

## Research Article

# Motor demands of cognitive testing may artificially reduce executive function scores in individuals with spinal cord injury

Jinhyun Lee , Shauna Dudley-Javoroski , Richard K. Shields 

Department of Physical Therapy and Rehabilitation Science, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA

**Objective:** To determine whether the motor demands of cognitive tests contribute to differences in cognitive function scores in participants with and without spinal cord injury (SCI).

**Design:** Cohort study.

**Setting:** Rehabilitation research laboratory.

**Participants:** 68 individuals without SCI (“NON”) and 22 individuals with motor complete SCI (“SCI”).

**Interventions:** None.

**Outcome Measures:** NIH Toolbox cognitive assessments, including two with motor demands and reaction-time based scoring (Dimensional Change Card Sort (DCCS), Flanker Inhibitory Control and Attention (Flanker) and two without timed scoring (List Sorting Working Memory (List Sorting), Picture Sequence Memory Test (Picture Sequence)). Tests were administered with and without the assistance of a proctor on two randomly-determined days (>24 hr interval). For DCCS and Flanker, the motor-task score offset was estimated as the difference between the proctored and non-proctored scores.

**Results:** For demographically-corrected data, proctoring reduced DCCS and Flanker scores ( $P < 0.001$ ) but mitigated apparent differences between SCI and NON (all  $P > 0.403$ ). SCI and NON did not differ for List Sorting ( $P > 0.072$ ) but did differ significantly for Picture Sequence ( $P < 0.001$ ). Significant practice effects existed for memory-based tests (List Sorting and Picture Sequence); all  $P < 0.015$ , effect size  $> 0.645$ .

**Conclusions:** DCCS and Flanker scores for individuals with SCI may be artificially reduced consequent to secondary motor demands of the tests. Proctoring and computation of a motor-response score offset enables comparisons to be made between individuals with SCI and a Non-SCI control cohort; however, further work is needed to determine whether offset-adjusted scores can be compared to standardized normative values.

**Keywords:** Spinal cord injuries, Cognition, Executive function, Memory

## Introduction

To maximize function and independence, people with recent spinal cord injury (SCI) must undertake a stringent course of rehabilitation therapy as rapidly as medically feasible. During rehabilitation, patients must assimilate large amounts of new information and they must often learn to use complex adaptive technology. Patients must achieve these cognitive milestones within the compressed time frame imposed by the current cost-restricted health care economy.<sup>1,2</sup> Patients with cognitive impairment, a common sequelae of SCI,<sup>3</sup> may

struggle to rapidly learn all the information and skills presented during intensive rehabilitation.

The etiology of cognitive impairment may reflect the unique medical history of each individual. Undiagnosed traumatic brain injury appears to be a common factor,<sup>4</sup> as does autonomic dysfunction and associated systemic hypotension.<sup>5,6</sup> Grey matter degeneration in brain regions that process emotional states and attention<sup>7</sup> and the development of intrusive/impulsive processing<sup>8</sup> may both undermine executive function. New findings from animal models suggest that SCI instigates systemic inflammation and progressive neurodegeneration in brain regions that govern learning and cognition.<sup>9–12</sup> “Crystallized” cognitive domains (e.g. word knowledge

Correspondence to: Richard K. Shields, Department of Physical Therapy and Rehabilitation Science, University of Iowa, 1-252 Medical Education Building, Iowa City, IA 52242, USA. Email: richard-shields@uiowa.edu

and vocabulary) appear to be relatively unaffected by SCI, but deficits appear common in process-oriented dimensions such as executive function, working memory, and episodic memory.<sup>13</sup> Impairments in cognitive processes such as motor sequence learning<sup>14</sup> may be especially problematic if they interfere with post-SCI rehabilitation activities.

The National Institutes of Health Toolbox for Neurological and Behavioral Function – Cognition Battery (NIH Toolbox) is a computerized testing platform designed to enable standardized assessment of cognitive function.<sup>15,16</sup> It has been validated for use in populations with neurologic impairment, including SCI.<sup>17</sup> Participants complete the Toolbox assessment via an iPad app that requires them to reach forward from a standardized start position to press a button on a keyboard. Several Toolbox test instruments have a reaction time scoring vector, which superimposes a motor demand upon the cognitive task. Respondents with trunk or upper extremity motor impairment may take longer to respond to test items than participants without motor impairment. A number of groups, including the developers of NIH Toolbox, have proposed accommodations for participants with SCI such as avoiding tests with reaction-time scoring vectors,<sup>18</sup> statistical correction for hand function,<sup>13</sup> or simply accepting a high test non-completion rate (~40%) for participants with tetraplegia.<sup>13</sup> None of these approaches provides a satisfactory way to test speed-based components of cognition in participants with trunk and upper extremity impairment.

In this study, we investigated an alternate accommodation approach designed to mitigate the motor penalty of timed Toolbox metrics. All participants, including control subjects without SCI, completed the timed tests with the assistance of a proctor. Respondents verbally answered test items and the proctors then executed the standard motor response. The proctor's auditory processing time and motor reaction time were therefore added to the speed vector of the respondent's score. By comparing proctored and non-proctored test scores in a cohort of subjects without SCI, we quantified the mathematical score offset contributed by the motor demand of the task. We then applied this same mathematical offset to proctored test scores from participants with SCI. In this way, the study placed participants with motor impairment on an equal footing with participants who did not have motor impairment. The purpose of this study was to determine whether the motor demands of NIH Toolbox tests contribute to differences in cognitive function scores between participants with and without spinal

cord injury (SCI). For timed tests, we expected that differences between groups would be negated by using a test administration method that mitigated test motor demands (proctoring and application of a motor-response score offset). We hypothesized that after this adjustment, scores for participants with SCI would cease to differ from Non-SCI scores obtained with the conventional un-proctored method.

