Test-Retest Reliability of Cognitive-Motor Interference Assessments in Walking With Various Task Complexities in Persons With Multiple Sclerosis

Peer-reviewed author version

VELDKAMP, Renee; Romberg, Anders; Hämäläinen, Paivi; GIFFROY, Xavier; MOUMDJIAN, Lousin; LEONE, Carmela; FEYS, Peter & BAERT, Ilse (2019) Test-Retest Reliability of Cognitive-Motor Interference Assessments in Walking With Various Task Complexities in Persons With Multiple Sclerosis. In: Neurorehabilitation and neural repair, 33 (8), p. 623-634.

DOI: 10.1177/1545968319856897
Handle: http://hdl.handle.net/1942/28857

Test-retest reliability of cognitive-motor interference assessments in walking with various task complexity in persons with Multiple Sclerosis.

R. Veldkamp[1] MSc, A. Romberg[2] PhD, P. Hämäläinen[2] PhD, X. Giffroy[3] PhD, L. Moumdjian[1,4] MSc, C. Leone[1] MSc, P. Feys[1] PhD and I. Baert[1] PhD

[1] Reval Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Belgium

[2] Masku Neurological Rehabilitation Centre, Finland

[3] Centre Hospitalier Universitaire de Liège, Belgium

[4] IPEM, Institute of Psychoacoustics and Electronic Music, Faculty of Arts and Philosophy, Gent University, Belgium.

Corresponding author (and first author)

Renee Veldkamp: Renee.Veldkamp@uhasselt.be, +31643016427

Count

- Number of words in the abstract: 240
- Number of words in the manuscript: 4856
- Number of references: 55
- Number of tables: 3 + 1 supplementary
- Number of figures: 4

Running title

Reliability of CMI assessments in MS

ABSTRACT

Background: Simultaneous execution of motor and cognitive tasks can result in worsened performance on one or both tasks, indicating cognitive motor interference (CMI). A growing amount of research on CMI in persons with Multiple Sclerosis (pwMS) is observed. However, psychometric properties of dual-task outcomes have been scarcely reported.

Objective: To investigate the between-day test-retest reliability of the motor and cognitive dual-task costs (DTC) during multiple CMI test conditions with various task complexity in pwMS and matched healthy controls (HC).

Methods: 34 pwMS (Expanded Disability Status Scale score 3.0±0.8) and 31 HCs were tested and retested on three single cognitive, four single motor and twelve cognitive-motor dual-tasks. Cognitive tasks included serial subtraction by seven, titrated digit span backwards and auditory vigilance. Motor tasks were walking: at self-selected speed, over obstacles, crisscross and while carrying a water-filled cup. Outcome measures were cognitive and motor DTC, calculated as percentage change of dual-task performance compared to single-task performance. Intraclass correlations (ICCs) and Spearman correlation coefficients were calculated as appropriate.

Results: For DTC_{motor} of gait speed, ICCs ranged from 0.45 to 0.81 and Spearman correlations from 0.74 to 0.82. For $DTC_{cognitive}$, ICCs ranged from -0.18 to 0.49 and Spearman correlations from -0.28 to 0.26. Reliability depended on the type of motor and cognitive task.

Conclusion: Reliability of the DTC_{motor} was, overall, good, while that of the $DTC_{cognitive}$ was poor. The ‘walking’ and ‘cup’ dual-task conditions were the most reliable regardless of the integrated cognitive task.

Keywords: Multiple Sclerosis, Reliability, Cognitive-motor interference, Walking, Gait, Dual task

Introduction

Dual tasking, such as walking while talking or browsing a smart phone while walking, is a common everyday act that can be defined as ‘the concurrent performance of two tasks that can be executed independently, measured separately and have distinct goals’ [1]. However, simultaneous performance of a motor (e.g. walking) and a cognitive (e.g. talking) task may be difficult and lead to worsening in performance on one or both tasks. Such deterioration in performance has been conceptualized as cognitive-motor interference (CMI).

CMI is usually investigated with a dual-task paradigm in which a motor task (e.g. walking, balance) and a cognitive task (e.g. subtracting, word list generation, etc.) are performed separately and concurrently. The dual task cost (DTC) quantifies this interference and is the percentage change of dual-task performance compared to single-task performance [2]. In healthy subjects and elderly, CMI assessment during walking is a common way to gauge the interaction between cognition and mobility in daily life [3]. Difficulties with dual tasking during walking were associated with a higher fall risk in elderly [4,5]. Also, in persons with Multiple Sclerosis (pwMS), fall risk may be related with DTC of walking (velocity, stride length and cadence) [6,7] and with DTC of standing balance [8]. These consequences highlight the need to assess and treat CMI.

