HHS Public Access

Author manuscript

Qual Life Res. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Qual Life Res. 2016 October ; 25(10): 2559–2564. doi:10.1007/s11136-016-1300-z.

Testing the measurement invariance of the University of Washington Self-Efficacy Scale short form across four diagnostic subgroups

Hyewon Chung[1], **Jiseon Kim**[2], **Ryoungsun Park**[3], **Alyssa M. Bamer**[2], **Fraser D. Bocell**[2], and **Dagmar Amtmann**[2]

[1] Department of Education, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Korea

[2] Department of Rehabilitation Medicine, University of Washington, Box 354237, Seattle, WA 98195, USA

[3] Theoretical and Behavioral Foundations Division, Wayne State University, 5425 Gullen Mall, Detroit, MI 48202, USA

Abstract

Purpose—The University of Washington Self-Efficacy Scale (UW-SES) was originally developed for people with multiple sclerosis (MS) and spinal cord injury (SCI). This study evaluates the measurement invariance of the 6-item short form of the UW-SES across four disability subgroups. Evidence of measurement invariance would extend the UW-SES for use in two additional diagnostic groups: muscular dystrophy (MD) and post-polio syndrome (PPS).

Methods—Multi-group confirmatory factor analysis was used to evaluate successive levels of measurement invariance of the 6-item short form, the UW-SES: (a) configural invariance, i.e., equivalent item-factor structures between groups; (b) metric invariance, i.e., equivalent unstandardized factor loadings between groups; and (c) scalar invariance, i.e., equivalent item intercepts between groups. Responses from the four groups with different diagnostic disorders were compared: MD (n = 172), MS (n = 868), PPS (n = 225), and SCI (n = 242).

Results—The results of this study support that the most rigorous form of invariance (i.e., scalar) holds for the 6-item short form of the UW-SES across the four diagnostic subgroups.

Conclusions—The current study suggests that the 6-item short form of the UW-SES has the same meaning across the four diagnostic subgroups. Thus, the 6-item short form is validated for people with MD, MS, PPS, and SCI.

Keywords

Measurement invariance; Multi-group confirmatory factor analysis; Self-efficacy; Muscular dystrophy; Multiple sclerosis; Post-polio syndrome; Spinal cord injury

Hyewon Chung hyewonchung7@gmail.com.

Introduction

Self-efficacy is the extent or strength of an individual's belief in his/her own ability to produce the desired effects of a task or activity [1]. Individuals with disabilities often report lower quality of life due to secondary conditions such as limitations in mobility, changes in bowel and bladder function, chronic pain, or ongoing medical complications [2]. These challenges may affect individuals' confidence in their ability to achieve desired goals, such as participating in meaningful relationships, promoting well-being, or managing health issues [3, 4].

Several scales have been developed to measure self-efficacy for specific chronic diseases such as multiple sclerosis (MS) [5], spinal cord injury (SCI) [6, 7], MS or SCI [8], epilepsy [9], arthritis [10], chronic disease in general [11], or general self-efficacy [12]. While these scales have been developed for people with specific chronic diseases, a question remains whether the construct measured (i.e., self-efficacy) has the same meaning across groups with different chronic diseases [13, 14]. Thus, further investigation is needed to answer the question of whether the scale items perform similarly across subgroups with different chronic diseases. One way to examine this question is through assessing the measurement invariance of a scale.

Testing measurement invariance has been primarily conducted using two approaches: multi-group confirmatory factor analysis (MG-CFA) and item response theory (IRT). Similarities and differences between the two approaches have been investigated; a simulation study by Kim and Yoon [15] pointed out that both MG-CFA and IRT showed reasonable power to identify differential item functioning (DIF). Stark et al. [16] reported both the MG-CFA and IRT methods showed similar results in detecting DIF across a majority of simulated conditions. Specifically, the authors note that the MG-CFA approach performed slightly worse than the IRT approach in dichotomous data, while MG-CFA approach performed better with polytomous data. Moreover, the previous literature suggests that testing measurement invariance via the IRT approach is less preferable as the MG-CFA approach is more advanced, simpler, and more user-friendly than IRT [16, 17].

