ISSN: 2583-5386
DOI: doi.org/10.51219/JIH/claudio-perret/49

URF PUBLISHERS
connect with research world

Journal of Integrated Health

https://urfpublishers.com/journal/integrated-health

Vol: 3 & Iss: 3

Research Article

Reliability of a Cognitive Test Battery for Individuals with a Spinal Cord Injury

Aglaia Emmenegger[1], Fabian Grossmann[2,3], Selin Scherrer[1,4] and Claudio Perret[5,6,*]

[1]Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

[2]Swiss Paraplegic Centre, Sports Medicine, Nottwil, Switzerland

[3]Human Physiology and Sports Physiotherapy Research Group, Vrije Universiteit Brussel, Brussel, Belgium

[4]Faculty of Science and Medicine, University of Fribourg, Fribourg, Switzerland

[5]Swiss Paraplegic Research, Nottwil, Switzerland

[6]Faculty of Health Sciences and Medicine, University of Lucerne, Lucerne, Switzerland

Citation: Emmenegger A, Grossmann F, Scherrer S, Perret C. Reliability of a Cognitive Test Battery for Individuals with a Spinal Cord Injury. *J Integrated Health* 2024;3(3): 266-271. DOI: doi.org/10.51219/JIH/claudio-perret/49

Received: 11 July, 2024; **Accepted:** 22 July, 2024; **Published:** 24 July, 2024

***Corresponding author:** Prof. Dr. Claudio Perret, Swiss Paraplegic Research, Guido A. Zäch-Strasse 4, CH-6207 Nottwil, Switzerland, E-mail: claudio.perret@paraplegie.ch

ABSTRACT

Individuals with spinal cord injury (SCI) have a higher incidence of cognitive deficits than the able-bodied (AB) population. In literature, there is only little reference about the reliability of cognitive test batteries in individuals with SCI. Therefore, the present study investigated the reliability of a test battery consisting of the Stroop Color and Word Test (SCWT) and four subtests of the Test of Attentional Performance (TAP) in individuals with SCI.

Thirteen volunteers with a traumatic SCI participated in this study. After a familiarization trial, each participant was tested twice on two different days, separated by seven days. The within and between-day intraclass correlation coefficient (ICC), the limits of agreement (LOA), the standard error of measurement (SEM) and the smallest real differences (SRD) were calculated for SCWT and TAP data over the four sessions.

SCWT presented excellent within-day (ICC: 0.92 / 0.98) and between-day (ICC: 0.92 / 0.97) reliability. In the TAP, mean overall reaction time showed excellent within-day reliability (ICC: 0.93 / 0.93) and good to excellent between-day reliability (ICC: 0.82 / 0.90). Therefore, the test battery showed encouraging results within the population of individuals with SCI. Even though the ICC results look promising, the test battery should be assessed with a higher number of participants to reduce LOA, SEM and SRD.

Keywords: paraplegia, cognition, reaction time, working memory, rehabilitation, attention

1. Introduction

Several scientists support the hypothesis that cognition is important for individuals with spinal cord injury (SCI) to gain independence and get reintegrated into the community for review see[1]. However, cognitive deficits seem to be present in

10-60% of individuals with SCI (for more details refer to the systematic review by Sachdeva, 2018)[2] which can lead to a more complicated and therefore prolonged rehabilitation process in this population.

The reasons for a higher prevalence of cognitive deficits in individuals with SCI are various. Possible causes can be mood disorders[3-6] chronic pain[7-9], poor sleep and fatigue[10] or medication side[11]. However, the most cited cause of cognitive impairment in individuals with SCI is traumatic brain injury (TBI). Concurrent severe TBIs often delay motor recovery after SCI, due to impairments in processing speed, language comprehension, memory and problem solving[1]. As there is a relatively high prevalence for cognitive deficits, it seems important to assess cognitive function during the rehabilitation process of individuals with SCI. For this purpose, it is crucial to have validated measures and tools. However, to our knowledge, there is a lack of studies investigating the reliability of a cognitive test battery in people with SCI.

To fill this research gap, the aim of the present study was to investigate the test-retest reliability of the outcome parameters from a cognitive test battery including the Stroop Color and Word Test (SCWT) and the computer-based Test of Attentional Performance (TAP) including the subtests for Alertness, Divided Attention, Go/Nogo and Working Memory in individuals with traumatic SCI.

