

Supplementary material

1. Methods

Motor tasks detailed description

Psychomotor vigilance test (PVT). During this 10-minutes sustained attention task, participants were placed in front of a computer screen showing a “+” sign. As soon as the symbol was substituted by a millisecond counter, they had to press on the response button (spacebar), to stop the counter. The counter appeared 100 times within 10 minutes, at random intervals between 2 and 10 seconds. In the event that the participants pressed the response button when the counter had not appeared, the “+” sign would automatically decrease its size by 10%, after the message “watch out the counter” appeared at the centre of the screen.

2. Performance Index supplementary results

Verification of performance plateau

To ensure that the necessary plateau of performance was reached before the intervention, we explored the differences in performance index (PI) among the four blocks of test pre-intervention with a Bayesian ANOVA, including group (i.e., people with Parkinson disease (PD) and healthy older adults (HOA)), intervention and block as fixed effects and subject as random effect. This was supplemented with a linear mixed effect model including the same fixed and random effects. Here we found decisive evidence for an effect of block ($BF_{10} > 1000$, $F_{(3,180)} = 15.62$, $p < 0.01$, partial $\eta^2 = 0.21$), and post-hoc analyses revealed decisive evidence for a difference between the first block and the other three (block 1-2: $BF_{10} = 322.29$, $V = 422$, $p < 0.01$; block 1-3: $BF_{10} > 1000$, $V = 338$, $p < 0.01$; block 1-4: $BF_{10} > 1000$, $V = 289$, $p < 0.01$), but weak and moderate evidence for no difference among the other three (block 2-3: $BF_{10} = 0.86$, $V = 839$, $p = 0.18$; block 2-4: $BF_{10} = 0.63$, $V = 812$, $p = 0.13$; block 3-4: $BF_{10} = 0.14$, $V = 1025$, $p = 0.92$). Weak to moderate evidence for no main effects of group and intervention (group: $BF_{10} = 0.98$, $F_{(1,60)} = 2.70$, $p = 0.11$, partial $\eta^2 = 0.04$; intervention: $BF_{10} = 0.62$, $F_{(1,60)} = 0.48$, $p = 0.49$, partial $\eta^2 < 0.01$), and for their interactions (group by intervention: $BF_{10} = 0.68$, $F_{(1,60)} = 0.26$, $p = 0.62$, partial $\eta^2 < 0.01$; group by block: $BF_{10} = 0.14$, $F_{(3,180)} = 1.31$, $p = 0.27$, partial $\eta^2 = 0.02$; intervention by block: $BF_{10} = 0.16$, $F_{(3,180)} = 1.43$, $p = 0.23$, partial $\eta^2 = 0.02$; group by intervention by block: $BF_{10} = 0.11$, $F_{(3,180)} = 0.28$, $p = 0.84$, partial $\eta^2 < 0.01$). These findings support evidence for a first block effect, therefore we decided to exclude this first block of test pre-intervention in the calculation of offline performance changes post-intervention, both in PD and in HOA.

Similarly, we explored the differences among the four blocks of test post-intervention with a Bayesian ANOVA in the PD and in the HOA group, including intervention and block as fixed effects and subject as random effect. Again we found decisive evidence in favour of an effect of block ($BF_{10} > 1000$, $F_{(3,179)} = 16.59$, $p < 0.01$, partial $\eta^2 = 0.22$). Post-hoc analyses revealed decisive evidence in favour of a difference between the first block and the other three ($BF_{10} > 1000$, $V = 339$, $p < 0.01$; block 1-3: $BF_{10} = 735.18$, $V = 426$, $p < 0.01$; block 1-4: $BF_{10} > 1000$, $V = 387$, $p < 0.01$), and weak to moderate evidence in favour of no difference for the other pairwise comparisons (block 2-3: $BF_{10} = 0.14$, $V = 1066$, $p = 0.69$; block 2-4: $BF_{10} = 0.31$, $V = 849$, $p = 0.20$; block 3-4: $BF_{10} = 0.39$, $V = 836$, $p = 0.24$). For the other factors of the Bayesian ANOVA, weak evidence was found for no group effect ($BF_{10} = 1.35$, $F_{(1,60)} = 3.74$, $p = 0.06$, partial $\eta^2 = 0.06$), and weak to moderate evidence in favour of no effect of intervention ($BF_{10} = 0.53$, $F_{(1,60)} = 0.16$, $p = 0.69$, partial $\eta^2 < 0.01$) nor of the interactions of group by intervention ($BF_{10} = 0.61$, $F_{(1,60)} = 0.06$, $p = 0.81$, partial $\eta^2 < 0.01$) group by block ($BF_{10} = 0.11$, $F_{(1,179)} = 1.12$, $p = 0.34$, partial $\eta^2 = 0.02$), intervention by block ($BF_{10} = 0.21$, $F_{(1,179)} = 1.78$, $p = 0.15$, partial $\eta^2 = 0.03$), group by intervention by block ($BF_{10} = 0.13$, $F_{(1,179)} = 0.58$, $p = 0.63$, partial $\eta^2 < 0.01$). Given the degree of evidence for the differences between blocks, we opted for excluding the first block of the test post-intervention

Supplementary material

also in the computation of post-night offline changes (i.e., retention). Note that one subject in the PD group had a missing value for block 3, as they failed to perform the sequence correctly at all the attempts.

