An ecological measure to screen executive functioning in MS: the Picture Interpretation Test (PIT) 360°

Olivia Realdon1, Silvia Serino2,3, Federica Savazzi4, Federica Rossetto4, Pietro Cipresso2,3, Thomas D. Parsons5, Giacomo Cappellini6, Fabrizia Mantovani1, Laura Mendozzi7, Raffaello Nemni8,9, Giuseppe Riva2,3 & Francesca Baglio4

Executive functions are crucial for performance of everyday activities. In Multiple Sclerosis (MS), executive dysfunctions can be apparent from the early onset of the disease. Technology-based time-efficient and resource-saving tools for early evaluation of executive functions using an ecological approach are needed to assess functional performance in real-life. The aim was to compare the efficiency of the Picture Interpretation Test 360° (PIT 360°) with traditional measures on executive dysfunction in Persons with Multiple Sclerosis (PwMS) and Healthy Controls (HC). Participants were 31 patients with Relapsing-Remitting MS (mean age = 44.32 ± 13.149; mean Expanded Disability Status Scale = 2) and 39 HC (mean age = 39.538 ± 15.728). All were tested with standard neuropsychological tests of executive functions, PIT 360°, and measures of user experience. While standard neuropsychological tests failed to differentiate between PwMS and HC group, the PIT 360° was successful in detecting executive dysfunction in PwMS. All participants reported the PIT 360° to be an engaging tool and endorsed positive reactions to their experience. Overall, the PIT 360° is a quick, sensitive, and ecological tool that captures real-world executive dysfunction in PwMS. This engaging measure is sensitive for the detection of executive deficits since the early phases of the disease.

Cognitive impairment in Multiple Sclerosis (MS) includes, among other deficits, executive dysfunction, multitasking difficulties, verbal fluency declines, and visuo–spatial deficits1–3. Cognitive impairment has been found in all disease subtypes4, including one-third of patients with early Relapsing–Remitting (RR) MS. Early onset of difficulties in simultaneous management of everyday activities is often reported in Persons with Multiple Sclerosis (PwMS)5. Altogether, these impairments have a disruptive impact on quality of life and the ability of PwMS to actively adapt to the changing demands of the physical and social environment2,4. Although conventional neuropsychological tests exist for assessing cognitive dysfunction in PwMS, they tend to be limited in their capacity for capturing the sorts of executive functioning deficits that are critical for functioning in real-world contexts6.

Evaluating functional performance across a range of real-life situations is the core of the function-led approach to the assessment of executive functions9. As highlighted by Chan and colleagues10, this approach, rather than fractionating the executive dysfunction, aims to incorporate the complexity of real-life challenges into tasks able to tap a number of executive domains simultaneously. Using this approach, executive functions can be captured in
Table 1. Sociodemographics, neuropsychological assessment and PIT 360° scores for PwMS and HC groups.

| Age (years) [mean ± SD] | PwMS group [N = 31] | HC group [N = 39] | Group comparison* p_value |
|-------------------------|---------------------|-------------------|--------------------------|
| 44.323 ± 13.149         | 39.538 ± 15.728     | 0.170             |
| Education (years) [mean ± SD] | 15.161 ± 2.922 | 16.385 ± 2.551 | 0.066                    |
| Gender (males:females)   | 15:16               | 15:24             | 0.405                    |
| EDSS [mean ± SD]         | 2.350 ± 1.750       | —                 | —                        |
| Disease Duration (years) [mean ± SD] | 12.387 ± 9.583 | —                 | —                        |
| MoCA [mean ± SD]         | 24.902 ± 1.971      | 26.145 ± 1.497    | 0.055                    |
| TMT-A [mean ± SD]        | 32.871 ± 12.099     | 36.154 ± 13.698   | 0.298                    |
| TMT-B [mean ± SD]        | 98.032 ± 36.892     | 100.128 ± 43.808  | 0.832                    |
| TMT-BA [mean ± SD]       | 64.258 ± 27.922     | 63.718 ± 33.114   | 0.942                    |
| Verbal Fluency (FAS form) [mean ± SD] | 35.229 ± 10.479 | 38.259 ± 10.361 | 0.231                    |
| PIT 360°, time [mean ± SD] | 65.249 ± 55.307     | 33.048 ± 19.565   | 0.024                    |
| PIT 360°, number [mean ± SD] | 8.355 ± 7.153       | 3.872 ± 3.299     | 0.063                    |

