25–4.25 in the working age adult sample, and 3.84–4.55 in the adults with work disability sample for the WD-FAB Physical Function scales (Table 2a). For the WD-FAB Behavioral Health scales, the SEM ranged from 6.28–6.97 in the working age adult sample, and 5.27–5.87 in the adults with work-disability (Table 2b). MDC_{90} ranged from 7.58–9.92 for the working age adult sample and 8.97–10.62 in the sample of adults with work-disability for the WD-FAB Physical Function scales; and 14.66–16.27 for the in the working age adult sample and 12.29–13.69 in the sample of adults with work-disability for the WD-FAB Behavioral Health scales.

Discussion

This study assessed the test-retest reliability of Physical Function and Behavioral Health CAT scales developed for use by the Social Security Administration. Our results show that the WD-FAB displayed acceptable test-retest reliability in both a working age adult sample and a sample of adults with work-disability. Other widely used self-reported measures of physical function and behavior health typically display ICC's above 0.7.[16–25] The Physical Function scales demonstrated somewhat better reliability than the Behavioral Health scales particularly in the working age adult sample. Test-retest reliability estimates were more similar between the two samples for the Physical Function scales than the Behavioral Health scales.

The SEM reflects the measurement error in an individual score, with lower SEM's indicating better accuracy. When comparing scores from different respondents on a given scale, the SEM sets a confidence interval (CI) around each score; if the CIs do not overlap we can be confident that the two respondents have different levels of functional ability. The SEM's obtained for the WD-FAB physical function scales all ranged within 3.2–4.5 points' range, indicating good discriminating ability. The behavioral health scales displayed larger SEM's, particularly in the working age adult sample, indicating less ability for the measure to discriminate between two different scores on these scales.

The MDC_{90} can be interpreted as the smallest detectable change that falls outside the measurement error of the instrument, and is therefore primarily a concern when evaluating change. For all scales in both samples, the MDC_{90} ranged from 7.58–16.27. As previously mentioned, the standard deviation of all the scales is 10; thus for several scales a respondent's score would have to change by approximately one standard deviation, a large amount of change, before one could be sure that change had truly occurred.

Some limitations of this study should be noted. We were unable to obtain sufficient sample size to estimate ICC for the wheelchair scale in the working age adult sample. Due to the nature of our sampling, we were also unable to verify work-disability status beyond self-report. The reliability of the WD-FAB scales, though adequate, could be strengthened, particularly in the Behavioral Health domain. We are currently engaged in item replenishment for all domains to better refine the scales. In addition to strengthening the reliability, this item replenishment could also give the scales better discriminating ability and sensitivity to change.

Conclusion

This study provides substantial evidence of the reliability of CAT-based assessments of Physical Function and Behavioral Health Function using the WD-FAB. Although the WD-FAB was initially developed for use by the Social Security Administration, these scales have demonstrated reliability in samples of adults with work disability and working age adults, and therefore could be used for assessment and measurement of work related functioning in other contexts.

Acknowledgments

References

1. Social Security Administration. Annual Statistical Report on the Social Security Disability Insurance Program. SSA Publication No. 13-11826. 2012. Available at http://www.ssa.gov/policy/docs/statcomps/di_asr/2012/di_asr12.pdf

2. Social Security Administration. Social Security handbook. Available at: http://www.ssa.gov/OP_Home/handbook/

3. Brandt DE, Houtenville AJ, Huynh MT, Chan L, Rasch EK. Connecting contemporary paradigms to the social security administration's disability evaluation process. Journal of Disability Policy Studies. 2011 doi:1044207310396509.

4. Marfeo EE, Haley SM, Jette AM, Eisen SV, Ni P, Bogusz K, Rasch EK. Conceptual Foundation for Measures of Physical Function and Behavioral Health Function for Social Security Work Disability Evaluation. Archives of physical medicine and rehabilitation. 2013; 94(9):1645–1652. [PubMed: 23548543]

5. McDonough CM, Jette AM, Ni P, Bogusz K, Marfeo EE, Brandt DE, Rasch EK. Development of a self-report physical function instrument for disability assessment: item pool construction and factor analysis. Archives of physical medicine and rehabilitation. 2013; 94(9):1653–1660. [PubMed: 23542402]

6. Marfeo EE, Ni P, Haley SM, Jette AM, Bogusz K, Meterko M, Rasch EK. Development of an instrument to measure behavioral health function for work disability: item pool construction and factor analysis. Archives of physical medicine and rehabilitation. 2013; 94(9):1670–1678. [PubMed: 23548542]