The challenges described above can lead to limited cognitive performance (also day-dependent) in individuals with SCI, which may contribute to a higher variability in test results compared to an able-bodied (AB) population. Therefore, we expected less reliable results than known from studies with AB individuals.

2. Materials and Methods

2.1. Participants

Thirteen volunteers were recruited (**Table 1**). One had to be excluded in the SCWT, because of color blindness. Inclusion criteria were (i) traumatic SCI between C5 and L1 (ii) age between 18 to 65 years (iii) good German language skills (iv) ability to push a button by hand and (v) no severe head- or brain-injury. Participants gave their written informed consent before the start of the measurement. Ethical approval for this study was obtained from the local ethics committee (EKNZ, Basel, Switzerland, project-ID: 2020-00057).

Table 1: Participants' characteristics.

ID	Age [years]	Gender	Lesion Level	AIS	Time since injury [years]	Highest Education
1	26	Male	TH3	A	7	University
2	31	Male	TH5	A	3	Vocational Training
3	21	Male	TH11	A	5	Vocational Training
4	27	Male	TH6	A	1	Vocational Training
5	32	Male	TH10	A	6	Vocational Training
6	48	Female	TH5	A	13	University
7	51	Female	TH6	A	42	Vocational Training
8	43	Female	TH10	A	1	University
9	63	Male	TH5	A	20	Vocational Training
10	21	Male	TH3	A	2	Vocational Training
11	29	Male	C6	A	4	Vocational Training
12	58	Male	C5	A	2	Vocational Training
13	50	Male	C5	D	1	University
Median IQR	32				4	
	23				5	

IQR, interquartile range; TH, thoracic lesion; C, cervical lesion; AIS, American Spinal Injury Association (AISA) impairment scale[12].

2.2. Test battery

The test battery in the present study started with an adapted version of Golden's SCWT[13]. Three different charts were presented to the participants: (i) color words printed in black (ii) color patches and (iii) color words printed in another color. In every task, the goal was to read (i) or name (ii and iii) as many colors as possible in 30 seconds. The outcome parameter for the SCWT was the number of correct answers for each task.

For the second part of the test battery, four subtests of the computer-based TAP (Version 2.3.1, Psytest, Vera Fimm, Herzogenrath, Germany), were chosen. The used subtests Alertness, Divided Attention, Go/Nogo and Working Memory are described in detail elsewhere[14]. The outcome parameters for all the subtests were mean reaction time (RT), as well as the number of mistakes made by the participant. For the overall performance, RT and mistakes were summed over all subtests.

2.3. Study design

Figure 1 shows the design that was chosen for this study. Each participant was tested twice on three different days, separated by one week. The first day served as familiarization trial and was not included in the data analysis.

For each individual, every appointment started at the same time in the morning between 8 and 12 am. The tests were conducted in a quiet room at the Swiss Paraplegic Centre, where temperature (21°C), light and distance to the screen (50 cm) were standardized.

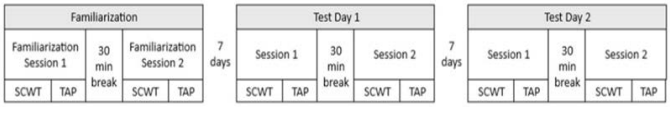

Figure 1: Study design. The Stroop Color and Word Test (SCWT) and the Test of Attentional Performance (TAP) were performed during a familiarization trial, followed by two test days, which were analysed for the study.

2.4. Statistical analysis

The software R (R Foundation for Statistical Computing version 4.0.0; Vienna, Austria) was used to analyze the data. First, all measured parameters were checked for normality using Shapiro Wilk's tests. As data were not normally distributed, median and interquartile range (IQR) were calculated for all parameters. Differences between the appointments were assessed with the Wilcoxon signed-rank test. Statistical significance was set at the 0.05 level. Interclass correlation coefficients ($ICC_{3,1}$) were used to evaluate the relative test-retest reliability of the outcome parameters. $ICC_{3,1}$ is a two-way mixed single measure of absolute agreement[15]. The ICC values were interpreted as poor (< 0.5), moderate (0.5 - 0.749), good (0.75 - 0.9), or excellent (> 0.9)[16]. Within-day reliability was calculated and interpreted for day 1 session 1 (ICC_11) versus day 1 session 2 (ICC_12) and for day 2 session 1 (ICC_21) versus day 2 session 2 (ICC_22). Additionally, between-day reliability for day 1 session 1 (ICC_11) versus day 2 session 1 (ICC_21) and for day 1 session 2 (ICC_12) versus day 2 session 2 (ICC_22) were calculated.

