Rasch analyses of the Wheelchair Use Confidence Scale

Brodie M. Sakakibara, PhD^{a,b}, William C. Miller, PhD, FCAOT^{a,b,c}, and Paula W. Rushton, PhD^d

^aGraduate Program in Rehabilitation Sciences, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

^bRehabilitation Research Program, Vancouver Coastal Health Research Institute, Vancouver, British Columbia, Canada

^cDepartment of Occupational Sciences and Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

^dSchool of Rehabilitation, Faculty of Medicine, University of Montréal, Québec, Canada

Abstract

Objective—To compare the functioning of the Wheelchair Use Confidence Scale's 101-point response format with shortened 11-point formats, and to evaluate the scale's measurement properties using principal components and Rasch analyses.

Design—Secondary analysis of cross-sectional data.

Setting—Community.

Participants—Volunteer participants from British Columbia, Ontario, Quebec, and Nova Scotia, Canada were manual wheelchair-users (n=220), 19 years of age, with 6 months experience with daily wheelchair-use, and no cognitive impairment.

Intervention—None.

Measurements—65-item Wheelchair Use Confidence Scale (WheelCon).

Results—The 11-point response format outperformed the original 101-point format. Principal components analyses confirmed the presence of two dimensions: 1) Mobility efficacy; and 2) Self-management efficacy. Thirteen items in the Mobility efficacy subscale, and eight items in the Self-management efficacy subscale fit the Rasch Rating Scale model. Five items misfit the model developed using the 21-items from both subscales. In each of the 13- and 8-item subscales, and the 21-item short form, the two lowest and highest scores had internal consistency reliability estimates below 0.70; all other scores had reliability estimates above 0.70.

Conclusion—The WheelCon is comprised of two dimensions. The recoded measurements using a 0 to 10 response scale from the 13-item mobility and 8-item self-management efficacy subscales have good reliability as do the measurements from the 21-item WheelCon Short Form. The use of

the subscales and/or the short form depends on the context in which they are being considered. Research to establish the reliability and validity of the measurements using the 0 to 10 response format is warranted.

Keywords

self-efficacy; wheelchairs; measurement; Rasch analyses; principal components analyses

The Wheelchair Use Confidence Scale (WheelCon)¹ is a newly developed measurement scale that assesses self-efficacy with manual wheelchair-use in six conceptual areas including: 1) the physical environment (34 items); 2) activities performed (11 items); 3) knowledge and problem solving (8 items); 4) advocacy (4 items); 5) social situations (7 items); 6) and emotions (1 item).¹ For each item, individuals are asked, “As of now, how confident are you ...” Each item is rated on a 0 (not confident) to 100 (completely confident) point response scale. A mean score is calculated with higher scores indicating higher self-efficacy. Evidence using Classical Test Theory methods indicates that the WheelCon measurements are both reliable and valid in manual wheelchair-users.² Furthermore, research using the WheelCon has shown that lower self-efficacy is associated with lower levels of participation frequency^{3,4} and life-space mobility⁵ in wheelchair-users aged 50 and older.

Despite foundational research supporting the WheelCon’s measurement properties, and the scale’s effective use in research and clinical settings,^{2–9} further evaluation of the WheelCon’s measurement properties using contemporary measurement methods (i.e., Rasch analyses) may address current limitations with the existing scale. For example, during the development of the WheelCon,¹ items were created that inquire about beliefs pertaining to both physical (e.g. move over carpet) and non-physical (e.g. problem solve how to get to a destination when there is an unexpected detour) abilities to use a wheelchair. Although these items are conceptually different, they are used together to derive a total score. Quantitative investigation into the WheelCon’s dimensionality through the use of factor analysis or item response modeling has yet to be performed. Next, because the WheelCon measurements have high internal consistency reliability (Cronbach alpha=0.92),⁶ there may be redundant items that could be eliminated. Identifying and eliminating redundant items will reduce the administrative and responded burden, which may in turn increase the WheelCon’s use, particularly in the clinical setting. Furthermore, the use of a 101-point response format with the WheelCon may be resulting in biased measurement. Evidence shows that individuals, especially older adults, have difficulty using response formats with many response options.^{11,12} An inability to differentiate between adjacent options leads to respondent bias arising from increases in subjectivity.¹³ A shortened response format, therefore, may improve measurement precision relative to the original format and may be tested using item-response procedures.

The purpose of this study is to compare the functioning of the WheelCon’s 101-point response format to shortened response formats, examine the dimensionality of the WheelCon, identify items not conforming to the Rasch Rating Scale model as well as

redundant items that could be considered for elimination, and determine the standard errors of measurement and reliability estimates for the entire range of scores.

It is hypothesized that a shortened response format will outperform the original response format. It is also hypothesized that the WheelCon with a shortened response format will result in more than one dimension with fewer items and good internal consistency reliability (Cronbach alpha 0.70) throughout the range of scores.

METHODS

Study Design and Participants

Data from this cross-sectional study (study 1; n=146) were combined with data from another study (study 2; n=74)² using the same inclusion/exclusion criteria: community-dwelling manual wheelchair-users; 19 years of age; 6 months experience with daily wheelchair-use; communicates in English or French; and no cognitive impairment (i.e. score ≥ 23 on the Mini-Mental State Examination).¹⁴ The samples from both studies were comprised of volunteer participants drawn from British Columbia, Ontario, Quebec, and Nova Scotia, Canada between 2010 and 2012.

