Statistical Analysis Plan

Protocol Identification Number: S61792
Brief Title: Role of sleep on motor learning in Parkinson's disease and healthy older adults
Official Title: Boosting motor learning through sleep and targeted memory reactivation in Parkinson's disease and healthy older adults
Acronym: TARGET-SLEEP
Study type: Interventional
Principal Investigator: Prof Alice Nieuwboer

Note on protocol amendment 27 April 2021:
It was noted that a formal power calculation was not provided in the original submission on clinicaltrials.gov. The a-priori determined power calculation has now been added. Importantly, the target sample for recruitment did not change to what was noted on the original submission. Moreover, a change was made to the secondary outcomes 7 and 8 regarding dual tasking for experiment 2. It was decided to not have participants perform the dual-task condition during learning, as to minimize learning effects on the primary outcome. Importantly, due to COVID-19 restrictions, recruitment for experiment 2 had not yet commenced and hence no subjects had been randomized prior to submission of this protocol amendment.

Brief summary:
People with Parkinson's disease (pwPD) often present difficulty consolidating newly learned skills into long-term memory. Sleep facilitates motor memory consolidation in healthy adults, especially in combination with targeted memory reactivation (TMR). TMR works by adding associated sounds during learning that are replayed during sleep and thus reinforce the recently formed neural connections. Importantly, recent work suggested that consolidation during sleep may be preserved in pwPD, but robust findings are lacking and have not involved TMR. The objective of the present study is to address this imperative question by investigating the effect of napping on motor memory consolidation by experimentally manipulating exposure to sleep (Experiment 1) and TMR (Experiment 2) for the first time. Concretely, the investigators will first compare the effect of a 2-hour nap to that of a wake control period in pwPD and healthy age-matched controls. A validated motor sequence learning task will be used to test for behavioral markers of motor learning and polysomnography with electroencephalography (EEG) will be conducted to study the neural correlates of sleep-related motor learning effects. In a second experiment, the investigators will then test the effects of adding TMR during post-learning sleep, by comparing performance on two motor sequences of which only one is reactivated during post-learning napping using auditory TMR.

Primary outcome:
Experiment 1: Participants perform a self-initiated MSL task by tapping a five-element finger sequence presented on screen as rapidly and accurately as possible with their non-dominant hand for 18 blocks during learning, immediately after the nap or wake intervention (Retest 1) and again 24hours

later without further practice (Retest 2). Each block consists of 50 key presses (ideally 10 sequences) and is followed by a rest block of 15-20 seconds without finger tapping. A two-minute rest period will be implemented after 14 blocks to further minimize the effects of fatigue on the last 4 blocks that are used to calculate the primary outcome. Performance on the MSL will be assessed using the 'Performance Index (PI)' [$PI=exp^{-(seqDur(x))} * exp^{-(Errors(x)/12)} * 100$], whereby (x) represents the blocks of the trials (King et al. 2017b). Given that both speed and accuracy determine the PI, these measures will be explored. During the single-task condition, participants view a fixation cross in the middle of the screen with the sequence presented above.

The primary outcome for experiment 1 will be the offline consolidation, measured as the change in PI between the first 4 blocks at Retest 1 immediately after the nap or wake intervention and the last 4 blocks of Learning immediately prior to the intervention. The primary outcome for experiment 1 will be analyzed with a repeated measures analysis of variance with the following factors and conditions: 4 blocks * 2 sessions (end of learning/begin of retest 1) * 2 groups (nap/wake) * 2 populations (pwPD/controls). Linear mixed models will be considered depending on the presence of missing values and interactions with covariates (age, gender, disease severity and levodopa equivalence dose) will be analyzed.

Experiment 2**:** The design of experiment 2 is similar to experiment 1, except that participants will learn two motor sequences that are visually and auditory cued by means of a serial reaction time task (SRT). After learning both sequences, participants will nap for 2 hours, but this time while the melody of one of the two sequences will be replayed using auditory TMR. Performance on both sequences will be re-assessed immediately after the intervention (Retest 1), and again at 24h retention (Retest 2) without cues.

The primary outcome for experiment 2 is the 'offline consolidation' defined as the change in PI between the first 4 blocks at retest immediately after the nap and the last 4 blocks at learning prior to the intervention, which will be compared between the sequence that was replayed during the nap (replay) and the sequence that was not replayed (no-replay). The primary outcome for experiment 2 will be analyzed with a repeated measures analysis of variance with the following factors and conditions: 4 blocks * 2 sessions (end of learning/ begin of retest 1) * 2 sequences (replay/no-replay) * 2 groups (pwPD/controls). Linear mixed models will be considered depending on the presence of missing values and interactions with covariates (age, gender, disease severity and levodopa equivalence dose) will be analyzed.

Secondary outcomes

All other comparisons for experiments 1 and 2 are considered secondary in importance. The PI is used as the main dependent variable of interest for all secondary outcomes listed below, which will be compared across the two intervention groups (Nap/Wake) and two study populations (pwPD/Controls). An MSL with consecutive dual task condition is performed for an additional 4 blocks at learning, immediately after the intervention (Retest 1) and again after

24hrs of retention (Retest 2). The following comparisons will be analyzed for both experiments.

