Reliability and Construct Validity of Three Self-report Questionnaires Assessing Dual-Task Difficulties in People With Multiple Sclerosis: An International Study

Peer-reviewed author version

ABASIYANIK, Zuhal; Pedulla, L; Kahraman, T; D'HOOGE, Mieke; Santoyo-Medina, C; Soler, B; Tacchino, A; VELDKAMP, Renee; Meza-Murillo, ER; Omar, A; Ciampi, E; KOS, Daphne; Ozakbas, S; Kalron, A & FEYS, Peter (2025) Reliability and Construct Validity of Three Self-report Questionnaires Assessing Dual-Task Difficulties in People With Multiple Sclerosis: An International Study. In: Archives of physical medicine and rehabilitation, 106 (2) , p. 230 -238.

DOI: 10.1016/j.apmr.2024.08.024
Handle: http://hdl.handle.net/1942/49379

Reliability and construct validity of three self-report questionnaires assessing dual-task difficulties in people with multiple sclerosis: an international study

Running title: Dual-task questionnaires in MS

Zuhal Abasıyanık[1,2,3] PhD, Ludovico Pedullà[4] PhD, Turhan Kahraman[3,5] PhD, Mieke D'Hooge[1,2,6] MSc, Carme Santoyo-Medina[7,8] MSc, Bernardita Soler[9] MD-PhD, Andrea Tacchino[4] PhD, Renee Veldkamp[1,2] PhD, Edwin-Roger Meza-Murillo[7] MD, Areen Omar[10,11] MSc, Ethel Ciampi[9] MD, Serkan Özakbaş[12] MD, Daphne Kos[6,13] PhD, Alon Kalron[10,11] PhD, Peter Feys[1,2] PhD

[1] REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Hasselt, Belgium
[2] UMSC, Hasselt-Pelt, Belgium
[3] Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Izmir Katip Celebi University, Izmir, Turkey
[4] Scientific Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy
[5] Department of Health Professions, Faculty of Health and Education, Manchester Metropolitan University, Manchester, UK
[6] National MS Center, Melsbroek, Belgium
[7] Department of Neurology-Neuroimmunology & Neurorehabilitation Unit, Multiple Sclerosis Centre of Catalonia (Cemcat), Vall d'Hebron University Hospital, Barcelona, Spain
[8] Department of Physiotherapy, Faculty of Medicine and Health Sciences, Universitat Internacional de Catalunya, Barcelona, Spain
[9] Hospital Doctor Sótero del Río, Pontificia Universidad Católica de Chile, Santiago, Chile
[10] Department of Physical Therapy, School of Health Professions, Faculty of Medicine, Tel-Aviv University, Tel-Aviv
[11] Multiple Sclerosis Center, Sheba Medical Center, Tel-Hashomer, Israel.
[12] Department of Neurology, Faculty of Medicine Izmir University of Economics, Izmir, Turkey
[13] KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium

Corresponding author: Zuhal Abasıyanık
REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Martelarenlaan 42, Agoralaan 1, 3500 Hasselt, Belgium; Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Izmir Katip Celebi University, Izmir, Turkey
e-mail: zuhal.abasiyanik@uhasselt.be, Tel.: +90 2323293535; Fax: +90 2323860888
ORCID ID: 0000-0003-3086-8102

Funding: This study was funded by a grant (RiGra 2022) provided by RIMS, European Network for best practice and research in MS Rehabilitation (www.eurims.org).

Financial disclosure and conflict of interest statement: Authors have nothing to disclose, neither financial interests nor non-financial interests.

Abstract

Objective: To evaluate the reliability and validity of the Dual-Tasking Questionnaire (DTQ), Dual-Task Screening-List (DTSL), and Dual-Task-Impact on Daily-life Activities Questionnaire (DIDA-Q).

Design: Multi-center, cross-sectional study

Setting: PwMS were recruited from seven MS centers across six countries (Belgium, Chile, Italy, Israel, Spain, and Turkey).