Participants were recruited using letters of information sent by clinicians and/or wheelchair vendors. Study information was also provided to advocacy and community groups. Individuals who consented to participate met with a trained research assistant to complete the questionnaires. Participants also had the option to complete the English questionnaires online. The ethics boards from all relevant institutions approved this study.

The Wheelchair Use Confidence Scale

Slightly different versions of the WheelCon were used in studies 1 and 2. The 65-item version 3.0 was used in study 1, and the 63-item version 2.4 was used in study 2². In this study, the 60 similar items in both versions of the WheelCon were retained for analyses, as were the 5 new items in the WheelCon version 3.0.

Response format category collapsing

Bandura advocates the use of either 0 to 100 or 0 to 10 point response formats with self-efficacy measures.¹⁵ In order to compare the functioning of the original 101-point response format with shorter 11-point response formats, the original dataset was rescored using two modified 11-point response formats (i.e. 0 was retained as a distinct option in one format: 0 = 0...91–100 = 10 and 100 was retained as a distinct option in the other format: 0–9 = 0...100 = 10) and the functioning of each was analyzed individually.

Data analyses

Descriptive statistics were used to characterize the sample. Results from categorical variables were calculated as percentages and continuous variables as means and standard deviations. The Barthel Index¹⁶ was used to assess functional independence.

Response format functioning—Two criteria were used to evaluate the response format functioning of each of the 65 items in the WheelCon. First, the Rasch Rating Scale Model's¹⁷ item-by-response option outfit statistics were examined. Outfit statistics greater than 2.0 were considered misfitting and an indication that the response option was being used in an unexpected manner.¹⁸ Second, average measure values were inspected to determine the extent to which the response options were properly ordered (i.e. higher response options should manifest higher self-efficacy).¹⁸ The format with the fewest misfitting and unordered response options was used in subsequent analyses.

Evaluating Rasch assumptions

Unidimensionality: Two separate Principal Components Analysis (PCA) with varimax rotation¹⁹ were used to investigate the dimensionality of the *60 similar items* in the two WheelCon versions, as well as the 5 new items in the WheelCon version 3.0. The use of varimax rotation was selected because the items in the WheelCon assess beliefs pertaining to both physical and non-physical abilities to use a wheelchair. Although similar in that self-efficacy is being assessed, conceptually, the beliefs are different, and independent, situation specific dimensions. Prior to conducting PCA, we examined the correlation matrix to confirm that more than 20% of the inter-item correlations were above 0.30, which indicated the presence of an underlying factor structure.¹⁹ A Kaiser-Meyer-Olkin Measure of Sampling Adequacy value >0.70 was also used as an indicator of adequate inter-item correlation.¹⁸

Items not meeting the minimal loading value of 0.35, calculated using the formula $5.15/(n-2)$,¹⁹ were considered for elimination. Items that loaded only on one dimension in which all other items were conceptually different were also considered for elimination.¹⁹ Those items having the strongest loading on a dimension in which all other items were conceptually different, but also loaded on other dimensions (i.e. factorially complex), were retained in all dimensions for analyses.¹⁹

Unidimensionality was considered satisfied if the first dimension explained 40% of the variance¹⁹ and if no other dimension had an eigenvalue greater than three.²⁰ Visual inspection of the scree plot was also used to determine dimensionality. In order to confirm robust results, the findings from the original PCA were verified using principal axis extraction methods with promax rotation. After establishing dimensionality of the 60 similar items (n=220), another PCA of the five new items in WheelCon version 3.0 was conducted (n=146). These items were then allocated to the dimension(s) originating from previous analyses using conceptual reasoning.

Local independence: This assumption is that a person's response to an item is only due to his/her ability/attribute that is being measured.¹⁰ If this assumption were not true, then something other than self-efficacy would be influencing the responses. In such cases, the test would be multidimensional. Examining the correlation coefficients of the inter-item residuals tested local independence. Correlation coefficient greater than 0.20 indicated a violation of the local independence assumption.²¹

Assessing item fit and item elimination—Subsequent Rasch analyses on the 60 similar items were performed using the entire sample (n=220), whereas analyses of the 5 new items were performed using the sample in study 1 (n=146). Because Rasch analyses examine the quality of information provided by each item, and not total scores, using different sample sizes to investigate the properties of individual items is not concerning.

Because the WheelCon uses the same response format for each item, the Rasch Rating Scale Model¹⁷ was used. This model is used with datasets derived from polytomous ordered response scales (i.e. response scales with more than two response options that are ordered by difficulty), and assumes that increases in the amount of the construct needed to move up adjacent response options is the same for all items.¹⁷

The Rasch Rating Scale model¹⁷ item infit and outfit statistics were used to assess the fit between the model and data.²⁰ Infit statistics provide information on unexpected responses to items that have a difficulty level ‘near’ one’s reported self-efficacy estimate, whereas outfit statistics provide information on unexpected responses to items with self-efficacy levels ‘far’ from one’s reported self-efficacy.²⁰ Fit statistics follow a chi-square distribution with a range from 0 to infinity, with an expectation of 1. Values substantially greater than 1 reflect more variability than expected, and values less than 1 reflect less variability than expected.¹⁰ A range of acceptable infit and outfit statistics was calculated using the formulas $1 \pm 2/\sqrt{N}$ and $1 \pm 6/\sqrt{N}$, respectively.¹⁰ The standardized fit statistics (ZSTD) were also used to identify misfitting items.¹⁰ Items with infit and outfit statistics outside the ranges of 0.87 to 1.13, and 0.60 to 1.40, respectively, and/or ZSTDs greater than ± 2 were considered for elimination. After removing the items with misfitting outfit statistics, another model was developed, and the infit statistics of the remaining items examined.²⁰ We used Joint Maximum Likelihood Estimation to estimate the item difficulty parameters.²⁰

We then assessed for item redundancy. Items were considered redundant if there was overlap in both item difficulty $\pm$ the standard error of measurement (SEM) and conceptual content. When redundancy was identified, one item was considered for elimination.