The standard error of measurement (SEM) and the smallest real difference (SRD) were chosen to determine absolute reliability[17]. The SEM was calculated with Equation (1) (*SD* = standard deviation; r_{xx} = reliability of the test). The percentage of SEM (SEM%) was calculated with Equation (2). Equation (3) was used to determine SRD. With Equation (4) percentage of SRD (SRD%) was calculated.

$$SEM = SD\sqrt{1 - r_{xx}} \quad (1)$$

$$SEM\% = SD\sqrt{1 - \frac{r_{xx}}{mean\ T1+T2}} \times 100 \quad (2)$$

$$SRD = 1.96 \times SEM \times \sqrt{2} \quad (3)$$

$$SRD\% = 1.96 \times SEM \times \sqrt{\frac{2}{mean\ T1+T2}} \times 100 \quad (4)$$

With the Bland-Altman method the differences between the test and retest were examined. Based on the mean difference between the test and retest and the standard deviation (SD) of the difference, the 95% limits of agreement (LoA95)[18] were calculated with the Equation (5).

$$LoA95 = mean\ difference \pm 2 \times SD\ difference \quad (5)$$

3. Results

Within-day and between-day reliability analyses for the SCWT parameters are summarized in **Table 2**, the analyses for the TAP parameters in **Table 3**. All SCWT parameters showed excellent reliability (ICC > 0.90). Mean overall reaction time showed excellent within-day reliability (ICC = 0.93). Between-day reliability was good as well (ICC_ 11 vs.. ICC_21 = 0.82; ICC_12 vs.. ICC_22 = 0.90). Bland-Altman plots for the SCWT parameters are shown in **Figure 2**, for the TAP parameters in **Figure 3**. Even tough ICCs showed good reliabilities, SEM and SRD demonstrated high numbers.

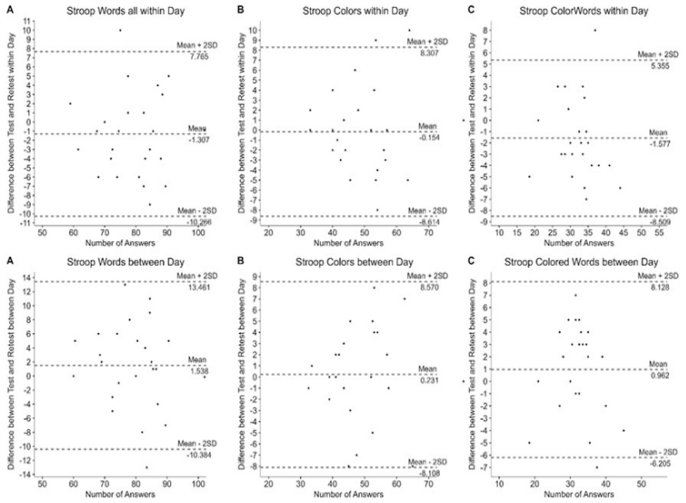

Figure 2: Bland-Altman plot for Stroop parameters A) Words, B) Colors and C) Color-words. On top the within-day reliability, on the bottom the between-day reliability is presented. The middle dashed line represents the mean difference between test and retest. The two outer dashed lines represent the limits of agreement for the number of answers. N = 12; One subject had to be excluded due to color-blindness.