Walking and cognitive impairments are common features in pwMS, with prevalence rates ranging between 41% and 75% [9] and 43% and 70% [10,11], respectively. Due to impaired sensory and motor system functions walking may become less automatic and require increased attention, while at the same time pwMS suffer from reduced attentional and executive function, possibly leading to worsening in dual task (DT) performance [11]. Recent reviews in MS reported diminished performance on motor performance (DTC_{motor}) and on cognitive performance ($DTC_{cognitive}$) under DT conditions ranging from ~6% to ~27% [2,12] and from ~6% to ~16% [12], respectively. It is unclear to which extent CMI in pwMS deviates from that of healthy controls (HC). Contradictory results were found in different reviews on whether pwMS showed greater decrements compared to HC in walking or balance [2,12,13] or in cognitive tasks performance [14-17] under DT conditions. Across studies however, diverse combinations of cognitive-motor dual tasks have been used, hampering direct comparisons of results. Moreover, cognitive tasks from various cognitive domains may result in different DTCs in persons with and without MS [13,18]. Additionally, it is scarcely investigated whether the DTC is dependent on the complexity of the motor task.

Despite the increasing interest in DTC as a new ecologic disease outcome measure, reported psychometric properties (reliability, validity and responsiveness) are still rare, halting the use of DTC as an experimental outcome measure [19]. As shown by two recent reviews, conducted in individuals with diverse neurological conditions and in the elderly, absolute measures of gait parameters (velocity, stride length, double support time, …) under DT conditions can be reliably measured [19,20]. In the cognitive domain, cognitive performances under DT conditions and cognitive DTCs have, conversely, shown lower reliability values than gait parameters did in stroke and Parkinsonian patients [19]. In pwMS the reliability of absolute gait parameters under DT conditions was investigated during normal walking [21] and walking over a narrow path way [22]. In daily life however, various complex motor tasks as walking over obstacles or carrying a cup with water are common acts. Therefore, there is a need to investigate the reliability of DT outcomes in more complex motor tasks as well. Most notably, although Prosperini et al. (2016) reported excellent test-retest reliability for outcomes of dual-task balance and cognition[23], reliability of relative CMI outcome measures as DTC_{motor} and $DTC_{cognitive}$ during walking is still unknown in pwMS [19].

It is important to study the psychometric properties of the motor and cognitive DTCs to ensure accurate measurements and to be able to allow for sound interpretations of (change) scores of CMI in pwMS for clinical purposes or for research. The aim of the present study was to examine the between-day test-retest reliability of CMI of both motor and cognitive performance during different DT conditions, with motor and cognitive tasks of various task complexity, in pwMS and age-gender matched HC. We hypothesized the reliability of the motor DTC to be good and expected lower reliability values for the cognitive DTC compared to the motor DTC [19].

Method

Participants

Between September 2016 and April 2017, 34 pwMS and 31 age and gender matched HC participated in the study in the Masku Neurological Rehabilitation Centre (Finland) and the Centre Hospitalier Universitaire de Liège (Belgium). The sample was a convenience sample. Prospective participants were recruited among the patients from the participating centres and via the Neuro Society in South-Western Finland. Inclusion criteria for pwMS and HC were age between 18 and 65 and minimal cognitive function as measured with the Mini Mental State Examination (MMSE) ≥ 26 [24]. Inclusion criteria for pwMS were: diagnosis of MS according to

McDonald criteria, Expanded Disability Status Scale (EDSS) ≥2 and ≤5 [25], no relapse within the last 30 days, no changes in immunomodulatory disease treatment and no corticoid-therapy within the last 50 days and presence of dual task interference (dual task screening list ≥1) [26]. Participants were excluded if there were other medical conditions interfering with mobility, other neurological diagnosis, MS-like syndromes such as neuromyelitis optica, or when participants were not able to understand instructions or had major problems with hearing or vision. The study size was based on Hobart et al. (2012) who reported that sample sizes of a minimum of 20 persons provided robust reliability estimates[27]. The study was approved by the ethical committees of the Centre Hospitalier Universitaire de Liège, Hasselt University and Hospital District of Southwest Finland and executed according to the Helsinki declaration. All participants received written information and signed informed consent.

Procedure

Testing was divided over three days. On day one, descriptive outcomes (demographics, cognition, mobility, quality of life measures and a dual task screening list) were assessed. On the second and third (test-retest) days the participants performed the experimental CMI assessments. A time window of three to five days between the testing days was chosen to balance between minimising practice effects from the first to the second test-day and minimising time in-between test-sessions to avoid possibility of changes in the patient's functioning, while fitting practical organisation of scheduling tests.

Instruction booklets with lots of details and internal agreements were provided and pilot trials with new equipment were conducted before the actual start of study.

Measures were assessed according to a standardized test protocol, including detailed test procedures, verbal instructions, and level of encouragement during testing, besides standardized electronic files to collect and transfer data. Data quality control was performed by the study coordinator.

Descriptive outcome measures

Data on age, gender, disease duration since diagnosis, type of MS and disability level (Expanded Disability Status Scale, EDSS) were collected. Cognition, mobility and quality of life measures are described below.

Cognitive function

The Brief Repeatable Battery of Neuropsychological Tests (BRB-NT) [28] was used to assess participants' cognitive function. The BRB-NT contains five tests in different cognitive domains: 1) the Selective Reminding Test (SRT) measuring verbal learning and delayed recall; 2) the 10/36 Spatial Recall Test (SPART) to assess visuospatial learning and delayed recall; 3) the Symbol Digital Modalities Test (SDMT) and 4) the Paced Auditory Serial Addition Test (PASAT), which are two tests of sustained attention, concentration and information processing speed and 5) the Word List Generation (WLG), or fluency task to measure semantic word retrieval.

