Supplementary Material

1. **Details about training programs**
2. **Supplementary Table 1**
3. **Supplementary Figure 1**
4. **Details about measuring freezing of gait ratio (FOG-ratio)**
5. **Details about measuring Frontal Assessment Battery (FAB)**
6. **Details about measuring Montreal Cognitive Assessment (MoCA)**
7. **Details about measuring Digit Symbol Substitution Test (DSST)**
8. **Details about measuring Trial Making Test part B-A (TMT_{B-A})**
9. **Details about measuring Stroop test**
10. **Supplementary Figure 2**
11. **Supplementary Figure 3**
12. **Supplementary Figure 4**
13. **Supplementary Table 2**
14. **Supplementary ANCOVA results**

1. Details about training programs

Adapted Resistance Training with Instability (ARTI) program

One-on-one training was provided three times a week for 12 weeks (36 training sessions). Each training session lasted between 80-90 minutes. ARTI is an adaptation of our previously published RTI program for non-freezers of PD patients.[1] ARTI consisted of seven lower-limb and upper-limb free weight exercises (half squat, plantar flexion, chest press, knee-lifting stand, lunge, reverse fly, and dual-task squat) performed on unstable devices (i.e., foam pad, dyna discs, balance disc, BOSU®, and Swiss Ball) as demonstrated in Figure 1. A rest interval of 90 seconds was allowed between exercises and sets. There was a progressive and concomitant increase in degree of instability and weight of the exercises during the three-month period. Regarding instability, unstable devices were changed throughout the experimental period from the least to the most unstable devices to increase the motor complexity, as demonstrated in Table 1. An unstable device was changed to a more unstable one whenever subjects decreased body sway considerably (i.e., ability to balance the body on the device, without presenting large excursions of the center of mass in the anteroposterior and mediolateral axes – visual inspection) and force production increased abruptly while performing the selected free weight exercise. If subjects were unable to increase exercise weight due to the high instability when performing the exercise on the new unstable device, training load from the session prior to changing devices was used. Regarding sets and repetitions, individuals performed 2–3 sets of 10–12 repetitions maximum in the first month; performed 3–4 sets of 8–10 repetitions maximum; in the second month and 4 sets of 6–8 repetitions maximum in the third month. Initial exercise weight was adjusted throughout the sets to allow individuals to perform the assigned repetitions. Afterward, exercise weight was systematically increased (5%–10%) whenever individuals were able to perform the predefined repetitions maximum for two consecutive training sessions (e.g., 10–12 repetitions maximum in the first month, 8–10 repetitions maximum in the second month, and 6–8 repetitions maximum in the third month). ARTI sessions were individualized and monitored by trainers knowledgeable in working with individuals with PD and with the RTI program.[1]

Traditional Motor Rehabilitation (TMR) program

The TMR group trained three times a week for 12 weeks (36 training sessions). Each training session lasted between 80-90 minutes. TMR sessions were group-based (up to 8 individuals), monitored by a physical therapist knowledgeable in working with individuals with PD. TMR consisted of exercises performed in the following order:

1) stretching - standing chest stretch, seated neck and chest stretch, seated rotation stretch, overhead stretch, standing back stretch, hamstring stretch, lying shoulder stretch, seated side stretch, standing shoulder stretch, calf stretch. Three times for 10 breath counts;

2) gait - individuals were instructed to cover as much distance as possible and walk up and down a 20-meter hallway “as fast as possible”, 10 times. There was a 90 seconds of rest (or more if needed);

3) balance and posture - weight shifts forward and backward 10 to 20 times while standing with feet placed hip width apart and single leg stance on each leg for 10 seconds. There was a 90 seconds of rest (or more if needed);

4) lower- and upper-limbs free weight exercises - half squat, plantar flexion, chest press, back extension, knee extension, triceps extension, and knee-lifting stand. The exercises included the addition of weights in which load was systematically increased (5%–10%) whenever individuals were able to perform the predefined repetitions (10 repetitions) for two consecutive training sessions. Subjects performed three sets of 10 repetitions during the 3 months. There was a 90 seconds of rest (or more if needed).