Participants: A total of 356 pwMS (mean age 47.5±11.5, EDSS 3.79±1.83) were enrolled.

Interventions: Not applicable

Main Outcome Measures: The reliability (internal consistency, test-retest, and measurement error) and construct validity (structural and convergent) were assessed.

Results: The DTQ, DTSL, and DIDA-Q demonstrated excellent test-retest reliability [Intraclass correlation coefficients (95% CI): 0.84 (0.80-0.87) to 0.90 (0.87-0.92)] and internal consistency (Cronbach's α: 0.86 to 0.96). As hypothesized, the three questionnaires showed a strong correlation with each other, moderate-to-strong correlations with other self-report questionnaires (perceived walking difficulties, fatigue, and fear of falling), and low-to-moderate correlations with cognitive information processing speed, manual dexterity, and dual-task walking performance (walking with word list generation task), showing convergent validity. The DIDA-Q exhibited systematically superior properties. These results were also verified in subsets from six different countries. In the structural validity analysis, all questionnaires displayed two main factors, allocated as ‘motor-driven’ and ‘cognitive-driven’ subscales.

Conclusions: The DTQ, DTSL, and DIDA-Q have good-to-excellent measurement properties, with the highest properties observed in DIDA-Q. The use of these self-reported questionnaires can be used in research and clinical practice to assess the impact of dual-task difficulties on the daily life of ambulatory pwMS.

Key Words: Patient-reported outcome measure, dual task, reliability, validity, gait, multiple sclerosis

List of abbreviations:

CMI: Cognitive-motor interference

COSMIN: Consensus-based standards for the selection of health status measurement instruments

DAQ: Divided attention questionnaire

DIDA-Q: Dual-task-impact on daily-life activities questionnaire

DTC: Dual-task cost

DTQ: Dual-tasking questionnaire

DTSL: Dual-task screening-list

EDSS: Expanded disability status scale

FES-I: 7-item falls efficacy scale international

ICC: Intraclass correlation coefficients

MFIS: Modified fatigue Impact Scale

MS: Multiple sclerosis

MSWS-12: 12-Item multiple sclerosis walking scale

NHPT: Nine-hole peg test

PROMs: Patient-reported outcome measures

PwMS: Persons with multiple sclerosis

QOC: Questionnaire by Cock et al.

RIMS: European network for best practice and research in MS rehabilitation network

SDC: Smallest detectable change

SDMT: Symbol digit modalities test

SEM: Standard error of measurement

T25FW: Timed 25-foot walk

Introduction

Daily life activities typically involve engaging in a secondary cognitive or motor task, known as a dual-task, such as walking while talking or planning.[1] In persons with MS (pwMS), the combination of motor and cognitive tasks often leads to decreased performance in one or both tasks, termed cognitive-motor interference (CMI). CMI has been attributed to MS-related cognitive and motor impairment and is associated with fatigue and disability levels in pwMS.[2–4]

Assessing dual-task walking is crucial as it reflects daily-life challenges such as employment.[5–7] Research has shown that CMI depends on task complexity[8,9] and is affected by MS-related symptoms such as fatigue, fatigability, pain, and spasticity.[4,10] Data on CMI in pwMS are primarily derived from controlled, laboratory-based studies focused on walking or balancing while concurrently engaging in cognitive tasks. Although these studies provide valuable insights into the basic mechanisms of CMI in pwMS, there is a need to understand how this phenomenon reflects daily life activities. Surprisingly, little is known about the perceived difficulties of dual-tasking in the daily life of pwMS. Thus, practical, accurate, valid, and reliable assessment tools to assess dual-task difficulties are needed in pwMS.