Motor function

Motor function was assessed using the Timed 25 Foot-Walk (T25FW) [29] with maximal walking speed, the Timed-Up-and-Go (TUG) test [30], the Dynamic Gait Index (DGI) [31] and the 2-minute walking test (2MWT) [32]. The Multiple Sclerosis Walking Scale-12 (MSWS-12) [33] and Falls Efficacy Scale (FES-I) [34], were completed to measure perception of the limitations in walking ability due to MS and the concern about the possibility of falling, respectively.

Quality of life, perceived fatigue and daily life dual tasking

The Multiple Sclerosis Impact Scale-29 (MSIS-29) [35] and the Modified Fatigue Impact Scale (MFIS) [36] were used to record participants' perceived impact of MS on day-to-day life and of fatigue on daily functioning, respectively. Furthermore, the Dual Task Questionnaire of Evans et al (2009) (DTQ) [37] asked how often participants experienced problems with dual tasks (e.g. needing to stop an activity to talk) over the past two-weeks.

Experimental outcome measures

Three single cognitive-, four single motor- and twelve integrated cognitive-motor dual tasks were performed twice. All tests had a duration of one minute. The order in which the blocks of single cognitive, single motor or cognitive-motor dual tasks were performed was computerized randomized between participants, as well as the sequence of each separate task within one block. The sequence of the tasks performed was the same during the test- and retest-session for a participant.

Cognitive tasks

The cognitive tasks used were: the titrated digit span backwards [38], the serial seven subtraction test [39] and the auditory vigilance with alphabets task. These cognitive tasks were chosen based on pilot-studies, compatibility with walking and on the cognitive domains, as information processing speed, attention and working memory are among the most affected cognitive domains in pwMS[40]. Cognitive stimuli were delivered by auditory speech via a headset microphone with noise cancelling, while participant responses were noted by the assessor and audio recorded on a specifically developed tablet application. For all cognitive tasks an example was given by the researcher and participants practiced with a short example before the test to make sure the participants fully understood the task.

The digit span backwards loads working memory. Participants listened to a titrated string of digits (e.g. 5-3-1-8), at the presented rate of one number per second and repeated them in the reverse order. The individual sequence length was assessed for each participant on the first CMI test-day, before the first trial, to determine the subject's digit span. Participants started at a digit length of three numbers and length was increased by one digit. CMI of the participant was tested with the last sequence length at which three out of four trials were performed correctly. Performance on the digit span task was quantified as 'number of correct answers'.

The serial sevens subtraction test loads working memory and information processing speed. Participants had to count backwards from a given number (between 152 and 198). Performance of this task was determined as 'number of correct subtractions'. Answers were counted as correct each time there was a subtraction of seven from the previous number.

The auditory vigilance task is a test of sustained attention. Vigilance refers to a 'state of readiness to detect and respond to certain small changes occurring at random time intervals' [41]. Participants listened to letters presented at the rate of one letter per 2,5 seconds and said aloud 'yes' every time they heard one of the two target letters (e.g. 'L' or 'R') and were asked not to respond when another letter was heard. 24 letters were randomly presented including 10 target letters. Performance of this task was determined as ‘number of correct answers’, as the reaction time could not be calculated due to technical problems.

Motor tasks

Walking was performed on a 30-meter, free-of-obstacles, quiet walkway that was marked with 80cm start and 80cm turning lines. Four different motor tasks were performed: 1) walking at self-selected speed (‘walk’), 2) walking while carrying a cup filled with water (‘cup’), 3) walking while stepping over obstacles (10cm height and width) placed every 3m (‘obstacles’) and 4) walking crisscross from cone to cone every 2m with 80cm fixed width in between (‘crisscross’). For all motor tasks, the researcher demonstrated how to walk over the walkway and participants walked a part of it. All tasks were performed for one minute at self-selected, comfortable, speed. For safety reasons, the examiner always walked closely but behind the participant. The motor task were chosen based on the criterion that they are all common walking activities carried out in daily life, but differing in motor complexity and required attention or adaptation. Besides, the motor tasks were based on previous studies investigating reliability in persons with neurological conditions during various walking tasks [21,42]. Spatio-temporal gait parameters were recorded by three wearable APDM sensors, placed on each foot and in the lower lumbar region, with the respective Mobility Lab Software (Portland, USA). For the ‘walking’, ‘cup’ and ‘obstacles’ conditions these parameters were gait speed (m/s), cadence (steps/minute), double support time (%), stride length (m) and for the ‘crisscross’ conditions this was turn velocity (degrees/s).

****Insert figure 1 near here****

Dual cognitive-motor task performance

Participants performed the motor and cognitive tasks described above simultaneously in 12 different DT conditions and were instructed to perform both tests at their best level (in order to avoid any task prioritization). The same procedures as described above for the single cognitive and single motor task conditions were used. To examine DT performance, DTCs were calculated for the diverse parameters for each DT condition as follows:

$$\mathrm{DTC}_{\mathrm{cognitive}}\ (\%) = \frac{(single-task\ cognitive\ score) - (dual-task\ cognitive\ score)}{single-task\ cognitive\ score} * 100$$

$$\mathrm{DTC}_{\mathrm{motor}}\ (\%) = \frac{(single-task\ motor\ score)\ - (dual-task\ motor\ score)}{single-task\ motor\ score} * 100$$

Statistical analysis

All analysis were conducted with IBM SPSS Statistics 25 and performed for pwMS and HC separately. Outliers were analysed with the DTCs of gait speed (‘walk’, ‘obstacles’, ‘cup’), turning velocity (‘crisscross’) and number of correct answers. DTCs were defined as outlier when above or below three standard deviations from the mean on test or retest moment, and excluded from further analysis per DT condition. Normality of the data was assessed with the Shapiro-Wilk test.