After eliminating misfitting and redundant items, a final Rating Scale model was developed. Because it is common for items to misfit in the more model-fitting context,²⁰ we considered item infit and outfit statistics in the range of 0.50 and 1.50 to be productive for measurement.²⁰

Reliability and validity—Both the SEM and reliability were estimated for the entire range of total scores. The SEMs derived via Rasch analyses were converted into coefficients similar to Cronbach alpha by using the formula: $1/(1 + \logit\ SEM^2)$.²² Cronbach alpha estimates <0.70 were considered poor.²³ For ease of interpretation, the Rasch parameters were transformed into a 0 to 100 standardized scale.

Winsteps version 3.75²⁰ and SPSS version 19.0 (SPSS Inc., Chicago, IL) were used to perform the analyses.

Results

Participants

The mean age of the combined sample ($n=220$) was 54.2 ($SD=13.0$) years, and the mean years of wheelchair-use experience was 17.9 ($SD=14.7$). One-hundred-thirty-nine (63.2%) individuals were male, 115 (52.3%) individuals had a spinal cord injury, and the mean Barthel Index score was 14.6/20 ($SD=2.8$). Table 1 presents the sample characteristics.

Response format functioning

In the original 101-point response format, eight (7.9%) response options had misfitting outfit statistic, and 34 response options were not consecutively ordered. In each of the 11-point response formats, one (9.1%) response option had a misfitting outfit statistic, and all response options were consecutively ordered. The 11-point format with 100 as a unique category was determined to outperform the other 11-point response format because its misfitting outfit statistic was less severe. The recoded data from the 11-point response format with 100 as a unique category was used in subsequent analyses. Table 2 details the two 11-point response formats.

Evaluating Rasch assumptions

Unidimensionality—The data had sufficient inter-item correlations to perform PCA. The initial PCA of 60 similar items resulted in 11 dimensions with eigenvalues greater than 1. The first dimension explained more than 40% of the variance, however, both the first and second dimensions had eigenvalues greater than three, thereby suggesting multidimensionality as per our criteria. Examination of the scree plot, shown in figure 1, also suggested the presence of two dominant dimensions. After running another PCA and forcing two dimensions, the amount of variance explained was 31.3% and 20.8% in each dimension with eigenvalues of 18.8 and 12.5, respectively. This indicated the presence of two primary dimensions, which were accordingly named Mobility, and Self-management efficacy.

Examination of the item loadings revealed that five items (items 3, 4, 10, 12, and 42) were factorially complex. These items, therefore, were retained in both the Mobility and Self-management efficacy dimensions for subsequent analyses. One item (item 2) inquiring about self-efficacy related to moving around furniture in the home was removed because it only loaded on the self-management efficacy dimension (i.e. the dimension in which all other items were conceptually different).

In the second PCA of the five new items (items 37, 53, 54, 58, 61), item 37 was included in the Mobility efficacy dimension, and items 58 and 61 were included in the Self-management efficacy dimension. Items 53 and 54 were factorially complex, and were included in both the Mobility and Self-management efficacy dimensions for subsequent analyses. Overall, as a result of both PCAs the Mobility and Self-management efficacy dimensions were comprised of 46 and 25 items, respectively.

Local independence—All inter-item residual correlations were below 0.20 in both dimensions thereby satisfying the local independence assumption.

Assessing item fit, reliability, and validity

Because two dimensions were found, separate Rasch analyses were performed on each of the Mobility and Self-management efficacy dimensions. For exploratory purposes, we also performed Rasch analyses on the two dimensions combined. Moreover, because not everyone in the combined sample completed the 5 new items in the WheelCon version 3.0, subsequent Rasch analyses used the combined sample ($n=220$) to investigate the properties of the 60 similar items, and the sample from study 1 ($n=146$) to examine the properties of the 5 new items.

Mobility efficacy subscale—Of the 46 items in the Mobility efficacy dimension, 32 items had misfitting outfit and infit statistics (appendix 1), and one item (item 39) was deemed to be redundant. After removing these 33 misfitting and redundant items, a final Rating Scale Model was developed using the 13 remaining items. The fit statistics of these 13 items were within the acceptable range.

With the exception of the two lowest and highest scores, all other standardized scores (97% of all scores) had reliability estimates ≥ 0.70 . Standardized scores in the range of 40.8 to 48.6 had the highest reliability estimates (0.99) and low SEM, as shown in figure 2a. The mean item difficulty of this subscale was 46.1 ($SD=5.2$), whereas the sample's mean self-efficacy was 59.6 ($SD=16.5$). Table 3 presents the 13 items organized by increasing difficulty.

Self-management efficacy subscale—Of the 25 items in the Self-management efficacy dimension, 17 items had misfitting outfit and infit statistics (appendix 2). After removing these items, a final Rating Scale Model was developed using the eight remaining items. The fit statistics of these eight items were within the acceptable range.

