



ARTICLE

# Validation of the World Health Organization Disability Assessment Schedule 2.0 in adults with spinal cord injury in Taiwan: a psychometric study

Tzu-Ying Chiu<sup>1</sup> · Monika E. Finger<sup>2</sup> · Carolina S. Fellinghauer<sup>2</sup> · Reuben Escorpizo<sup>2,3</sup> · Wen-Chou Chi<sup>4</sup> · Tsan-Hon Liou<sup>5,6</sup> · Chia-Feng Yen<sup>7</sup>

Received: 10 July 2018 / Revised: 23 November 2018 / Accepted: 28 November 2018 / Published online: 14 January 2019  
© International Spinal Cord Society 2019

## Abstract

**Study design** Psychometric study including exploratory factor analysis and Rasch analysis.

**Objective** The aim of the present study was to examine the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) in terms of its dimensionality and metric properties in a sample of people with spinal cord injury (SCI).

**Setting** Two hundred and thirty-nine hospitals in Taiwan

**Methods** Secondary analysis of cross-sectional data from the National Disability Determination System in Taiwan, including data of individuals with more than 1-year chronic spinal cord injury and over 18 years of age. We would ask the all 6 domains of WHODAS 2.0, except those participants who were not working in present, based on the WHODAS 2.0 manual.

**Results** Data from 521 persons were included. The internal consistency of WHODAS 2.0 was high for all six domains (Cronbach's  $\alpha$  between 0.87–0.99). The exploratory factor analysis supported the original six domain structure of WHODAS 2.0 to a large extent. Rasch analysis provided domain scores usable for measurement at the individual level and an overall WHODAS 2.0 score that takes into account the multidimensionality of the instrument.

**Conclusions** WHODAS 2.0 provides a reliable and valid instrument to measure relevant aspects of “activity and participation” in the context of functioning in people with SCI in Taiwan and may guide their rehabilitation.

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41393-018-0231-7>) contains supplementary material, which is available to authorized users.

✉ Chia-Feng Yen  
mapleyeng@gmail.com

- <sup>1</sup> Graduate Institute of Long-term Care, Tzu Chi University of Science and Technology, Hualien, Taiwan
- <sup>2</sup> Swiss Paraplegic Research, Nottwil, Switzerland
- <sup>3</sup> Department of Rehabilitation and Movement Science, College of Nursing and Health Sciences, University of Vermont, Burlington, USA
- <sup>4</sup> School of Occupational Therapy, Chung Shan Medical University, Taichung, Taiwan
- <sup>5</sup> Department of Physical Medicine and Rehabilitation, Shuang Ho Hospital, Taipei Medical University, Taipei, Taiwan
- <sup>6</sup> Graduate Institute of Injury Prevention and Control, Taipei Medical University, Taipei, Taiwan
- <sup>7</sup> Department of Public Health, Tzu Chi University, Hualien, Taiwan

## Introduction

To better assess and subsequently enhance the social participation of adults with disabilities in Taiwan, the Taiwanese government developed a new tool of disability evaluation; the Functioning Disability Evaluation Scale (FUNDES) [1]. FUNDES is based on the International Classification of Functioning, Disability and Health (ICF) of the World Health Organization (WHO) [2], and the WHO Disability Assessment Schedule 2.0 (WHODAS 2.0), which was chosen as the main instrument to assess the activity and participation domains in FUNDES [1, 3]. All domains of FUNDES based on WHODAS 2.0 address the performance and capability dimensions in this study. The difficulty levels of the performance dimensions were judged with typical assistive technology and personal assistances, and that of the capability dimensions were without the aid of devices and personal assistances [4].

In 2014, the FUNDES group developed its own norm values of WHODAS 2.0 for adults with disabilities in

Taiwan [5]. The results, using classical test theory (CTT) approach, showed that the WHODAS 2.0 was valid and reliable in evaluating the activity and participation function (AP function) problems of adults with a disability [4]. However, we still lack information about the validity of WHODAS 2.0 in the group of people with spinal cord injury (SCI) in a cultural setting such as Taiwan, including information about the measurement properties of the WHODAS 2.0 when analyzed with modern test theory. Modern test theory in addition to CTT, provides information on how and why participants responded to an item in the way they did [6]. Such information would be useful, especially to know whether WHODAS 2.0 can capture the needs of people with a high level of disability as in SCI [7].

Persons with SCI living in Taiwan have major challenges in the area of health care, financial and family support, and community integration, which includes finding suitable employment, social prejudices, and negative attitudes from others [8]. Thus far, most SCI studies in Asian countries have focused on epidemiology, quality of life, and comparison with other health conditions [9, 10]. Therefore, integrating WHODAS 2.0 into FUNDES in Taiwan represents relevant additional information to gain a better and more comprehensive understanding of the impact of SCI on individual functioning. However, the reliability and validity of WHODAS 2.0 in the SCI population in Taiwan remain unexplored. To help investigate the utility of WHODAS 2.0 for the disability assessment of persons with SCI in Taiwan, this study aims to examine the psychometric properties of WHODAS 2.0 in terms of its reliability and construct validity in the SCI population in Taiwan based on classical and modern test theory.