Test-Retest reliability was examined for all 12 DT conditions for the DTCs of one cognitive parameter (number of correct answers) and five motor parameters (gait speed, cadence, double support time, stride length and turn velocity). To assess reliability, intraclass correlation coefficients for single measurements (ICCs) with a two-way random effect with absolute agreement ($\mathrm{ICC}_{2,1}$) and Spearman correlation coefficients were calculated, for normally and non-normally distributed data respectively. For ICC-values reliability was categorized according to Shrout and Fleiss' cut-offs of >0.75 as 'excellent', >0.6 as 'good', >0.4 as 'fair' or ≤0.4 as 'poor' [43]. Landis and Koch's cut-offs categorizing reliability as 'excellent' >0.8, 'good' >0.6, 'fair' >0.2 or 'poor' ≤0.2 were used for Spearman correlation coefficients [44]. Listwise deletion was used in case of missing data, meaning that a participant was excluded from data-analyses only for the DT-condition in which a DTC-value was missing.

Agreement between repeated measures was analyzed using Bland-Altman plots. The mean difference between the two test-moments was plotted against the mean of the two test-moments. 95% Limits of Agreement (LOA) were defined as 1.96 x SD above and below the mean difference. The plots were only given for pwMS for the motor DTCs of gait speed over the 9 diverse dual 'walk', 'cup' and 'obstacles' conditions and for the cognitive DTCs in the dual 'walk' conditions [45,46].

Means and standard deviations of absolute cognitive and motor performance during the single and dual 'walk' conditions and DTCs were provided for gait speed and number of correct answers as supplementary data. A paired t-test or the Wilcoxon signed rank test was executed to assess systematic differences between both test sessions, for normally and non-normally distributed data respectively. The standard error of measurement ($SEM = SD_{pooled} \times \sqrt{(1-ICC)}$) was calculated to provide an absolute value of reliability to be able to quantify the precision of individual scores on a test [47]. The minimal detectable change ($MDC = SEM \times 1.96 \times \sqrt{2}$) was calculated to be able to determine whether a change in score can be considered without measurement error [47].

Results

Descriptive characteristics

Descriptive data of the participants are presented in table 1. In total, 35 pwMS and 33 HC were recruited to participate in the study. One pwMS and two HC did not perform the retest due to sickness and lack of interest. Therefore, 34 pwMS and 31 HC completed the whole study. Age and gender distribution were equal in the two groups, with around two-thirds female participants. For pwMS, disease duration was on average 11.9±10.6 years, with most of them having a relapsing remitting disease form (82.4%). HC and pwMS did not perform different on the Selective Reminding, Spatial Recall, Word List Generation and PASAT-2sec tests. PwMS performed worse on all other tests of cognition and mobility and on the questionnaires for quality of life and dual tasking in daily life.

Table 2 shows the motor and cognitive DTCs on test and retest for all DT conditions for the parameters gait speed (‘walk’, ‘obstacles’, ‘cup’), turning velocity (‘crisscross’) and number of correct answers. The DTC_{motor} ranged from 8.7% to 15.8%, from 11.6% to 18.1% and from 3.7% to 12.1% in the digit span, subtraction and vigilance DT conditions, respectively. The $DTC_{cognitive}$ ranged from -0.8% to 18.2% on the digit span task, from -5.0% to 16.9% on the subtraction task and from -0.0% to 3.9% on the vigilance task.

****Insert table 1 and 2 near here****

Reliability of motor dual task costs

Table 3 and figure 2 show the test-retest reliability of the DTCs for all motor and cognitive outcome parameters in all DT conditions for both pwMS and HC (figure 2 visualises the reliability of the various DTCs). In total 0.90% of the 6630 DTCs are missing due to technical errors or a zero score on the cognitive single task. In pwMS, the reliability of the $\text{DTC}_{\text{motor}}$ for all gait-parameters in the ‘walk’ conditions was mostly excellent and in the ‘cup’ conditions mostly good. In the ‘obstacles’ conditions the reliability of the motor DTCs was mostly fair, and ranged from poor to good. In the ‘crisscross’ conditions the reliability of the $\text{DTC}_{\text{motor}}$ of turn velocity was mostly poor for pwMS.

For HC, the reliability of the $\text{DTC}_{\text{motor}}$ was good for all gait-parameters in the ‘walk’ conditions and mostly good to excellent in the ‘cup’ conditions. In the ‘obstacles’ conditions the reliability of the $\text{DTC}_{\text{motor}}$ of gait speed, cadence and double support time was mostly good to excellent, while the reliability of the $\text{DTC}_{\text{motor}}$ of stride length was fair. The reliability of the $\text{DTC}_{\text{motor}}$ in the ‘crisscross’ conditions was good to excellent for HC.