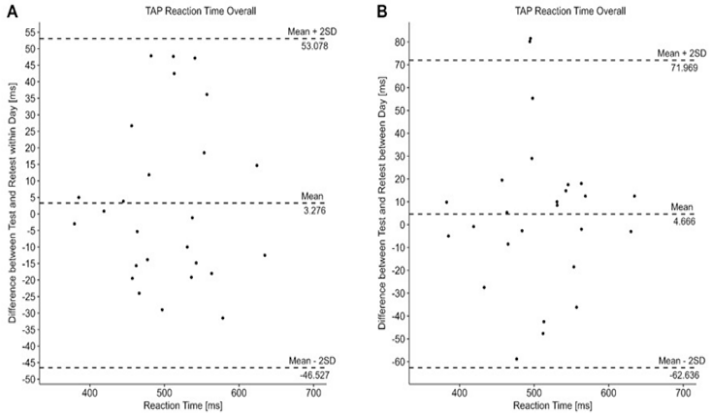

Figure 3: Bland-Altman plot for Overall Test of Attentional Performance (TAP) reaction times A) Within-day and B) Between-day. The middle dashed line represents the mean difference between test and retest. The two outer dashed lines represent the limits of agreement for the reaction times.

Table 2: Relative and absolute reliability for the Stroop test (N = 12)*.

			ICC				Bland-Altman LOA		
			ICC	p-value	SEM	SRD	Bias	Upper	Lower
			[95% CI]		**[SEM%]**	**[SRD%]**	**[N]**	**[N]**	**[N]**
Word	**Within-day**	11 vs.. 12	0.98 [0.92;0.99]	<0.001	4.71	13.05	-2.31	6.4	-11.01
		21 vs. 22	0.98 [0.95;0.99]	<0.001	4.1	11.32	-0.31	8.44	-9.05
Word	**Between-day**	11 vs. 21	0.96 [0.87;0.99]	<0.001	6.47	17.93	-2.54	10.05	-15.13
		12 vs. 22	0.97 [0.92;0.99]	<0.001	5.1	14.11	-0.54	10.3	-11.38
Color	**Within- day**	11 vs. 12	0.96 [0.86;0.99]	<0.001	7.47	20.71	0.77	10.38	-8.84
		21 vs. 22	0.98 [0.93;0.99]	<0.001	5.32	14.75	-1.07	5.54	-7.7
Color	**Between-day**	11 vs. 21	0.97 [0.9;0.99]	<0.001	6.21	17.21	0.69	8.67	-7.29
		12 vs. 22	0.97 [0.89;0.99]	<0.001	6.58	18.21	-1.15	7.12	-9.43
Color-Word	**Within- day**	11 vs. 12	0.92 [0.77;0.98]	<0.001	10.02	27.77	-1.31	6.87	-9.48
		21 vs. 22	0.96 [0.81;0.99]	<0.001	7.36	20.4	-1.85	3.51	-7.21
Color-Word	**Between-day**	11 vs. 21	0.92 [0.77;0.98]	<0.001	9.69	26.85	-0.69	7.4	-8.79
		12 vs. 22	0.97 [0.87;0.99]	<0.001	7.3	20.24	-1.23	4.82	-7.28

ICC: Intra class Correlation Coefficient; CI: Confidence Interval; SEM, Standard Error of Measurement; SRD, Smallest Real Difference; LOA, Limits of Agreement; 11: First test on test day 1; 12: Second test on test day 1; 21: First test on test day 2; 22: Second test on test day 2.

*One subject had to be excluded due to colour-blindness.

Table 3: Relative and absolute reliability for the overall mean reaction time in the TAP test.

		ICC				Bland-Altman LOA		
		ICC	p-value	SEM	SRD	Bias	Upper	Lower
		[95% CI]		**[SEM%]**	**[SRD%]**	**[ms]**	**[ms]**	**[ms]**
Within-day	11 vs. 12	0.93 [0.80;0.98]	<0.001	3.34	9.25	3.19	52.75	-46.37
	21 vs. 22	0.93 [0.78;0.98]	<0.001	3.46	9.56	3.36	53.42	-46.71
Between-day	11 vs. 21	0.82 [0.52;0.94]	<0.001	5.54	15.35	12.53	90.58	-65.53
	12 vs. 22	0.90 [0.71;0.97]	<0.001	4.08	11.32	12.69	67.24	-41.58

ICC: Intra class Correlation Coefficient; CI: Confidence Interval; SEM: Standard Error of Measurement; SRD: Smallest Real Difference; LOA: Limits of Agreement; 11: First test on test day 1; 12: Second test on test day 1; 21: First test on test day 2; 22: Second test on test day 2.