A recent systematic review has identified five patient-reported outcome measures (PROMs) assessing dual-task difficulties and their psychometric properties in the elderly and persons with Parkinson's disease, pwMS, traumatic and acquired brain injury.[11] These included the Divided Attention Questionnaire (DAQ), Dual-Task-Impact on Daily-life Activities Questionnaire (DIDA-Q), Questionnaire by Cock et al. (QOC), Dual-Tasking Questionnaire (DTQ), and the Dual-Task Screening-List (DTSL). Nevertheless, only few studies have investigated their psychometric properties, generally reporting on very-low-to-low reliability and validity, except for the DIDA-Q, which demonstrated moderate-to-high methodological quality, particularly in pwMS.[11] Moreover, research on the relationship between dual-task difficulties and cognition, physical functioning, patient-reported outcome measures and demographic outcomes is scarce, and the evidence is inconclusive.

This study examines the DTQ, DTSL, and DIDA-Q, which include commonly performed cognitive-motor dual-task activities in daily life. Our aim was to investigate their psychometric properties, including internal consistency, test-retest reliability, measurement error, and construct validity of in pwMS. Data were collected in six countries (Belgium, Chile, Italy, Israel, Spain, and Turkey) within the European Network for Best Practice and Research in MS Rehabilitation (RIMS) network, special interest group 'Mobility' (SIG Mobility).

Methods

This study was approved by local ethical committees from each participating center. All participants provided written informed consent. The protocol of the study was designed according to the Consensus-Based Standards for The Selection of Health Status Measurement Instruments (COSMIN) guideline.[12,13]

Translation and adaptation from English were performed to Spanish, Flemish, Italian, Hebrew, and Turkish, according to the proposed guideline. [14] The details of the translations and adaptation are provided in Appendices.

Data collection was conducted from January 2022 until March 2023. All performance-based tests and questionnaires were administered during a single session. To assess the test-retest reliability, the dual-task questionnaires were applied twice with an interval of 7-14 days.

Participants

PwMS were recruited from 7 MS centers from 6 countries (see Appendices). Each country recruited a minimum convenience sample of 50 pwMS, in line with the COSMIN guidelines.[15] Inclusion criteria: age ≥18 years, confirmed diagnosis of definite MS[16], relapse-free for 30 days at least, and Expanded Disability Status Scale (EDSS) score lower than 7.0. Exclusion criteria: a neurological disease other than MS, other medical conditions interfering with mobility, and major cognitive decline limiting the patient to understand the questionnaires. The cognitive status was based on a clinical observation by an experienced clinician specialized in MS.

Outcome measures

Socio-demographic and clinical information included sex, age, education, MS course[17], disease duration, and level of disability defined by the EDSS score.

PROMs assessing dual-tasking difficulties

- DTQ was developed for patients with acquired brain injury.[18] DTQ includes 10 items on cognitive-motor dual tasks. The subject indicates how often the dual-task difficulties occurred in the past few weeks based on a 5-point scale [0(never)-4(very often)]. Higher scores indicate greater perceived dual-task difficulties. A 'Not applicable' option is included to indicate tasks that are not commonly performed by the subject; this was not included in the

total score which was calculated by adding up the score of all responses ranging between 'Never' and 'Very often' divided by that number of items. The DTQ has been validated in Turkish pwMS.[19] The internal consistency is acceptable (Cronbach's alpha=0.78), test-retest reliability is excellent (intraclass correlation coefficients (ICC)>0.90)), demonstrating low-to-moderate correlations with clinical assessments.

- DTSL was developed for people with Parkinson's disease.[20] DTSL includes 13 items on walking and simultaneous cognitive and motor tasks. Responses for each item include ''yes, having difficulty (1 point)'', ''no (0 point)'', or ''not applicable''. No reliability and validity study has been found for this measurement tool. In the present analysis, the total score (%) was calculated by summing the scores of all responses and dividing by the number of items selected, except for 'Not applicable'.

- DIDA-Q was developed for pwMS.[21] The questionnaire includes 19 items and three subscales on cognitive, balance and mobility, and upper-limb tasks. Each item is rated using a Likert-type scale ranging from 0 (not difficult) to 4 (extremely difficult), with a total score ranging from 0 to 76. The DIDA-Q has demonstrated high internal consistency (Cronbach's alpha: 0.90, 0.93, and 0.90) and good test-retest reliability (ICC: 0.76, 0.89, and 0.81).