7. Ni P, McDonough CM, Jette AM, Bogusz K, Marfeo EE, Rasch EK, Chan L. Development of a computer-adaptive physical function instrument for Social Security Administration disability determination. Archives of physical medicine and rehabilitation. 2013; 94(9):1661–1669. [PubMed: 23578594]

8. Marfeo EE, Ni P, Haley SM, Bogusz K, Meterko M, McDonough CM, Jette AM. Scale refinement and initial evaluation of a behavioral health function measurement tool for work disability evaluation. Archives of physical medicine and rehabilitation. 2013; 94(9):1679–1686. [PubMed: 23542404]

9. Meterko M, Marfeo EE, McDonough CM, Jette AM, Ni P, Rasch EK, Brandt DE, Chan L. Work Disability Functional Assessment Battery: Feasibility and Psychometric Properties. Archives of physical medicine and rehabilitation. 2014:11.025. Epub ahead of print.

10. Rubin D. Matched Sampling for Casual Effects. AMC. 2006; 10:12.

11. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Statistics in medicine. 2012; 31(29):3972–3981. [PubMed: 22764084]

12. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological bulletin. 1979; 86(2):420. [PubMed: 18839484]

13. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychological methods. 1996; 1(1):30.

14. Monti KL. Folded empirical distribution function curves—mountain plots. The American Statistician. 1995; 49(4):342–345.

15. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. Journal of clinical epidemiology. 1999; 52(9):861–873. [PubMed: 10529027]

16. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? Journal of biopharmaceutical statistics. 2004; 14(1):97–110. [PubMed: 15027502]

17. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and Reliability of Patient-Reported Outcomes Measurement Information System Instruments in Osteoarthritis. Arthritis care & research. 2013; 65(10):1625–1633. [PubMed: 23592494]

18. Andres PL, Haley SM, Ni PS. Is patient-reported function reliable for monitoring postacute outcomes? American Journal of Physical Medicine & Rehabilitation. 2003; 82(8):614–621. [PubMed: 12872018]

19. Rejeski WJ, Ip EH, Marsh AP, Miller ME, Farmer DF. Measuring disability in older adults: the International Classification System of Functioning, Disability and Health (ICF) framework. Geriatrics & gerontology international. 2008; 8(1):48–54. [PubMed: 18713189]

20. Brazier J1, Harper R, Jones NM, O'cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ: British Medical Journal. 1992; 305(6846):160. [PubMed: 1285753]

21. Bjorner JB, Rose M, Gandek B, Stone AA, Junghaenel DU, Ware JE Jr. Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. Journal of clinical epidemiology. 2014; 67(1):108–113. [PubMed: 24262772]

22. Veit CT, Ware JE. The structure of psychological distress and well-being in general populations. Journal of consulting and clinical psychology. 1983; 51(5):730. [PubMed: 6630688]

23. Andresen EM. Criteria for assessing the tools of disability outcomes research. Archives of physical medicine and rehabilitation. 2000; 81:S15–S20. [PubMed: 11128900]

24. Fitzpatrick R, Davey C, Buston MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technology Assessment. 1998; 2(14)

25. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Medical care. 2003; 41(5):582–592. [PubMed: 12719681]
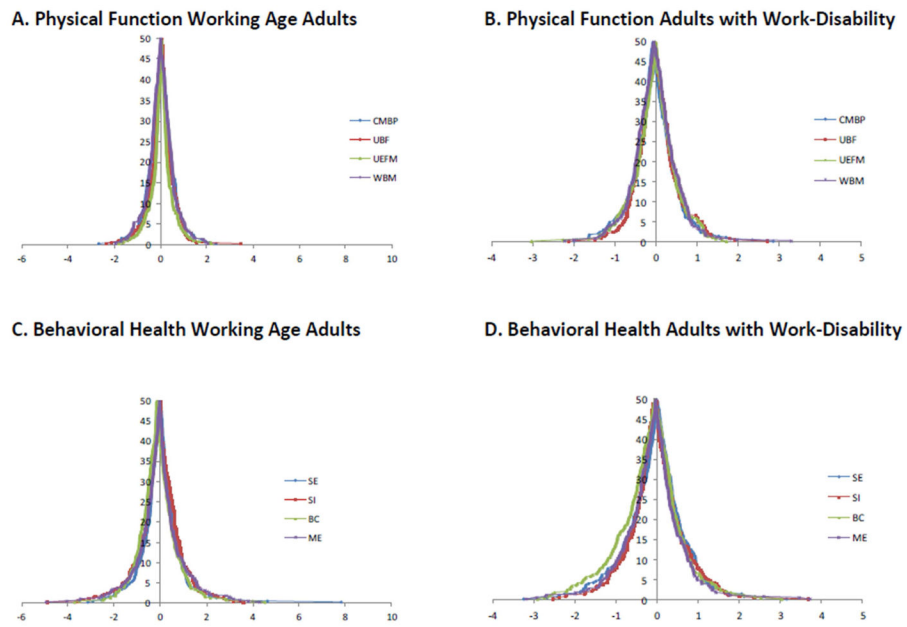
Figure 1.
Mountain plot for direction of changes in CAT scores from Time 1 to Time 2 by sample. Changing & Maintaining Body Position (CMBP), Upper Body Function (UBF), Upper Extremity Fine Motor (UEFM), Whole Body Mobility (WBM) Self-Efficacy (SE), Social Interaction (SI), Behavioral Control (BC), Mood and Emotions (ME)