Motor sequence learning and general motor execution

Firstly, to clarify whether any change in performance was related to actual learning and consolidation, or to mere motor execution improvement, we tested for differences in performance between motor sequence learning (MSL) and the motor execution test (MET). We thus compared the relative changes in PI between the start of the first session and the end of the last session, and the relative changes between MET performed at the start of the first session and the one performed at the end of the last session. The Bayesian ANOVA including group, intervention and task revealed strong evidence in favour of an effect of task ($BF_{10} = 66.21$, $F_{(1,60)} = 12.19$, $p < 0.01$, partial $\eta^2 = 0.17$) whereas moderate evidence in favour of no effect of group ($BF_{10} = 0.24$, $F_{(1,60)} = 0.13$, $p = 0.72$, partial $\eta^2 < 0.01$), intervention ($BF_{10} = 0.27$, $F_{(1,60)} = 0.39$, $p = 0.53$, partial $\eta^2 < 0.01$) group by task interaction ($BF_{10} = 0.30$, $F_{(1,60)} = 0.35$, $p = 0.56$, partial $\eta^2 < 0.01$) and intervention by task interaction ($BF_{10} = 0.27$, $F_{(1,60)} = 0.15$, $p = 0.70$, partial $\eta^2 < 0.01$) was found. Weak evidence was found for no group by intervention interaction ($BF_{10} = 0.45$, $F_{(1,60)} = 1.14$, $p = 0.29$, partial $\eta^2 = 0.02$) and the three-way interaction of group, intervention and task ($BF_{10} = 0.46$, $F_{(1,60)} = 0.60$, $p = 0.44$, partial $\eta^2 < 0.01$). Post-hoc comparisons revealed strong evidence for a difference in performance change between the MSL and the MET ($BF_{10} = 18.86$, $V = 213$, $p < 0.01$, $r_{rb} = 0.41$, $p < 0.01$). Overall, these results suggest that the performance change was not related to general motor execution improvement only, but to actual learning, irrespective of group and intervention.

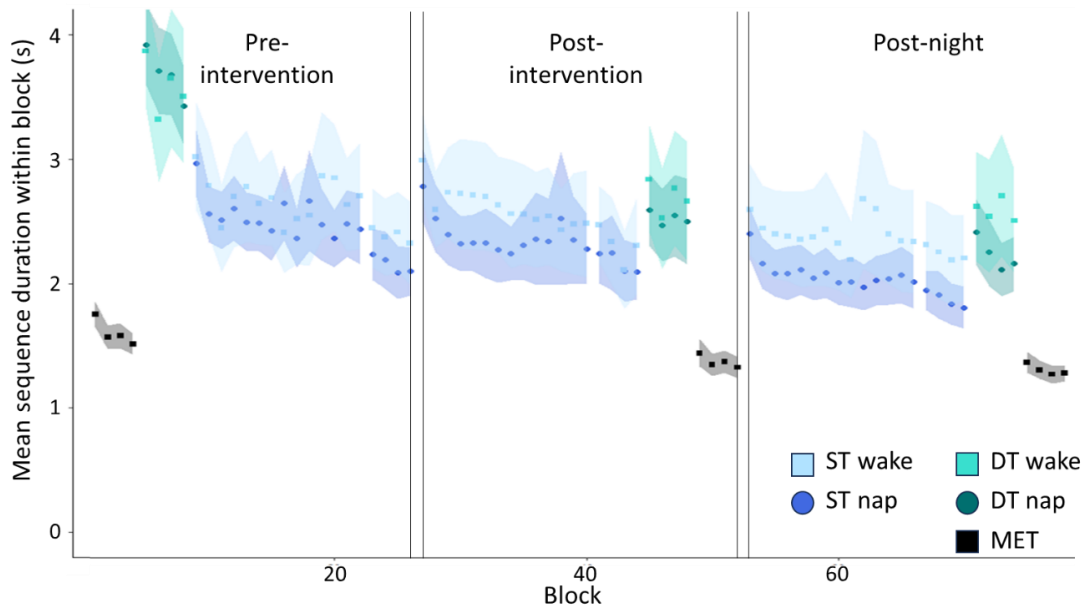
Supplementary material

3. Sequence Duration analysis

The overall performance of the PD and HOA groups across the three sessions is represented in Figure S1.