## Methods

### Participants

The participants who registered from November 2013 to January 2015 in the National Disability Determination System in Taiwan were selected. Data were collected in 239 authorized hospitals by trained interviewers in face-to-face interviews, as part of the official disability evaluation in Taiwan.

The sample was defined using the SCI-related diagnostic codes 344, 806, and 952 from the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). We included 521 adults with SCI more than 1-year chronic injury and over 18 years of age. We excluded those cases where missing data exceeded 50% in every domain of WHODAS 2.0. If the missing data were less than 50%, we used that domain's mean scores for imputation, which is the approach recommended by Garin (2009) to handle missing

values in WHODAS [11]. The study was approved by the Tzu-Chi research ethics committee (IRB-104-4-A).

### Variables

Sociodemographic information included sex, age, work status, education, living status, and injury characteristics. The living status was defined as living in the community or institution. The variable severity of disability was defined as degree of disability resulting from the lesion base on the body function and structure of ICF which is the physician's rating of the level of disability.

The criteria of neurological status about sustained paraplegia, tetraplegia, incomplete lesion, and complete lesion are based on the codes of body function and body structure from ICF, which were developed by professionals and physicians during 2009–2012. The criteria are the following: the definition code of tetraplegia was b730a (Muscle power functions, Upper limbs) and b730b, and the paraplegia was b730b (Muscle power functions, Lower limbs). For the complete type, we have defined that the complete was b730b.3 (MRC classification grade was 0 or 1), the incomplete was b730b.1 or b730b.2 (MRC classification grade was 2 or 3) [12].

WHODAS 2.0 was used to evaluate the difficulty experienced in an individual's daily life over the past 30 days, and the difficulty levels of the performance dimensions were judged with typical assistive technology and personal assistances. WHODAS 2.0 contains six domains with a total of 36 items, including D1 "Understanding and communicating" (6 items), D2 "Getting around" (5 items), D3 "Self-care" (4 items), D4 "Getting along with people" (5 items), D5 "Life activities" (household and school/work, 8 items), and D6 "Participation in society" (8 items). In fact, D5 "Life activities", in the present study included only the items D5.1 to D5.4 that relate to "Household activities". We would not ask the D5.5–D5.8 "Work" if the participants were not working, the 32-item WHODAS 2.0 score can be used as full 36-item version based on the WHODAS 2.0 manual, because the participants were not engaged in work activities at this moment [13]. Response options for the WHODAS 2.0 to the stem question "In the past 30 days, how much difficulty did you have...." ranged from 1 = "non" to 5 = "Extreme or cannot do". Domain raw scores were calculated and converted into an ordered metric ranging from 0 to 100, using the algorithm provided by WHO [13].

### Data analysis

Data were analyzed with IBM SPSS 20.0 for classical statistical tests and the software R with package eRM for the Rasch analyses.

The descriptive analysis of the sample distribution by domains includes the description of ceiling and floor effect and normality test. The ceiling effect was defined as the proportion of participants with maximum scores of 100, and the floor effect as the proportion of minimum scores of 0. Then the 15% of maximum and minimum score reflect floor and ceiling effects [14]. Skewness and kurtosis were used to test the normality of WHODAS 2.0 scores distribution. The distribution of skewness and kurtosis values are both 0 which means the scores is normal distribution, the skewness value less than 0, the distribution is negatively skewed that means more cases with higher scores. The kurtosis value less than 0 which is low peak, lower than normal distribution.

The internal consistency was analyzed using Cronbach's  $\alpha$ , a statistical coefficient which estimates the degree to which the items of a questionnaire are interrelated. The values of Cronbach's  $\alpha$  above 0.9 were excellent, 0.8–0.9 were good, and 0.6–0.8 were acceptable [15].

We use the exploratory factor analysis to identify the underlying relationships between the measured variables of the WHODAS 2.0, the exploratory factor analysis of WHODAS 2.0 in the sample of individuals with SCI in Taiwan allowed to determine its factorial structure and hence its construct validity. We assumed that the results of the exploratory factor analysis would confirm the original factor structure provided by WHO. As a prerequisite, we first checked the sample adequacy according to Kaiser–Meyer–Olkin (KMO) and Bartlett's test ( $p$ -value < 0.05). KMO values between 0.8 and 1 are considered acceptable.

Subsequently, a Rasch analysis taking into consideration the findings of the exploratory factor analysis were performed. The Rasch analysis provides more information on metric properties at item level in addition to the model reliability and scale dimensionality questions [16]. Rasch's probabilistic measurement approach assumes that the ability of individuals is a function of the difficulty of questionnaire items. The degree to which the data complies to a series of Rasch assumptions determines the quality of its measurement properties. An important assumption is that items in a summary score are locally independent, with correlations of the standardized analysis residuals below 0.2 [17]. Locally dependent items may lead to biased and inflated analysis estimates. In the presence of local item dependencies (LID), correlating items can be aggregated to enter the analysis as a so-called testlet [18]. Lesion level and type are not included in the differential item functioning analysis as they are on the pathway of functioning and it can be expected that the WHODAS 2.0 items are in favor of participants with paraplegia.