A recently introduced self-efficacy measure called the University of Washington Self-Efficacy Scale (UW-SES) was developed for people living with MS and SCI [18]. The initial UW-SES development study investigated whether measurement invariance of UW-SES items existed between people with MS and SCI, using the IRT method, by detecting DIF. An absence of meaningful DIF between MS and SCI was found, indicating that the items performed similarly between the two groups [18].

However, this scale has not yet been tested for measurement invariance across other diagnostic groups. Thus, the current study will extend the previous study by Amtmann et al. [18] by testing the measurement invariance of the 6-item short form of UW-SES across people with muscular dystrophy (MD) and post-polio syndrome (PPS). The evidence of measurement invariance would provide support for the construct of self-efficacy related to management of chronic conditions as measured by the UW-SES being the same in people with MD and PPS as in the people with MS and SCI. Establishing measurement invariance

would allow users to attribute differences in self-efficacy scores to actual differences between groups. The previous literature supports the use of MG-CFA over IRT with polytomous items [16]. Thus, the current study employs the MG-CFA approach to test the existence of measurement invariance for the UW-SES.

Methods

Instruments

The UW-SES short form is a unidimensional instrument that includes 6 items developed for individuals with disabilities and chronic conditions. Responses to items are made on a five-point scale (1 = “not at all,” 2 = “a little,” 3 = “quite a bit,” 4 = “a lot,” and 5 = “completely”); higher scores indicate higher self-efficacy. High reliability has been demonstrated for UW-SES with a Cronbach's alpha coefficient of .90 for the 6-item short form. Construct and discriminant validity have also been supported [18]. The items appear in Appendix 1. The concordance table appears in Appendix 2.

Participants

Two data sets obtained from different sources were pooled: (1) a longitudinal study of symptoms and quality-of-life indicators of persons with MS and (2) a longitudinal study investigating the role of secondary conditions in aging individuals with a disability. Previous publications have described recruitment procedures for the two studies in more detail [19–22]. Briefly, study 1 participants were recruited through the Greater Washington chapter of the US National Multiple Sclerosis Society. Eligibility requirements included a self-reported diagnosis of MS, the ability to read and write English, being at least 18 years old, and providing written consent. While the baseline survey included 1271 subjects, a random subset of 584 subjects from the seventh time point (collected between June and December 2009) was used. The second data set (n = 923) was administered from July 2009 to April 2010 and required participants to have a self-reported diagnosis of MD, MS, PPS, or SCI; be able to read and write English; be at least 18 years old; and provide written consent. Recruitment strategies included advertisements in organization newsletters and Web sites (e.g., Muscular Dystrophy Foundation), invitation letters from the University of Washington registry of individuals with disabilities, and other specific condition registries (e.g., SCI Model Systems). The Institutional Review Board at the University of Washington, Seattle, reviewed and approved both studies’ procedures.

Analysis

Three levels of measurement invariance were tested via MG-CFA. The first and weakest level, *configural invariance*, tests that the same pattern of item-factor loadings exists across the groups being compared, which requires that the same items have nonzero loadings on the same factors. The second and more constrained level, *metric invariance*, additionally requires that unstandardized factor loadings are the same across groups. Finally, *scalar invariance*, the most constrained model, requires meeting not only the assumptions of configural and metric invariance but also that unstandardized item thresholds be invariant across groups [14]. There is no consensus regarding the level of measurement invariance necessary to confidently compare scores across groups. In practice, either metric invariance

or scalar invariance has been recommended as a prerequisite for group comparisons [23, 24]. Invariance models were estimated using Mplus software 7.0 [25] with a weighted least-squares mean and variance-adjusted estimation.