Behavioural results: Sequence duration

A. Performance PD group across sessions



B. Performance HOA group across sessions

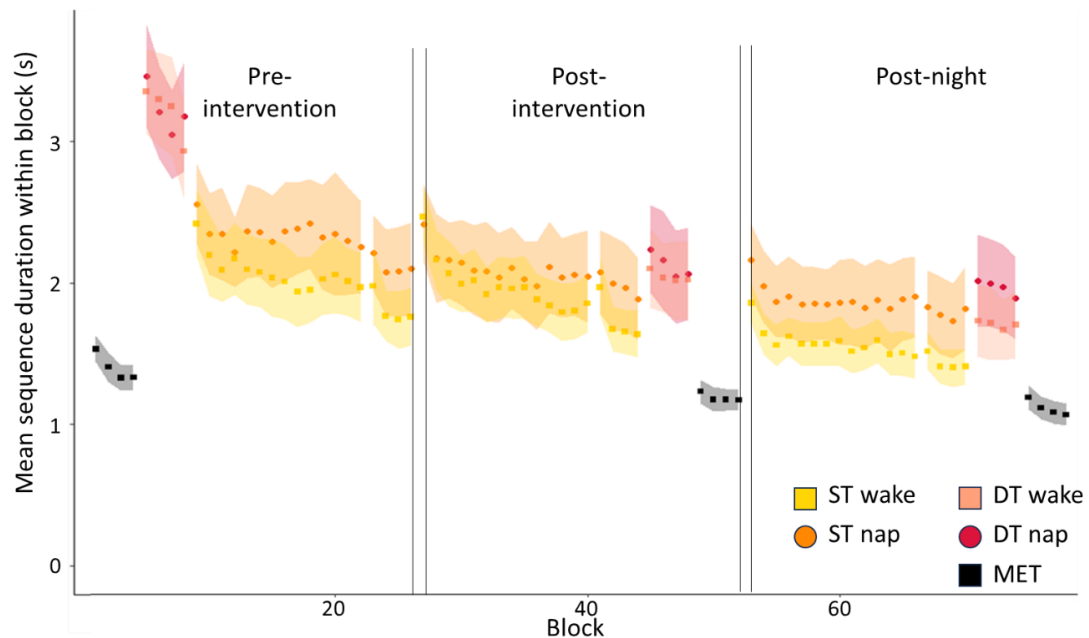


Figure S1. (A) Representation of sequence duration over the three sessions, expressed as mean and standard error for the PD group (in cold colours) and (B) in the HOA (in warm colours). MET: motor execution test; ST: single task; DT: dual-task.

Supplementary material

Pre-requisites

Plateau verification

We explored the differences in sequence duration among the four blocks of test pre-intervention with a Bayesian ANOVA on the PD and HOA group, including intervention and block as fixed effects and subject as random effect.

Here we found moderate evidence for an effect of block ($BF_{10} = 5.42$, $F_{(3,180)} = 4.54$, $p < 0.01$, partial $\eta^2 = 0.07$), and post-hoc analyses revealed decisive evidence for a difference between blocks 1 and 3 ($BF_{10} = 456.33$, $V = 1745.5$, $p < 0.01$) and moderate between blocks 1 and 4 ($BF_{10} = 3.43$, $V = 1657$, $p < 0.01$). However, this was weakly to moderately in favour of no difference for the other comparisons (block 1-2: $BF_{10} = 0.91$, $V = 1571$, $p < 0.01$; block 2-3: $BF_{10} = 0.32$, $V = 1288$, $p = 0.10$; block 2-4: $BF_{10} = 0.15$, $V = 1173$, $p = 0.38$; block 3-4: $BF_{10} = 0.19$, $V = 983$, $p = 0.71$). Weak evidence was found for the other main effects (group: $BF_{10} = 0.77$, $F_{(1,60)} = 1.15$, $p = 0.29$, partial $\eta^2 = 0.02$; intervention: $BF_{10} = 0.66$, $F_{(1,60)} = 0.01$, $p = 0.91$, partial $\eta^2 < 0.01$), and for the group by intervention interaction ($BF_{10} = 0.67$, $F_{(1,60)} = 1.04$, $p = 0.31$, partial $\eta^2 = 0.02$), while moderately in favour of no effect of the interactions group by block ($BF_{10} = 0.21$, $F_{(3,180)} = 1.45$, $p = 0.23$, partial $\eta^2 = 0.02$), intervention by block ($BF_{10} = 0.06$, $F_{(3,180)} = 0.31$, $p = 0.82$, partial $\eta^2 < 0.01$) and group by intervention by block ($BF_{10} = 0.25$, $F_{(3,180)} = 1.19$, $p = 0.31$, partial $\eta^2 = 0.02$). These findings suggest differences among blocks, but these were not linked to a specific block.

Similarly, we explored the differences among the four blocks of test post-intervention, to verify the plateau at post-intervention test, with a Bayesian ANOVA in the PD and in the HOA group, including intervention and block as fixed effects and subject as random effect. Again we found decisive evidence in favour of an effect of block ($BF_{10} = 862.03$, $F_{(3,179)} = 8.67$, $p < 0.01$, partial $\eta^2 = 0.13$). Post-hoc analyses revealed decisive evidence in favour of a difference between blocks one and three, and between blocks one and four (block 1-3: $BF_{10} = 115.94$, $V = 1631$, $p = 0.02$; block 1-4: $BF_{10} = 321.32$, $V = 1773$, $p < 0.01$), weak evidence for no difference between blocks one and two and blocks two and four ($BF_{10} = 1.37$, $V = 1522$, $p < 0.01$, block 2-4: $BF_{10} = 1.68$, $V = 1465$, $p < 0.01$) and weak to moderate evidence in favour of no difference for the other comparisons (block 2-3: $BF_{10} = 0.74$, $V = 1338.5$, $p < 0.01$; block 3-4: $BF_{10} = 0.15$, $V = 1240.5$, $p = 0.11$). For the other factors and interactions of the Bayesian ANOVA, weak to moderate evidence was found in favour of no effect (group: $BF_{10} = 0.65$, $F_{(1,60)} = 2.06$, $p = 0.16$, partial $\eta^2 = 0.03$; intervention: $BF_{10} = 0.64$, $F_{(1,60)} = 0.03$, $p = 0.86$, partial $\eta^2 < 0.01$; group by intervention: $BF_{10} = 0.83$, $F_{(1,60)} = 0.61$, $p = 0.44$, partial $\eta^2 = 0.01$; group by block: $BF_{10} = 0.11$, $F_{(1,179)} = 0.95$, $p = 0.42$, partial $\eta^2 = 0.02$; intervention by block $BF_{10} = 0.15$, $F_{(1,179)} = 1.31$, $p = 0.27$, partial $\eta^2 = 0.02$; group by intervention by block: $BF_{10} = 0.10$, $F_{(1,179)} = 0.18$, $p = 0.91$, partial $\eta^2 < 0.01$). Given that performance was overall not stable across blocks, we opted for including all four in the computation of post-night offline changes.