Bland-Altman plots for $\text{DTC}_{\text{motor}}$ of gait speed for pwMS are presented in figure 3. LOA for $\text{DTC}_{\text{motor}}$ gait speed were ±8.6 to ±11.5% in the ‘walk’, ±12.3 to ±15.8% in the ‘cup’ and ±12.9 to ±18.6% in the ‘obstacles’ conditions. Points were equally distributed around zero for most conditions. However, the ‘digit span obstacles’ and ‘subtraction obstacles’ conditions showed a general pattern of decrease in $\text{DTC}_{\text{motor}}$ with repeated measurements of -3.9% and -2.6%, respectively. Furthermore, the ‘vigilance obstacles’ and all dual ‘cup’ conditions, showed a trend of dependency of the difference scores on the magnitude of the $\text{DTC}_{\text{motor}}$. It seems that pwMS with a low mean $\text{DTC}_{\text{motor}}$ had a greater $\text{DTC}_{\text{motor}}$ on retest- than on test-moment, while in contrast, pwMS with a high mean $\text{DTC}_{\text{motor}}$ showed smaller $\text{DTC}_{\text{motor}}$ on retest- than on test-moment.

****Insert figure 2 near here***

****Insert table 3 near here****

***Insert figure 3 near here**

Reliability of cognitive dual task costs

Table 3 and figure 2 show the test-retest reliability of the cognitive DTCs in all DT conditions. For pwMS, reliability of the DTC$_{\text{cognitive}}$ was mostly poor, independent of the DT condition. For HC, the reliability of the DTC$_{\text{cognitive}}$ was fair in the ‘walk’ conditions, poor in the ‘obstacles’ conditions and poor to fair in the ‘cup’ and ‘crisscross’ conditions. Bland-Altman plots for DTC$_{\text{cognitive}}$ are presented in figure 4. LOA were ±80.3%, ±78.3% and ±9.6% in the digit span, subtraction and vigilance dual ‘walk’ conditions, respectively.

***Insert figure 4 near here**

Absolute motor and cognitive performance in ‘walk’ conditions and its reliability

Supplementary table 1 shows absolute single- and dual-task performance and reliability of the absolute measures during the ‘walk’ conditions. For pwMS, significant differences between test and retest in number of correct answers were found for performance on the digit span in the single and dual task condition and on subtraction in single task condition. For HC, significant differences between test and retest in number of correct answers were found for performance on the subtraction task in the single task condition and on the digit span, subtraction and vigilance tasks in the DT conditions. In all these cases scores were higher on retest- than on test-moment (supplementary data table 1).

Reliability of single task and dual task gait speed (absolute value) was excellent for both groups (MS: ICC = .87-.94; HC: ICC = .88-.93). The SEM for gait speed during single and dual task performance ranged from 0.04 to 0.07 m/s, resulting in a MDC ranging from 0.12 to 0.20 m/s for pwMS and HC.

For performance on the cognitive tasks (number of correct answers), reliability in the single task conditions was poor to fair for vigilance (MS: Spearman's *r* = .05, HC: Spearman's *r* = .55), fair to good for digit span (MS: ICC = .71, HC: ICC = .60) and excellent for subtraction (MS: Spearman's *r* = .82, HC: ICC = .79). In the dual 'walk' conditions, reliability for the digit span and subtraction tasks was good for pwMS (ICC = .63-.64) and excellent for HC (ICC = .77-.77). Reliability of the vigilance task under dual 'walk' conditions was poor (MS: Spearman's *r* = .31, HC: *r* = .09). The SEM for number of correct answers on the digit span task during single and dual task performance ranged from 1.29 to 1.63, resulting in a MDC ranging from 3.58 to 4.52 for pwMS and HC. For number of correct answers on the subtraction task during single and dual task performance ranged from 3.36 to 3.89, resulting in a MDC ranging from 9.32 to 10.79 for pwMS and HC.

Discussion

This study, in persons with MS, is the first to show that the reliability of the DTCs, especially $\mathrm{DTC}_{\mathrm{motor}}$, depends strongly on the type and complexity of the motor dual task and to some extent on the choice of gait parameter, regardless of the cognitive task. In contrast, the $\mathrm{DTC}_{\mathrm{cognitive}}$ showed poor test-retest reliability in pwMS and HC.

The 'walk' and 'cup', but not the 'obstacles' and 'turning' conditions showed to be reliable dual-task conditions to measure the $\mathrm{DTC}_{\mathrm{motor}}$ in pwMS. None of the cognitive tasks used in the DT conditions resulted in a systematic higher or lower reliability of the $\text{DTC}_{\text{motor}}$. For pwMS, reliability of the $\text{DTC}_{\text{motor}}$ seemed to decrease with increasing motor task complexity as reliability was the highest in the ‘walk’ and ‘cup’, and the lowest in the ‘obstacles’ and ‘crisscross’ DT conditions. This pattern was not clearly observed in HC. This pattern was also confirmed by inspecting B-A plots as the ‘walk’ conditions showed acceptable agreement, while a learning effect was present in the ‘obstacles’ conditions and agreement tended to depend on the magnitude of the $\text{DTC}_{\text{motor}}$ in the ‘cup’ conditions. The significantly higher motor impairment in pwMS compared to HC may have resulted in less confidence in DT conditions with larger motor complexity, leading to higher variation in motor performance and consequentially to a lower reliability. For example, turn parameters have been found to be important markers for self-perceived balance confidence of pwMS [48]. While in HC motor performance on the ‘crisscross’ DT conditions could be reliable measured, this was not true in pwMS.