Monotonicity of the item response options is another assumption that requires items to display strictly increasing levels of difficulty. The Rasch analysis may reveal disordered thresholds that can be repaired by collapsing the

reversed response options into one single option. In general, collapsing of disordered response options of testlets is not required. While strict monotonicity is required, equidistance of the intervals between response thresholds is not expected, as this can be handled by the specific Rasch model applied in this study, i.e., the partial credit model [19].

Unidimensionality of the instrument requires that the loading structure of standardized analysis residuals indicates the presence of one single dimension with eigenvalues below 2 [20]. It is expected that component loadings of a principal component analysis (PCA) of the residuals confirm the dimensional structure found with the factor analytical approach. Finally, The Person Separation Index (PSI) gives the reliability of the model. It can be interpreted similarly to Cronbach alpha, ranging between 0 and 1. Usually, a PSI reliability of 0.70 is required for analysis at the group level, and values of 0.85 and higher for individual use [21].

The fit of the WHODAS 2.0 items to the Rasch model is given with the infit statistic derived from the mean squared standardized residuals. In the present setting, infit values within 0.8–1.2 are considered sufficient for high stakes assessment purposes [22]. The Rasch analysis was first performed for the entire WHODAS 2.0, mainly to verify and confirm the dimensional structure revealed by the exploratory factor analysis. Then the relevant dimensions or factors were analyzed separately and adjustments were undertaken towards creating reliable summary scores for WHODAS 2.0. Finally, a conjoint analysis of all domains of the WHODAS 2.0 was performed, using a testlet per domain approach to overcome multidimensionality issues and to come up with one single interval-scaled WHODAS 2.0 summary score. Based on the metric information derived from the Rasch analysis, the quality and validity of domain and summary scores will be discussed.

Finally, to explore the sensitivity of the WHODAS 2.0 scores in the light of the level of disability, the 0–100 Rasch-transformed person scores for each WHODAS 2.0 domain and the WHODAS 2.0 summary score (testlet analysis) were compared to the physician's rating of the level of disability (1 = mild, 2 = moderate, 3 = severe, or 4 = profound).

## Results

### Characteristics of participants and the distribution of AP function

The study population was 59 years old on average with 67% of the participants being male. About 54% had sustained paraplegia and 46% a tetraplegia, 86% had an incomplete lesion, and 14% a complete lesion. A little more than

twenty-five percent of the study population experienced mild disability of severity, 31% moderate, 37% severe, and 7% profound disability. Further characteristics of the study population are shown in Table 1. Table 2 outlines the average scores, median of the six WHODAS 2.0 domains, and ceiling and floor effects for the overall SCI population. The internal consistency (Cronbach's  $\alpha$ ) of the six domains ranged from 0.87 to 0.99. The skewness and kurtosis values were  $-0.091$  and  $-0.459$  that were negatively skewed and the distribution was lower peak than normal distribution.

**Table 1** Sociodemographic of study population ( $n = 521$ )

Variables		All $N$ (%)
Age		$59.03 \pm 16.43^a$
Sex	Male	350 (67.2)
	Female	171 (32.8)
Injury level	Paraplegia	283 (54.3)
	Tetraplegia	238 (45.7)
Injury type	Incomplete	448 (86.2)
	Complete	72 (13.8)
Severity of disability	Mild	132 (25.3)
	Moderate	161 (30.9)
	Severe	192 (36.9)
	Profound	36 (6.9)
Work status	Employed	30 (5.8)
	Unemployed	338 (65.5)
	Others	148 (28.7)
Education <sup>b</sup>	<Elementary school	80 (15.4)
	Elementary school	159 (30.5)
	Junior high school	167 (32.1)
	Senior high school	79 (15.2)
	University	36 (6.9)
Living status	Community	478 (91.7)
	Institution	43 (8.3)

<sup>a</sup>Mean $\pm$ SD

<sup>b</sup>Education is categorized according to Taiwan's education system

## Construct validity: factor analysis

The KMO measure of sampling adequacy was  $>0.91$  and the significance of Bartlett's test of sphericity was  $p < 0.05$ , meaning that exploratory factor analysis could safely be applied to the obtained dataset. As hypothesized, the factor loadings of exploratory factor analysis confirmed the original WHODAS 2.0 scale for the most part, explaining a total variance of 72.5%. Yet three items cross-loaded on different scales: D4.3 "Getting along with close people" loaded strongly on D1 "Understanding and communicating", and two questions from D2 "Getting around" loaded on D3 "Self-care" (see Table 3).

## Measurement properties: Rasch analysis

The Rasch analysis was performed in several steps. First the Rasch analysis of the entire WHODAS 2.0, provided further support for the findings of the factor analysis. The PCA indicated multidimensionality with at least 5 dimensions (Supplement 1). The dependency structure, based on the correlation matrix of the standardized residuals, showed that mainly the items from a same domain were strongly related ( $r > 0.2$ ). The only exception was the item D6.4 "How much time did you spend on your health condition, or its consequences" which was uncorrelated to any other WHODAS 2.0 item. Also items of domain D1 "Understanding and communicating", i.e., D1.1 Concentration, D1.3 Problem-solving, D1.4 Learning a new task, and D1.6 Conversation appeared associated to items from domain D4 "Getting along with people", and items of domain D2 "Getting around", i.e., D2.3 Moving around inside home and D2.4 Getting out of home, to those of the domain D3 "Self-care" (Fig. 1).