Table 1 exhibits results from the comparison between groups (PwMS vs Healthy Control, HC) on baseline characteristics and neuropsychological assessment scores. No significant difference was detected between the two groups for gender [χ² = 0.695; df = 1, p = 0.405], age [t(67.804) = 1.386, p = 0.170], or education [t(68) = −1.869, p = 0.066]. Findings obtained from Independent Student’s t-tests indicated no significant differences between PwMS and HC group with respect to the three sub-tests of TMT [TMT-A, t(68) = −1.048; TMT-B, t(68) = −0.213; TMT-BA, t(68) = 0.073] or Verbal Fluency [t(68) = −1.209]. When compared with HC, PwMS obtained a significantly lower MoCA score [t(54,730) = −2.906].

Performance on PIT 360°. Results in both indices obtained from the PIT 360° (see Table 1) revealed greater performance deficits for PwMS group when compared to the HC group. The PwMS group took more time (mean = 65.249; SD = 55.307) when compared to the HC group (mean = 33.048; SD = 19.565) for interpreting the PIT 360° scene [F(1,66) = 5.329; p = 0.024; Partial η² = 0.075]. Moreover, PwMS reported a higher number of scene elements in comparison to the HC group [MS, mean = 8.355; SD = 7.153; HC, mean = 3.872, SD = 3.299; F(1,34) = 4.449; p = 0.039; Partial η² = 0.063].

User experience assessment. Findings from comparison of the number of self-reported felt emotions and their intensities (in each of the four Geneva Emotion Wheel - GEW - quadrants) between the two groups
PwMS (vs. HC), revealed a significant difference in the number of emotions with negative valence and high coping potential. In this specific quadrant, in comparison to PwMS group, the HC group reported a significantly higher number of self-reported emotions.

No significant differences were observed between groups with respect to the overall number of self-reported felt emotions and their intensities (Table 2).

However, findings obtained from the Friedman Test revealed a significant difference among the four quadrants of GEW in terms of the number of self-reported emotions \(\chi^2(3) = 155.285; p < 0.001\).

Wilcoxon tests (with Bonferroni adjustment) indicated that all participants experienced a higher number of emotions with positive valence and high coping potential (Table 3).

Classification of PwMS or HC. Table 4 and 5 show the classification results for discriminating between the HC and the PwMS groups. Naïve Bayes and Support Vector Machine algorithms emerged as the best algorithms for classifying HC and PwMS in their respective groups. Using the scores from conventional executive functions tests as input, the machine learning algorithms showed a classification accuracy between 52.9% and 65.7%. In contrast, the indices from PIT 360° achieved a higher classification accuracy, ranging from 65.7% to 72.9%.
represents the proportion of true positives among the positive instances in our data. In our case, the proportion of a condition was identified correctly; Recall represents the proportion of true positives among all the examples that were classified correctly; F1 represents the weighted harmonic average of the precision and recall (defined below); Precision represents a proportion of true positives among all the instances classified as positive. Table 5. Stratified 10-fold Cross validation for the neuropsychological assessment battery. *AUC (Area under the ROC curve) is the area under the classic receiver-operating curve. CA (Classification accuracy) represents the proportion of the examples that were classified correctly; F1 represents the weighted harmonic average of the precision and recall (defined below); Precision represents a proportion of true positives among all the instances classified as positive. In our case, the proportion of a condition was identified correctly; Recall represents the proportion of true positives among the positive instances in our data.