Multiple spatiotemporal gait parameters were measured in the present study. For the calculation of a $\text{DTC}_{\text{motor}}$, the choice of spatiotemporal gait parameter seems to have some influence on its reliability. For both groups, reliability of the motor DTCs of gait speed, cadence, stride length and double support time was mostly good to excellent in the ‘walk’ and ‘cup’ conditions. Conversely, for the ‘obstacles’ conditions the reliability of the $\text{DTC}_{\text{motor}}$ varied according to gait parameter. For example, in the obstacles conditions, stride length has to be adapted to the encountered obstacles, this probably leads to more variability and a lower reliability of the DTC of stride length. Only one study explored the reliability of DTC in an obstacle course and showed a low reliability of the DTC of walking time in stroke patients [42].

To our knowledge, this is the first study evaluating the reliability of the $\text{DTC}_{\text{cognitive}}$ during walking in pwMS. The poor reliability of the $\text{DTC}_{\text{cognitive}}$ is in accordance with a previous study in people with stroke [42]. Factors such as fatigue, psychological state and attention may affect cognitive performance more than gait performance [19]. Learning effects were greater for cognitive tasks than for gait tasks, as gait is more habitual behaviour. Indeed, pwMS performed significantly better on the subtraction and digit span backwards tasks during retest assessment than during the first assessment. This is in accordance with Strouwen et al. (2016) who also found the digit span task to be subject to learning effects [43]. The lack of sufficient practice might have led to a more pronounced learning effect over the two sessions and a lower reliability for the $DTC_{cognitive}$ of number of correct answers.

The low test-retest reliability of the $DTC_{cognitive}$ may also relate to methodological factors. First, the number of correct answers was used in the present study as cognitive performance outcome. Among other possible parameters are accuracy, error rate, correct response rate, and reaction time. In previous studies, fair to moderate reliability has been found for absolute DT error rate in PD and MS [22,43], fair to excellent reliability for absolute DT correct response rate in chronic stroke [42] and excellent reliability for absolute DT reaction time in PD [43]. Including reaction time into cognitive performance measures seems to enhance reliability. Especially for the vigilance task reaction time could be of more value, as almost none of the participants made any or more than one error during the task. Second, the lower reliability of especially the cognitive DTCs may also be related to the instructions given to the participants. In the current study, the instruction was to perform both tasks at best level to prohibit task prioritization. However, previous studies in pwMS comprising both cognitive and motor DTCs showed different prioritization strategies when no instruction to prioritise one of the tasks was given [14,15,49]. Various task- and patient-related factors may affect prioritization-strategies [14,50-53], illustrating the importance of measuring both the motor and cognitive DTC [54]. The present sample showed motor impairments in walking, but no clear cognitive impairment except for processing speed.

For both motor and cognitive measures, absolute DT performance showed to be more reliable than DTC. Absolute measure of gait speed under DT conditions demonstrated excellent reliability in pwMS and HC, which is in accordance with two previous studies on DT reliability in pwMS showing excellent reliability for various absolute spatiotemporal gait parameters [21,22]. Also comparable to previous findings [22,42,43], good to excellent reliability was found for the absolute cognitive performance under DT conditions for the digit span and subtraction tasks. However, large SD and very broad ranges in the ICC 95% confidence intervals were found, indicating large variability in cognitive performance. When calculating the DTC, measurement errors inherent to the single and the dual task condition are both taken into account. This could partly explain the lower reliability of the DTCs. However, as Plummer and Eskes (2015) described, it is important to take both, absolute and relative measures, into account to be able to conclude improvement in overall dual-task performance [54]. For example, gait speed might improve in both single and dual task conditions, resulting in the same DTC despite improvements. Notwithstanding, DTCs are relevant to compare different populations and diverse DT paradigms. Moreover, as described previously, to be able to analyse whether someone has improved in dual tasking, both a motor and cognitive DTC are needed in order to determine whether a new strategy, i.e. gait-priority or cognitive-priority trade-off has occurred or whether indeed an overall improvement has taken place [54].

Limitations

There are some limitations to this study. First of all, the sample size of 31 pwMS might be insufficient to give generalized conclusions. The results are therefore confined to mildly motor and cognitive impaired pwMS. Second, one may comment on limited sample size. We are however confident in the observed results as our sample size matches previous studies on dual-task walking reliability in persons with MS [21,22] and recommendations on reliability in neurological patients made by Hobart et al. (2012)[27]. Importantly, current results are comparable to previous studies in neurological populations. The three cognitive tasks used in the current study were based on the cognitive domains involved, namely information processing speed, working memory and attention as these are frequently impaired in pwMS[40], and on previous studies. Tasks of executive functioning as the Stroop colour word task and verbal fluency tasks have however also shown to be promising concurrent cognitive tasks in pwMS [13,55] and showed excellent test-retest reliability in a dual balancing task[23].