## Methods

68 individuals without SCI (“NON”) and 22 individuals with motor complete SCI (American Spinal Injury Association Impairment Scale-A<sup>19</sup>) (“SCI”) participated in a cohort study in a rehabilitation research laboratory. Demographic data appear in Table 1. All subjects were a convenience sample recruited from the healthy general population in a university community. 56 of the 68 participants were college graduates, 8 were college students, and 4 were high school graduates expecting to enter college in the next year. 34 were male and 59 were right handed. 60 of the participants were Caucasian, 5 were Asian, and 3 were African American. The SCI participants duration of injury was from 8 months to 32 years (mean = 13.4 years) and 50% of participants had quadriplegia. 12 of the SCI participants were high school graduates, 4 college graduates, and 2 college students. Exclusion criteria were a medical diagnosis of cognitive decline (e.g. dementia), traumatic brain injury, or a neurodegenerative condition (e.g. multiple sclerosis). All participants signed an informed consent document approved by our institution's Human Subjects Institutional Review Board.

### *Nih Toolbox testing*

Participants completed four NIH Toolbox standardized tests via an iPad app (software version 1.17.1650) with a Bluetooth keyboard interface. Two Toolbox tests included reaction time-dependent scoring vectors: Dimensional Change Card Sort (DCCS) and Flanker Inhibitory Control and Attention (Flanker). Two tests did not incorporate time-based scoring and depended upon accuracy alone: List Sorting Working Memory

**Table 1. Participant demographic data. Values are mean (SD).**

	Age	Height (cm)	Weight (kg)	SCI duration (yr)
NON-SCI (n = 68)	23.6 (2.3)	174.0 (9.5)	74.2 (14.0)	N/A
SCI (n = 22)	43.3 (15.9)	175.3 (34.5)	80.2 (23.7)	13.4 (8.3)

(List Sorting) and Picture Sequence Memory Test (Picture Sequence). Reliability and responsiveness of the NIH Toolbox instruments has been previously established.<sup>20–22</sup>

The DCCS is an executive function metric that provides an estimate of set shifting, an indicator of cognitive flexibility.<sup>20</sup> Participants view two images that differ according to two dimensions (color, shape). Participants then match a series of test images (e.g. blue truck, yellow ball) first by one dimension (e.g. color) and then by the other (e.g. shape). “Switch” trials are also given, in which participants must employ cognitive flexibility to change the dimension being matched. The participant’s score is automatically calculated as a composite of test accuracy and speed.<sup>23</sup>

The Flanker test is an executive function metric that estimates attention (allocating cognitive resources toward external stimuli) and inhibitory control (ignoring superfluous stimuli).<sup>20</sup> Participants identify the left-right orientation of a central arrow situated amidst four other arrows; these “flankers” may point in the same (congruent) or different (incongruent) direction as the center arrow. Incongruent trials test a participant’s capacity for inhibitory control. The participant’s score is automatically calculated as a composite of test accuracy and speed.<sup>24</sup>

The List Sorting test evaluates working memory, the capacity to store and manipulate information over a brief time.<sup>25</sup> The participant repeats a list of items (animals or foods) in order from smallest to largest according to one category (e.g. foods) and then according to both categories (e.g. first foods, then animals). Scores reflect test accuracy with no speed component.<sup>26</sup>

The Picture Sequence Memory test evaluates episodic memory, the cognitive capacity to acquire, store, and retrieve information.<sup>27</sup> The participant recalls sequences of up to 18 items, with scoring based on the number of correctly-recalled pairs of items. Scores reflect test accuracy with no speed component.<sup>28</sup>

### *Experimental procedures*

Ten proctors administered the NIH Toolbox metrics to between 6 and 10 non-SCI participants each. All proctors (4 women, 6 men) were college-educated adults with a mean age of 24.5 years (1.3). Each proctor received 3 h of training prior to administering any formal testing. The training consisted of a) general reading and familiarization with the NIH toolbox; b) review of the test order; and c) practice administering test on three other proctors followed by general reflection and discussion of differences as feedback. For ease of utility, we limited the extent of training to just

3 h. A single proctor administered the Toolbox metrics to all participants with SCI. This proctor was shown to be internally consistent with the other proctors with scores that fell within a 95% confidence interval of the other proctors (see results section). Instructions to the proctors emphasized the need to 1) enter the participant’s verbal response as rapidly as possible, and 2) to return the hand to the standardized “home” position between each response item. Non-SCI participants underwent one proctored session and one un-proctored session, separated by at least 24 h (mean = 5 days). The order of sessions was counter-balanced; that is, for every participant that received a proctored session first, there was a subsequent participant that received the un-proctored session first. All participants with SCI underwent one proctored session. In addition, a subgroup of participants with low paraplegia, and thus minimal trunk/upper extremity motor impairment, underwent both a proctored and un-proctored session. As with the non-SCI group, these repeat bouts were separated by at least 24 h and were counterbalanced.