4. Discussion

Existing literature on the reliability of cognitive test batteries in individuals with SCI is scarce. Only Nightingale, et al.[19] investigated test-retest reliability of different neuropsychological tests across a range of cognitive domains in this population. However, they used a different test battery, which makes direct comparison to the present study results difficult. Nevertheless, most ICCs (0.77-0.93) were found to be similar compared to the present study. The aim of the present study was to remedy that lack of knowledge by assessing the reliability of the TAP subtests Alertness, Divided Attention, Go/Nogo and Working Memory, as well as of the SCWT in this population.

5. Stroop Colour and Word Test (SCWT)

The SCWT presents good reliability in AB populations[20-23]. We hypothesized to find less reliable results in the SCI population. However, the present study showed excellent within-day and between-day reliability (ICC > 0.9) for all SCWT parameters. Nevertheless, the LOAs shown in the Bland-Altman plots (**Figure 2)** are rather big. Therefore, a big variability was present in the dataset. The same can be seen related to SEM and SRD; both values present rather high numbers in our sample. Those findings might partly be explained by the number of participants: The greater the number of samples, the narrower the LOAs will be[24] and the better the population would be represented, which would lower the SEM as well. Moreover, inconsistency of test performance might be associated with broader LOAs. As there is evidence that individuals with SCI have some deficits in cognitive performance[2], it could be possible that performance varies between tests. Different psychological and somatic comorbidities in individuals with SCI such as pain, fatigue, anxiety, and others are associated with a decline in cognitive ability[2]. Those comorbidities can appear as poor performance in general, or as inconsistent performance. In the present study, the conditions were standardized as much as possible. Nevertheless, it seems to be difficult to compare performances within the same individual, as some factors, such as fatigue, pain, or medication side effects may vary daily. The population with SCI could be more vulnerable to those daily changes than an able-bodied population. This addresses the importance of standardization, also for factors like sleep quality, pain, or medication.

6. Test of Attentional Performance (TAP)

Zimmermann and Fimm[14] assessed reliability for the TAP within the AB population. They observed good reliabilities for reaction times. In the present study good to excellent within-day (ICC = 0.93) and between-day (ICC = 0.82 and 0.90) reliability was demonstrated in participants with SCI. Within-day reliability was better than between-day reliability. Therefore, the performance seems to be slightly inconsistent between different days. The reasons for this finding are already discussed in the section above for the SCWT and are thought to be the same for the TAP test.

7. Limitations

As already addressed in the discussion, it is believed to get better values for LOAs, SEM and SRD if the number of participants would be higher. Especially because individuals with SCI seem to be sensitive to various factors that can influence cognitive function. Furthermore, the population itself is very diverse, as different lesion levels are associated with different complications. Therefore, to represent the population accurately, one would need a much higher number of participants and divide the population into smaller subgroups, for example according to the level of lesion. However, the population in question is relatively small, which makes it challenging to recruit many volunteers for a study protocol that requires a long time investment and is not particularly appealing for the participants.

Another issue is the missing randomization. The different tests were done in the same order for every session: SCWT, TAP Alertness, TAP Divided Attention, TAP Go/Nogo and TAP Working Memory. Because long and demanding cognitive activity can result in mental fatigue and can therefore deteriorate performance[25], the order should probably be randomized in future research. For now, we can only prove reliability for the given order of the test battery.

8. Conclusion

To conclude, the test battery consisting of the SCWT and the four subtests Alertness, Divided Attention, Go/Nogo and Working Memory of the TAP seems to be a reliable tool to assess cognitive function in the population of SCI. However, one has to be aware that individuals with SCI can be vulnerable to daily changes of performance as a result of pain, poor sleep or other factors. This can lead to inconsistent performance between days, which emphasizes the importance of standardization for factors like sleep quality, pain, or medication, which might be difficult to apply in a clinical setting.

9. Authors Contribution

AE, SS and CP were responsible for the study conception and study protocol. AE and SS collected and analysed the data. FG provided statistical expertise. AE and CP prepared the first draft of the manuscript. All authors reviewed and revised the manuscript before submission and approved its content.