Dual-task performance measure

Dual-task performances were assessed through a single motor task of walking and a cognitive-motor task. The single motor walking task was performed on a 15 m walkway for 30 s and the participants were instructed to walk as fast as possible. The total distance in meters walked during the task was recorded.

For the cognitive-motor task the participants were requested to walk for 30s while performing a semantic word list generation (WLG) test, a common test for verbal fluency[3]. Namely, the participants were instructed to say aloud as many words as possible belonging to a certain category (fruits). The total distance (meter) and the number of correct answers were recorded during the trial. No instructions were provided to the participant regarding task prioritization (walking or the cognitive task).

Walking performances were assessed through the motor dual-task cost (DTC), a measurement based on the total distance during each task, calculated by the following formula:[2,22,23]

DTC_{motor} (%) =

Walking, cognition, and disease-related outcome measures

To assess validity, various common outcome measures were employed, including the *Timed 25-Foot Walk (T25FW)*[24], *Nine-Hole Peg Test (NHPT)*[25], *Symbol Digit Modalities Test (SDMT)*[26], *12-Item Multiple Sclerosis Walking Scale (MSWS-12)*[27], *Modified Fatigue Impact Scale (MFIS)* [28], and *7-item Falls Efficacy Scale International (FES-I).*[29] The written version of SDMT was used in Spain and Israel while other countries used oral version. Additionally, the number of falls in the last 6-months was recorded retrospectively. A fall was described as "an event where the participant unintentionally landed on the ground or a lower level."[30]

Statistical analysis

The data were analyzed using IBM SPSS (Version 25.0. Armonk, NY: IBM Corp.). Normality was checked using Kolmogorov-Smirnov and Shapiro-Wilk tests and visually investigating histograms and outliers. Statistical significance was set at $p<0.05$. Differences in clinical and demographic data between the countries were analyzed using one-way analysis of variance or Kruskal-Wallis test for continuous data, and chi-squared test for nominal data.

Reliability

Internal consistency for each PROM was determined using Cronbach's α coefficient and the average inter-item correlation, and a value of 0.8 or above is considered excellent; 0.70 to 0.8 is adequate.[31] Test-retest reliability was assessed via ICC with 95% CI. Categories were established as follows: >0.75 excellent, >0.6 good, >0.4 fair, or ≤0.4 poor.[32] Bland-Altman plots evaluated absolute test-retest reliability. Measurement error was calculated using Error of Measurement (SEM) and smallest detectable change (SDC). To explore disability level's impact on reliability, participants were categorized as "mild" (EDSS<4.0) or "moderately-severely" (EDSS≥4.0).

Validity

Structural validity was assessed via factor analysis, hypothesis testing, and known-groups validity.

Principal component exploratory factor analysis employed varimax rotation and Kaiser normalization. Hypotheses testing involved predefined correlations assessed by Spearman or Pearson coefficients. Four main priori hypotheses were formulated:

1. A strong correlation was anticipated between DTQ, DTSL, and DIDA-Q, as they aim to assess the same construct (i.e., dual-task-related difficulties in daily life).
2. A moderate correlation was expected between the lab-based measure of DTC and patients' perceptions reflected in DTQ, DTSL, and DIDA-Q.
3. Moderate to strong correlations were expected between DTQ, DTSL, DIDA-Q, and other self-reported measures (FES-I, MFIS, MSWS-12, number of falls).
4. Weak to moderate correlations were expected between DTQ, DTSL, DIDA-Q, and performance-based measures (T25FW, NHPT, SDMT).

Floor and ceiling effects were checked, considering metrics present if >15% scored the minimum or maximum possible in questionnaires.[33] Additional details on statistical analyses can be found in the Appendices.

Results

Descriptive data of the participants are presented in Table 1. In total 356 pwMS participated in the study and completed all the assessments except for retests of dual-task PROMs. The majority of the sample was female (66%), with a relatively higher level of education (14 years) and disease duration (11 years).