Motor Sequence Learning vs Motor Execution Test

Contrary to the PI and accuracy metrics, lower values of sequence duration indicate improved performance. Therefore, for the comparisons between MET and MSL, and for all the main outcomes, offline changes in sequence duration were expressed as the negative values of the relative difference in sequence duration, to allow for consistent interpretation of the results.

When comparing the sequence duration change between the beginning of pre-intervention and end of post-night of the MET and MSL sequences, we found strong evidence for an effect of task ($BF_{10} = 22.33$, $F_{(1,60)} = 9.27$, $p < 0.01$, partial $\eta^2 = 0.13$), but weak to moderate evidence for no effect of all the other factors and their interactions (group: $BF_{10} = 0.46$, $F_{(1,60)} = 2.08$, $p = 0.15$, partial $\eta^2 = 0.03$; intervention: $BF_{10} = 0.24$, $F_{(1,60)} = 0.02$, $p = 0.89$, partial $\eta^2 < 0.01$; group by intervention: $BF_{10} = 0.89$, $F_{(1,60)} = 3.57$, $p = 0.06$, partial $\eta^2 = 0.06$; group by task: $BF_{10} = 0.25$, $F_{(1,60)} = 0.005$, $p = 0.94$, partial $\eta^2 <$

Supplementary material

0.01; intervention by task: $BF_{10} = 0.28$, $F_{(1,60)} = 0.20$, $p = 0.66$, partial $\eta^2 < 0.01$; group by intervention by task: $BF_{10} = 0.58$, $F_{(1,60)} = 1.22$, $p = 0.27$, partial $\eta^2 = 0.02$). Post-hoc comparisons showed moderate evidence in favour of a difference between the changes in MSL compared to MET ($BF_{10} = 9.35$, $V = 1481$, $p < 0.01$, $r_{rb} = -0.26$, $p < 0.01$). These results suggest that the performance change in speed was not related to general motor improvement only, but to actual learning, irrespective of group and intervention.

Main analyses

Offline changes post-intervention

When comparing offline changes post-intervention measured with sequence duration in people with PD allocated to nap and to wake, we found weak evidence for no difference between interventions ($BF_{10} = 0.34$, $t_{(30)} = -0.08$, 95% CI [-10.34–9.61], $p = 0.94$, Cohen's $d = -0.03$; Figure S2A, in cold colours). In HOA, offline changes post-intervention suggested weak evidence for a difference between nap and wake interventions ($BF_{10} = 1.62$, $W = 174$, $p = 0.09$, $r_{rb} = -0.31$, $p = 0.08$; Figure S2A, in warm colours). We found weak evidence for similarity between PD and HOA allocated to nap ($BF_{10} = 0.70$, $t_{(30)} = -1.40$, 95% CI [-16.81–3.15], $p = 0.17$, Cohen's $d = 0.49$; Figure S2A), and to wake interventions ($BF_{10} = 0.68$, $W = 136$, $p = 0.78$, $r_{rb} = 0.05$, $p = 0.77$; Figure S2A).

Offline memory consolidation: sequence duration

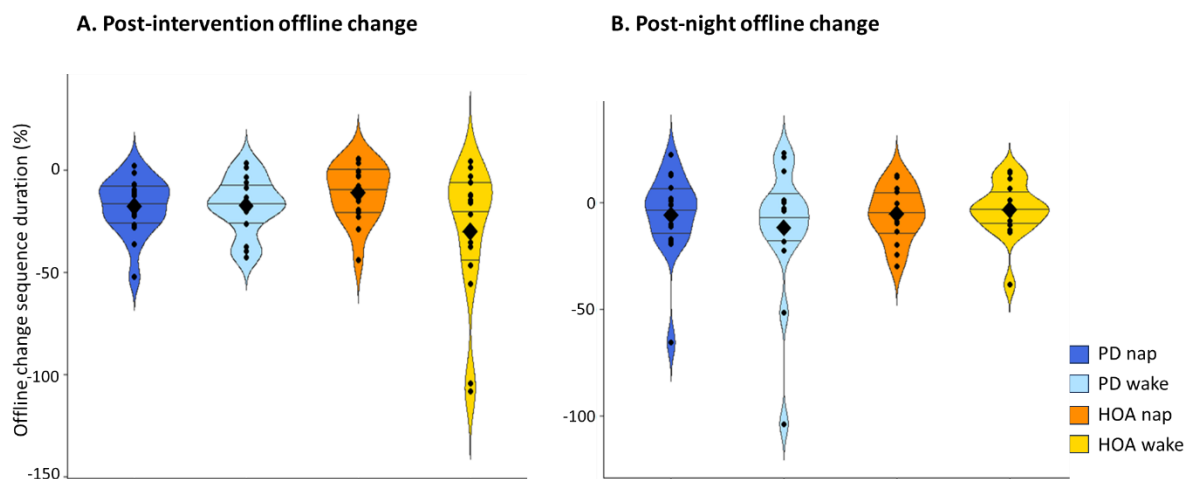


Figure S2. Offline changes of single-task sequence duration of the MSL sequence in the PD group (blue shades) and HOA (orange shades) (A) post-intervention and (B) post-night.