Goodness of fit was evaluated using χ^2, the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA). CFI and TLI assess the relative improvement in fit of the hypothesized model compared with a baseline model. Values above .95 are considered to indicate acceptable model fit [26]. RMSEA is a badness-of-fit index in that a value of 0 indicates the best fit and higher values indicate worse fit. Values less than .08 are considered to indicate adequate fit, and any values above .10 are considered to indicate poor fit [27].

In the MG-CFA approach, the fit of a baseline model is compared to the fit of increasingly constrained models. The χ^2 difference test is typically used to compare the fit of two nested models. When the χ^2 difference is not statistically significant (i.e., the addition of constraining parameters does not significantly decrease model fit), researchers have evidence favoring the less parameterized model. However, the χ^2 test statistic is inflated by sample size. Thus, other fit indexes are typically used to evaluate model fit with larger sample sizes [28]. As with the χ^2 test statistic, the χ^2 difference test is also sensitive to sample size. To account for this, we calculated the ΔCFI index. A difference of less than .01 in the ΔCFI index supports the less parameterized model. Model fit is compared only when both models of interest individually fit the data.

Results

Descriptive analysis

Demographic characteristics (age, duration of disease, gender, ethnicity, education, and marital status) of the diagnostic samples (individuals with MD, MS, PPS, or SCI) are described in Table 1. There were statistically significant differences among the groups in age ($F(3, 1502) = 109.254$, $p < .001$); the individuals with post-polio syndrome were statistically significantly older than the other three groups. There were also statistically significant differences in race (χ^2 ($df = 6$) = 24.76, $p < .001$), education (χ^2 ($df = 12$) = 66.135, $p < .001$), and marital status (χ^2 ($df = 15$) = 66.135, $p < .001$). There were no statistically significant differences among the groups in disease duration ($F(3, 1462) = .410.254$, $p = .746$). In addition, we ran correlation analyses between each demographic variable and the UW-SES scale across the four groups. Given that the demographic variables were not substantially associated with UW-SES, we did not extend the current study to run the MG-CFA model to include demographic variables.

Measurement invariance

One-factor CFA with each of the 6 items loading onto the factor was modeled across the four groups and tested in this study. With the exception of the RMSEA statistic, the configural invariance models produced acceptable fit indexes, which supports configural invariance across the fourdiagnostic subgroups: χ^2 ($df = 69$) = 393.127, $p < .001$, CFI = .981, TLI = .983, and RMSEA = .112 (.101–.123). The metric invariance models also had

good fit, with the exception of RMSEA: χ^2 (df = 87) = 378.522, p < .001, CFI = .983, TLI = .988, and RMSEA = .094 (.085–.104). The result of the χ^2 difference test of model fit between the configural and metric invariance models was statistically significant: $\Delta\chi^2$ (Δdf = 18) = 62.202 and p < .001. However, the CFI difference test supported metric invariance as there was a decrease of less than .01 in the CFI value (Δ CFI = −.002) in the nested model comparison. In consideration of the relatively large sample size employed in the current study, we concluded that metric invariance was supported based on the CFI difference test and that the χ^2 difference test was overly sensitive.

Scalar invariance was supported across the four subgroups for all model fit indexes: χ^2 (df = 105) = 343.743, p < .001, CFI = .986, TLI = .992, and RMSEA = .078. The comparison between the metric and scalar invariance models showed nonsignificant χ^2 difference: $\Delta\chi^2$ (Δdf = 18) = 22.413, p = .214, and the CFI difference, ΔCFI = −.003, supported the existence of scalar invariance across MD, MS, PPS, and SCI (See Table 2).

Conclusion

The findings of the current study support configural, metric, and scalar invariance for the 6-item short form of UW-SES across the four diagnostic subgroups. The results of this study provide evidence that the measurement invariance requirement for valid group comparisons has been satisfied, and further suggest that self-efficacy, as measured by this scale, has the same meaning across the four subgroups. The findings of the current study also support the results of a previous study using DIF for meaningful group comparison in people with MS and SCI. Future studies should examine whether our results are supported using IRT approaches in comparing across the four different diagnostic groups. Comparing results based on MG-CFA, as used in the current study, and results based on the IRT method would extend our understanding of the level of measurement invariance for UW-SES items.