Conclusion

The present study provides valuable information on CMI assessments for further studies. It was shown that motor DTCs and absolute DT performance of various spatiotemporal gait parameters can be reliably measured in pwMS with motor impairments, if the complexity of the dual motor task is not too difficult. ‘Walking’ and ‘cup’ conditions showed to be the most reliable DT conditions regardless of the integrated cognitive task. Cognitive performance during DT could not reliably be measured in pwMS and HC, but the cognitive DTC is important for a complete understanding of DT performance. Testing multiple cognitive-motor dual-task combinations reliably is viable in pwMS and gives a better understanding of the dual-task performance when including both absolute and relative measures.

Acknowledgements

The study was supported by RIMS by Novartis AG, MS Society Flanders, Swedisch Promobilia Foundation and SailingChallenge4MS.

Conflict of interest statement

The Authors declare that there is no conflict of interest.

Reference list

1. McIsaac TL, Lamberg EM, Muratori LM. Building a framework for a dual task taxonomy. *Biomed Res Int.* 2015;2015:591475.
2. Leone C, Patti F, Feys P. Measuring the cost of cognitive-motor dual tasking during walking in multiple sclerosis. *Mult Scler.* 2015;21(2):123-131.
3. Montero-Odasso M, Verghese J, Beauchet O, Hausdorff JM. Gait and cognition: a complementary approach to understanding brain function and the risk of falling. *J Am Geriatr Soc.* 2012;60(11):2127-2136.
4. Lundin-Olsson L, Nyberg L, Gustafson Y. "Stops walking when talking" as a predictor of falls in elderly people. *Lancet.* 1997;349(9052):617.
5. Beauchet O, Annweiler C, Dubost V, et al. Stops walking when talking: a predictor of falls in older adults? *Eur J Neurol.* 2009;16(7):786-795.
6. Wajda DA, Motl RW, Sosnoff JJ. Dual task cost of walking is related to fall risk in persons with multiple sclerosis. *J Neurol Sci.* 2013;335(1-2):160-163.
7. Etemadi Y. Dual task cost of cognition is related to fall risk in patients with multiple sclerosis: a prospective study. *Clin Rehabil.* 2017;31(2):278-284.
8. Wajda DA, Motl RW, Sosnoff JJ. Correlates of dual task cost of standing balance in individuals with multiple sclerosis. *Gait Posture.* 2014;40(3):352-356.
9. Larocca NG. Impact of walking impairment in multiple sclerosis: perspectives of patients and care partners. *Patient.* 2011;4(3):189-201.
10. Rao SM, Leo GJ, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology.* 1991;41(5):685-691.
11. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. *Lancet Neurol.* 2008;7(12):1139-1151.
12. Wajda DA, Sosnoff JJ. Cognitive-motor interference in multiple sclerosis: a systematic review of evidence, correlates, and consequences. *Biomed Res Int.* 2015;2015:720856.
13. Learmonth YC, Ensari I, Motl RW. Cognitive Motor Interference in Multiple Sclerosis: Insights From a Systematic Quantitative Review. *Arch Phys Med Rehabil.* 2017;98(6):1229-1240.

14. Hamilton F, Rochester L, Paul L, Rafferty D, O'Leary CP, Evans JJ. Walking and talking: an investigation of cognitive-motor dual tasking in multiple sclerosis. *Mult Scler.* 2009;15(10):1215-1227.
15. Allali G, Laidet M, Assal F, Armand S, Lalive PH. Walking while talking in patients with multiple sclerosis: the impact of specific cognitive loads. *Neurophysiol Clin.* 2014;44(1):87-93.
16. Negahban H, Mofateh R, Arastoo AA, et al. The effects of cognitive loading on balance control in patients with multiple sclerosis. *Gait Posture.* 2011;34(4):479-484.
17. Downer MB, Kirkland MC, Wallack EM, Ploughman M. Walking impairs cognitive performance among people with multiple sclerosis but not controls. *Hum Mov Sci.* 2016;49:124-131.
18. Al-Yahya E, Dawes H, Smith L, Dennis A, Howells K, Cockburn J. Cognitive motor interference while walking: a systematic review and meta-analysis. *Neurosci Biobehav Rev.* 2011;35(3):715-728.
19. Yang L, Lam FMH, Liao LR, Huang MZ, He CQ, Pang MYC. Psychometric properties of dual-task balance and walking assessments for individuals with neurological conditions: A systematic review. *Gait Posture.* 2017;52:110-123.
20. Yang L, Liao LR, Lam FM, He CQ, Pang MY. Psychometric properties of dual-task balance assessments for older adults: a systematic review. *Maturitas.* 2015;80(4):359-369.
21. Monticone M, Ambrosini E, Fiorentini R, et al. Reliability of spatial-temporal gait parameters during dual-task interference in people with multiple sclerosis. A cross-sectional study. *Gait Posture.* 2014;40(4):715-718.
22. Rosenblum U, Melzer I. Reliability and Concurrent Validity of the Narrow Path Walking Test in Persons With Multiple Sclerosis. *J Neurol Phys Ther.* 2017;41(1):43-51.
23. Prosperini L, Castelli L, De Luca F, Fabiano F, Ferrante I, De Giglio L. Task-dependent deterioration of balance underpinning cognitive-postural interference in MS. *Neurology.* 2016;87(11):1085-1092.
24. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-198.
25. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology.* 1983;33(11):1444-1452.
26. Strouwen C, Molenaar EA, Keus SH, et al. Protocol for a randomized comparison of integrated versus consecutive dual task practice in Parkinson's disease: the DUALITY trial. *BMC Neurol.* 2014;14:61.
27. Hobart JC, Cano SJ, Warner TT, Thompson AJ. What sample sizes for reliability and validity studies in neurology? *J Neurol.* 2012;259(12):2681-2694.
28. Boringa JB, Lazeron RH, Reuling IE, et al. The brief repeatable battery of neuropsychological tests: normative values allow application in multiple sclerosis clinical practice. *Mult Scler.* 2001;7(4):263-267.
29. Jill S. Fischer AJJ, Judith E. Kniker, Richard A. Rudick, Gary Cutter. Multiple Sclerosis Functional Composite Administration and Scoring Manual. 2001.
30. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39(2):142-148.
31. Shumway-Cook A WM. Motor Control Theory and Applications. *William and Wilkins Baltimore.* 1995:323-324.
32. Goldman MD, Marrie RA, Cohen JA. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Mult Scler.* 2008;14(3):383-390.
33. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology.* 2003;60(1):31-36.
34. Yardley L, Beyer N, Hauer K, Kempen G, Piot-Ziegler C, Todd C. Development and initial validation of the Falls Efficacy Scale-International (FES-I). *Age Ageing.* 2005;34(6):614-619.
35. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain.* 2001;124(Pt 5):962-973.