Violin plot: mean (diamond) median (central horizontal bar), and 25th (lower bar) and 75th (higher bar) percentiles.

Post-night offline changes

Our secondary interest was the post-night offline change of the MSL sequence. For this analysis we found weak evidence for no difference between interventions within the PD group ($BF_{10} = 0.40$, $W = 143$, $p = 0.58$, $r_{rb} = -0.10$, $p = 0.58$; Figure S2B, in cold colours) and within the HOA ($BF_{10} = 0.36$, $t_{(30)} = -0.42$, 95% CI [-11.04–7.25], $p = 0.68$, Cohen's $d = -0.15$; Figure S2B, in warm colours). Weak evidence for no difference between the people with PD and the HOA allocated to nap ($BF_{10} = 0.34$, $W = 137$, $p = 0.75$, $r_{rb} = 0.06$, $p = 0.74$; Figure S2B), as well as for those allocated to wakefulness ($BF_{10} = 0.50$, $W = 111$, $p = 0.54$, $r_{rb} = -0.12$, $p = 0.53$; Figure S2B) was also found.

Supplementary material

Dual-task cost

Dual-task cost: Sequence Duration

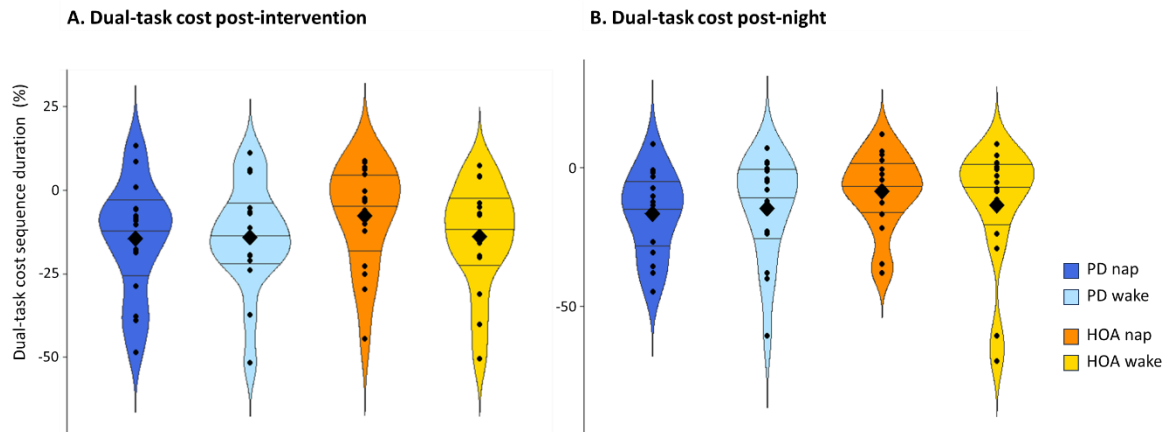


Figure S3. (A) Post-intervention dual-task cost of sequence duration performance of the MSL sequence in the PD (green shades) and HOA (red shades) groups. (B) Post-night dual-task cost of sequence duration performance of the MSL sequence in the PD and HOA groups

Violin plot: mean (diamond), median (central horizontal bar), and 25th (lower bar) and 75th (higher bar) percentiles.

The findings on automaticity measured with sequence duration suggested weak evidence for no difference in dual-task cost between people with PD allocated to nap or wake intervention, neither post-intervention ($BF_{10} = 0.34$, $t_{(30)} = -0.09$, 95% CI [-12.40–11.38], $p = 0.93$, Cohen's $d = -0.03$; Figure S3A, in cold colours) nor post-night ($BF_{10} = 0.35$, $t_{(30)} = -0.34$, 95% CI [-14.13–10.08], $p = 0.73$, Cohen's $d = -0.12$; Figure S3B, in cold colours). Similar findings were evidenced in the HOA allocated to the nap and to the wake interventions, with weak evidence for no difference at post-intervention ($BF_{10} = 0.54$, $t_{(30)} = 1.12$, 95% CI [-5.14–17.56], $p = 0.27$, Cohen's $d = -0.40$; Figure S3A, in warm colours) and at post-night ($BF_{10} = 0.42$, $W = 134$, $p = 0.84$, $r_{rb} = -0.04$, $p = 0.83$; Figure S3B, in warm colours).

Analyses of the dual-task costs at post-intervention between people with PD and HOA allocated to the nap intervention showed weak evidence for no difference between groups ($BF_{10} = 0.57$, $t_{(30)} = -1.17$, 95% CI [-18.51–5.02], $p = 0.25$, Cohen's $d = 0.41$; Figure S3A) and for the participants allocated to the wake intervention ($BF_{10} = 0.34$, $t_{(30)} = -0.004$, 95% CI [-11.50–11.46], $p = 0.99$, Cohen's $d = 0.05$; Figure S3A). At post-night, dual-task costs showed weak evidence for no difference between people with PD and HOA allocated to the nap ($BF_{10} = 0.85$, $t_{(30)} = -1.57$, 95% CI [-18.33–2.42], $p = 0.13$, Cohen's $d = 0.55$; Figure S3B), and to the wake intervention ($BF_{10} = 0.34$, $W = 113$, $p = 0.59$, $r_{rb} = -0.10$, $p = 0.58$; Figure S3B).