### *Cognitive test scoring systems*

The iPad software calculated cognitive test scores as Uncorrected and Fully-Corrected T scores. Uncorrected scores contained no mathematical adjustment for demographic factors<sup>29</sup> and reflected a participant’s raw cognitive capacity for a particular cognitive domain.<sup>30</sup> Fully-Corrected T scores incorporated mathematical adjustment for age, sex, race, ethnicity, and educational attainment.<sup>29</sup> Higher values for both scales indicated better cognitive performance. Uncorrected scale scores are anchored to a nationally-representative normative sample with mean = 100 and SD = 15. Fully-Corrected T scores are anchored to a nationally-representative normative sample with mean = 50 and SD = 10. For DCCS and Flanker, reaction time for each participant was separately calculated as the mean response time in seconds for all presented test items.

Using data from participants without SCI, the motor-performance offset was calculated as the difference between a participant’s un-proctored test and their proctored test. For both DCCS and Flanker, mean (SD) offset was calculated across all 68 non-SCI participants for both the Uncorrected score and the Fully-Corrected score. These offsets were then applied to the SCI proctored scores, yielding an Adjusted score.

The offset represents the score penalty imposed by the motor demand of the DCCS and Flanker tasks. Converting the DCCS and Flanker tests into a verbal-response task reduces differences in test demands

between people with and without SCI. Calculation of the proctoring offset in the non-SCI group reveals the additional reaction time required for the proctor's response and the associated score penalty caused by the prolonged reaction time. This offset can then be added to proctored SCI scores to provide a motor-free estimate of cognitive function.

### Statistical procedures

Pilot testing of participants with SCI indicated that an SCI cohort of  $n > 18$  would provide  $>80\%$  power to detect significant differences from expected NON values (U.S. normative values). Differences between SCI and NON groups and between Proctored and Un-proctored tests were examined via the Kruskal-Wallis H Test for non-normally distributed data, followed by Dunn's post-hoc tests of pairwise differences with Benjamini-Hochberg false discovery rate adjustment for multiple comparisons. Based on the study aims, we *a priori* identified the key pairwise comparisons to be used during Benjamini-Hochberg adjustment. For DCCS and Flanker (timed tests that included calculation of the Adjusted score), pairwise comparisons of interest were: 1) NON Proctored vs NON Un-Proctored; 2) NON Proctored vs SCI Proctored; 3) NON Un-Proctored vs. SCI Un-Proctored; 4) NON Un-proctored vs SCI Adjusted (to determine whether apparent NON vs SCI differences persist after applying the mathematical offset); and 5) SCI Un-proctored vs SCI Adjusted (to determine whether application of the offset yields SCI values that are different from a group of people with SCI who have minimal arm/trunk impairment). For List Sorting and Picture Sequence (untimed tests with no calculation of an Adjusted score), pairwise comparisons of interest were: 1) NON Proctored vs NON Un-Proctored; 2) SCI Proctored vs SCI Un-Proctored; 3) NON Proctored vs SCI Proctored; 3) NON Un-Proctored vs SCI Un-Proctored. Significance was  $P < 0.05$  for all between-group analyses.

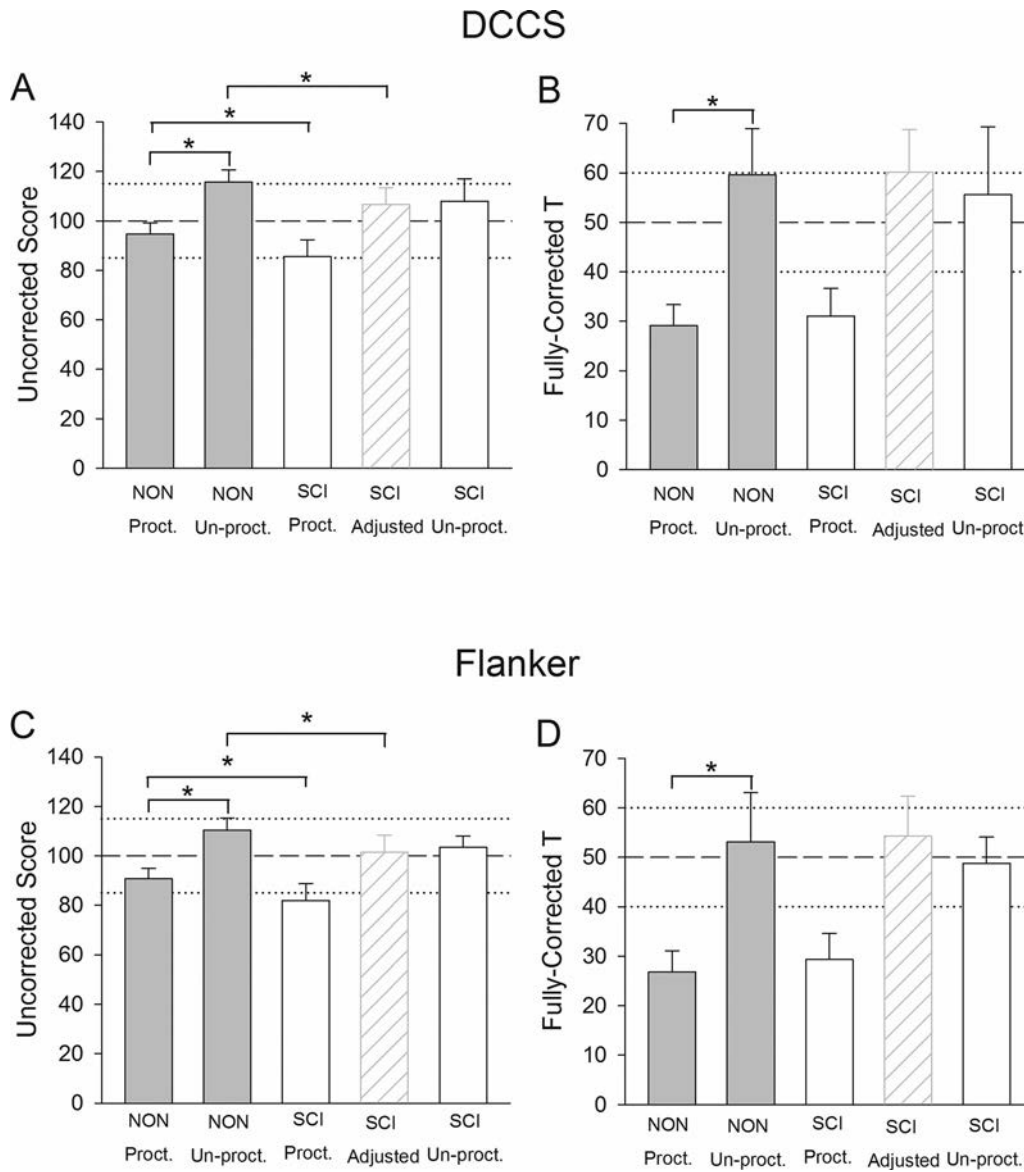
Previous studies have observed a significant practice effect for repeat assessment with NIH Toolbox.<sup>20–22</sup> We carried out a 2-way (day x proctor) ANOVA to assess if there was a day versus proctor interaction ( $P < 0.05$ ), followed by *post hoc* testing as indicated. We carried out a 1-way ANOVA on the proctor's offset values to determine if scores were different among the 10 proctors. We also conducted an analysis to assure that the single proctor used for all SCI bouts was consistent with the other proctors by using a Mann-Whitney rank sum test for non-parametric data ( $P < 0.05$ ) and assessing if the single proctor values

