

Reliability of the Wheelchair Skills Test Version 4.1

Trent Magis, Patricia Walton, Jessica Rispin & Noelle Lindquist

Department of Physical Therapy, University of Alberta, Edmonton, AB Canada



Background / Purpose: The Wheelchair Skills Test (WST) was developed to evaluate functional wheelchair skills required in daily life. The WST consists of thirty-two skills which are listed in Table 2. Previous studies have examined the reliability of the WST, but no reliability studies exist for the current version 4.1, which includes a safety component. The purpose of our study was to examine the inter-rater, intra-rater and test-retest reliability for performance and safety scores of the WST version 4.1.

Materials/Methods: The WST was administered according to the WST 4.1 test manual (<http://www.wheelchairskillsprogram.ca/>). All raters completed training prior to testing. Performances were videotaped. Each participant was tested twice with a minimum of one week between trials. Following completion of all participants' trials, four raters scored the WST from the video recordings. Two raters viewed the first trial of all subjects twice for intra-rater analysis. Two raters viewed the first and second trials of all subjects for test-retest analysis. Scoring of the first trial by all four raters was used for inter-rater analysis. There was a minimum of two weeks elapsed time between first and second viewings. Intra-class correlation coefficients (ICC) were used to calculate inter-rater, intra-rater, and test-retest reliability for both performance and safety scores. Table 1 displays ICC values, while Table 2 displays percent rater agreement for each skill.

Participants: Eleven participants volunteered, including nine males and two females, with ages ranging from 17 to 66 years. Participant diagnoses included stroke, spinal cord injury, spina bifida, and transverse myelitis. The participants' years of experience with a wheelchair ranged from one to thirty-seven years. All participants used their wheelchairs at least fifty percent of the day.



Table 1. ICCs for Inter-rater, Intra-rater, and Test-retest Reliability of WST v4.1

	Performance Component			Safety Component		
	ICC	95% confidence limit	p-values	ICC	95% confidence limit	p-values
Inter-rater	0.855	0.683-0.953	< 0.001	0.061	-0.086-0.384	0.243
Intra-rater	0.950	0.880-0.984	<0.001	0.228	-0.034-0.609	0.048
Test-retest	0.901	0.768-0.971	<0.001	0.254	-0.026-0.651	0.041

Table 2. Percent Rater Agreement for Trial One

Item Skill	Performance	Safety
	% Rater Agreement	% Rater Agreement
1 Rolls forward 100m	100	100
2 Rolls forward 100m in 30s	100	100
3 Rolls backward 5m	68	68
4 Turns 90° moving forward	100	100
5 Turns 90° moving backward	100	100
6 Turns 180° in place	95	100
7 Maneuvers sideways	85	89
8 Gets through hinged door	88	88
9 Reaches 1.5m object	100	100
10 Picks object from floor	100	100
11 Relieves weight from buttocks	91	91
12 Transfers (WC to bench and back)	82	76
13 Folds and unfolds WC	92	92
14 Rolls 100m	100	100
15 Avoids moving obstacles	100	100
16 Ascends 7.5° incline	95	95
17 Descend 7.5° incline	100	100
20 Rolls 2m across 5° side slope	100	100
21 Rolls 2m on soft surface	100	100
22 Gets over 15cm pothole	95	95
23 Gets over 2cm threshold	95	95
24 Ascends 5cm level change	95	100
25 Descends 5cm level change	100	100
26 Ascends 15cm curb	91	92
27 Descends 15cm curb	95	95
28 Performs 30s stationary wheelie	94	94
29 Turns 180° in place in wheelie	91	86
30 Gets from ground into WC	83	86
31 Ascends stairs	91	91
32 Descends stairs	82	85

WC = wheelchair

Skills 18 and 19 were omitted from the table because the test was modified to include a 7.5° ramp instead of a 5° and 10° ramp.

Discussion: Reliability of the performance component was excellent. The safety component had significantly lower ICC scores, which may be a result of homogeneous safety scores. Discrepancies in agreement of performance and safety scores were organized into one of four categories: Video Error, Safety Misinterpretation, Performance Interpretation and Scoring Rule Interpretation. Four skills seemed to be most problematic and had less agreement between raters: Skill 3 - Rolling Backwards, Skill 12 - Transferring from W/C to Bench, Skill 30 - Getting from the Ground to W/C, and Skill 32 - Descending Stairs.

Conclusion: Suggestions for improvement in future studies include shorter time frame between tester training and scoring, larger sample size, and more practice scoring the WST. Clearer guidelines regarding the safety component and number of attempts allowed may also result in improved reliability scores.

Acknowledgements

Supervisor: Patricia Manns

External Advisor: Lee Kirby

Special Thanks: Jutikarn Charoensuk, Manu Singla, Karen Slater