Acknowledgments

The contents of this publication were developed in part under grants from the U.S. Department of Education, National Institute of Disability and Rehabilitation Research, Grant Numbers H133B080024 and H133B080025. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. This research was also supported in part by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number 5U01AR052171. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Appendix 1

University of Washington Self-Efficacy Scale 6-item short form

How confident are you that	Not at all	A little	Quite a bit	A lot	Completely
1. You can keep the physical discomfort related to your health condition or disability from interfering with the things you want to do?	☐	☐	☐	☐	☐
2. You can keep your health condition or disability from interfering with your ability to deal with unexpected events?	☐	☐	☐	☐	☐

How confident are you that	Not at all	A little	Quite a bit	A lot	Completely
3. You can keep your health condition or disability from interfering with your ability to interact socially?	☐	☐	☐	☐	☐
4. You can keep your health condition or disability from being the center of your life?	☐	☐	☐	☐	☐
5. You can bounce back from frustration, discouragement or disappointment that your health condition or disability may cause you?	☐	☐	☐	☐	☐
6. You can figure out effective solutions to issues that come up related to your health condition or disability?	☐	☐	☐	☐	☐

Items can be summed and then transformed to generate a total score for each form. Responses to items are made on a five-point scale: 1 = not at all, 2 = a little, 3 = quite a bit, 4 = a lot, 5 = completely

Appendix 2

6-Item short form summary score to T-score concordance table

Summary	Theta score	T-score
6	–3.00	20.0
7	–2.62	23.8
8	–2.35	26.5
9	–2.11	28.9
10	–1.90	31.0
11	–1.71	32.9
12	–1.53	34.7
13	–1.36	36.4
14	–1.20	38.0
15	–1.04	39.6
16	–.89	41.1
17	–.74	42.6
18	–.59	44.1
19	–.44	45.6
20	–.29	47.1
21	–.14	48.6
22	.02	50.2
23	.18	51.8
24	.34	53.4
25	.51	55.1
26	.69	56.9
27	.90	59.0
28	1.13	61.3
29	1.41	64.1
30	1.89	68.9

All 6 items are summed as a first step. After summing, scores are transformed to a T-score metric using the concordance table provided above. For comparison purposes, the mean in the development sample was 49.9 with a SD of 9.3

Abbreviations

DIF	Differential item functioning
IRT	Item response theory
MD	Muscular dystrophy
MG-CFA	Multi-group confirmatory factor analysis
MS	Multiple sclerosis
PPS	Post-polio syndrome
SCI	Spinal cord injury
UW-SES	The University of Washington Self-Efficacy Scale