36. Kos D, Kerckhofs E, Ketelaer P, et al. Self-report assessment of fatigue in multiple sclerosis: a critical evaluation. *Occup Ther Health Care.* 2004;17(3-4):45-62.
37. Evans JJ, Greenfield E, Wilson BA, Bateman A. Walking and talking therapy: improving cognitive-motor dual-tasking in neurological illness. *J Int Neuropsychol Soc.* 2009;15(1):112-120.
38. Iverson GL, Tulsky DS. Detecting malingering on the WAIS-III. Unusual Digit Span performance patterns in the normal population and in clinical groups. *Arch Clin Neuropsychol.* 2003;18(1):1-9.
39. Smith A. The serial sevens subtraction test. *Arch Neurol.* 1967;17(1):78-80.
40. Trenova AG, Slavov GS, Manova MG, Aksentieva JB, Miteva LD, Stanilova SA. Cognitive Impairment in Multiple Sclerosis. *Folia Med (Plovdiv).* 2016;58(3):157-163.
41. Mackworth NH. Vigilance. *The Advancement of Science.* 1957;53:389-393.
42. Yang L, He C, Pang MY. Reliability and Validity of Dual-Task Mobility Assessments in People with Chronic Stroke. *PLoS One.* 2016;11(1):e0147833.
43. Strouwen C, Molenaar EA, Keus SH, Munks L, Bloem BR, Nieuwboer A. Test-Retest Reliability of Dual-Task Outcome Measures in People With Parkinson Disease. *Phys Ther.* 2016;96(8):1276-1286.
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
45. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59(10):1033-1039.
46. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb).* 2015;25(2):141-151.
47. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19(1):231-240.
48. Adusumilli G, Lancia S, Levasseur VA, et al. Turning is an important marker of balance confidence and walking limitation in persons with multiple sclerosis. *PLoS One.* 2018;13(6):e0198178.
49. Goverover Y, Sandroff BM, DeLuca J. Dual Task of Fine Motor Skill and Problem Solving in Individuals With Multiple Sclerosis: A Pilot Study. *Arch Phys Med Rehabil.* 2018;99(4):635-640.
50. Lemmens J, Ferdinand S, Vandenbroucke A, Ilsbroukx S, Kos D. Dual-task cost in people with multiple sclerosis: A case-control study. *Brit J Occup Ther.* 2018;81(7):384-392.
51. Sosnoff JJ, Socie MJ, Sandroff BM, et al. Mobility and cognitive correlates of dual task cost of walking in persons with multiple sclerosis. *Disabil Rehabil.* 2014;36(3):205-209.
52. Motl RW, Sosnoff JJ, Dlugonski D, Pilutti LA, Klaren R, Sandroff BM. Walking and cognition, but not symptoms, correlate with dual task cost of walking in multiple sclerosis. *Gait Posture.* 2014;39(3):870-874.
53. Yogev-Seligmann G, Hausdorff JM, Giladi N. Do we always prioritize balance when walking? Towards an integrated model of task prioritization. *Mov Disord.* 2012;27(6):765-770.
54. Plummer P, Eskes G. Measuring treatment effects on dual-task performance: a framework for research and clinical practice. *Front Hum Neurosci.* 2015;9:225.
55. Coghe G, Pilloni G, Zucca E, et al. Exploring cognitive motor interference in multiple sclerosis by the visual Stroop test. *Mult Scler Relat Disord.* 2018;22:8-11.