fell within the 95% confidence interval for the rest of the proctor pool. For Un-Proctored sessions (NON cohort only), we carried out a 1-way ANOVA to determine whether scores differed for Day 1 vs Day 2. When data were non-normally distributed we substituted the Kruskal-Wallis non-parametric H test with Dunn's post-hoc testing ( $P < 0.05$ ). When significant day effects were present we calculated the magnitude of the practice effect via effect size ([Day 2 mean – Day 1 mean] / Day 1 SD) and Cohen's D.<sup>31</sup>

### Results

All recruited participants completed both days of testing. Fully-Corrected scores were unavailable for one participant with SCI because the participant declined to provide demographic data needed for computation of the scale. Data for timed tests (DCCS and Flanker) appear in Fig. 1. Mean (SD) proctored reaction time for participants with SCI was 1.57 (0.38) seconds for DCCS and 1.43 (0.36) seconds for Flanker. NON reaction time and the motor-response offsets used to calculate the SCI Adjusted scores appear in Table 2. For the DCCS Uncorrected score, there was a significant effect of participant group ( $P < 0.001$ ). For the five pairwise comparisons of interest: 1) NON Proctored differed significantly from NON Un-Proctored ( $P < 0.001$ ); 2) NON Proctored differed significantly from SCI Proctored ( $P = 0.008$ ); 3) NON Un-Proctored did not differ from SCI Un-Proctored ( $P = 0.122$ ); 4) NON Un-Proctored differed significantly from SCI Adjusted ( $P = 0.006$ ); and 5) SCI Un-Proctored did not differ from SCI Adjusted ( $P = 0.957$ ) (Fig. 1A). For the DCCS Fully-Corrected score, a significant effect of participant group existed ( $P < 0.001$ ), and a pairwise difference existed between NON Proctored and NON Un-Proctored ( $P < 0.001$ ) (Fig. 1B). None of the four other pairwise comparisons of interest demonstrated significant differences (all  $P > 0.650$ ).

For the Flanker Uncorrected score, there was a significant effect of participant group ( $P < 0.001$ ). For the five pairwise comparisons of interest: 1) NON Proctored differed significantly from NON Un-Proctored ( $P < 0.001$ ); 2) NON Proctored differed significantly from SCI Proctored ( $P = 0.007$ ); 3) NON Un-Proctored did not differ from SCI Un-Proctored ( $P = 0.225$ ); 4) NON Un-Proctored differed significantly from SCI Adjusted ( $P = 0.003$ ); and 5) SCI Un-Proctored did not differ from SCI Adjusted ( $P = 0.710$ ) (Fig. 1C). For the Flanker Fully-Corrected score, a significant effect of participant group existed ( $P < 0.001$ ) and a pairwise difference existed between



**Figure 1.** Mean (SD) DCCS and Flanker scores for participants with and without SCI. Dashed and dotted lines represent U.S. population normative mean and SD, respectively. \*P < 0.05.