In the next steps, each WHODAS 2.0 domain was analyzed separately with the Rasch model and then, if necessary, testlet based solutions for the six domain scores and the summary score were proposed. The results of the analysis by domain before (start model) and after Rasch based

**Table 2** The distribution of WHODAS 2.0 scores with floor and ceiling effects based on interval scores (0–100) for the people with spinal cord injury

	All $n = 521$			
	Mean (95% CI) <sup>a</sup>	Cronbach's $\alpha$	Floor effect ( $n$ , %)	Ceiling effect ( $n$ , %)
D1 Cognition	28.43 (26.1–30.8)	0.92	133(25.5)	12(2.3)
D2 Mobility	64.64 (62.3–67.0)	0.88	13(2.5)	88(16.9)
D3 Self-care	36.51 (33.9–39.2)	0.87	109(20.9)	34(6.5)
D4 Getting along	46.47 (43.7–49.2)	0.89	84(16.1)	37(7.1)
D5(1) Household activities	69.77 (66.4–7.1)	0.99	90(17.3)	267(51.5)
D6 Participation	53.69 (51.5–55.9)	0.88	11(2.1)	19(3.6)
Global score	53.42 (51.7–55.2)	0.95	0	3(0.6)

<sup>a</sup>where 0 = no disability; 100 = full disability)

**Table 3** The results of exploratory factor analysis for people with SCI ( $n = 521$ )

WHODAS item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
D1.1 Concentration	<b>0.75</b>	0.21	0.04	0.12	0.21	0.09
D1.2 Remembering to do important things	<b>0.81</b>	0.16	0.06	0.18	0.10	0.05
D1.3 Problem-solving	<b>0.80</b>	0.17	0.08	0.15	0.20	0.10
D1.4 Learning a new task	<b>0.75</b>	0.15	0.07	0.08	0.33	0.13
D1.5 Understanding	<b>0.81</b>	0.12	0.07	0.15	0.08	0.02
D1.6 Conversation	<b>0.80</b>	0.18	0.06	0.07	0.14	0.08
D4.3 Getting along with close people	<b>0.50</b>	0.31	0.06	0.19	0.49	0.03
D6.1 Joining in community activities	0.18	<b>0.57</b>	0.23	0.15	0.27	0.31
D6.2 Because of environmental barriers	0.20	<b>0.60</b>	0.13	0.37	−0.05	0.18
D6.3 Others affects one's dignity	0.23	<b>0.74</b>	0.07	0.21	0.13	−0.01
D6.4 Health affects time consumption	0.14	<b>0.59</b>	0.16	0.10	0.19	0.16
D6.5 Health affects one's emotion	0.26	<b>0.71</b>	0.03	0.08	0.22	0.12
D6.6 Health affects family finances	0.05	<b>0.75</b>	0.00	0.15	0.05	0.08
D6.7 Health affects family	0.13	<b>0.75</b>	0.03	0.12	0.11	0.04
D6.8 Doing things for relaxation or pleasure	0.23	<b>0.62</b>	0.09	−0.02	0.27	0.28
D5.1 Household responsibilities	0.08	0.09	<b>0.94</b>	0.21	0.08	0.19
D5.2 Do important household tasks well	0.08	0.11	<b>0.94</b>	0.21	0.08	0.17
D5.3 Do all needed household work	0.08	0.11	<b>0.94</b>	0.21	0.09	0.18
D5.4 Household work performed as quickly as needed	0.07	0.09	<b>0.95</b>	0.21	0.08	0.15
D2.3 Moving around inside home	0.12	0.19	0.28	<b>0.66</b>	0.07	0.44
D2.4 Getting out of home	0.10	0.20	0.25	<b>0.67</b>	0.05	0.41
D3.1 Washing whole body	0.15	0.17	0.18	<b>0.89</b>	0.02	0.10
D3.2 Getting dressed	0.14	0.15	0.17	<b>0.89</b>	0.05	0.06
D3.3 Eating	0.24	0.23	0.13	<b>0.75</b>	0.16	0.02
D3.4 Staying by oneself for a few days	0.12	0.08	0.19	<b>0.54</b>	0.28	0.28
D4.1 Dealing with strangers	0.53	0.25	0.00	0.15	<b>0.60</b>	0.05
D4.2 Maintaining a friendship	0.40	0.27	0.01	0.01	<b>0.73</b>	0.14
D4.4 Making new friends	0.31	0.22	0.08	0.07	<b>0.80</b>	0.18
D4.5 Sexual activities	0.20	0.17	0.23	0.17	<b>0.69</b>	0.13
D2.1 Standing for long periods	0.12	0.21	0.24	0.10	0.19	<b>0.77</b>
D2.2 Standing up from sitting	0.11	0.15	0.30	0.32	0.14	<b>0.68</b>
D2.5 Walking a long distance	0.09	0.24	0.16	0.27	0.09	<b>0.71</b>
Factor variance	15.80%	13.89%	13.00%	12.88%	9.19%	7.82%

KMO: 0.92; Bartlett's test &lt;0.05; total variance: 72.5%

The bold values defines the relationship of each variable to the underlying factor

adjustments (final model) are shown in table 1\_supplementary material.