Reliability

Test-retest reliability analyses was based on 50 pwMS from Italy, 50 from Spain, 50 from Chile, 34 from Belgium, 43 from Israel, and 39 from Turkey (Table 2). In the total sample, all PROM's showed excellent internal consistency (Cronbach's α 0.86 to 0.96) and excellent test-retest reliability (ICC 0.84 to 0.90). The mildly disabled pwMS showed higher test-retest reliability and internal consistency and lower SEM and SDC values, in all questionnaires, compared to the moderate-severe disability subgroup. However, all values were good to excellent in both groups (Table 2).

Each language demonstrated excellent internal consistency, except for DTSL in Spain, which demonstrated adequate internal consistency (Cronbach's α=0.78). Test-retest reliability of the questionnaires in subsamples replicate the total sample, except for DTQ in Israel (ICC=0.73),

DTSL in Belgium (ICC=0.71), and DTSL in Spain (ICC=0.62) which showed good test-retest reliability. Bland-Altman plots of each questionnaire based on the total sample are presented in Figure 2; and provided separately in Supplementary Material 1.

All subscales of the questionnaires showed good to excellent test-retest reliability and internal consistency scores, except for one DTSL's subscale (cognitive-driven subscale) in Spain. This subscale also showed the highest SEM and SDC values in all countries, except for Chile (Table 3).

Validity

Table 4 displays correlation coefficients with related constructs for the total sample. The coefficients largely meet a priori hypotheses for construct validity. Strong correlations were observed among the three questionnaires. DIDA-Q and DTSL showed moderate-to-strong correlations with T25FW, NHPT, and SDMT, while DTQ exhibited weak correlations. All three questionnaires demonstrated strong correlations with fear of falling (FES-I), perceived mobility (MSWS-12), and perceived fatigue (MFIS). A moderate correlation was found between the number of falls and DTSL and DIDA-Q, while the correlation with DTQ was weak. Supplementary Material 2 provides correlation coefficients for each country. Overall, a priori hypotheses were largely confirmed. Additionally, DIDA-Q showed significantly stronger correlations than DTQ and DTSL with the T25FW, NHPT, SDMT, falls, FES-I, MSWS-12, and MFIS ($p<0.05$).

Factor analysis revealed Kaiser-Meyer-Olkin values of 0.957, 0.889, and 0.895 for DIDA-Q, DTSL, and DTQ, respectively. Bartlett's Test of Sphericity ($p<0.001$) indicated the suitability of respondent data for factor analysis. A two-factor structure showed the best fit for all scales, with loadings ranging from 0.625 to 0.798 (DIDA-Q), 0.567 to 0.792 (DTSL), and 0.545 to 0.921 (DTQ). This structure explained 66.90%, 49.91%, and 67.26% of the variance in DIDA-Q, DTSL, and DTQ, respectively. Item clustering into "cognitive-driven" and "motor-driven" subscales revealed 11 motor-driven and 8 cognitive-driven items for DIDA-Q, 3 and 7 for DTQ, and 7 and 6 for DTSL. Supplementary material includes the English version, item distribution, and sub-score and total score calculations for all questionnaires.

No floor and ceiling effects were observed for the DIDA-Q and DTQ. The DTSL showed a floor effect (15.7%), but no ceiling effect.

Discussion

This study explored the measurement properties of dual-task PROMs in five languages across six countries. Test-retest reliability and internal consistency of DIDA-Q, DTSL, and DTQ proved adequate for all languages. Furthermore, they demonstrated good construct validity, revealing two factors: 'motor-driven' and 'cognitive-driven' activities. No floor or ceiling effects were observed for DIDA-Q and DTQ. However, DTSL exhibited a floor effect, with 15.84% of pwMS reporting the lowest score, possibly attributable to its three-point scale limiting scoring.