Extended practice post-intervention and post-night

We then tested for the effect of the nap intervention on sequence duration over extended practice post-intervention and post-night. For the analysis we used Bayesian t-tests.

When comparing people with PD allocated to the nap intervention, and those allocated to the wake intervention, we found weak evidence in favour of no difference ($BF_{10} = 0.34$, $W = 126$, $p = 0.96$, $r_{rb} = -0.01$, $p = 0.94$), and similar findings were evidenced in the HOA group ($BF_{10} = 0.40$, $W = 143$, $p = 0.59$, $r_{rb} = 0.10$, $p = 0.58$).

The analyses comparing people with PD and HOA allocated to the nap intervention and to the wake intervention showed weak evidence for no difference between the populations of study (nap: $BF_{10} = 0.40$, $W = 110$, $p = 0.51$, $r_{rb} = -0.12$, $p = 0.51$; wake: $BF_{10} = 0.84$, $W = 92$, $p = 0.18$, $r_{rb} = -0.24$, $p = 0.18$).

Supplementary material

At post-night, we found weak evidence for a difference between people with PD allocated to the two interventions ($BF_{10} = 0.42$, $W = 127$, $p = 0.99$, $r_{rb} = 0.0006$, $p = 0.97$), and similar findings were evidenced in the HOA group ($BF_{10} = 0.49$, $W = 95$, $p = 0.22$, $r_{rb} = 0.22$, $p = 0.22$).

Comparisons between people with PD and HOA also showed weak evidence in favour of no difference for the nap ($BF_{10} = 0.37$, $W = 151$, $p = 0.40$, $r_{rb} = -0.16$, $p = 0.39$) and for the wake intervention ($BF_{10} = 1.27$, $W = 178$, $p = 0.06$, $r_{rb} = -0.34$, $p = 0.06$).

In summary, the rate of learning post-intervention and post-night was not different between the populations of study, and it appears that it was not affected by the sleep intervention.

Aggregated effect across sessions

We also explored the differences among the mean performance of the first four blocks of the training of each session. Here we found a significant effect of session ($BF_{10} > 1000$, $F_{(2,120)} = 26.61$, $p < 0.01$, partial $\eta^2 = 0.31$), but weak evidence for no effect of group ($BF_{10} = 0.98$, $F_{(1,60)} = 2.77$, $p = 0.10$, partial $\eta^2 = 0.04$), intervention ($BF_{10} = 0.52$, $F_{(1,60)} < 0.01$, $p = 0.99$, partial $\eta^2 < 0.01$), and group by intervention interaction ($BF_{10} = 0.65$, $F_{(1,60)} = 0.42$, $p = 0.52$, partial $\eta^2 < 0.01$). Moderate evidence for no effect was found for the interactions of group by session ($BF_{10} = 0.13$, $F_{(2,120)} = 3.64$, $p = 0.70$, partial $\eta^2 < 0.01$), intervention by session ($BF_{10} = 0.14$, $F_{(2,120)} = 0.93$, $p = 0.40$, partial $\eta^2 = 0.02$) and group by intervention by session ($BF_{10} = 0.26$, $F_{(2,120)} = 1.36$, $p = 0.26$, partial $\eta^2 = 0.02$). Bayesian t-tests for the post-hoc analyses revealed weak evidence for a difference between pre-intervention and post-intervention ($BF_{10} = 1.41$, $V = 1337$, $p = 0.05$, $r_{rb} = 0.05$, $p = 0.60$), decisive evidence for a difference between post-intervention and post-night ($BF_{10} > 1000$, $V = 1845$, $p < 0.01$, $r_{rb} = 0.18$, $p = 0.04$), and between pre-intervention and post-night ($BF_{10} > 1000$, $V = 1831.5$, $p < 0.01$, $r_{rb} = 0.22$, $p = 0.01$).

Correlation analysis: sequence duration change and sleep micro-architecture

Correlation parameters	Correlation test, r, p-value	BF_{10}
Sequence duration– NREM2 (%) + NREM3 (%)	PD: $S = 820$, $r = -0.21$, $p = 0.44$	$BF_{10} = 0.79$
	HOA: $S = 506$, $r = 0.10$, $p = 0.73$	$BF_{10} = 0.53$
Sequence duration– Spindle density	PD: $S = 802$, $r = -0.18$, $p = 0.51$	$BF_{10} = 0.85$
	HOA: $S = 486$, $r = 0.29$, $p = 0.28$	$BF_{10} = 0.98$
Sequence duration– Spindle amplitude	PD: $S = 708$, $r = -0.04$, $p = 0.88$	$BF_{10} = 0.52$
	HOA: $S = 472$, $r = 0.31$, $p = 0.25$	$BF_{10} = 0.76$
Sequence duration– Spindle frequency	PD: $S = 754$, $r = -0.11$, $p = 0.69$	$BF_{10} = 0.54$
	HOA: $S = 1000$, $r = -0.47$, $p = 0.07$	$BF_{10} = 1.31$
Sequence duration– Slow wave density	PD: $S = 480$, $r = 0.29$, $p = 0.27$	$BF_{10} = 0.93$
	HOA: $S = 416$, $r = 0.26$, $p = 0.35$	$BF_{10} = 0.60$
Sequence duration– Slow wave amplitude	PD: $S = 500$, $r = 0.26$, $p = 0.32$	$BF_{10} = 1.29$
	HOA: $S = 342$, $r = 0.25$, $p = 0.39$	$BF_{10} = 1.30$
Sequence duration– Slow wave slope	PD: $S = 554$, $r = 0.19$, $p = 0.49$	$BF_{10} = 1.09$
	HOA: $S = 382$, $r = 0.16$, $p = 0.58$	$BF_{10} = 0.73$
Sequence duration– ndPAC	PD: $S = 660$, $r = 0.03$, $p = 0.92$	$BF_{10} = 0.53$
	HOA: $S = 586$, $r = -0.29$, $p = 0.32$	$BF_{10} = 0.65$

ndPAC = normalized direct phase-amplitude coupling

Offline changes expressed as the negative values of the relative change in sequence duration