With the exception of the two lowest and highest scores, all other standardized scores (95% of all scores) had reliability estimates ≥ 0.70 . Scores in the range of 38.1 to 49.0 had the highest reliability estimates (0.98) and low SEM, as shown in figure 2b. The mean item difficulty of this subscale was 45.5 ($SD=2.7$), whereas the sample's mean self-management efficacy was 68.5 ($SD=19.0$). Table 4 presents the 8 items organized by increasing difficulty.

The 21-item WheelCon short form—After combining the items from both subscales and developing another Rating Scale Model, one item had an infit statistic greater than 1.5, and four items had outfit statistics greater than 1.5, as shown in table 5. With the exception of the two lowest and highest scores, all other standardized scores (98% of all scores) had reliability estimates ≥ 0.70 . Scores in the range of 37.4 to 54.2 had the highest reliability estimates (0.99) and low SEM, as shown in figure 2c. The mean item difficulty was 46.5 ($SD=3.8$), whereas the sample's mean self-efficacy was 59.0 ($SD=14.3$).

Discussion

This study evaluated the measurement properties of the WheelCon using Rasch analyses. Evidence was first found in support of the hypothesis that a shorter response format outperformed the original 101-point format. Response formats are intended to capture different degrees of a construct from individuals completing the measurement scale.¹⁸

Nunnally states that there is always an advantage with using more rather than fewer response options up until the point where the number of options begins to confuse subjects.²³ In the evaluation of the WheelCon's original 101-point response format, response options were likely being used in unexpected contexts, which may have led to a disordering of response options. For example, if individuals provide a response of 70 to an item of certain difficulty, and a response of 75 to a slightly more difficult item, the response format is not being used as intended because probabilistically the more difficult item should elicit lower responses relative to the easier item. One explanation for this type of discrepancy is that individuals are not differentiating between adjacent response options.²⁴ A consequence is respondent bias arising from increases in subjectivity that may confound the interpretation of total scores. Therefore, the use of a 0 to 10 response format with the WheelCon may improve the precision of the self-efficacy estimates.

Through the use of PCA, two broad dimensions were found to comprise the 65-items in the WheelCon. The dimension loadings depended on whether the items required self-efficacy related to maneuvering the wheelchair or not, which is consistent with observations made by Rushton et al. during the development of the WheelCon.¹ Moreover, several items in both dimensions were eliminated due to model misfit and redundancy. This resulted in a 13-item Mobility efficacy, 8-item Self-management efficacy subscales, and a 21-item WheelCon Short Form. Overall, these results provide evidence in support of the hypothesis that the WheelCon with a shortened response format would result in more than one dimension with fewer items. Interestingly, the mobility dimension included items originally conceptualized as moving around the physical environment, and performing activities, and the self-management dimension included items originally conceptualized as knowledge and problem solving, advocacy, and social situations. The only conceptual area in the original WheelCon that failed to have items retained was emotions. This is likely because there was only 1 item in this area to begin with.

Because situation specific self-efficacy measures provide more explanatory and predictive value than global measures of self-efficacy,¹⁵ the Mobility and Self-management subscales may provide more accurate self-efficacy estimates for certain clinical situations or research questions than the 21-item WheelCon Short Form. For example, for outcomes related to mobility or physical abilities, the use of the 13-item Mobility efficacy subscale may be more predictive than the 8-item Self-management efficacy subscale. Similarly, the reverse may be true for outcomes related to problem solving, advocacy, or other self-management tasks. Although the 21-item short form does not function as well as the individual subscales, as indicated by the misfitting items, it may still be of value, particularly for the prediction of multifaceted outcomes such as participation in social and personal roles or other life situations. Research investigating this hypothesis is needed.

Contrary to the hypothesis that the entire range of scores would have reliability estimates 0.70, the scores at the lower and upper ranges for each measure are not as reliable or precise as the scores in the middle range for each scale. Nonetheless, more than 95% of all possible scores in each of the two subscales and 21-item short form have good reliability. Offering reliable and valid situation specific measurements along with a more global

measurement may facilitate the assessment of wheelchair-use self-efficacy both clinically and in research.

Finally, it is interesting to note that the sample's mean self-efficacy in each of the scales was higher than mean 'difficulty' of the items. This is an indication that the sample was responding to many of the items with high self-efficacy. The scales, therefore, may assess the self-efficacy constructs more reliably among wheelchair-users with lower self-efficacy than the individuals in this sample. For example, older, female wheelchair-users or other groups of wheelchair-users at risk of having lowered self-efficacy.²⁵ Research investigating differential item functioning among different subgroups of wheelchair-users will be an important next step to further establish each scales' utility.

Limitations

This study has several limitations. First, the modifications made to the response format and subsequent analyses were completed using the same dataset. As a result, the reliability estimates are based on recoded measurements. Furthermore, the analyses only examined the internal structure of the measure. Prospective studies with *a priori* hypotheses relative to other constructs are necessary, as are test-retest investigations. In addition, the sample size may be considered small for Rasch analyses, thereby producing less precise and robust estimates, and having less power for fit analyses. In general, the more item and response option parameters estimated the larger the sample size required. At present, however, there are no guidelines to orient researchers to sample size requirements.²¹ However, research has shown through the use of Monte Carlo simulations that the Rating Scale Model produces negligible item parameter bias in measures with at least 20 items with samples of at least 100 individuals.²⁶ Finally, the results are limited to individuals living in similar geographic/environmental regions as this study's sample.