| Method                      | AUC   | CA    | F1    | Precision | Recall |
|-----------------------------|-------|-------|-------|-----------|--------|
| Logistic Regression         | 0.494 | 0.529 | 0.653 | 0.554     | 0.795  |
| Random Forest               | 0.627 | 0.643 | 0.706 | 0.652     | 0.769  |
| Support Vector Machine (SVM)| 0.636 | 0.657 | 0.727 | 0.653     | 0.821  |
| Naive Bayes                 | 0.617 | 0.629 | 0.683 | 0.651     | 0.718  |

Table 4. Stratified 10-fold Cross validation for the indexes of PIT360°. *AUC (Area under the ROC curve) is the area under the classic receiver-operating curve. CA (Classification accuracy) represents the proportion of the examples that were classified correctly; F1 represents the weighted harmonic average of the precision and recall (defined below); Precision represents a proportion of true positives among all the instances classified as positive. In our case, the proportion of a condition was identified correctly; Recall represents the proportion of true positives among the positive instances in our data.

| Method                      | AUC   | CA    | F1    | Precision | Recall |
|-----------------------------|-------|-------|-------|-----------|--------|
| Logistic Regression         | 0.639 | 0.657 | 0.721 | 0.660     | 0.795  |
| Random Forest               | 0.707 | 0.729 | 0.787 | 0.700     | 0.897  |
| Support Vector Machine (SVM)| 0.668 | 0.700 | 0.779 | 0.661     | 0.949  |
| Naive Bayes                 | 0.678 | 0.700 | 0.764 | 0.680     | 0.872  |

Figure 1 shows the confusion matrix of all classifiers used for classifying individuals into PwMS Group and HC Group. Results revealed that indices from the PIT 360° had a higher capability for correctly classifying PwMS in their group.

**Discussion**

We aimed to evaluate the efficacy of a 360° version of the PIT for detecting executive dysfunction in PwMS through a function-led approach that combined experimental control with a real-world engaging background. In line with research findings on MS that reveal cognitive impairments that can be characterized as executive dysfunction, the PwMS performed significantly worse on the PIT 360° than did persons in the HC group.

While the mean global cognitive level of PwMS in our study was lower than that of HC, it was still in a non-pathological range. This suggests initial subclinical global dysfunction in PwMS. This initial dysfunction detected with a renowned test sensitive to MS-related cognitive impairment is in line with the prior work detecting cognitive impairment in PwMS. It is important to note that although verbal fluency is a sensitive tool for assessing executive impairment in PwMS and is part of the minimal assessment of cognitive function in MS (MACFIMS battery), the assessment with this test and the TMT failed in showing differences in executive functions between groups.

Different from standard neuropsychological tests used, the PIT 360° differentiated successfully between the pathological and the control conditions both in terms of time to give an answer and in number of elements in the scene. This result showed that PIT 360° is an ecological tool that is highly sensitive to MS pathology—even in its initial phases (EDSS, range 1–3). These findings were also confirmed by the higher accuracy in the Random Forest classification of participants to the clinical or non-clinical conditions (when using indices from PIT 360°) with respect to those from neuropsychological assessment. These robust findings demonstrated the efficacy of PIT 360° for detecting impairment of executive functions at an early clinical stage of MS. Moreover, they suggest that this ecological tool can be used for prompt diagnosis and early enrollment of PwMS in targeted rehabilitation.

The importance of an early management of cognitive impairment in MS is highlighted by the fact that it can precede the onset of physical disability and slow cognitive decline.

Although the findings of the present study are promising, in the comparison to standard neuropsychological assessment, PIT 360° is only a very sensitive screening tool not covering the need for a full and analytical examination of executive functions. In addition, it is a technology-based test implying the use of a VR headset with potential side-effects (e.g. nausea) in some patients.

Considering findings related to users’ experiences, we found that PIT 360° was considered to be an engaging tool both by the HC and the PwMS groups. Firstly, all participants reported a good sense of presence in the 360° scene (SUS Questionnaire) showing that they actively experienced the task in a context perceived as a real-life place. Both groups rated the challenge of the PIT 360° task as feasible, in the sense that it was considered balanced with respect to their skills (FSS scale). Furthermore, PwMS and HC positively assessed participant appreciation for and interest in performing PIT 360° task (IMI scale).