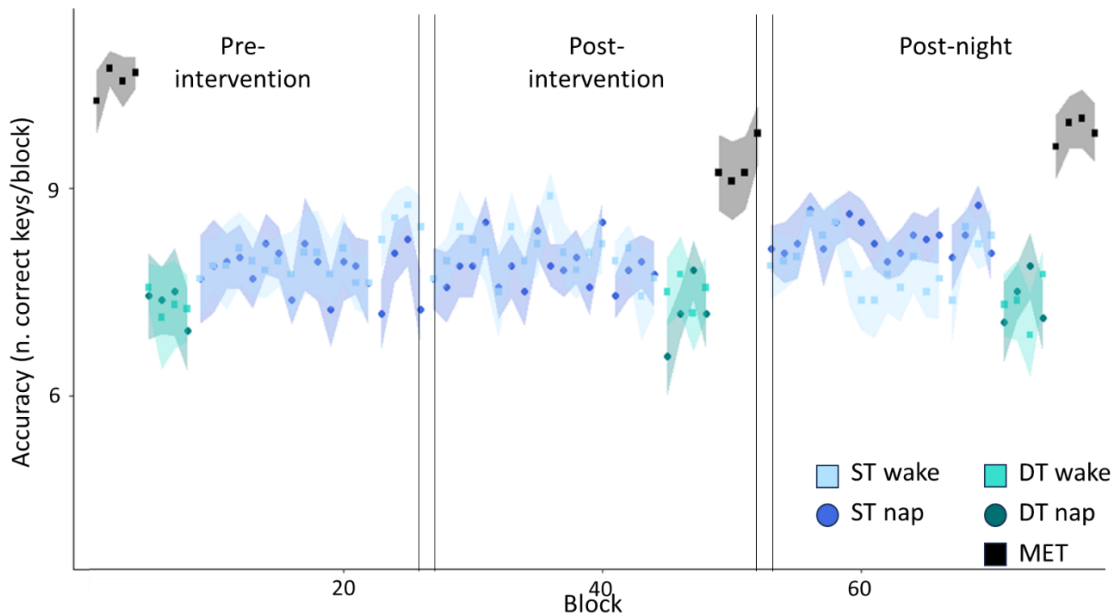
Supplementary material

4. Accuracy analysis

Figure S4 shows overall performance of the PD and HOA groups across the three sessions.

Behavioural results: Accuracy

A. Performance PD group across sessions



B. Performance HOA group across sessions

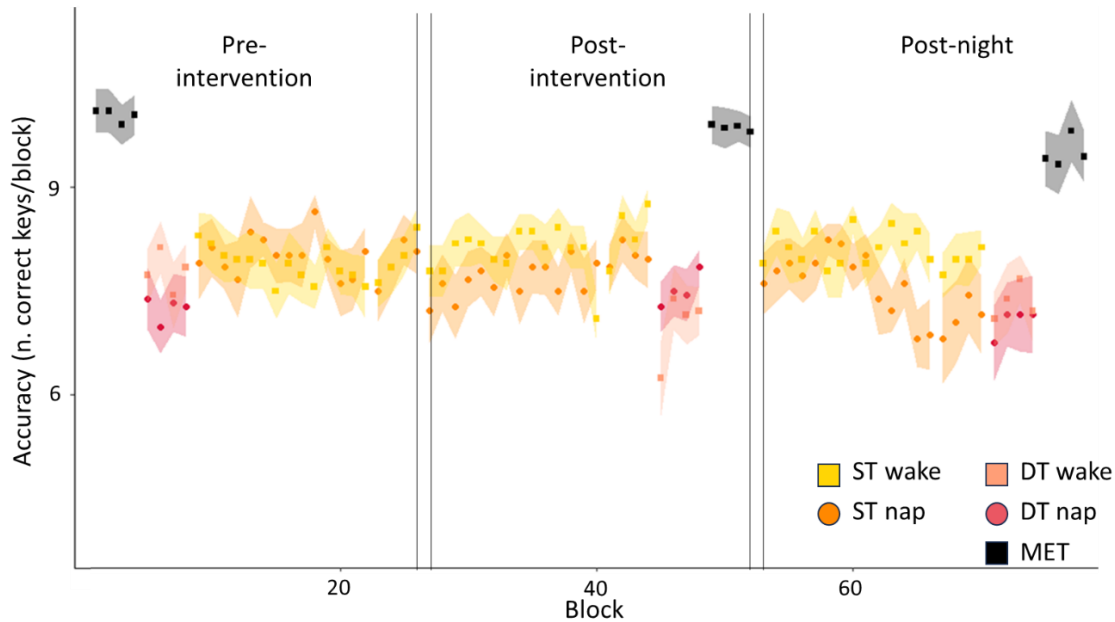


Figure S4. Representation of the non-transformed values of accuracy over the three sessions, expressed as mean and standard error (A) for the PD group (in cold colours) and (B) for the HOA (in warm colours). MET: motor execution test; ST: single task; DT: dual-task.