References

1. Bandura, A. Social foundations of thought and action: A social-cognitive view. Prentice-Hall; Englewood Cliffs: 1986.
2. Rothrock NE, Hays RD, Spritzer K, Yount SE, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). Journal of Clinical Epidemiology. 2010; 63(11):1195–1204. [PubMed: 20688471]
3. Schmitt MM, Goverover Y, DeLuca J, Chiaravalloti N. Self-efficacy as a predictor of self-reported physical, cognitive, and social functioning in multiple sclerosis. Rehabilitation Psychology. 2014; 59(1):27–34. [PubMed: 24320946]
4. Marks R, Allegrante JP. A review and synthesis of research evidence for self-efficacy-enhancing interventions for reducing chronic disability: Implications for health education practice (part II). Health Promotion Practice. 2005; 6(2):148–156. [PubMed: 15855284]
5. Sherer M, Maddux JE, Mercandante B, Prentice-Dunn S, Jacobs B, Rogers RW. The self-efficacy scale: Construction and validation. Psychological Reports. 1982; 51(2):663–671.
6. Airlie J, Baker GA, Smith SJ, Young CA. Measuring the impact of multiple sclerosis on psychosocial functioning: The development of a new self-efficacy scale. Clinical Rehabilitation. 2001; 15(3):259–265. [PubMed: 11386395]
7. Middleton JW, Tate RL, Geraghty TJ. Self-efficacy and spinal cord injury: Psychometric properties of a new scale. Rehabilitation Psychology. 2003; 48(4):281–288.
8. Horn W, Yoels W, Wallace D, Macrina D, Wrigley M. Determinants of self-efficacy among persons with spinal cord injuries. Disability and Rehabilitation. 1998; 20(4):138–141. [PubMed: 9571380]
9. Tedman S, Thornton E, Baker G. Development of a scale to measure core beliefs and perceived self efficacy in adults with epilepsy. Seizure. 1995; 4(3):221–231. [PubMed: 7582658]
10. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis and Rheumatism. 1989; 32(1): 37–44. [PubMed: 2912463]
11. Shnek ZM, Foley FW, LaRocca NG, Gordon WA, DeLuca J, Schwartzman HG, et al. Helplessness, self-efficacy, cognitive distortions, and depression in multiple sclerosis and spinal cord injury. Annals of Behavioral Medicine. 1997; 19(3):287–294. [PubMed: 9603703]
12. Jerusalem, M.; Schwarzer, R. Self-efficacy as a resource factor in stress appraisal processes.. In: Schwarzer, R., editor. Self-efficacy: Thought control of action. Hemisphere; Washington, DC: 1992. p. 195-213.
13. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling. 2002; 9(2):233–255.

14. Milsap, RE. Statistical approaches to measurement invariance. Routledge; New York: 2012.
15. Kim ES, Yoon M. Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. Structural Equation Modeling. 2011; 18(2):212–228.
16. Stark S, Chernshenko OS, Drasgow F. Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. Journal of Applied Psychology. 2006; 91(6):1292–1306. [PubMed: 17100485]
17. Reise SP, Widaman KF, Pugh RH. Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. Psychological Bulletin. 1993; 114(3):552–566. [PubMed: 8272470]
18. Amtmann D, Bamer AM, Cook KF, Askew RL, Noonan VK, Brockway JA. University of Washington self-efficacy scale: A new self-efficacy scale for people with disabilities. Archives of Physical Medicine and Rehabilitation. 2012; 93(10):1757–1765. [PubMed: 22575393]
19. Noonan VK, Cook KF, Bamer AM, Choi SW, Kim J, Amtmann D. Measuring fatigue in persons with multiple sclerosis: Creating a crosswalk between the Modified Fatigue Impact Scale and the PROMIS Fatigue Short Form. Quality of Life Research. 2012; 21(7):1123–1133. [PubMed: 22048931]
20. Matsuda PN, Shumway-Cook A, Bamer AM, Johnson SL, Amtmann D, Kraft GH. Falls in multiple sclerosis. PM&R. 2011; 3(7):624–632. [PubMed: 21777861]
21. Cook KF, Molton IR, Jensen MP. Fatigue and aging with a disability. Archives of Physical Medicine and Rehabilitation. 2011; 92(7):1126–1133. [PubMed: 21704793]
22. Molton I, Cook KF, Smith AE, Amtmann D, Chen WH, Jensen MP. Prevalence and impact of pain in adults aging with a physical disability: Comparison to a US general population sample. The Clinical Journal of Pain. 2014; 30(4):307–315. [PubMed: 23887334]
23. Little, TD.; Card, NA.; Slegers, DW.; Ledford, EC. Representing contextual effects in multiple-group MACS models.. In: Little, TD.; Bovaird, JA.; Card, NA., editors. Modeling contextual effects in longitudinal studies. Erlbaum; Mahwah, NJ: 2007. p. 121-147.
24. Widaman, KF.; Reise, SP. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain.. In: Bryant, KJ.; Windle, M.; West, SG., editors. The science of prevention. American Psychological Association; Washington, DC: 1997. p. 281-324.
25. Muthén, LK.; Muthén, BO. Mplus User's Guide. 7th ed.. Muthén & Muthén; Los Angeles, CA: 1998–2013.
26. Kline, RB. Principles and practice of structural equation modeling. 3rd ed.. Guilford Press; New York: 2011.
27. Browne, M.; Cudeck, R. Alternative ways of assessing model fit. Sage; London: 1993.
28. Brown, TA. Confirmatory factor analysis for applied research. 2nd ed.. Guilford Press; New York: 2015.