Table 4 shows the general fit of the Rasch models by domain or across domains, including the model before (start model) and after Rasch based adjustments to achieve good fit (final model). Two domains, D1 “Understanding and communicating” and D5(1) “Household activities”, did not need to be adjusted at all as the model and their items complied to all Rasch assumptions. All analyses by domain showed a high PSI, ranging between 0.87 for D2 “Mobility” and 0.98 for D5(1) “Household activities”.

Uniform differential item functioning for age was found for the testlet aggregating items D6.6 and D6.7 in the

adjusted analysis of domain D6” Participation in society”. Uniform differential item functioning for gender was not found in any of the final models. All items of the final adjusted models showed good fit, with a tendency towards overfit, meaning lower discrimination for scores in the middle of the range compared to more extreme scores, especially in domain D3 “Self-Care” and D5(1) “Household activities”.

Finally, a testlet approach for the domain analysis was undertaken to derive one single summary score as a measure of health and disability. This analysis confirmed the association between the D1 “Understanding and communicating” and D4 “Getting along with people”, as well as



the D2 “Getting around” and D3 “Self-Care” domains. In a final step, the correlated domains were aggregated. In the basic analysis across the original six WHODAS 2.0 domains, only D5(1) “Household activities” showed an insufficient fit. The fit of D5(1) increased in the final testlet

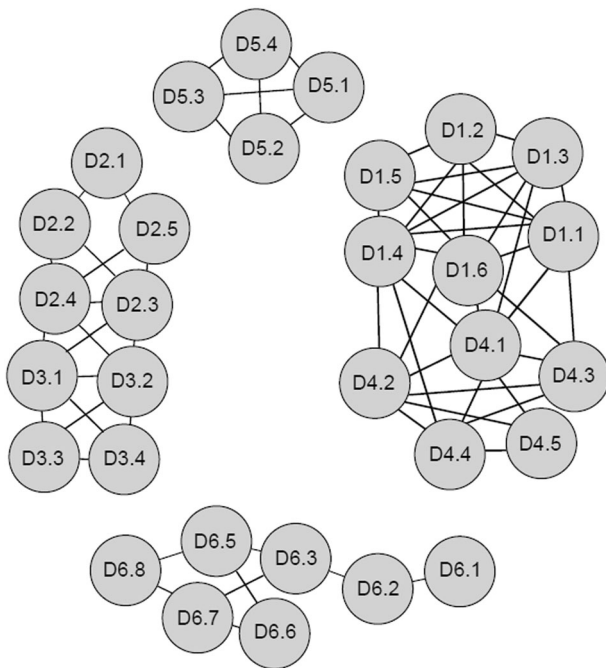
solution. Unsurprisingly, the PSI dropped for the testlet models, e.g., the first approach with the six original domains had a PSI of 0.85, while the final testlet approach, free of local dependencies, showed a PSI good enough for group level measurement ( $>0.7$ ). PSI for WHODAS 2.0 aggregated across domains was only 0.79.

The analysis of the scores per WHODAS 2.0 domain, as well as a Rasch based WHODAS 2.0 total score matched significantly with the physician’s general disability rating. The changes in the distribution of WHODAS 2.0 scores for the entire SCI-sample by disability level for each domain and across domains are illustrated in Fig. 2.

## Discussion

The current study has examined reliability and validity of WHODAS 2.0 with a classical and a modern psychometric approach, and provided evidence from both approaches that the instrument is reliable and valid to measure disability in a Taiwanese SCI population. It is indeed important to discuss the psychometric properties of WHODAS 2.0 in Taiwan, especially for the rare subgroups whose demographic characteristics were different with our total disabled population, like SCI population. Because WHODAS 2.0 is the main instrument to evaluate the people with disability under the national social welfare policy, eligibility and the severity of disability are strongly correlated with the welfare allocation and employment-oriented services in Taiwan.