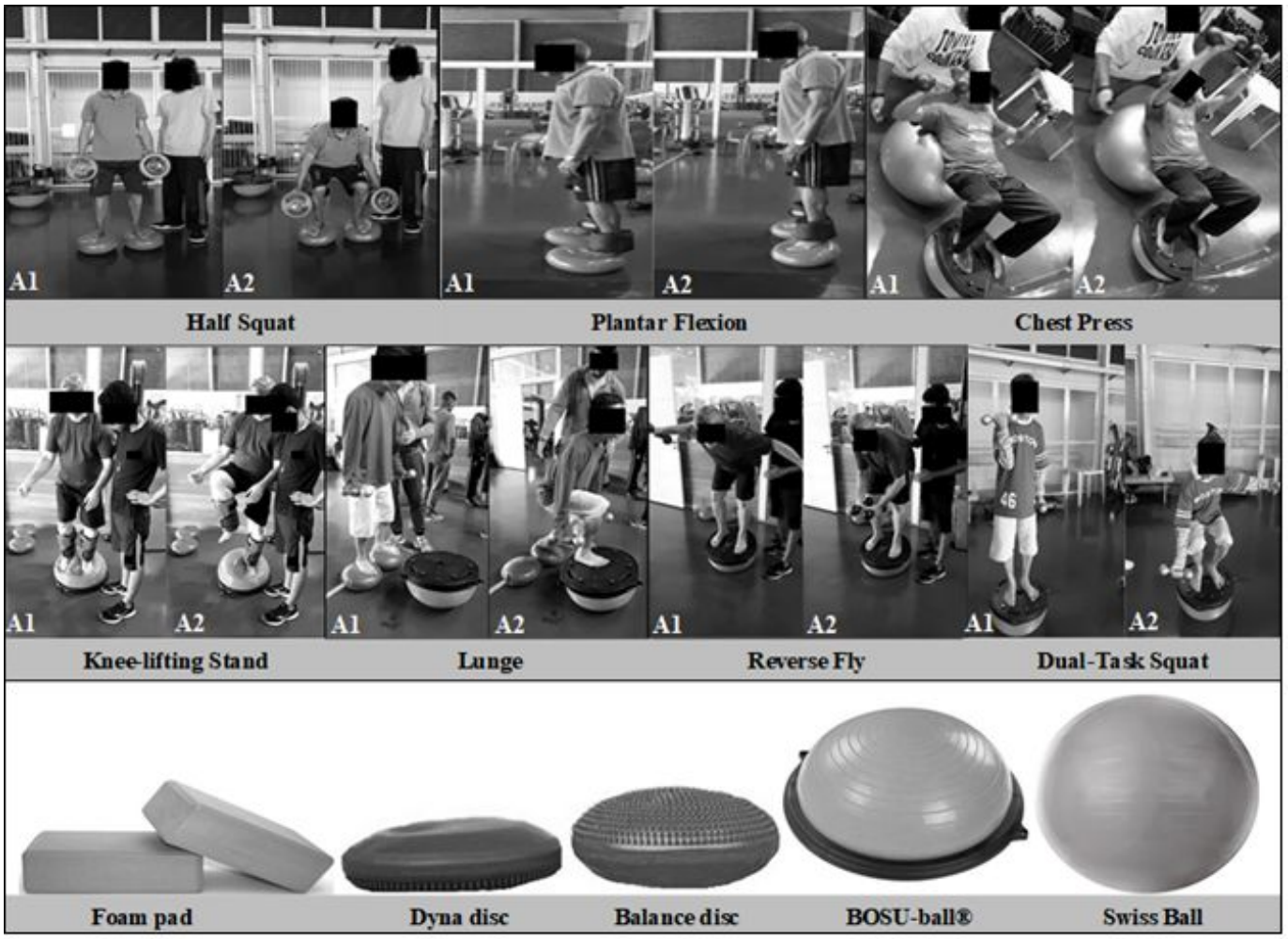
2. **Supplementary Table 1.** Location and progression of the unstable devices throughout the experimental protocol (12 weeks) for each free exercise.