Table 1

Respondent Demographics

Demographic Characteristics	Working Age Adults (n=376)		Adults with Work-Disability (n=316)	
	Mean	SD	Mean	SD
Age (Years)	46.25	13.19	52.09	9.36
Work Disability Duration (Years)	-	-	10.51	8.75
	Count	Percent	Count	Percent
Sex				
Male	171	45.48	141	44.62
Race				
White	309	82.10	252	79.75
Black/African American	30	7.98	28	8.86
Other or Multiple	32	8.51	32	10.13
Missing	5	1.33	4	1.27
Hispanic				
Yes	27	7.18	24	7.59
No	347	92.29	289	91.46
Refused	2	0.53	3	0.95
Education				
High school diploma or less	111	29.52	109	34.49
Associate degree; vocational training	65	17.29	50	15.82
Some college	67	17.82	94	29.75
College graduate or more	133	35.37	63	19.94
Relationship status				
Never married	97	25.80	62	19.62
Married or living with partner	226	60.11	145	45.89
Divorced, Separated or Widowed	51	13.56	108	34.18
Refused	2	0.53	1	0.32
Change in past week: Physical Health				
No change	305	81.12	217	68.67
Yes, got better	42	11.17	22	6.96
Yes, got worse	24	6.38	76	24.05
Missing	5	1.33	1	0.32
Change in past week: Mental Health				
No change	293	77.93	224	70.89
Yes, got better	58	15.43	39	12.34
Yes, got worse	22	5.85	52	16.46
Missing	3	0.80	1	0.32

Table 2

Reliability of WD-FAB Scales in Two Samples

a. Physical Function

a. Physical Function	Working Age Adults						Adults with Work-Disability					
	n	Mean (sd)	ICC (3,1)	95% CI	SEM	MDC$_{90}$	n	Mean (sd)	ICC (3,1)	95% CI	SEM	MDC$_{90}$
Changing & Maintaining Body Position	376	50.85 (9.17)	0.88	0.85 0.90	4.24	9.88	316	35.68 (8.56)	0.86	0.83 0.89	4.23	9.88
Upper Body Function	376	51.10 (8.11)	0.82	0.78 0.85	3.87	9.03	316	35.58 (7.70)	0.80	0.76 0.84	3.84	8.97
Upper Extremity Fine Motor	376	52.99 (5.92)	0.76	0.71 0.80	3.25	7.58	316	44.28 (7.85)	0.77	0.72 0.81	4.23	9.87
Whole Body Mobility	371[*]	50.31 (8.74)	0.89	0.86 0.91	4.25	9.92	297[*]	34.14 (7.03)	0.83	0.79 0.86	4.18	9.76
Wheelchair[**]							57	44.98 (8.50)	0.79	0.66 0.87	4.55	10.62

b. Behavioral Function

b. Behavioral Function	Working Age Adults						Adults with Work-Disability					
	n	Mean (sd)	ICC (3,1)	95% CI	SEM	MDC$_{90}$	n	Mean (sd)	ICC (3,1)	95% CI	SEM	MDC$_{90}$
Self-Efficacy	376	48.50 (12.71)	0.66	0.60 0.72	6.47	15.11	316	41.27 (14.36)	0.80	0.76 0.84	5.66	13.21
Social Interactions	376	52.20 (11.0)	0.66	0.60 0.71	6.91	16.12	316	39.52 (10.24)	0.77	0.72 0.81	5.27	12.29
Behavioral Control	376	48.14 (11.5)	0.70	0.64 0.75	6.28	14.66	316	46.83 (12.20)	0.78	0.72 0.81	5.87	13.69
Mood and Emotions	376	49.52 (13.40)	0.68	0.62 0.73	6.97	16.27	316	39.51 (12.62)	0.78	0.74 0.82	5.42	12.65

[*] Respondents who always used a Wheelchair did not complete Whole Body Mobility scale.

[**] ICC for Wheelchair scale not calculated in the general US adult sample due to small sample size (n=8).