**WHODAS - Local Dependencies**

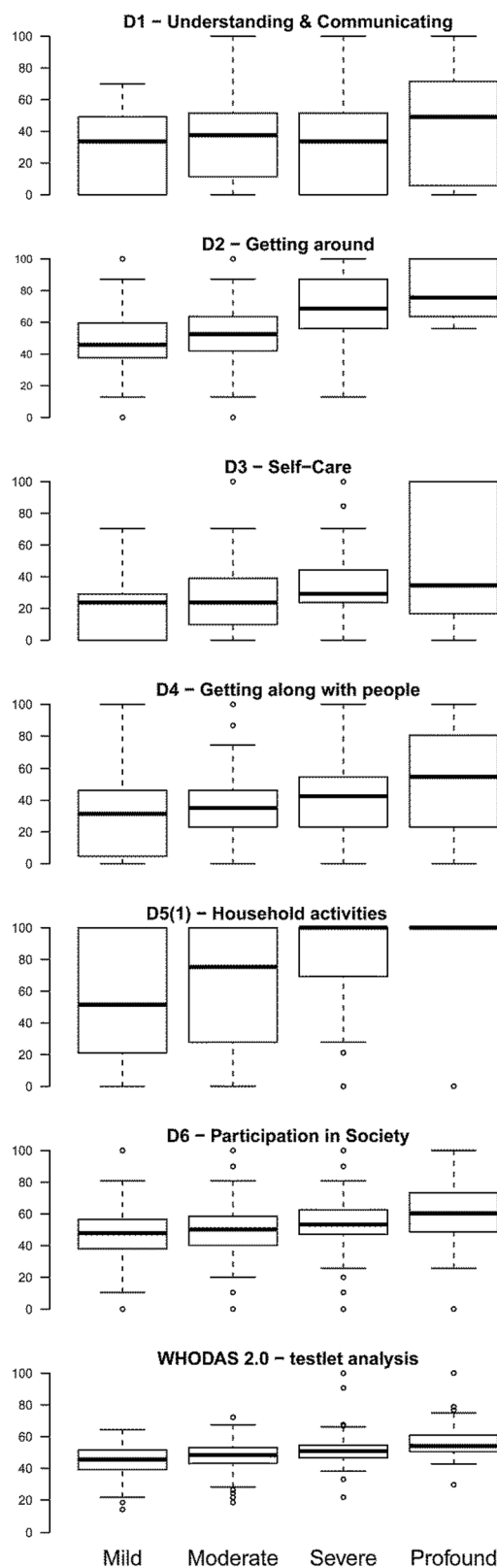


**Fig. 1** Local Item dependency structure for the WHODAS 2.0

**Table 4** Start and final model targeting fit of entire WHODAS 2.0, each subscale, and the calibration of domains as items

			Item difficulty		Person ability		Reliability				
			Mean	SD	Mean	SD	PSI	Cronbach alpha			
Dimension		Stage	Mean	SD	Mean	SD	PSI	Cronbach alpha	LID	Uniform DIF	Non-uniform DIF
All	WHODAS 2.0	Start	0.05	0.71	−0.13	0.78	0.95	0.95	Yes	Yes	No
D1	Understanding and communicating	Start & Final	0.44	1.26	−0.58	1.34	0.91	0.91	No	No	No
D2	Getting around	Start	0.35	1.23	0.59	1.18	0.91	0.88	Yes	No	No
		Final	0.37	1.35	0.73	1.25	0.87	0.84	No	No	Yes
D3	Self-care	Start	0.54	1.90	−0.33	1.32	0.92	0.87	Yes	Yes	No
		Final	0.46	1.83	−0.36	1.11	0.89	0.67	No	Yes	No
D4	Getting along with people	Start	0.31	1.10	0.01	1.18	0.91	0.89	No	No	No
		Final	0.41	1.62	0.05	1.47	0.90	0.87	No	No	No
D5(1)	Household activities	Start & Final	2.15	5.00	2.39	4.04	0.98	0.99	No	No	No
D6	Participation in society	Start	0.25	0.73	0.26	1.01	0.90	0.88	Yes	Yes	No
		Final	0.26	0.93	0.27	1.05	0.89	0.83	No	Yes	No
Testlet		Start	0.02	0.96	−0.03	0.27	0.85	0.83	Yes	Yes	No
		Final	0.01	0.93	−0.02	0.22	0.79	0.75	No	Yes	No

PSI Person separation index, LID Local item dependency, DIF Differential item functioning



**Fig. 2** Boxplot for WHODAS 2.0 0–100 transformed scores by disability level

Therefore, how to measure the disabilities status precisely would be helpful for the people with SCI population.

Nevertheless, it must be noted that items related to work in the domain D5(1) “Household activities” were excluded from the analysis because less than 6% of participants reported being employed. Not being employed may be a characteristic of SCI population in Taiwan in the first few years after SCI, but international data suggest that the employment rate of persons with SCI increases over the years [23]. Therefore, the consequence of not assessing the work items must be carefully evaluated in the light of disability eligibility in the long-term evaluation of functioning of people with SCI.

WHODAS 2.0 showed good to excellent internal consistency for chronic adults with SCI with a Cronbach’s alpha ranging from 0.85 to 0.99, which is slightly above the findings of other studies [24]. The PSI of WHODAS 2.0 total score was 0.79 that was below 0.85 in this study that implied the total score of WHODAS 2.0 should only be used on population or group level [25]. The observed difference between PSI and Cronbach’s  $\alpha$  may result from the fact that the PSI is sensitive to a skewed distribution of scores and increases as the scores become more extreme, while there is no such effect for Cronbach’s  $\alpha$  [26]. Even though the total score of WHODAS 2.0 distribution among SCI patients was negatively skewed (The skewness =  $-0.091$ ) that does not affect the applicability of this measurement tool. The exploratory factor analysis confirmed the original 6-factor structure in most parts and therefore all items were included in the following Rasch analysis [27].