	Half squat	**Plantar flexion**	**Chest press**	**Knee-lifting stand**	**Lunge**	**Reverse fly**	**Dual-task squat**
Week 1 and 2	Swiss Ball–back*	Foam pad–feet*	Swiss Ball–back*	Foam pad–feet*	Foam pad–feet* and forward stepping on foam pad	Foam pad–feet*	Swiss Ball–back* and alternating arm movements while squatting
Week 3	Foam pad–feet* and Swiss Ball–back*	Dyna discs –feet*	Foam pad–feet* and Swiss Ball–back*	Dyna discs –feet*	Dyna discs –feet* and forward stepping on foam pad	Dyna discs –feet*	Foam pad–feet* and Swiss Ball–back* and alternating arm movements while squatting
Week 4 and 5	Dyna discs–feet* and Swiss Ball–back*	Dyna discs –feet*	Dyna discs–feet* and Swiss Ball–back*	Dyna discs–feet*	Dyna discs–feet* and forward stepping on balance discs	Dyna discs – feet*	Dyna discs–feet* and Swiss Ball–back* and alternating arm movements while squatting
Week 6	Balance discs–feet* and Swiss Ball– back*	Balance discs–feet*	Balance discs–feet* and Swiss Ball– back*	Balance discs–feet*	Dyna discs–feet* and forward stepping on balance discs	Balance discs– feet*	Balance discs–feet* and Swiss Ball–back* and alternating arm movements while squatting
Week 7 and 8	Balance discs–feet*	Balance discs–feet*	BOSU-ball® –feet* and Swiss Ball–back*	Balance discs–feet*	Balance discs– feet* and forward stepping on BOSU-ball®	Balance discs– feet*	Balance discs–feet* and alternating arm movements while squatting
Week 9 and 10	BOSU-ball®–feet*	Balance discs–feet*	BOSU-ball® –feet* and Swiss Ball–back*	Balance discs–feet*	Balance discs– feet* and forward stepping on BOSU-ball®	BOSU-ball®–feet*	BOSU-ball®–feet* and alternating arm movements while squatting

Week 11 and 12	BOSU-ball® flipped upside down–feet*	Balance disc–feet*	BOSU-ball® flipped upside down–feet* and Swiss Ball – back*	BOSU-ball®–feet*	Balance discs– feet* and forward stepping on BOSU-ball® flipped upside down	BOSU-ball® flipped upside down–feet*	BOSU-ball® flipped upside down–feet* and alternating arm movements while squatting

* Indicates the body part placement of the unstable devices for each free weight exercise.

3. **Supplementary Figure 1**. Initial (A1) and final (A2) phase of motion in the free weight exercises performed with unstable devices in the Adapted Resistance Training with Instability.



4 Details about measuring FOG-ratio

The FOG-ratio was calculated during a 2-minute turning task, in which individuals made 360° turns on the spot, alternating between clockwise and anti-clockwise turns as fast as they could do safely. Inertial sensors (Physilog, Gait Up, Lausanne, Switzerland) were placed on the shins and at the lumbar level. Data were sampled at 128Hz and stored for offline analysis using Matlab 2016b (Mathworks Inc.). Analysis was based on power spectral density from the anteroposterior acceleration data. The FOG-ratio was then calculated as the ratio between the square of the total power in the frequency band corresponding with freezing episodes (3-8 Hz) and the total power in the frequency band corresponding to locomotion (0.5-3 Hz). Higher FOG-ratio scores indicate greater FOG severity. See more details in Mancini et al.[2]

5 Details about measuring FAB

A researcher, trained by a neuropsychologist, blind to the experimental design applied the FAB in a quiet room without distractions. The FAB test consisted of the following six subtests: (1) questions about the similarities between two objects to evaluate conceptualization and abstract thinking; (2) questions asking the participant to list as many words as possible starting with a specific letter within 1 minute to assess their mental flexibility and verbal fluency; (3) questions asking individuals to perform the Luria maneuver and fist-edge-palm patterns to determine if their programming and motor acts are correctly executed; (4) questions requesting the participants to provide an opposite response to the examiner's signals (conflicting instructions) to assess sensitivity and interference; (5) questions asking the participants to inhibit their response to a stimulus that was previously administered (go-no-go test) to assess their inhibitory control; and (6) questions assessing the involuntary behavior that is triggered by sensory stimulation (prehension behavior) to detect deficits in the environmental autonomy. Each subtest had a score of 0 to 3, with a total possible score of 18; a higher score indicated a better function [3]. Thus, the FAB score, the proportion of patients with frontal executive dysfunction, and the FAB cognitive domains were used for analysis.

6 Details about measuring MoCA

A researcher, trained by a neuropsychologist, blind to the experimental design applied the MoCA in a quiet room without distractions. The maximum score is 30 and a score of ≤ 25 indicates mild cognitive impairment [4]. A point is added to the total score for those with 12 yr or fewer years of education. The MoCA assesses seven cognitive domains, such as visuospatial and executive functions (5 points), naming (3 points), attention (6 points), language (3 points), abstraction (2 points), delayed recall (5 points), and orientation (6 points). Thus, the MoCA score, the proportion of patients with mild cognitive impairment, and the MoCA cognitive domains were used for analysis.