Assessing CMI using performance-based measures and PROMs is crucial to understanding daily life challenges, as they mirror real-world challenges. Despite numerous performance-based CMI measures using various motor and cognitive tasks, the limited use of PROMs has hindered understanding perceived difficulties. This study could contribute to the increased use of PROMs by clarifying their measurement properties. All three questionnaires exhibited excellent test-retest reliability and internal consistency in a large sample of pwMS, with DIDA-Q having the highest level. The reliability of DIDA-Q has been investigated in the Italian MS population and showed excellent test-retest reliability and internal consistency in line with our findings.[21] However, measurement error was first described in this present study, and DIDA-Q showed the lowest SEM and SDC values among the three questionnaires (SDC is 7.04, which is 9.2% relative to the maximum score of 76). The measurement error of DTQ (SDC is 14.1% of the maximum score) was lower than DTSL (SDC is 19.7% of the maximum score). The reliability of DTQ has been explored in mildly disabled Turkish pwMS demonstrating excellent test-retest reliability and adequate internal consistency.[19] When examined separately, the Chilean sample (with the lowest disability level) displayed the highest ICC values, while samples with higher disability levels (Spain, Belgium, Israel, and Italy) showed slightly lower ICC values, though still within an acceptable limit. Despite using the same translations, differences in Spain and Chile may be linked to different disability levels of participants. However, the internal consistency of the three questionnaires was quite similar across countries, suggesting that age and disability did not affect the internal consistency.

Previous research on DIDA-Q conducted factor analysis, revealing three factors in the Italian MS population: balance-mobility, upper limb, and cognitive sub-scores.[21] In our sample, all questionnaires yielded two factors, labeled "cognitive-driven activities" and "motor-driven activities." We categorized items from DIDA-Q's upper-limb sub-score into the "motor-driven

subscale" since these activities occur during walking. This aligns with all questionnaires, where motor-driven activities primarily involve walking combined with other motor tasks, while cognitive-driven activities involve cognitive tasks with motor components. Each factor comprised three to 11 items, meeting the recommended minimum of three items per subscale.[34]

A priori hypotheses regarding the construct validity of the questionnaires were confirmed. Strong correlations among the three PROMs support the notion that they measure the same construct. As anticipated, moderate-to-strong correlations were found between dual-task questionnaires and self-reported measurements (fatigue, fear of falling, and walking disability). Our findings on fatigue align with a previous study that identified a strong positive correlation between DIDA-Q and MFIS.[21] Given the prevalence of perceived fatigue in pwMS, significantly impacting daily activities,[35] a strong correlation with dual-task difficulties is expected. The high correlations among PROMs may also be attributed to the interrelation of patient experiences. Aweak-to-moderate correlation emerged between dual-task questionnaires and cognitive processing speed, consistent with a previous study.[21] Considering the demand for divided attention and concentration in dual-task activities, these results support our hypothesis.[36]

Although the association between falls and dual-task walking and balance is not fully understood, a recent study demonstrated that perceived dual-task difficulty assessed by the DTQ could predict fall risk in pwMS.[37] We observed low to moderate correlations between self-reported dual-task impairment and fall history. The DIDA-Q demonstrated the strongest correlation with falls, fear of falling, walking, and manual dexterity, while the DTQ showed the weakest correlation with these variables. Notably, the DTQ had the strongest correlation with fatigue, potentially due to its inclusion of cognitive-cognitive activities, unlike DTQ and DTSL, which involve at least one motor activity in all their items.

Interestingly, questionnaires showed a weak relation with motor DTC during walking, likely because only one walking and cognitive task were used to calculate DTC, possibly not capturing the full spectrum of daily dual-tasking activities. Acknowledging inconclusive evidence on DTC correlates is crucial, limited by a small number of studies and methodological heterogeneity.[4] Additionally, the correlation values with dual-task performance increase when considering absolute dual-task walking, represented by the motor-driven subscale of DIDA-Q (highest value of 0.607).