Conclusion

The WheelCon is comprised of two dimensions that assess self-efficacy related to mobility and self-management. The measurements from the 13-item Mobility and 8-item Self-management efficacy subscales, and 21-item WheelCon Short Form using a 0 to 10 response scale have good reliability, and may provide accurate and precise measurements of different forms of wheelchair-use self-efficacy. The usefulness of the subscales and/or the short form depends on the context in which they are being considered. For individuals who present with high mobility or self-management abilities but lead sedentary lifestyles or are uncertain about their problem solving, and/or action planning should be assessed for low self-efficacy. Remediation of low self-efficacy may lead to achievement of desired mobility and self-management outcomes. Finally, research to establish the reliability and validity of the measurements using the 0 to 10 response scale is warranted.

Acknowledgments

Supported by the Canadian Institutes of Health Research (doctoral scholarship and operating grant no. 107848-1).

List of Abbreviations

WheelCon	Wheelchair Use Confidence Scale
PCA	Principal Components Analysis
ZSTD	standardized fit statistics
SEM	standard error of measurement
SD	Standard deviation

References

1. Rushton PW, Miller WC, Kirby RL, Eng JJ, Yip J. Development and content validation of the Wheelchair Use Confidence Scale: a mixed-methods study. *Disabil Rehabil Assist Technol*. 2011;57–66. [PubMed: 20809874]
2. Rushton PW, Miller WC, Kirby RL, Eng JJ. Measure for the assessment of confidence with manual wheelchair use (WheelCon-M) version 2. 1: Reliability and validity. *J Rehab Med*. 2013; 45:61–67.
3. Sakakibara BM, Miller WC, Eng JJ, Backman CL, Routhier F. Preliminary examination of the relationship between participation and confidence in older manual wheelchair-users. *Arch Phys Med Rehabil*. 2013; 94(4):791–4. [PubMed: 23022451]
4. Sakakibara BM, Miller WC, Routhier F, Backman CL, Eng JJ. The association between self-efficacy and participation in community-dwelling manual wheelchair-users, aged 50 and over. *Phys Ther*. 2014; 94(5):664–674. [PubMed: 24415774]
5. Sakakibara BM, Miller WC, Eng JJ, Backman CL, Routhier F. Influences of wheelchair-related efficacy on life-space mobility in adult, community-dwelling manual wheelchair-users. *Phys Ther*. 2014 in-press.
6. Miller WC, Sakakibara BM, Rushton PW. The prevalence of low confidence with using a wheelchair and its relationship to wheelchair skills. *Gerontologist*. 2012; 52(S1):205.
7. Phang SH, Martin-Ginis KA, Routhier F, Lemay V. The role of self-efficacy in the wheelchair skills-physical activity relationship among manual wheelchair-users with spinal cord injury. *Disabil Rehabil*. 2012; 34(8):625–632. [PubMed: 21981243]
8. Sakakibara BM, Miller WC, Souza M, Nikolova V, Best KL. Wheelchair skills training to improve confidence with using a manual wheelchair among older adults: A pilot study. *Arch Phys Med Rehabil*. 2013; 94:1031–7. [PubMed: 23385110]
9. Rushton PW, Routhier F, Miller WC, Auger C, Lavoie M-P. French-Canadian translation of the WheelCon-M (WheelCon- M-F) and evaluation of its validity evidence using telephone administration. *Disabil Rehabil Assist Technol*. Jul.2014 in-press.
10. de Ayala, RJ. The theory and practice of item response theory. New York, NY: The Guildford Press; 2009.
11. Lachman ME, Howland J, Tennstedt S, Jette A, Assmann S, Peterson EW. Fear of falling and activity restriction: the Survey of Activities and Fear of Falling in the Elderly (SAFE). *J Gerontol B Psychol Sci Soc Sci*. 1998; 53:43–50.
12. Arnadottir SA, Lundin-Olsson L, Gunnarsdottir ED, Fisher AG. Application of Rasch analysis to examine the psychometric aspects of the Activities-Specific Balance Confidence Scale when used in a new cultural context. *Arch Phys Med Rehabil*. 2010; 91:156–163. [PubMed: 20103411]
13. Embretson, SE., Reise, SP. Item Response Theory for Psychologists. 2. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2000.
14. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”: a practical method for grading the state of patients for the clinician. *J Psychiatr Res*. 1975; 12:189–198. [PubMed: 1202204]
15. Bandura, A. Guide for creating self-efficacy scales. In: Pajares, F., Urdan, T., editors. Self-efficacy beliefs of adolescents. Greenwich, CT: Information Age Publishing; 2006.
16. Gompertz P, Pound P, Ebrahim S. A postal version of the Barthel Index. *Clin Rehabil*. 1994; 8:233.

17. Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978; 43:561–573.
18. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002; 3:85–106. [PubMed: 11997586]
19. Norman, GR., Streiner, DL. *Biostatistics: The bare essentials*. 3. Hamilton, ON: BC Decker Inc; 2008.
20. Linacre, JM. *A User's Guide to WINSTEPS*. Chicago, IL: MESA Press; 2009.
21. Velozo CA, Seel RT, Magasi S, Heinemann AW, Romero S. Improving measurement methods in rehabilitation: Core concepts and recommendations for scale development. *Arch Phys Med Rehabil*. 2012; 93(8 Suppl 2):S154–S163. [PubMed: 22840881]
22. Mâsse LC, Heesch KC, Eason KE, Wilson M. Evaluating the properties of a stage-specific self-efficacy scale for physical activity using classical test theory, confirmatory factor analysis and item response modeling. *Health Edu Res*. 2006; 21(S1):i33–i46.
23. Nunally, JC., Bernstein, IH. *Psychometric Theory*. 3. New York, NY: McGraw-Hill; 1994.
24. Streiner, DL., Norman, GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 4. Oxford: Oxford University Press; 2008.
25. Sakakibara, BM. PhD thesis. University of British Columbia; 2013. Towards an understanding of self-efficacy with using a manual wheelchair.
26. Wang WC, Chen CT. Item parameter recovery, standard error estimates, and fit statistics of the Winsteps program for the family of Rasch models. *Edu Psychol Measur*. 2005; 65:376–404.