7 Details about measuring DSST

A researcher, trained by a neuropsychologist, blind to the experimental design applied the DSST in a quiet room without distractions. The DSST is a pencil and paper test of psychomotor performance in which the subject is given a key grid of numbers and matching symbols and a test section with numbers and empty boxes. The test consists of filling as many empty boxes as possible

with a symbol matching each number. The score is the number of correct number symbol matches achieved in 90 s.[5]

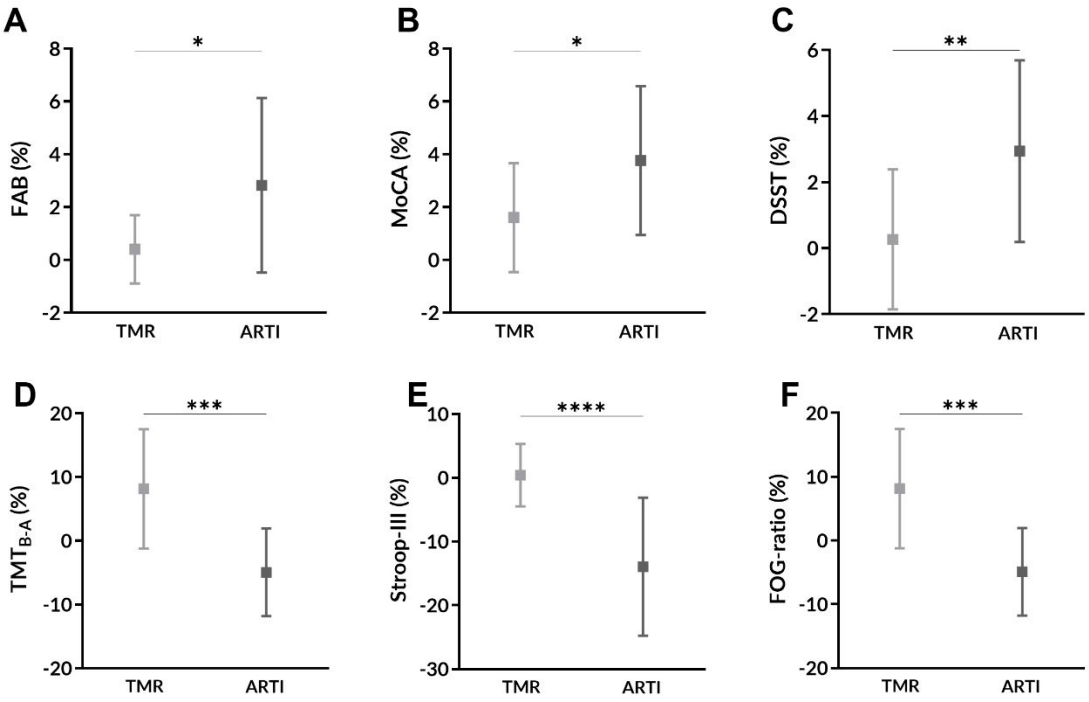
8 Details about measuring TMT_{B-A}

A researcher, trained by a neuropsychologist, blind to the experimental design applied the TMT_{B-A} in a quiet room without distractions. The test is divided into two sections: section A requires connecting 25 numbers within circles randomly arranged on an A4 sheet in ascending order; and section B requires connecting 12 letters and 13 numbers in alphabetical and ascending order alternately. The test should be performed as fast as possible without lifting the pen or pencil from the paper. The assessment was conducted in a quiet room without distractions. As has been previously described, the difference in time to completion between section B and section A (time between Trail Making Test part B and A [TMT_{B-A}]) measures the individual's ability to rapidly set-shift.[6, 7]