While each PROM exhibits good psychometric properties, variations in instrument characteristics may influence preferences in clinical utility. DTSL, designed as a checklist with straightforward responses, contrasts with DTQ and DIDA-Q, which employ Likert-type scales. Despite all three questionnaires showing good reliability and validity, DIDA-Q outperforms, making it a recommended choice for trials and clinical settings to assess dual-tasking difficulty. However, considering time constraints, DTQ may be preferable, or DTSL for a quick scan of dual-task-related difficulties. Caution is advised when using DTSL in moderately-severely disabled Spanish pwMS due to its high measurement error.

Study Limitations

We acknowledge several limitations in this study. Some participants did not respond, resulting in data loss in certain countries (11.3% in total) for test-retest reliability analysis. Potential demographic differences may arise due to the study sites. Data collection was conducted during routine outpatient clinic visits in Turkey and Chile, while in Belgium, Italy, Israel, and Spain, it was conducted in rehabilitation centers. This may explain the lower age and EDSS scores observed in Turkey and Chile. However, the multicenter design also contributed to the generalizability of the results by recruiting different patient profiles. The discrepancy in using the written or oral version of SDMT in different countries might have affected our cognitive data. The 30-s walking time may seem short, but there is no standardization of dual-task walking assessments yet, and the duration of our study is relatively long compared to other assessments in the literature. Furthermore, psychological factors such as anxiety and mood can influence the perception of dual-task difficulties, however these factors were not evaluated in our study. Future research should investigate the responsiveness and minimum detectable change values of these questionnaires and further explore their psychometric properties.

Conclusions

This study supports the clinical utility of DIDA-Q, DTSL, and DTQ by showcasing their overall good reliability and validity in a large sample of pwMS across all ambulatory disability levels. Notably, DIDA-Q demonstrates superior measurement properties. Additionally, we refined previous factorial analysis, introducing two subscales for future use. The three questionnaires, their subscales, and calculation methods are accessible in Supplementary Material.

References

1. McIsaac TL, Lamberg EM, Muratori LM. Building a framework for a dual task taxonomy. *Biomed Res Int*. 2015
2. Leone C, Patti F, Feys P. Measuring the cost of cognitive-motor dual tasking during walking in multiple sclerosis. *Multiple Sclerosis Journal*. 2015;21:123–131.
3. Postigo-Alonso B, Galvao-Carmona A, Benítez I, et al. Cognitive-motor interference during gait in patients with Multiple Sclerosis: a mixed methods Systematic Review. *Neurosci Biobehav Rev*. 2018;94:126–148.
4. Rooney S, Ozkul C, Paul L. Correlates of dual-task performance in people with multiple sclerosis: A systematic review. *Gait Posture*. 2020;81:172–182.
5. Shema-Shiratzky S, Hillel I, Mirelman A, et al. A wearable sensor identifies alterations in community ambulation in multiple sclerosis: contributors to real-world gait quality and physical activity. *J Neurol*. 2020;267:1912–1921.
6. Hillel I, Gazit E, Nieuwboer A, et al. Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *European Review of Aging and Physical Activity*. 2019;16:1–12.
7. Kahraman T, Temiz H, Abasiyanik Z, Baba C, Ozakbas S. Dual task difficulties as a risk factor for unemployment in people with multiple sclerosis. *Brain Behav*. 2023;13.
8. Raats J, Lamers I, Baert I, Willekens B, Veldkamp R, Feys P. Cognitive-motor interference in persons with multiple sclerosis during five upper limb motor tasks with different complexity. *Multiple Sclerosis Journal*. 2019;25:1736–1745.
9. Veldkamp R, Baert I, Kalron A, et al. Associations between clinical characteristics and dual task performance in Multiple Sclerosis depend on the cognitive and motor dual tasks used. *Mult Scler Relat Disord*. 2021;56:103230.
10. Wolkorte R, Heersema DJ, Zijdewind I. Reduced dual-task performance in ms patients is further decreased by muscle fatigue. *Neurorehabil Neural Repair*. 2015;29:424–435.
11. Abasıyanık Z, Veldkamp R, Fostier A, et al. Patient-Reported Outcome Measures for Assessing Dual-Task Performance in Daily Life: A Review of Current Instruments, Use, and Measurement Properties. *Int J Environ Res Public Health*. 2022;19:1–13.
12. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63.
13. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*. 2018;27:1159–1170.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186–3191.
15. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research*. 2018;27:1171–1179.
16. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17:162–173.
17. Lublin FD, Reingold SC, Cohen JA, et al. MS clinical definitions. *Neurology*. 2014.
18. Evans JJ, Greenfield E, Wilson BA, et al. Walking and talking therapy: Improving cognitive-motor dual-tasking in neurological illness. *Journal of the International Neuropsychological Society*. 2009;15:112–120.