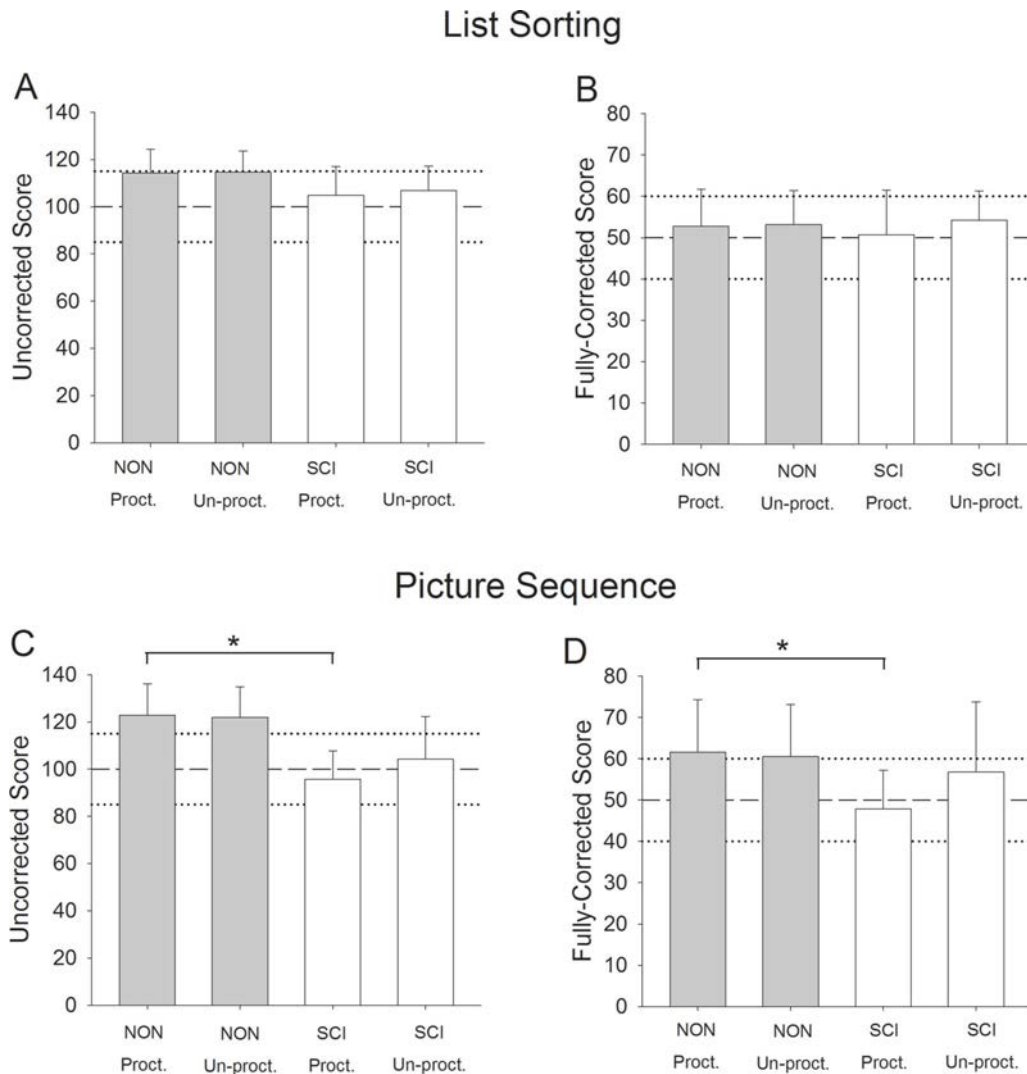
NON Proctored and NON Un-Proctored ( $P < 0.001$ ) (Fig. 1D). None of the four other pairwise comparisons of interest demonstrated significant differences (all  $P > 0.403$ ).

Data for untimed tests (List Sorting and Picture Sequence) appear in Fig. 2. For the List Sorting

Uncorrected score and Fully-Corrected score, there was no significant effect of participant group (both  $P > 0.072$ ). For the Picture Sequence Uncorrected score, there was a significant effect of participant group ( $P = 0.010$ ). Among the four pairwise comparisons of interest only NON Proctored differed significantly from SCI Proctored ( $P < 0.001$ ). No differences existed for the other three comparisons of interest (all  $P > 0.054$ ). Similarly for the Picture Sequence Fully-Corrected scale, there was a significant effect of group ( $P = 0.042$ ) and a significant pairwise difference between NON Proctored and SCI Proctored ( $P < 0.001$ ), but no differences for the remaining three pairwise comparisons of interest (all  $P > 0.42$ ).

**Table 2.** Motor-performance offset calculated from the NON cohort. Values are mean (SD).

	DCCS	Flanker
Reaction Time (seconds)	-0.57 (0.14)	-0.49 (0.13)
Uncorrected Score	21.09 (4.96)	19.60 (4.53)
Fully-Corrected Score	30.51 (9.12)	26.25 (9.53)



**Figure 2.** Mean (SD) List Sorting and Picture Sequence scores for participants with and without SCI. Dashed and dotted lines represent U.S. population normative mean and SD, respectively. \* $P < 0.05$ .

### Day and proctor effects

For DCCS and Flanker there was no significant Day x Proctor Interaction for the Uncorrected and Fully-Corrected scores, respectively (all  $P > 0.758$ ), and no differences between Day 1 and Day 2 scores (all  $P > 0.144$ ). There was no Proctor Effect for Uncorrected or Fully-Corrected motor score offset values (all  $P > 0.135$ ).

One proctor collected all SCI bouts; we examined the degree to which this proctor was representative of the other 9 proctors. The average DCCS motor score offset obtained for this individual (33.0) fell within the 95% confidence interval for the remaining 9 proctors (23.9–36.3), and no significant difference was present between this proctor and the remaining 9 proctors ( $P = 0.487$ ). Likewise, the average Flanker motor score offset obtained for this individual (25.3) fell within the

95% confidence interval for the remaining 9 proctors (20.2–32.6) and no significant difference was present between this proctor and the remaining 9 proctors ( $P = 0.696$ ).

Day 1 vs Day 2 differences existed for both List Sorting and Picture Sequence; data on between-day effect sizes are shown in Table 3.