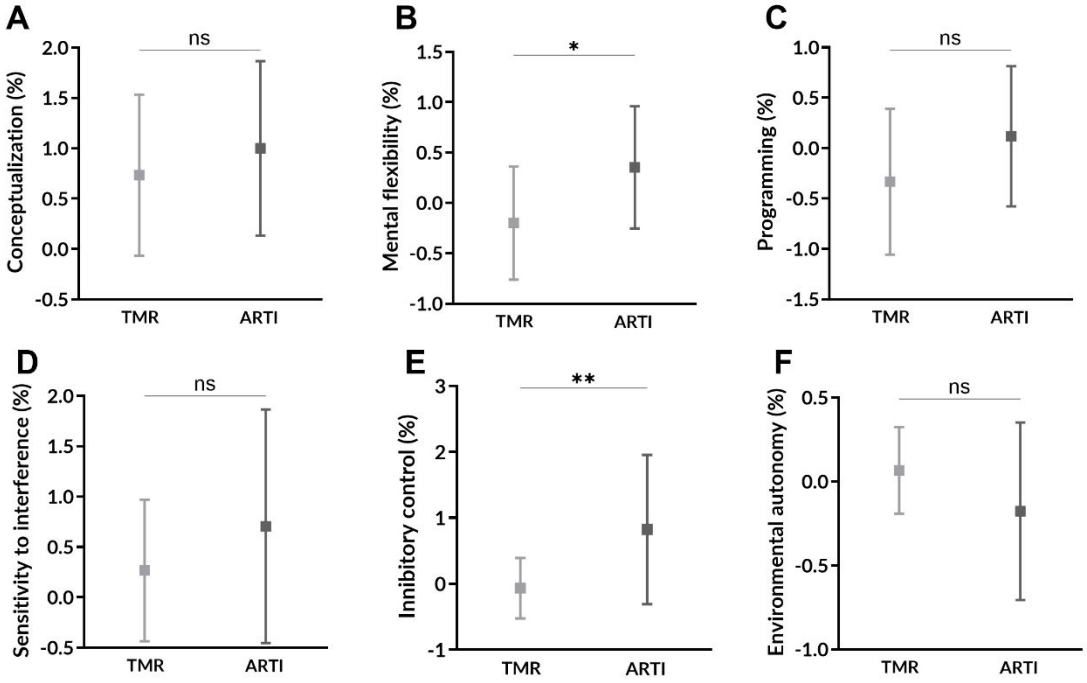
9 Details about Stroop test

A researcher, trained by a neuropsychologist, blind to the experimental design applied the Stroop test in a quiet room without distractions. The Stroop test is a neuropsychological test extensively used to assess the ability to inhibit cognitive interference[8]. We used the Stroop Color-Word Test-Victoria version,[8] which comprises three cards containing 24 stimuli each and a white background. Card A is composed of rectangles printed in green, pink, blue and brown, randomly arranged. Card B is organized similarly to Card A, but with rectangles replaced by the words "every, never, today and all" printed in capital letters in the four colors mentioned. Card C was similarly organized to Card A, representing the interference card in which the name of the colors (brown, blue, pink and green) printed and written on each card never matched (e.g. brown word printed in pink, green or blue). For the first card (Stroop-I), participants have to state the colors of the rectangles as quickly as possible. For cards B (Stroop-II) and C (Stroop-III), participants must state the color of the printed words and not the written name of the colors. The Conflict condition of the Stroop test (Stroop-III) assesses patients' ability to inhibit their dominant tendency to read words in the context of instructions instead of the name of written colors. As has been previously recommended[9], both time and error were used to estimate the cognitive interference scores on Stroop-III test.

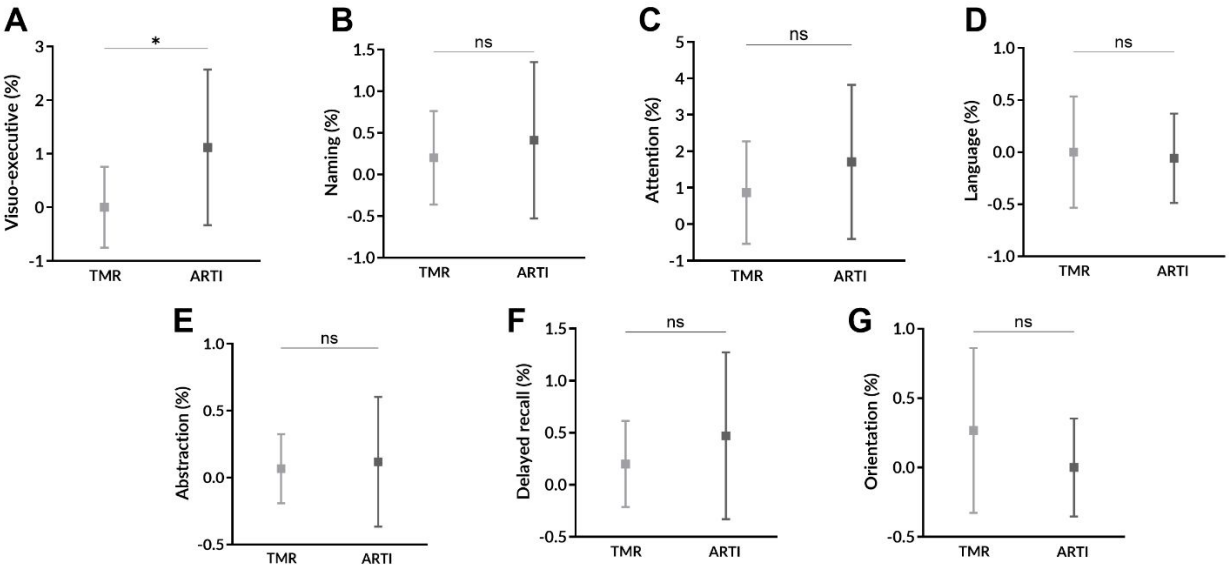
Supplementary Figure 2. Mean ± standard deviation for the change in secondary cognitive outcomes and FOG-ratio of the TMR and the ARTI groups. **P*<0.05; ***P*<0.01; ****P*=0.0001; and *****P*<0.0001.