To a large extent, both analyses, the exploratory factor analysis and the Rasch analysis, support the multidimensionality of the WHODAS and the validity of the construct used by WHO. Similar to the Rasch analysis, which indicates a certain independence of the item D6.4 “How much time did you spend on your health condition, or its consequences”, the exploratory factor analysis loadings were lowest on its domain factor and low on the other factors. The exploratory factor analysis and Rasch analysis both support the strong association of items D2.3 and D2.4 with the items from D3, the Self-Care domain. Interestingly, the exploratory factor analysis showed that item D4.3 loaded on Factor D1 “Understanding and communicating”. Also, the Rasch analysis showed high local dependency for all items of D1 and D4 “Getting along with people”, suggesting that the items from both domains could be joined to form one common domain. The association between D1 and D4 items could be explained by the Taiwanese culture, in which the items from the domain D4 are mostly perceived as a cognitive ability [4]. If researchers want to use

WHODAS 2.0 as an assessment instrument to evaluate adults with SCI in Taiwan or other Asia countries, it could be sensible to consider a formal cross-cultural adaption to ensure proper understanding of the items.

The decrease of the reliability from the WHODAS 2.0 analysis with and without the testlets can be explained; the local dependencies caused a reliability inflation in the first Rasch analysis which included all WHODAS 2.0 domains [28]. This inflation was controlled in the testlet approach, which treats each domain like a super-item. Further, the testlet strategy allows overcoming the dimensionality issues found in the first analysis and provides the true reliability of the WHODAS 2.0 scale, as well as information about the reliability of the raw scores. A WHODAS 2.0 summary score for the entire SCI population could be confirmed. This testlet approach was also successfully applied in a stroke population [29].

In the light of disability eligibility determination, it will be crucial to see whether our solutions can be replicated in a chronic SCI population or in different diagnostic groups. Although we expected a substantial influence of the cultural background no differential item functioning for gender was found in item D5.1 “Taking care of your household responsibilities”, e.g., social roles of males and females in Asian countries, where household tasks that are traditionally performed by women in Taiwan. This means that the item was understood and answered the same by both male and female participants. Only the domain D6: “Participation in society” showed uniform differential item functioning for age in the overall score, which would be reasonable given the changing participation patterns in different life phases. Typically, differential item functioning can be solved by applying a so-called item split [30]. Splitting an item for differential item functioning would result in estimating different item difficulties across levels of person factor subgroups that cause differential item functioning. However, one must keep in mind that adjusting the item difficulties would then favor the subgroup with more difficulties on an item to adjust for the difference. This adjustment, of course, would impact the comparability of the person ability across groups. Also, different cut-off scores must then be calculated as the subgroups may not be comparable anymore. For these reasons, we decided to report for differential item functioning without undertaking any adjustments.

## Study limitation

Our participants first registered in the National Disability Determination System in Taiwan between 2013 and 2015. In other words, the study period was limited. Therefore, it will be applicable to provide more information about their work status, especially the work items of D5-2.

Nevertheless, our findings indicate that WHODAS 2.0 is valid to assess and document the level of activity and participation of adults with SCI in Taiwan. The external validity issue would be most careful because the results might not apply to a sub population that has not adequately participated in the study.

## Conclusion

The present study is the first to examine the psychometric properties of WHODAS 2.0 in a population of persons with chronic SCI in Taiwan using a classical and modern test approach. Our study could provide reference information for the Taiwanese government to assess functioning based on WHODAS 2.0 for disability policy in the future.

## Data archiving

The data that support the findings of this study are available from Ministry of Health and Welfare in Taiwan but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publically available. Data are however available from the authors upon reasonable request and with permission of Ministry of Health and Welfare in Taiwan.

**Acknowledgements** We thank the Swiss Paraplegic Research (SPF) in Switzerland for its kind support and the ICF team in Taiwan. Special thanks also to Christine Richards of the University of Vermont's Department of Rehabilitation and Movement Science, and Melissa Selb from SPF for their assistance in the editing of this manuscript.

**Funding** This research was financially supported by the Ministry of Science and Technology of Taiwan, MOST (104-2917-I-320 -001). All authors declare sources of funding material not available commercially.

**Author Contributions** T-YC was responsible for writing the manuscript, conducting the statistical analysis, interpreting the results, and the discussion. RE and CFY were responsible supervised for the whole study, screening potentially eligible studies, and providing feedback on the manuscript. MEF and CSF conducted the statistical analyses, supported the writing of the manuscript of methods, results, and discussion and provided feedback on the entire manuscript. T-HL and C-FY contributed to the data extraction and provided feedback on the manuscript.