Figure 1.
Scree plot showing two dominant dimensions

Figure 2.
Reliability and SEM of the Mobility efficacy, Self-management efficacy, and WheelCon Short Form scores

Table 1

Sample characteristics

	Study 1 (n=146)	Study 2 ⁶ (n=74)	Combined (n=220)
	mean±sd/frequency (%)		
Age	57.1±10.4	48.5±15.5	54.2±13.0
Wheelchair-use experience (years)	20.4±15.7	13.0±10.9	17.9±14.7
Functional independence (0–20)	14.3±2.9	14.1±5.4	14.6±2.8 *
Sex:			
Male	89 (61.0)	50 (67.6)	139 (63.2)
Female	57 (39.0)	24 (32.4)	81 (36.8)
Diagnosis:			
Spinal cord injury	77 (52.7)	38 (51.4)	115 (52.3)
Multiple sclerosis	17 (11.6)	8 (10.8)	25 (11.4)
Stroke	14 (9.6)	3 (4.1)	17 (7.7)
Lower extremity amputation	9 (6.2)	10 (13.5)	19 (8.6)
Other (spina bifida, Cerebral Palsy, Parkinson's disease, arthritis)	29 (19.9)	15 (20.3)	44 (20.0)
Married/Common law	69 (47.3)	29 (39.2)	98 (44.6)
Education:			
Some high school	15 (10.3)	5 (6.8)	20 (9.1)
Graduated high school	26 (17.8)	29 (39.2)	55 (25.0)
Some university	32 (21.9)	0 (0.0)	32 (14.6)
Graduated university	61 (41.8)	17 (23.0)	78 (35.5)
Post-graduate studies	9 (6.2)	4 (5.4)	13 (5.9)
Other	3 (2.1)	19 (25.7)	22 (10.0)

*
n=217

Table 2

11-point response format outfit statistics and average measure values

Response option	0 unique category		100 unique category	
	Average measure	Outfit statistic	Average measure	Outfit statistic
0	-0.23	2.62 [*]	-0.26	2.56 [*]
1	-0.19	1.53	-0.21	1.65
2	-0.06	1.59	-0.08	1.53
3	-0.02	1.07	-0.06	0.87
4	0.04	1.12	0.02	1.15
5	0.17	1.26	0.14	1.24
6	0.21	0.95	0.18	0.94
7	0.31	0.70	0.30	0.75
8	0.43	0.63	0.41	0.71
9	0.62	0.79	0.63	0.79
10	1.01	1.00	1.03	0.98

^{*} response option being used unexpectedly

Table 3

Mobility efficacy subscale items in order by difficulty and fit statistics

Mobility efficacy			Infit		Outfit	
Items (n=13)	Logits (SE)	Std score (SE)	Mnsq (ZSTD)	Mnsq (ZSTD)		
1 ...move over carpet?	-0.60 (0.06)	38.23 (0.75)	1.13 (0.80)	1.01 (0.10)		
4 ...move in small spaces, such as a bathroom?	-0.57 (0.06)	38.64 (0.73)	1.27 (1.60)	1.10 (0.60)		
13 ...open, go through, and close a standard door?	-0.44 (0.05)	40.35 (0.67)	1.32 (2.00)	1.00 (0.10)		
21 ...move down a curb cut?	-0.26 (0.05)	42.67 (0.61)	1.06 (0.50)	0.94 (-0.30)		
32 ...press the crosswalk button and cross the street?	-0.10 (0.04)	44.83 (0.55)	1.15 (1.10)	0.84 (-1.00)		
18 ...move down a dry steep slope?	-0.03 (0.04)	45.76 (0.54)	1.27 (2.00)	1.16 (1.00)		
20 ...move up a curb cut?	-0.02 (0.04)	45.93 (0.54)	1.06 (0.50)	0.90 (-0.60)		
33 ...cross a street at a crosswalk with no traffic lights?	0.09 (0.04)	47.32 (0.51)	1.22 (1.70)	1.40 (2.40)		
34 ...move across flat, freshly mowed, dry grass?	0.10 (0.04)	47.40 (0.51)	1.18 (1.50)	1.18 (1.20)		
23 ...move down a curb cut then over drainage grate?	0.19 (0.04)	48.57 (0.49)	1.06 (0.50)	0.90 (-0.70)		
22 ...move over a drainage grate, then up a curb cut?	0.26 (0.04)	49.56 (0.48)	0.88 (-1.10)	0.73 (-2.00)		
28 ...move down a curb cut then through snow?	0.63 (0.04)	54.35 (0.47)	1.09 (0.80)	1.08 (0.60)		
29 ...move through snow then up a curb cut?	0.76 (0.04)	56.12 (0.47)	1.03 (0.30)	0.96 (-0.20)		
Mean	0.00 (0.04)	46.13 (0.56)	1.13 (0.90)	1.02 (0.10)		

Note: item 39 (not shown) was removed due to redundancy

Std=standardized; SE=standard error; CI=confidence interval

Table 4

Self-management efficacy subscale items in order by difficulty and fit statistics