All participants endorsed positive reactions to the task, showing that their experience of the PIT 360° was highly pleasant and under control. This was apparent in the higher number of self-reported emotions in the first quadrant of the GEW, which includes interest, joy, happiness, satisfaction, elation and pride. Interestingly, HC...
vs. PwMS reported a higher number of self-reported emotions with high coping potential and negative valence. This finding can be related to the higher level of challenges perceived by the PwMS group when faced with the proposed activity in comparison to HC (FSS scale). A possible interpretation of these results is that PwMS exerted greater attentional effort when attempting to complete the task which, most probably, was higher than that of HC.

Interestingly, the efficiency of PIT 360° in detecting executive dysfunctioning was observed also in PD but was lower compared to SM. This is not surprising because executive function disorders are defined in functional terms and not as a topographic syndrome and are assessed by functional-led approach in PIT 360°. The difference in the accuracy of classification in the two clinical conditions with respect to HC may be due to several reasons. The aging can represent a factor: people with PD were older than PwMS for the natural history of the disease. Moreover, it is well known that aging is associated with decline in executive function and in the two studies were included different HC groups due to the age-related demographics of the two neurological conditions. Then, it is reasonable to expect the different sensitivity in classification accuracy in a middle-aged vs. older adults sample. Furthermore, the overall brain profile vary in the two conditions involving fronto-subcortical degeneration in PD and white matter frontal pathway disconnection in MS. Therefore, the degree of severity of the brain damage can impact executive functioning differently in the two diseases.

Finally, PIT 360° offers a promise for answering the need for time-efficient and resource-saving tools that can screen PwMS for executive deficits. This reduces patient stress at the first evaluation and orients clinicians to perform subsequent clinical investigations using longer neuropsychological assessment batteries and the prompt inclusion in targeted rehabilitation programs.

Future studies should examine PIT 360° efficacy in detecting executive dysfunctioning with a larger cohort and with other clinical populations. Moreover, it will be important that the PIT 360° be investigated using neuroimaging to establish neural correlates. Additionally, it will be of major importance to proceed with the validation of PIT 360° parallel forms to make possible a short-term re-evaluation of executive functions.

In conclusion, the PIT 360° is a quick and ecological measure that demonstrated effective and sensitive screening of real-world deficits related to executive functioning in the early stages of MS. These findings support, within Parsons’ theoretical framework for the assessment of executive functions, the methodological note advanced by Sumowski and colleagues on the need of advancing effective, evidence-based, clinically feasible understanding and measurement of dysexecutive functioning.

**Materials and Methods**

**Participants.** Seventy participants took part in the study: thirty-one PwMS (51.6% female; mean age = 44.32 ± 13.149; mean years of education = 15.16 ± 2.922; PwMS group) and thirty-nine healthy controls (61.5% female; mean age = 39.53 ± 15.728; mean years of education = 16.38 ± 2.551; HC group). Outpatients meeting the diagnostic criteria for clinically definite MS with a RR disease course were consecutively recruited from the MS Unit of Don Carlo Gnocchi Foundation, IRCCS. All patients were at a mild stage of the disease, scoring between 1 and 3 of the Expanded Disability Status Scale (EDSS).
Exclusion criteria were as follows: less than 6 months from diagnosis, documented relapses within the last 3 months, severe psychiatric and neurological disorders other than MS.

The study was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2013 and approved by the Local Ethics Committee (IRCCS Don Carlo Gnocchi Foundation). Written informed consent was obtained for all participants before study initiation.

**Procedure of the study and measures.** The study was carried out in three subsequent steps, as in Serino and colleagues. After conventional neuropsychological assessment, we administered the PIT 360° session. Next, we evaluated participants' experiences relative to their subjective feelings, intrinsic motivations, balance between resources and demands while performing the task. Additionally, their sense of presence in the 360° environment was assessed.