Single task performance (PI):

1. Retention, measured as the change in PI between the first 4 blocks at Retest 2 compared to the last 4 blocks at Retest 1.
2. Learning slope over 14 blocks during learning, prior to the intervention.
3. Learning slope over 14 blocks after the 24hr retention period (Retest 2)
4. End of learning performance assessed over the last 4 blocks during learning prior to the intervention (Learning)
5. Post-extensive practice performance assessed over the last 4 blocks immediately after the intervention (Retest 1)

Dual tasking performance (PI):

1. Offline consolidation measured as the change in PI between the 4 blocks of dual tasking at Retest 1 immediately after the intervention compared to the 4 blocks of dual tasking during learning for experiment 1, and differences on PI between sequences A and B during the 4 blocks at Retest 1 for experiment 2.
2. Retention, measured as the change in PI between the 4 blocks of dual tasking at Retest 2 compared to the 4 blocks of dual tasking at Retest 1 for experiment 1, and differences on PI between sequences A and B during the 4 blocks at Retest 2 for experiment 2.
3. Learning performance over the 4 blocks of dual tasking during learning prior to the intervention for experiment 1 only.
4. Learning performance over the 4 blocks of dual tasking immediately after the intervention (Retest 1) for both experiments.
5. Learning performance over the 4 blocks of dual tasking after the 24hr retention period (Retest 2) for both experiments.

Dual task cost (ΔPI), i.e. the difference in PI between 4 blocks of dual tasking and 4 blocks of single tasking.

1. Offline consolidation measured as the change in ΔPI between Retest 1 immediately after the intervention (4 blocks of dual tasking versus last 4 blocks of single tasking) and the ΔPI during learning (4 blocks of dual tasking versus the first 4 blocks of single tasking) for experiment 1, and difference in ΔPI between sequences A and B (4 blocks of dual tasking versus last 4 blocks of single tasking) at Retest 1 for experiment 2.
2. Retention, measured as the change in ΔPI between Retest 2 (4 blocks of dual tasking and last 4 blocks of single tasking) and Retest 1 (4 blocks of dual tasking and last 4 blocks of single tasking) for experiment 1, and difference in ΔPI between sequences A and B (4 blocks of dual tasking versus last 4 blocks of single tasking) at Retest 2 for experiment 2.
3. Learning performance assessed as the ΔPI of the 4 blocks of dual tasking and first 4 blocks of single tasking during learning for experiment 1 only.
4. Post extensive practice assessed as the ΔPI of the 4 blocks of dual tasking and last 4 blocks of single tasking immediately after the intervention (Retest 1) for both experiments.

5. Post extensive practice assessed as the ΔPI of the 4 blocks of dual tasking and last 4 blocks of single tasking after the 24hours retention period (Retest 2) for both experiments.

Control comparisons (PI): For experiment 2, four additional comparisons will be analyzed to ensure that participants could learn both sequences equally well. The two sequences can be labeled as either (sequence A / sequence B), irrespective of which sequence was replayed during the nap period, or as (replay / no-replay) respective of the sequence that was replayed during the nap period.

1. End of learning performance on sequence A and B will be assessed as the difference in PI between the last 4 blocks of sequence A and the last 4 blocks of sequence B during learning prior to the intervention.
2. Learning slopes of sequence A and B will be assessed as the difference in PI between the first 14 blocks of sequence A and the first 14 blocks of sequence B during learning prior to the intervention.
3. End of learning performance on the replay and no-replay sequences will be assessed as the difference in PI between the last 4 blocks of the TMR sequence and the last 4 blocks of the control sequence during learning prior to the intervention.
4. Learning slopes of the replay and no-replay sequences will be assessed as the difference in PI between the first 14 blocks of sequence A and the first 14 blocks of sequence B during learning prior to the intervention.

Tertiary outcomes

Experiment 1 - Motor Execution Test (MET)
Participants perform a self-initiated MET by tapping a simple 4-element (i.e. 4-3-2-1) finger sequence presented on screen as rapidly and accurately as possible with their non-dominant hand for 4 blocks during learning and again for 4 blocks at each retest assessment. Each block consists of 52 key presses (ideally 13 sequences) and is followed by a rest block of 15-20 seconds without finger tapping. Performance on the MET will be compared between Retest 1 and learning and between Retest 2 and Retest 1.

Experiment 2 - Random Serial Reaction Time task (random SRT)
Participants will perform random sequences that are visually cued by means of a SRT for 4 blocks at learning and again at each retest assessment. Each block consists of 50 key presses and is followed by a rest block of 15-20 seconds without finger tapping. Performance on the random SRT will be compared between Retest 1 and learning and between Retest 2 and Retest 1.

Psychomotor Vigilance Test (PVT):
The MSL (Experiment 1) and SRT (Experiment 2) tests will be preceded by a 10-minute PVT as an objective measure of the participants' vigilance prior to testing. The PVT will be presented on a laptop and subjects will be instructed to press a button as quickly as possible each time a visual stimulus is presented in the middle of the screen. Changes in PVT performance (number of lapses in attention defined as response times ≥ twice the median response time) will be

assessed between Retest 1 and the period prior to the intervention (i.e. Learning), as well as between Retest 2 and Retest 1.

Planned data exclusion

Anticipated data exclusion criteria are the performance measures of the first block after each NAP/WAKE or 24-hour Retention period, given these may be confounded by start-up effects. We will perform a sensitivity analysis without these trials to control for these confounding effects.

Power calculation

Based on the findings by Terpening et al. (2013) and Dan et al. (2015), a minimum of 16 subjects per group (NAP, WAKE) will be required according to our power analysis based on the MSL-outcomes using $\boldsymbol{\beta}$=0.20 and $\boldsymbol{\alpha}$=0.05 to detect a significant group difference. To account for potential dropouts, the recruitment target is set 20% higher to ensure adequate power in our final analysis. As such, a total of 40 PD patients and 40 healthy elderly controls will be recruited for experiment 1 (i.e. 20 in each NAP/WAKE group). The best sample estimation at this time for experiment 2 is based on previous TMR studies in younger adults also recruiting 16 subjects per nap/wake group (Antony et al. 2012). Therefore, we will target to recruit a total of 20 PD and 20 healthy elderly controls for Experiment 2, again accounting for 20% potential dropout.