Table 1

Demographic characteristics of the four chronic disease subgroups used to examine measurement invariance of the University of Washington Self-Efficacy Scale

	MD (*n* = 172)		**MS (*n* = 868)**		**PPS (*n* = 225)**		**SCI (*n* = 242)**	
Age (year)	53.38 ± 12.68 (20–85)		53.60 ± 10.74 (21–84)		66.99 ± 8.28 (41–91)		50.03 ± 13.33 (21–88)	
Duration of disease (year)	15.51 ± 11.71 (0–52)		15.27 ± 9.64 (0–62)		15.76 ± 10.41 (0–70)		16.04 ± 10.90 (0–50)	
Gender								
Male	78	(45.3 %)	151	(17.4 %)	56	(24.9 %)	164	(67.8 %)
Female	94	(54.7 %)	716	(82.5 %)	169	(75.1 %)	78	(32.2 %)
Missing			1	(.1 %)				
Ethnicity								
White	164	(95.3 %)	810	(93.3 %)	210	(93.3 %)	207	(85.5 %)
Nonwhite	6	(3.5 %)	53	(6.1 %)	15	(6.7 %)	34	(14.0 %)
Missing	2	(1.2 %)	5	(.6 %)			1	(.4 %)
Education								
Less than high school grade	4	(2.3 %)	8	(.9 %)	2	(.9 %)	10	(4.1 %)
High school grade/GED	21	(12.2 %)	105	(12.1 %)	19	(8.4 %)	38	(15.7 %)
Some college/technical degree/AA	44	(25.6 %)	335	(38.6 %)	59	(26.2 %)	83	(34.3 %)
College degree	51	(29.7 %)	253	(29.1 %)	62	(27.6 %)	73	(30.2 %)
Grad/professional school	52	(30.2 %)	166	(19.1 %)	83	(36.9 %)	38	(15.7 %)
Missing			1	(.1 %)				
Marital status								
Married/lives with significant other	129	(75.0 %)	613	(70.6 %)	144	(64.0 %)	120	(49.6 %)
Separated/divorced	20	(11.6 %)	146	(16.8 %)	34	(15.1 %)	58	(24.0 %)
Never married	17	(9.9 %)	45	(5.2 %)	16	(7.1 %)	57	(23.6 %)
Widowed	6	(3.5 %)	34	(3.9 %)	30	(13.3 %)	6	(2.5 %)
Single			28	(3.2 %)				
Other			2	(.2 %)				
Missing					1	(.4 %)	1	(.4 %)

MD Muscular dystrophy, *MS* multiple sclerosis, *PPS* post-polio syndrome, *SCI* spinal cord injury, *GED* certificate of high school equivalency, *AA* undergraduate academic degree

Table 2

Results of testing measurement invariance of the 6-item short form of the University of Washington Self-Efficacy Scale

	Overall fit indexes					Comparative fit indexes			
	χ^2	*df*	RMSEA	CFI	TLI	Model comparison	$\Delta\chi^2$	Δdf	ΔCFI
1. Configural	393.127***	69	.112 (.101, .123)	.981	.983				
2. Metric	378.522***	87	.094 (.085, .104)	.983	.988	1 versus 2	62.202***	18	–.002
3. Scalar	343.743***	105	.078 (.069, .087)	.986	.992	2 versus 3	22.413	18	–.003

The Chi-square difference ($\Delta\chi^2$) was tested with using the Mplus DIFFTEST option [25]

*** $p < .001$