**Pre-task evaluation: neuropsychological measures.** We used the same battery as in Serino et colleagues: global cognitive level was assessed with the Montreal Cognitive Assessment (MoCA); executive functioning was assessed using the Trail Making Test (in two specific sub-tests: TMT-A and TMT-B) as a visuo-spatial examination with an index of time; PIT 360°; and a measure of phonemic verbal fluency, the controlled oral word association test (FAS form).

**PIT 360° session.** The PIT 360° is the 360° version of the Picture Interpretation Test (PIT). The PIT 360° environment consists of a scene in a contemporary real-world room with three frightened girls standing on chairs and a boy who is searching for something on the floor. Although not visible, it is apparent that there is a mouse (or some other small animal) hidden behind a piece of furniture. This scene is a present-day adaptation of the painting “Il Sorcio” (The Mouse), 1878, by Giacomo Favretto. Participants undergo a visual exploration task in which they are asked to interpret what is happening in a limited time frame. Time to correct interpretation of the scene ("There is a mouse/small animal") and number of scene elements before correct interpretation are the outcome metrics. Session components and their unfolding over time are illustrated in Fig. 2.

**Post-task evaluation: user experience assessment.** After task completion, we evaluated a) self-reported subjective feelings through the Geneva Emotion Wheel (GEW). This tool provides a wheel shaped arrangement of 20 emotion words. Emotion labels are considered as indices "reflecting a unique experience of mental and bodily changes in the context of being confronted with a particular event". The wheel is displayed on a space formed by the underlying dimensions of valence (negative to positive) and control/coping potential (low to high). The orthogonal combination of these dimensions generates four quadrants: negative valence - low control; negative valence - high control; positive valence - low control; positive valence - high control. Subjective feelings about performing the task were rated through the mean number of emotion labels (range 0–5) and the respective reported intensity (range 1–5) within each quadrant; b) we also evaluated the skill-demands compatibility through the Flow Short Scale (FSS); assessing the perceived level of skills in coping with the task ("Perceived coping skills"), the perceived level of challenges of the task ("Perceived challenge"), and the perceived challenge-skill balance ("Perceived challenge-skill balance") on a 5-points Likert scale; c) intrinsic motivation in performing the task was measured using the Interest/Enjoyment subscale of the Intrinsic Motivation Inventory (IMI; Deci). The mean of the item scores (N = 5, 7-points Likert scale) was considered; d) finally, we measured the sense of presence experienced in the 360° environment through the Slater-Ushoh-Steed Questionnaire (SUS Questionnaire). The scale evaluated participants’ sense of being present in the 360° scene, and the extent to which experiencing the scene using the PIT 360° became the dominant reality and recall as a place, through three items on 7-point scale.

**Data analyses.** First, the Kolmogorov-Smirnov test was used to check for the normality of data distribution for all the variables. Independent Student’s t-tests and chi-square tests were used to compare group baseline characteristics. Then, independent Student’s t-tests were carried out to explore between-group differences in the conventional assessment of executive functions (i.e., MoCA, TMT and phonemic fluency task). Two univariate analyses of variance as covariates was carried out to investigate differences between HC group and PwMS group on the two performance indices (i.e., Correct Interpretation and Number of Scene Element). Since the distribution of these two variables differed moderately from normal, a square root transformation was tried. With this procedure, data were closer to the normal distribution as assessed with the Kolmogorov-Smirnov test.

Next, differences in conventional tests of executive functions between the two groups were evaluated using non-parametric tests (Wilcoxon tests). A univariate analysis of covariance (ANCOVA) with age and education as covariates was carried out to investigate differences between HC and PwMS groups in the indexes of PIT 360° (i.e., Correct Interpretation and Number of Scene Element). To investigate potential differences between the HC group and the PwMS group in user experience variables (i.e., GEW, FSS, IMI, and SUS Questionnaire), we performed independent Student’s t-tests (for normal variables) and Mann-Whitney U tests (for not normal variables). As specifically concerns the number of self-reported emotions, the Friedman test was used to explore differences within the four quadrants of the GEW. A series of Bonferroni adjusted Wilcoxon tests were subsequently computed to explore significant effects. All these statistical analyses were conducted using the Statistical Package for the Social Sciences for Windows (SPSS Inc., Chicago, IL, USA), version 23.