19. Eldemir S, Ozkul C, Eldemir K, Saygili F, Guclu-Gunduz A, Irkec C. The reliability and validity of the Turkish version of the dual-task questionnaire in patients with multiple sclerosis. *Mult Scler Relat Disord*. 2022;64:103942.
20. Strouwen C, Molenaar EALM, Keus SHJ, et al. Protocol for a randomized comparison of integrated versus consecutive dual task practice in Parkinson's disease: The DUALITY trial. *BMC Neurol*. 2014;14:1–12.
21. Pedullà L, Tacchino A, Podda J, et al. The patients' perspective on the perceived difficulties of dual-tasking: development and validation of the Dual-task Impact on Daily-living Activities Questionnaire (DIDA-Q). *Mult Scler Relat Disord*. 2020;46.
22. 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. *Multiple Sclerosis*. 2009;15:1215–1227.
23. Baddeley A, Sala S Della, Gray C, Papagno C, Spinnler H. Testing central executive functioning with a pencil-and-paper test. In: *Methodology of Frontal and Executive Function.*; 2004.
24. Fischer JS, Rudick RA, Cutter GR, Reingold SC. The multiple sclerosis functional composite measure (MSFC): An integrated approach to MS clinical outcome assessment. In: *Multiple Sclerosis*.Vol 5.; 1999:244–250.
25. Feys P, Lamers I, Francis G, et al. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Multiple Sclerosis*. 2017;23:711–720.
26. Benedict RHB, Deluca J, Phillips G, LaRocca N, Hudson LD, Rudick R. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple Sclerosis*. 2017;23:721–733.
27. 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;60:31–36.
28. Larson RD. Psychometric properties of the modified fatigue impact scale. *Int J MS Care*. 2013;15:15–20.
29. Van Vliet R, Hoang P, Lord S, Gandevia S, Delbaere K. Falls efficacy scale-international: A cross-sectional validation in people with multiple sclerosis. *Arch Phys Med Rehabil*. 2013;94:883–889.
30. Lamb SE, Jørstad-Stein EC, Hauer K, Becker C. Development of a common outcome data set for fall injury prevention trials: The Prevention of Falls Network Europe consensus. *J Am Geriatr Soc*. 2005;53.
31. Price PC, Jhangiani RS, Chiang I-CA, Leighton DC, Cuttler C. *Research Methods in Psychology*. 3rd American Edition. 2017.
32. Cicchetti D V. The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol*. 2001;23.
33. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60.
34. Comrey AL, Lee HB. A first course in factor analysis, 2nd ed. . *APA PsycNet*. 1992.
35. Vucic S, Burke D, Kiernan MC. Fatigue in multiple sclerosis: Mechanisms and management. *Clinical Neurophysiology*. 2010;121:809–817.
36. Bayot M, Dujardin K, Tard C, et al. The interaction between cognition and motor control: A theoretical framework for dual-task interference effects on posture, gait initiation, gait and turning. *Neurophysiologie Clinique*. 2018;48:361–375.

37. VanNostrand M, Sogoloff B, Giroux C, Bergmans L, Kasser SL. Predicting falls in adults with multiple sclerosis using patient-reported measures: Are perceptions of dual-tasking missing? *Mult Scler Relat Disord*. 2022;68:104115.

Figure legends

Figure 1. Overview of the PROMs' translation procedure

Figure 2. Bland-Altman plots with limits of agreement of three dual-task questionnaires