10. References

1. Cohen Matthew L, David S Tulsky, James A Holdnack, et al. Cognition among community-dwelling individuals with spinal cord injury. Rehabil Psychol, 2017; 62: 425-434.
2. Sachdeva Rahul, Feng Gao, Chetwyn CH. et al. Cognitive function after spinal cord injury: A systematic review. Neurology, 2018; 91: 611-621.
3. Boyer Bret, Michelle Knolls, Christina Kafkalas, et al. Prevalence of posttraumatic stress disorder in patients with pediatric spinal cord injury: Relationship to functional independence. Topics in Spinal Cord Injury Rehabilitation, 2000.
4. Dryden Donna M, L Duncan Saunders, Brian H Rowe, et al. Depression following traumatic spinal cord injury. Neuroepidemiology, 2005; 25: 55-61.
5. Krause James S, Lee L Saunders, Susan Newman. Posttraumatic stress disorder and spinal cord injury. Archives of Physical Medicine and Rehabilitation, 2010; 1182-1187.
6. Bonanno George A, Paul Kennedy Isaac R, Galatzer-Levy, et al. Trajectories of resilience, depression, and anxiety following spinal cord injury. Rehabilitation Psychology, 2012; 57: 236-247.
7. Dijkers Marcel, Thomas Bryce, Jeanne Zanca. Prevalence of chronic pain after traumatic spinal cord injury: A systematic review. Journal of Rehabilitation Research and Development, 2009.
8. Masri Radi, Asaf Keller. Chronic pain following spinal cord injury. In Advances in Experimental Medicine and Biology, 2012; 760.
9. Hassanijirdehi Marzieh, Mohammad Khak, Sohrab Afshari-Mirak, et al. Evaluation of pain and its effect on quality of life and functioning in men with spinal cord injury. Korean J Pain, 2015; 28: 129-136.
10. Avluk Ozlem Celik, Eda Gurcay, Ahmet Gurhan Gurcay, et al. Effects of chronic pain on function, depression, and sleep among patients with traumatic spinal cord injury. Annals of Saudi Medicine, 2014; 34.
11. Macciocchi Stephen N, Ronald T Seel, Nicole Thompson. The impact of mild traumatic brain injury on cognitive functioning following co-occurring spinal cord injury. Archives of Clinical Neuropsychology, 2013; 28: 684-691.
12. Maynard Frederick, Michael Bracken, Graham Creasey, et al. International standards for neurological and functional classification of spinal cord injury. Spinal Cord, 1997; 35: 266-274.

13. Golden Ch J. Stroop color and word test: A manual for clinical and experimental uses. Chicago: Stoelting, 1978.

14. Zimmermann Peter, Bruno Fimm. A test battery for attentional performance. In Applied Neuropsychology of Attention: Theory, Diagnosis, and Rehabilitation, 2002.

15. Shrout Patrick E, Joseph L Fleiss. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 1979; 86: 420-428.

16. Koo Terry K, Mae Y Li. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine. 2015; 15: 155-163.

17. Beckerman H, ME Roebroeck, GJ Lankhorst, et al. Smallest real difference, a link between reproducibility and responsiveness. Quality of Life Research, 2001; 10: 571-578.

18. Martin Bland J, Douglas G Altman. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, 1986; 327: 307-310.

19. Nightingale Tom Edward, Chloe Ah Ryung Lim, Rahul Sachdeva, et al. Reliability of cognitive measures in individuals with a chronic spinal cord injury. PM and R, 2019; 11: 1278-1286.

20. Santos JF, JR Montgomery. Stability of performance on the color-word test. Perceptual and Motor Skills, 1962.

21. Jensen AR. Scoring the stroop test. Acta Psychologica, 1965; 24: 398-408.

22. Golden Charles J. A group version of the stroop color and word test. Journal of Personality Assessment, 1975; 39.

23. Franzen Michael D, Amy C Tishelman, Brian H Sharp, et al. An investigation of the test-retest reliability of the stroop colorword test across two intervals. Archives of Clinical Neuropsychology, 1987; 2: 265-272.

24. Giavarina, Davide. Understanding Bland Altman Analysis. Biochemia Medica, 2015; 25: 359-372.

25. Lorist Monicque M, Maarten AS Boksem, K Richard Ridderinkhof. Impaired Cognitive Control and Reduced Cingulate Activity during Mental Fatigue. Cognitive Brain Research,2005;24: 199-205.