To compare the classification accuracy of traditional tests of executive functions and indices from PIT 360°, nonlinear stochastic approximation (i.e., machine learning) methods were employed. In particular, a leave-one-out cross-validation was carried out with the following methods (as in our previous study): (a) a Logistic Regression classification algorithm with ridge regularization; (b) a Random Forest classification to classify features using an ensemble of decision trees; (c) a Support Vector Machine (SVM) to map inputs to
higher-dimensional feature spaces that best separate different features; (d) a naïve Bayes classification. All these analyses were computed using Python 3.4 with the Orange 3.3.5 data mining suite, which was available free in open source code (https://github.com/biolab/orange3).

Data Availability
The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References
1. Benedict, R. H. et al. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). Journal of the International Neuropsychological Society 12, 549–558 (2006).
2. Deloire, M. S. et al. MRI predictors of cognitive outcome in early multiple sclerosis. Neurology 76, 1161–1167 (2011).
3. Rao, S. M., Lee, G. J., Bernardin, L. & Unverzagt, F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. Neurology 41, 685–691 (1991).
4. Amato, M. P. et al. Cognitive impairment in early stages of multiple sclerosis. Neurological sciences 31, 211–214 (2010).
5. Amato, M. P. et al. Relevance of cognitive deterioration in early relapsing-remitting MS: a 3-year follow-up study. Multiple Sclerosis. Journal 16, 1474–1482 (2010).
6. Sumowski, J. J. et al. Cognition in multiple sclerosis: State of the field and priorities for the future. Neurology 90, 278–288 (2018).
7. Amato, M. P. et al. Treatment of cognitive impairment in multiple sclerosis: position paper. J. Neurol. 260, 1452–1468 (2013).
8. Mitolo, M., Venneri, A., Wilkinson, I. D. & Sharrack, B. Cognitive rehabilitation in multiple sclerosis: a systematic review. J. Neurol. Sci. 354, 1–9 (2015).
9. Parsons, T. D., Carlew, A. R., Magtoto, J. & Stonecipher, K. The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. Neuropsychological rehabilitation, 1–31 (2015).
10. Chan, R. C., Shim, D., Tzolopoulou, T. & Chen, E. Y. Assessment of executive functions: Review of instruments and identification of critical issues. Archives of clinical neuropsychology 23, 201–216 (2008).
11. Rouaud, O. et al. Contribution of ecological evaluation of executive disorders in multiple sclerosis. Rev. Neurol. (Paris) 162, 964–969 (2006).
12. Roet, A. & Broche, B. Cognitive assessment in patients with multiple sclerosis: From neuropsychological batteries to ecological tools. Annals of physical and rehabilitation medicine (2018).
13. Parsons, T. D. Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. Frontiers in human neuroscience 9, 660 (2015).
14. Cipresso, P. et al. Virtual multiple errands test (VMET): a virtual reality-based tool to detect early executive functions deficit in Parkinson’s disease. Frontiers in behavioral neuroscience 8, 405 (2014).
15. Cipresso, P. et al. Break in volition: A virtual reality study in patients with obsessive–compulsive disorder. Experimental brain research 229, 443–449 (2013).
16. Grewe, P. et al. Real-life memory and spatial navigation in patients with focal epilepsy: ecological validity of a virtual reality supermarket task. Epilepsy & Behavior 31, 57–66 (2014).
17. Lamargue-Hamel, D. et al. Cognitive evaluation by tasks in a virtual reality environment in multiple sclerosis. J. Neurol. Sci. 359, 94–99 (2015).
18. Langdon, D. et al. Recommendations for a brief international cognitive assessment for multiple sclerosis (BICAMS). Multiple Sclerosis. Journal 18, 891–898 (2012).