## Compliance with ethical standards

**Ethical approval** The study was approved by the Tzu-Chi research ethics committee (IRB-104-4-A)

**Conflict of interest** The authors declare that they have no conflict of interest.



**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Teng SW, Yen CF, Liao HF, Chang KH, Chi WC, Wang YH, et al. Evolution of system for disability assessment based on the International Classification of Functioning, Disability, and Health: A Taiwanese study. *J Formos Med Assoc.* 2013;112:691–8.
2. World Health Organization W. International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization; 2001.
3. Chiu WT, Yen CF, Teng SW, Liao HF, Chang KH, Chi WC, et al. Implementing disability evaluation and welfare services based on the framework of the international classification of functioning, disability and health: experiences in Taiwan. *BMC Health Serv Res.* 2013;13:416.
4. Yen CF, Hwang AW, Liou TH, Chiu TY, Hsu HY, Chi WC, et al. Validity and reliability of the Functioning Disability Evaluation Scale-Adult Version based on the WHODAS 2.0—36 items. *J Formos Med Assoc.* 2014;113:839–49.
5. Chiu TY, Yen CF, Chou CH, Lin JD, Hwang AW, Liao HF, et al. Development of traditional Chinese version of World Health Organization disability assessment schedule 2.0 36—item (WHODAS 2.0) in Taiwan: validity and reliability analyses. *Res Dev Disabil.* 2014;35:2812–20.
6. Tractenberg RE. Classical and modern measurement theories, patient reports, and clinical outcomes. *Contemp Clin Trials.* 2010;31:1–3.
7. Gabbe BJ, Lyons RA, Simpson PM, Rivara FP, Ameratunga S, Polinder S, et al. Disability weights based on patient-reported data from a multinational injury cohort. *Bull World Health Organ.* 2016;94:806C–16.
8. World Health Organization W. International perspectives on spinal cord injury. Geneva: World Health Organization; 2014.
9. Ning GZ, Yu TQ, Feng SQ, Zhou XH, Ban DX, Liu Y, et al. Epidemiology of traumatic spinal cord injury in Tianjin, China. *Spinal Cord.* 2011;49:386–90.
10. Kuo C-Y, Liou T-H, Chang K-H, Chi W-C, Escorpizo R, Yen C-F, et al. Functioning and disability analysis of patients with traumatic brain injury and spinal cord injury by using the world health organization disability assessment schedule 2.0. *Int J Environ Res Public Health.* 2015;12:4116–27.
11. Garin O, Ayuso-Mateos JL, Almansa J, Nieto M, Chatterji S, Vilagut G, et al. Validation of the "World Health Organization Disability Assessment Schedule, WHODAS-2" in patients with chronic diseases. *Health Qual Life Outcomes.* 2010;8:51.
12. Chang K-H, Liao H-F, Yen C-F, Hwang A-W, Chi W-C, Escorpizo R, et al. Association between muscle power impairment and WHODAS 2.0 in older adults with physical disability in Taiwan. *Disabil Rehabil.* 2015;37:712–20.
13. Üstün TB, Kostanjsek N, Chatterji S, Rehm J. Measuring health and disability: manual for WHO disability assessment schedule WHODAS 2.0. Geneva: World Health Organization; 2010.
14. Wolf AC, Tate RL, Lannin NA, Middleton J, Lane-Brown A, Cameron ID. The World Health Organization Disability Assessment Scale, WHODAS II: reliability and validity in the measurement of activity and participation in a spinal cord injury population. *J Rehabil Med.* 2012;44:747–55.
15. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297–334.
16. Rasch GW. Probabilistic Models for some Intelligence and Attainment Tests, Copenhagen, 1960.
17. Yen W. Performance assessments: strategies for managing local item dependence. *J Educ Meas.* 1993;30:187–213.
18. Wainer H, Kiely GL. Item clusters and computer adaptive testing: a case for testlets. *J Educ Meas.* 1987;24:185–210.
19. Masters GN, Wright BD. The Partial Credit Model. In: van der Linden WJ, Hambleton RK, (eds.) *Handbook of Modern Item Response Theory.* New York, NY: Springer New York; 1997. p. 101–21.
20. Raïche G. Critical eigenvalue sizes (variances) in standardized residual principal components analysis. *Rasch Meas Trans.* 2005;19:1012.
21. Streiner DL, Norman GR. Health measurement scales : a practical guide to their development and use. Oxford: Oxford University Press; 2008.
22. Bond T, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. UK: Taylor & Francis; 2015.
23. Reinhardt JD, Post MW, Fekete C, Trezzini B, Brinkhof MW. Labor Market integration of people with disabilities: Results from the Swiss Spinal Cord Injury Cohort Study. *PLoS ONE.* 2016;11: e0166955.
24. Posl M, Cieza A, Stucki G. Psychometric properties of the WHODASII in rehabilitation patients. *Qual Life Res.* 2007;16:1521–31.
25. Duncan PW, Bode RK, Min Lai S, Perera S. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil.* 2003;84:950–63.
26. Andrich D. An index of person separation in latent trait theory, the traditional KR-20 index and the Guttman scale response pattern. *Educ Res Pers.* 1982;9:95.
27. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate data analysis. NJ: PearsonPrentice Hall Upper Saddle River; 2006.
28. Baghaei P. Local dependency and Rasch measures. *Trans Rasch Meas.* 2007;21:1105–6.
29. Kucukdeveci AA, Kutlay S, Yildizlar D, Oztuna D, Elhan AH, Tennant A. The reliability and validity of the World Health Organization Disability Assessment Schedule (WHODAS-II) in stroke. *Disabil Rehabil.* 2013;35:214–20.
30. Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care.* 2004;42(1 Suppl): 137–48.