### Discussion

The purpose of this study was to determine whether the motor demands of NIH Toolbox tests contribute to differences in cognitive function scores between participants with and without spinal cord injury (SCI). We hypothesized that Adjusted scores for participants with SCI would not differ from Non-SCI scores collected via the conventional Un-Proctored method. This hypothesis was supported for the Fully-Corrected scale

**Table 3. Between-day effect size for list sorting and picture sequence.**

	Bout	Scale	P	Effect size	Cohen's D
List Sorting	Proctored	Uncorrected	NS		
		Fully-Corrected	NS		
	Un-Proctored	Uncorrected	0.003	0.814	0.763 ("medium")
		Fully-Corrected	0.005	0.720	0.713 ("medium")
Picture Sequence	Proctored	Uncorrected	0.009	0.653	0.718 ("medium")
		Fully-Corrected	0.015	0.645	0.669 ("medium")
	Un-Proctored	Uncorrected	0.003	0.687	0.791 ("medium")
		Fully-Corrected	0.004	0.663	0.727 ("medium")

Note: NS, non-significant. Text-based interpretation of Cohen's D is provided in parentheses.

for both DCCS and Flanker. These data showed that executive function scores did not differ between participants with and without SCI when demographic factors and test motor demands were mitigated as sources of variation. However, evidence did emerge for post-SCI cognitive impairment in episodic memory, underscoring the multidimensionality of cognitive function after SCI.

Data from the present study support that DCCS and Flanker scores for individuals with SCI and arm/trunk impairment may be artificially reduced because of the secondary motor demands of the tests, aside from any potential cognitive impairment. When these tests were administered in a way that created equivalent motor demands (proctoring) and that corrected for demographic factors (Fully-Corrected scale), no differences in executive function were apparent between SCI and NON participants. Application of a motor-response offset yielded SCI values that did not differ statistically from Un-Proctored non-SCI values or from scores obtained from people with SCI who had minimal hand/trunk impairment (low paraplegia; SCI Un-Proctored sub-cohort). Particularly for SCI cohorts with a significant proportion of individuals with quadriplegia, such as in the present study, low SCI cognitive scores appear to be partially driven by motor demands of the NIH Toolbox tests.

The NIH Toolbox "Reasonable Accommodation Guidelines" acknowledges the difficulty of administering timed tests to respondents with motor impairment, but offers no solution beyond requiring researchers to document non-standard test administration circumstances.<sup>32</sup> Non-standard test administration was common (25%) in the NIH Toolbox validation study for individuals with neurologic impairment.<sup>33</sup> 89% of these non-standard administrations involved accessibility issues such as motor limitations. Additionally, 24.4% of enrolled participants with SCI in the validation study failed to complete the Toolbox cognition battery.<sup>17</sup> This number rose to 39.6% among enrollees with tetraplegia.<sup>13</sup> Investigators note that this occurred "... particularly on speed-based tests. This was mainly

because it was not possible to make accommodations or because participants were unable to complete the task even with recommended accommodations".<sup>13</sup> A "motor free" composite cognition score has recently been suggested for the NIH Toolbox battery.<sup>18</sup> However, this composite score includes no metrics of executive function and several sub-scales are not currently available in the NIH Toolbox iPad app.

In the validation study for Toolbox metrics in populations with neurologic compromise, trained test administrators spontaneously provided proctoring assistance to respondents in 8.1% of the total sample.<sup>33</sup> End-users of NIH Toolbox who do not receive rigorous standardization training may make similar adaptations when their research participants encounter motor-related difficulties. This is particularly likely because admonitions against proctoring, while present in the Toolbox Reasonable Accommodations Guidelines, are not integrated into the iPad test administration platform.<sup>33</sup> We believe that proctoring will occur frequently "in the wild" as Toolbox end-users strive to achieve full datasets and to help their participants avoid feelings of frustration and failure. The results of the present study provide a strategy by which proctored data may be used when studies compare participants with SCI to a non-SCI control cohort. This study represents a step toward "careful and ongoing scrutiny of test accommodations and empirical evaluation of the consequences that accommodations have on cognitive test scores".<sup>33</sup>

#### *Clinical utility and implications*

DCCS and Flanker motor offset scores were not different among the proctors. Because we counter-balanced our study design by equally distributing the proctored and un-proctored sessions across days, we demonstrated no day by proctor interaction and no day to day change on these timed tests. These findings do not suggest that people with equivalent levels of cognitive function were assessed by each proctor. However, given a particular baseline assessment of cognitive function, we show a consistent motor offset score across the limited sample tested in this

study. Importantly, the motor offset scores were grounded by each person's own processing speed that would occur prior to either moving themselves or articulating the answer to the proctor. Future studies are needed, across a wide range of demographics, to capture when delay in articulation by the participant is a larger part of the overall motor impairment. People with impaired upper extremity movement and impaired motor speech, like individuals with stroke, will likely not benefit from the motor offset score developed in this study. Conversely, people with impaired upper extremity movement, but without impaired motor speech, like people with SCI, will have utility for the offset score presented in this study.

It is noteworthy that we did not design this study to assess whether specific proctor demographic factors influenced the offset scores. Rather, we intentionally chose a young, healthy, educated group of proctors in an effort to minimize proctor variance. We are unable to generalize our proctor's performance to other proctors with a more diverse set of demographic factors (age, education, behavioral factors). We encourage future investigators to explore the factors that influence variance among proctors, and, a method to account for this variance to improve proctor generalizability.