19. Serino, S. et al. Picture Interpretation Test (PIT) 360: An Innovative Measure of Executive Functions. Scientific reports 7, 16000 (2017).
20. Bisach, E., Cappa, S. & Vallar, G. Guida all’esame neuropsicologico (R. Cortina, 1983).
21. Rosci, C., Sacco, D., Laiacona, M. & Capitani, E. Interpretation of a complex picture and its sensitivity to frontal damage: a reappraisal. Neurological Sciences 25, 322–330 (2005).
22. Luria, A., Karpov, B. & Yarbuss, A. Disturbances of active visual perception with lesions of the frontal lobes. Cortex 2, 202–212 (1966).
23. Carlesimo, G. A., Caltagirone, C. & Gainotti, G. The Mental Deterioration Battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. The Group for the Standardization of the Mental Deterioration Battery. Eur. Neurol. 36, 378–384 (1996).
24. Gromisch, E. S., Portnoy, J. G. & Foley, F. W. Comparison of the abbreviated minimal assessment of cognitive function in multiple sclerosis (aMACFIMS) and the brief international cognitive assessment for multiple sclerosis (BICAMS). J. Neurol. Sci. 388, 70–75 (2018).
25. Nabavi, S. M. & Sangelaji, B. Cognitive dysfunction in multiple sclerosis: Usually forgotten in the clinical assessment of MS patients. J. Res. Med. Sci. 20, 533–534 (2015).
26. Godefroy, O. et al. Dysexecutive disorders and their diagnosis: A position paper. Cortex, 109, 322–335 (2018).
27. Polman, C. H. et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol. 69, 292–302 (2011).
28. Conti, S., Bonazzi, S., Laiacona, M., Masina, M. & Coralli, M. V. Montreal Cognitive Assessment (MoCA)-Italian version: regression based norms and equivalent scores. Neurological Sciences 36, 209–214 (2015).
29. Freitas, S. et al. The Montreal Cognitive Assessment (MoCA) as a screening test for cognitive dysfunction in multiple sclerosis. Applied Neuropsychology: Adult 25, 70–77 (2018).
30. Giovanoli, A. R. et al. Trail making test: normative values from 287 normal adult controls. The Italian journal of neurological sciences 17, 305–309 (1996).
31. Scherer, K. R., Shuman, V., Fontaine, J. R. J. & Soriano, C. The GRID meets the Wheel: Assessing emotional feeling via self-report. Components of emotional meaning: A sourcebook, 281–298 (2013).
32. Scherer, K. R. What are emotions? And how can they be measured? Social science information 44, 695–729 (2005).
33. Rheinberg, F., Vollmeyer, R. & Engeser, S. In Diagnostik von Motivation und Selbstkonzept (eds Stiensmeier-Pelster, J. & Rheinberg, F.) 261–279, (2003).
34. Engeser, S. & Rheinberg, F. Flow, performance and moderators of challenge-skill balance. Motiv. Emotion 32, 158–172 (2008).
35. Deci, E. L., Eghrari, H., Patrick, B. C. & Leone, D. R. Facilitating internalization: The self-determination theory perspective. J. Pers. 62, 119–142 (1994).
36. Usóh, M., Catena, E., Arman, S. & Slater, M. Using presence questionnaires in reality. Presence: Teleoperators & Virtual Environments 9, 497–503 (2000).

Acknowledgements
The authors wish to thank all the participants for taking part in the study.

Author Contributions
O.R., S.S., F.S., G.R. and F.B. developed the study concept. All authors contributed to the study design. F.R. was in charge of the data collection. G.C. was responsible for the technical development of the PIT. O.R., S.S. and F.S. performed the data analysis and interpretation under the supervision of F.B., F.M., R.N., L.M., T.D.P. P.C. performed computational data analysis. F.S., F.B., O.R. and S.S. wrote the first draft of the manuscript. All authors were involved in a critical revision of the manuscript for important intellectual content. All the authors approved the final version of the manuscript for submission.

Additional Information
Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019