Self-management efficacy				Infit		Outfit	
Items (n=8)		Logits (SE)	Std score (SE)	Mnsq (ZSTD)	Mnsq (ZSTD)	Mnsq (ZSTD)	
57	...tell someone how to move your wheelchair if it gets stuck?	-0.32 (0.06)	41.00 (0.88)	1.11 (0.70)		1.03 (0.20)	
47	...use strategies, such as humour, that will help people feel comfortable if they are unsure how to act because you use a wheelchair?	-0.14 (0.06)	43.56 (0.78)	1.00 (0.00)		1.10 (0.60)	
64	...advocate for your needs at work or school?	-0.10 (0.05)	44.03 (0.76)	1.25 (1.60)		1.31 (1.70)	
48	...correct others' mistaken beliefs about people who use wheelchairs?	-0.01 (0.05)	45.32 (0.72)	0.91 (-0.60)		0.98 (-0.10)	
52	...solve how to get to your destination when there is an unexpected situation, such as detours?	0.00 (0.05)	45.50 (0.72)	1.25 (1.70)		1.23 (1.40)	
50	...present yourself as you wish to be seen when you are in public and feel people are watching you?	0.03 (0.05)	45.90 (0.71)	0.84 (-1.10)		0.69 (-2.10)	
51	...present yourself as you wish to be seen when you want to impress others?	0.16 (0.05)	47.73 (0.66)	0.85 (-1.10)		0.91 (-0.50)	
54	...move in situations making you anxious/nervous?	0.38 (0.05)	50.84 (0.76)	1.31 (1.80)		1.48 (2.40)	
Mean		0.00 (0.05)	45.49 (0.75)	1.06 (0.40)		1.09 (0.50)	

Std=standardized; SE=standard error; CI=confidence interval

Table 5

WheelCon short form items in order by difficulty and fit statistics

WheelCon Short Form					Infit	Outfit
Items (n=21)	Logits (SE)	Std score (SE)	Mnsq (ZSTD)	Mnsq (ZSTD)		
57 ...tell someone how to move your wheelchair if it gets stuck?	-0.43 (0.05)	41.16 (0.68)	1.06 (0.40)	1.36 (1.70)		
1 ...move over carpet?	-0.37 (0.05)	41.89 (0.64)	1.08 (0.50)	0.90 (-0.40)		
4 ...move in small spaces, such as a bathroom?	-0.34 (0.05)	42.21 (0.63)	0.99 (0.00)	0.80 (-1.00)		
47 ...use strategies, such as humour, that will help people feel comfortable if they are unsure how to act because you use a wheelchair?	-0.29 (0.05)	42.87 (0.60)	1.22 (1.40)	2.12 (4.50) *		
64 ...advocate for your needs at work or school?	-0.27 (0.05)	43.18 (0.59)	1.51 (3.00) *	2.48 (5.70) *		
13 ...open, go through, and close a standard door?	-0.24 (0.05)	43.53 (0.57)	0.92 (-0.50)	0.71 (-1.70)		
48 ...correct others' mistaken beliefs about people who use wheelchairs?	-0.20 (0.04)	44.03 (0.55)	1.25 (1.70)	1.65 (3.00) *		
52 ...solve how to get to your destination when there is an unexpected situation, such as detours?	-0.19 (0.04)	44.16 (0.55)	0.93 (-0.50)	0.81 (-1.00)		
50 ...present yourself as you wish to be seen when you are in public and feel people are watching you?	-0.17 (0.04)	44.42 (0.54)	1.28 (1.90)	1.03 (0.20)		
21 ...move down a curb cut?	-0.10 (0.04)	45.30 (0.52)	0.94 (-0.40)	0.88 (-0.70)		
51 ...present yourself as you wish to be seen when you want to impress others?	-0.07 (0.04)	45.66 (0.51)	1.11 (0.90)	1.05 (0.40)		
32 ...press the crosswalk button and cross the street?	0.04 (0.04)	46.95 (0.47)	1.17 (1.30)	0.89 (-0.70)		
54 ...move in situations making you anxious/nervous?	0.08 (0.05)	47.55 (0.58)	0.92 (-0.50)	1.07 (0.40)		
18 ...move down a dry steep slope?	0.09 (0.04)	47.65 (0.46)	1.02 (0.20)	1.04 (0.30)		
20 ...move up a curb cut?	0.10 (0.04)	47.78 (0.46)	0.96 (-0.30)	0.90 (-0.60)		
33 ...cross a street at a crosswalk with no traffic lights?	0.19 (0.03)	48.84 (0.43)	1.17 (1.50)	1.45 (2.60)		
34 ...move across flat, freshly mowed, dry grass?	0.19 (0.03)	48.90 (0.43)	1.05 (0.40)	1.13 (0.90)		
23 ...move down a curb cut then over drainage grate?	0.26 (0.03)	49.79 (0.42)	0.96 (-0.30)	0.86 (-1.00)		
22 ...move over a drainage grate, then up a curb cut?	0.33 (0.03)	50.56 (0.41)	0.85 (-1.40)	0.74 (-1.90)		
28 ...move down a curb cut then through snow?	0.62 (0.03)	54.26 (0.40)	1.03 (0.30)	1.01 (0.10)		
29 ...move through snow then up a curb cut?	0.73 (0.03)	55.62 (0.40)	1.01 (0.20)	1.08 (0.60)		
Mean	0.00 (0.04)	46.49 (0.52)	1.07 (0.50)	1.14 (0.50)		