In the present study, one proctor conducted all SCI testing. This could have introduced systematic bias to the Adjusted SCI scores if the motor offset obtained for that proctor's bouts differed from the motor offset obtained for the rest of the proctor pool. Offsets calculated for this proctor's bouts did not differ significantly from the other proctors, and the mean offset for this proctor fell within the 95% CI for the offsets obtained from the rest of the group. Offsets obtained by this proctor were representative of the entire pool of proctors; thus the motor offset calculated via the NON data could be confidently applied to the SCI scores. A final limitation of this study is that we do not know whether cognitive function changes according to SCI acuity; this potential change would be confounded by the normal maturational changes in cognition that occur with aging and we did not have the sample size and range of acuities needed to address this question.

While the participants with SCI in the present study did not demonstrate evidence of executive function impairment, numerous other studies indicate that this cognitive domain may indeed be at risk after SCI.<sup>3,34</sup> The SCI cohort enrolled to the present study were healthy, independent volunteers from the local community who may not have experienced the same psychosocial risk factors for cognitive impairment as the broader SCI population.<sup>3,35</sup> The majority of studies do indicate

that because of psychosocial factors, medical factors,<sup>4,6</sup> and/or neuro-inflammatory processes,<sup>10,36</sup> executive function is at risk after SCI. Importantly, one recent study found post-SCI executive function deficits via neuropsychological tests that did not contain motor demands.<sup>34</sup> Rather than contradicting this body of evidence, the findings of the present study suggest that previously-reported rates for post-SCI cognitive impairment should be revisited. The Toolbox SCI validation study reported cognitive impairment rates of 26% via DCCS and 25% via Flanker.<sup>17</sup> The results of the present study indicate that at least a portion of this relative risk may have been contributed by the motor demands of DCCS and Flanker. Deficits in Picture Sequence observed in the present study were in accord with the elevated risk for impaired episodic memory noted for the validation study.<sup>17</sup> Also consistent with that study, we observed no evidence for impaired working memory via the List Sorting test. These findings underscore the multi-dimensionality of cognitive function after SCI and the potential for impairment in some domains and not others.

## Conclusions

NIH Toolbox DCCS and Flanker scores for individuals with SCI and arm/trunk impairment may be artificially reduced because of the secondary motor demands of the tests. Proctoring and computation of a motor-response score offset enables comparisons to be made between individuals with SCI and a Non-SCI control cohort; however, the specific offset values reported in the present study reflect the characteristics of this particular study cohort and should not be considered to be universally-applicable to the broader SCI population. Moreover, further work is needed to determine whether offset-adjusted scores can be compared to standardized normative values. For cognitive domains that do not involve a processing speed component (eg. episodic memory), this study is in accord with previous studies that identify cognitive impairment after SCI.

## Disclaimer statements




**Contributors** None.

**Funding** This work was supported by National Center for Medical Rehabilitation Research [grant number R01 HD082109]; National Center for Medical Rehabilitation Research [grant number R01 HD084645]; National Center for Advancing Translational Sciences Research [grant number UL1TR002537].



**Conflicts of interest** Authors have no conflict of interests to declare.

## ORCID

Jinhyun Lee  <http://orcid.org/0000-0001-8692-4594>  
 Shauna Dudley-Javoroski  <http://orcid.org/0000-0002-0164-6212>  
 Richard K. Shields  <http://orcid.org/0000-0002-5838-9317>