Supplementary Figure 3. Mean ± standard deviation for the change in FAB subscores of the TMR and the ARTI groups. *P<0.05; **P<0.01; and ns=non-significant.



Supplementary Figure 4. Mean ± standard deviation for the change in MoCA subscores of the TMR and the ARTI groups. *P<0.05; and ns=non-significant.



13 Supplementary Table 2. A linear multiple regression (forward stepwise method) with included independent factor, dependent variable, and baseline freezing of gait ratio as a covariate.

Independent factor	Partial R^2	Adjusted R^2	SE of the Estimate	F change	95% CI (lower)	95% CI (upper)	*P* value
FAB (%)	0.68	0.43	5.16	13.3	-2.26	-0.59	P<0.01

FAB = Frontal Assessment Battery; SE = standard error; CI = confidence interval.

Changes in FAB (r=-0.68, *P*=0.003), Stroop (r=-0.65, *P*=0.002), and TMT_{B-A} (r=0.56, *P*=0.02) values were associated with changes in FOG ratio. These three cognitive variables entered in the regression model but only changes in FAB scores explained changes in FOG ratio as demonstrated in Supplementary Table 2.

14 Supplementary ANCOVA results

ANCOVA analyses showed that FOG-ratio at baseline as a covariate did not influence the effects of training protocols on FAB ($F_{[1, 28]}$=0.59, *P*=0.44), MoCA ($F_{[1, 28]}$=3.59, *P*=0.07), DSST ($F_{[1, 30]}$=2.32, *P*=0.13), Stroop-III ($F_{[1, 28]}$=0.83, *P*=0.37), TMT_{B-A} ($F_{[1, 28]}$=0.33, *P*=0.57), and FOG ratio values ($F_{[1, 28]}$=0.32, *P*=0.57).

REFERENCES

1. Silva-Batista C, Corcos DM, Roschel H, et al. Resistance Training with Instability for Patients with Parkinson's Disease. *Med Sci Sports Exerc*. Sep 2016;48(9):1678-87. doi:10.1249/MSS.0000000000000945
2. Mancini M, Smulders K, Cohen RG, Horak FB, Giladi N, Nutt JG. The clinical significance of freezing while turning in Parkinson's disease. *Neuroscience*. Feb 20 2017;343:222-228. doi:10.1016/j.neuroscience.2016.11.045
3. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology*. Dec 12 2000;55(11):1621-6. doi:10.1212/wnl.55.11.1621
4. Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. Apr 2005;53(4):695-9. doi:10.1111/j.1532-5415.2005.53221.x
5. Wechsler D. WAIS-III: administration and scoring manual. *Psychol Corp*. Nov 1 1997;
6. Naismith SL, Shine JM, Lewis SJ. The specific contributions of set-shifting to freezing of gait in Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society*. Jun 15 2010;25(8):1000-4. doi:10.1002/mds.23005
7. Shine JM, Naismith SL, Palavra NC, et al. Attentional set-shifting deficits correlate with the severity of freezing of gait in Parkinson's disease. *Parkinsonism & related disorders*. Mar 2013;19(3):388-90. doi:10.1016/j.parkreldis.2012.07.015
8. Troyer AK, Leach L, Strauss E. Aging and response inhibition: Normative data for the Victoria Stroop Test. *Neuropsychology, development, and cognition Section B, Aging, neuropsychology and cognition*. Mar 2006;13(1):20-35. doi:10.1080/138255890968187
9. Gardner RW, Holzman PS, Klein GS, Linton HP, Spence DP. Cognitive control: a study of individual consistencies in cognitive behavior. *Psychological Issues*. Nov 1959;1(4):1-186.