Std=standardized; SE=standard error; CI=confidence interval

* misfitting item

Appendix 1

Items in the Mobility efficacy Rating Scale models with outfit and infit statistics outside the acceptable ranges

	Rating Scale model assessing outfit statistics	Infit	Outfit
	Items removed (n=22/46)	Mnsq (ZSTD)	Mnsq (ZSTD)
5	...transfer to your bed?	1.56 (3.00)	1.53 (2.30)
6	...transfer to your toilet?	1.64 (3.80)	1.58 (2.70)
7	...transfer into your bathtub?	1.74 (4.80)	1.55 (2.80)
8	...transfer from the floor to your wheelchair?	2.71 (9.90)	3.02 (8.60)
9	...transfer to your vehicle?	1.77 (5.00)	2.47 (6.20)
10	...make a light meal? [^]	1.65 (3.70)	1.68 (2.90)
14	...open, go through a spring loaded door?	0.61 (−3.80)	0.54 (−3.40)
15	...move up a standard ramp?	0.91 (−0.60)	0.65 (−2.10)
16	...move down a standard ramp?	0.89 (−0.60)	0.57 (−2.40)
17	...move up a dry steep slope?	0.72 (−2.90)	0.61 (−3.00)
19	...move down a dry steep slope and stopping?	0.86 (−1.20)	0.67 (−2.30)
24	...move through a puddle then up a curb cut?	0.69 (−3.00)	0.59 (−3.00)
25	... move down a curb cut then through a puddle?	0.74 (−2.30)	0.59 (−2.80)
26	...move through slush then up a curb cut?	0.73 (−2.90)	0.71 (−2.20)
27	...move down a curb cut then through slush?	0.78 (−2.20)	0.72 (−2.10)
30	...move up curb without curb cut?	1.99 (6.20)	2.52 (5.90)
31	...move down curb without curb cut?	1.73 (5.90)	1.53 (3.20)
36	...move along a sidewalk that is cracked/uneven?	0.70 (−2.90)	0.63 (−2.60)
40	...move through a crowd without hitting anyone?	1.24 (1.80)	3.00 (7.80)
43	...manage toileting activities in a public bathroom?	1.54 (3.80)	1.69 (3.50)
44	...use public transportation in your town?	1.97 (6.80)	2.47 (6.80)
53	...negotiate a challenging and unusual obstacle? [^]	1.11 (0.70)	1.55 (2.20)
	Rating Scale model assessing infit statistics		
	Items removed (n=10/24)		
3	...move over thresholds, such as between rooms? [^]	1.19 (1.00)	0.87 (−0.60)
11	...carry a hot drink while in your wheelchair?	1.28 (2.20)	1.24 (1.50)
12	...move through a door that opens automatically? [^]	1.49 (2.30)	0.74 (−1.20)
35	... move through a pothole?	1.14 (1.30)	1.36 (2.40)
37	...move along a flat dirt path with tree roots/rocks?	0.76 (−1.90)	0.80 (−1.10)
38	...move across flat, unpacked gravel?	1.18 (1.70)	1.34 (2.30)
42	...move in store aisle without knocking items over? [^]	1.29 (2.0)	1.27 (1.50)
45	...do your chosen leisure activities?	1.85 (4.80)	1.33 (1.70)
46	...transport items in a backpack?	1.88 (5.20)	1.37 (1.90)
54	...move in situations making you anxious/nervous? [^]	1.15 (0.90)	1.26 (1.20)

[^] factorially complex items

Appendix 2

Items in the Self-management efficacy Rating Scale models with outfit and infit statistics outside the acceptable ranges

	Rating Scale model assessing outfit statistics	Infit	Outfit
	Items removed (n=8/25)	Mnsq (ZSTD)	Mnsq (ZSTD)
10	...make a light meal? [^]	1.84 (4.40)	1.76 (3.20)
42	...move in store aisle without knocking items over? [^]	1.11 (0.90)	2.31 (5.50)
49	...present yourself as you wish to be seen around acquaintances, colleagues, or peers?	0.72 (-1.90)	0.56 (-2.60)
56	...know what your wheelchair can and can't do, separate from your own abilities?	1.43 (2.90)	1.80 (3.60)
58	...ask someone for help?	0.59 (-1.90)	0.53 (-1.90)
61	...know what to do if you fall out of your wheelchair?	1.61 (2.90)	1.58 (2.20)
63	...advocate for changes in your home?	1.07 (0.50)	1.50 (2.20)
65	...advocate for changes in your community?	1.01 (0.20)	1.42 (2.20)
	Rating Scale model assessing infit statistics		
	Items removed (n=9/17)		
3	...move over thresholds, such as between rooms? [^]	1.31 (1.50)	1.33 (1.40)
4	...move in small spaces, such as a bathroom? [^]	1.17 (1.00)	1.03 (0.20)
12	...move through a door that opens automatically? [^]	1.46 (2.10)	0.99 (0.00)
41	...ask people to move out of your way?	1.30 (1.70)	0.96 (-0.10)
53	...negotiate a challenging and unusual obstacle? [^]	0.69 (-2.10)	1.39 (1.70)
55	...know when your wheelchair is not working properly?	1.31 (1.70)	1.40 (1.80)
59	...tell a cab driver how to fold/unfold your wheelchair?	1.41 (2.30)	1.15 (0.80)
60	...tell a stranger how to help you safely get back into your wheelchair if you tip over?	1.25 (1.60)	0.96 (-0.10)
62	...advocate for changes to your wheelchair?	1.21 (1.20)	1.16 (0.80)

[^] factorially complex items