## References

- Shields RK. 48th Mary McMillan lecture: turning over the hour-glass. *Phys Ther*. 2017;97(10):949–63.
- Qu H, Shewchuk RM, Chen Y, Deutsch A. Impact of medicare prospective payment system on acute rehabilitation outcomes of patients with spinal cord injury. *Arch Phys Med Rehabil*. 2011;92(3):346–51.
- Craig A, Guest R, Tran Y, Middleton J. Cognitive impairment and mood states after spinal cord injury. *J Neurotrauma*. 2017;34(6):1156–63.
- Sharma B, Bradbury C, Mikulis D, Green R. Missed diagnosis of traumatic brain injury in patients with traumatic spinal cord injury. *J Rehabil Med*. 2014;46(4):370–3.
- Bassi A, Bozzali M. Potential Interactions between the autonomic nervous system and higher level functions in neurological and neuropsychiatric conditions. *Front Neurol*. 2015;6:182.
- Jegade AB, Rosado-Rivera D, Bauman WA, Cardozo CP, Sano M, Moyer JM, et al. Cognitive performance in hypotensive persons with spinal cord injury. *Clin Auton Res*. 2010;20(1):3–9.
- Nicotra A, Critchley HD, Mathias CJ, Dolan RJ. Emotional and autonomic consequences of spinal cord injury explored using functional brain imaging. *Brain*. 2006;129(Pt 3):718–28.
- Lazzaro I, Tran Y, Wijesuriya N, Craig A. Central correlates of impaired information processing in people with spinal cord injury. *J Clin Neurophysiol*. 2013;30(1):59–65.
- Wu J, Stoica BA, Luo T, Sabirzhanov B, Zhao Z, Guanciale K, et al. Isolated spinal cord contusion in rats induces chronic brain neuroinflammation, neurodegeneration, and cognitive impairment. Involvement of cell cycle activation. *Cell Cycle*. 2014;13(15):2446–58.
- Wu J, Zhao Z, Kumar A, Lipinski MM, Loane DJ, Stoica BA, et al. Endoplasmic reticulum stress and disrupted neurogenesis in the brain are associated with cognitive impairment and depressive-like behavior after spinal cord injury. *J Neurotrauma*. 2016;33(21):1919–35.
- Faden AI, Wu J, Stoica BA, Loane DJ. Progressive inflammation-mediated neurodegeneration after traumatic brain or spinal cord injury. *Br J Pharmacol*. 2016;173(4):681–91.
- Jure I, Labombarda F. Spinal cord injury drives chronic brain changes. *Neural Regen Res*. 2017;12(7):1044–7.
- Cohen ML, Tulskey DS, Holdnack JA, Carlozzi NE, Wong A, Magasi S, et al. Cognition among community-dwelling individuals with spinal cord injury. *Rehabil Psychol*. 2017;62(4):425–34.
- Bloch A, Tamir D, Vakil E, Zeilig G. Specific deficit in implicit motor sequence learning following spinal cord injury. *PLoS One*. 2016;11(6):e0158396.
- Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ. Nih toolbox for assessment of neurological and behavioral function. *Neurology*. 2013;80(11 Suppl 3):S2–6.
- Weintraub S, Dikmen SS, Heaton RK, Tulskey DS, Zelazo PD, Bauer PJ, et al. Cognition assessment using the NIH Toolbox. *Neurology*. 2013;80(11 Suppl 3):S54–64.
- Carlozzi NE, Goodnight S, Casaletto KB, Goldsmith A, Heaton RK, Wong AW, et al. Validation of the NIH Toolbox in individuals with neurologic disorders. *Arch Clin Neuropsychol*. 2017;1–19.
- Carlozzi NE, Goodnight S, Umlauf A, Heaton RK, Heinemann AW, Schalet BD, et al. Motor-free composites from the national institutes of health Toolbox cognition battery (NIHTB-CB) for people with disabilities. *Rehabil Psychol*. 2017;62(4):464–73.
- Association ASI. International Standards for Neurological Classification of SCI. Atlanta (GA): American Spinal Injury Association. 2002 Revised 2011, Updated 2015.
- Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Conway KP, et al. Nih Toolbox cognition battery (CB): validation of executive function measures in adults. *J Int Neuropsychol Soc*. 2014;20(6):620–9.
- Tulskey DS, Carlozzi N, Chiaravalloti ND, Beaumont JL, Kisala PA, Mungas D, et al. Nih Toolbox cognition battery (NIHTB-CB): list sorting test to measure working memory. *J Int Neuropsychol Soc*. 2014;20(6):599–610.
- Dikmen SS, Bauer PJ, Weintraub S, Mungas D, Slotkin J, Beaumont JL, et al. Measuring episodic memory across the life-span: NIH Toolbox picture sequence memory test. *J Int Neuropsychol Soc*. 2014;20(6):611–9.
- Slotkin J, Kallen M, Griffith J, Magasi S, Salsman J, Nowinski CJ, et al. Nih Toolbox Technical Manual - Dimensional Change Card Sort test. Chicago (IL): National Institutes of Health and Northwestern University; 2012.
- Slotkin J, Kallen M, Griffith J, Magasi S, Salsman J, Nowinski CJ, et al. Nih Toolbox Technical Manual - Flanker Inhibitory Control and Attention Test. Chicago (IL): National Institutes of Health and Northwestern University; 2012.
- Baddeley A. Working memory. *Science*. 1992;255(5044):556–9.
- Slotkin J, Kallen M, Griffith J, Magasi S, Salsman J, Nowinski CJ, et al. Nih Toolbox Technical Manual - List Sorting Working Memory Test. Chicago (IL): National Institutes of Health and Northwestern University; 2012.
- Tulving E. Episodic memory: from mind to brain. *Annu Rev Psychol*. 2002;53:1–25.
- Slotkin J, Kallen M, Griffith J, Magasi S, Salsman J, Nowinski CJ, et al. Nih Toolbox Technical Manual - Picture Sequence Memory Test. Chicago (IL): National Institutes of Health and Northwestern University; 2012.
- Slotkin J, Nowinski CJ, Hays R, Beaumont J, Griffith J, Magasi S, et al. Nih Toolbox Scoring and Interpretation Guide. Chicago (IL): National Institutes of Health and Northwestern University; 2012.
- Casaletto KB, Umlauf A, Beaumont J, Gershon R, Slotkin J, Akshoomoff N, et al. Demographically Corrected normative Standards for the English version of the NIH Toolbox cognition battery. *J Int Neuropsychol Soc*. 2015;21(5):378–91.
- Cohen JD. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed Mahwah (NJ): Lawrence Erlbaum Associates; 1988.
- Nih Toolbox Reasonable Accommodation Guidelines. Chicago (IL): National Institutes of Health and Northwestern University, 2012 September.
- Magasi S, Harniss M, Tulskey DS, Cohen ML, Heaton RK, Heinemann AW. Test accommodations for individuals with neurological conditions completing the NIH Toolbox-cognition battery: An evaluation of frequency and appropriateness. *Rehabil Psychol*. 2017;62(4):455–63.
- Molina B, Segura A, Serrano JP, Alonso FJ, Molina L, Perez-Borrego YA, et al. Cognitive performance of people with traumatic spinal cord injury: a cross-sectional study comparing people with subacute and chronic injuries. *Spinal Cord*. 2018;56(8):796–805.
- Craig A, Nicholson Perry K, Guest R, Tran Y, Dezarnaulds A, Hales A, et al. Prospective study of the occurrence of psychological disorders and comorbidities after spinal cord injury. *Arch Phys Med Rehabil*. 2015;96(8):1426–34.
- Allison DJ, Josse AR, Gabriel DA, Klentrou P, Ditor DS. Targeting inflammation to influence cognitive function following spinal cord injury: a randomized clinical trial. *Spinal Cord*. 2017;55(1):26–32.