Supplementary material

Plateau verification

Weak to moderate evidence favouring no effect was found for the main effects (group: $BF_{10} = 0.34$, $F_{(1,60)} = 0.73$, $p = 0.40$, partial $\eta^2 = 0.01$; intervention: $BF_{10} = 0.60$, $F_{(1,60)} = 1.98$, $p = 0.17$, partial $\eta^2 = 0.03$; block: $BF_{10} = 1.93$, $F_{(3,180)} = 3.70$, $p = 0.01$, partial $\eta^2 = 0.06$), and for the interactions (group by intervention: $BF_{10} = 0.57$, $F_{(1,60)} = 1.45$, $p = 0.27$, partial $\eta^2 = 0.02$; group by block: $BF_{10} = 0.29$, $F_{(3,180)} = 1.75$, $p = 0.16$, partial $\eta^2 = 0.03$; intervention by block: $BF_{10} = 0.12$, $F_{(3,180)} = 0.99$, $p = 0.40$, partial $\eta^2 = 0.02$; group by intervention by block: $BF_{10} = 0.09$, $F_{(3,180)} = 0.09$, $p = 0.96$, partial $\eta^2 < 0.01$). These findings suggest that performance was not stable, but the differences across blocks did not follow a pattern of learning, therefore we decided to include all the blocks in the calculation of offline performance changes post-intervention, both in PD and in HOA.

With a similar analysis at post-intervention test we found decisive evidence in favour of an effect of group ($BF_{10} = 10.63$, $F_{(1,60)} = 9.02$, $p < 0.01$, partial $\eta^2 = 0.13$; post-hoc analysis: $BF_{10} = 76.92$, $W = 5682$, $p < 0.01$). For the other factors and interactions of the Bayesian ANOVA, weak to moderate evidence was found in favour of no effect (intervention: $BF_{10} = 0.34$, $F_{(1,60)} = 0.55$, $p = 0.46$, partial $\eta^2 < 0.01$; block: $BF_{10} = 0.15$, $F_{(3,179)} = 1.60$, $p = 0.19$, partial $\eta^2 = 0.03$; group by intervention: $BF_{10} = 0.33$, $F_{(1,60)} = 0.27$, $p = 0.60$, partial $\eta^2 < 0.01$; group by block: $BF_{10} = 0.09$, $F_{(1,179)} = 0.55$, $p = 0.65$, partial $\eta^2 < 0.01$; intervention by block $BF_{10} = 0.06$, $F_{(1,179)} = 0.56$, $p = 0.64$, partial $\eta^2 < 0.01$; group by intervention by block: $BF_{10} = 0.24$, $F_{(1,179)} = 1.18$, $p = 0.32$, partial $\eta^2 = 0.02$). Given the degree of evidence for the differences between blocks, we opted for including all four blocks of the test post-intervention in the computation of post-night offline changes. The evidence for a difference between groups does not represent an issue in the calculation of offline changes, as these are computed as relative differences of means.

Pre-requisites

Motor Sequence Learning vs Motor Execution Test

For the comparison between the performance of the MSL sequence and of the MET, we found overall weak to moderate evidence for an effect of all the factors and their interactions (group: $BF_{10} = 0.42$, $F_{(1,60)} = 1.31$, $p = 0.26$, partial $\eta^2 = 0.02$; intervention: $BF_{10} = 0.27$, $F_{(1,60)} = 0.16$, $p = 0.69$, partial $\eta^2 < 0.01$; group by intervention: $BF_{10} = 0.40$, $F_{(1,60)} = 0.57$, $p = 0.45$, partial $\eta^2 < 0.01$; task: $BF_{10} = 1.07$, $F_{(1,60)} = 3.84$, $p = 0.05$, partial $\eta^2 = 0.06$; group by task: $BF_{10} = 1.32$, $F_{(1,60)} = 3.80$, $p = 0.06$, partial $\eta^2 = 0.06$; intervention by task: $BF_{10} = 0.26$, $F_{(1,60)} = 0.13$, $p = 0.72$, partial $\eta^2 < 0.01$; group by intervention by task: $BF_{10} = 0.36$, $F_{(1,60)} = 0.23$, $p = 0.63$, partial $\eta^2 < 0.01$).

Supplementary material

Main analyses

Offline changes post-intervention

Weak evidence for a difference between nap and wake was found both in PD ($BF_{10} = 0.48$, $t_{(30)} = 0.96$, 95% CI [-6.78–18.84], $p = 0.34$, Cohen's $d = 0.34$; Figure S5A, in cold colours), and in HOA ($BF_{10} = 0.73$, $W = 101.5$, $p = 0.33$, $r_{rb} = 0.18$, $p = 0.33$; Figure S5A, in warm colours). When comparing people with PD and HOA on post-intervention offline changes, we found weak evidence for a difference both in the nap groups ($BF_{10} = 0.63$, $W = 159$, $p = 0.25$, $r_{rb} = 0.21$, $p = 0.25$; Figure S5A) and in the wake groups ($BF_{10} = 0.49$, $W = 112.5$, $p = 0.57$, $r_{rb} = -0.10$, $p = 0.57$; Figure S5A).

Offline memory consolidation: accuracy

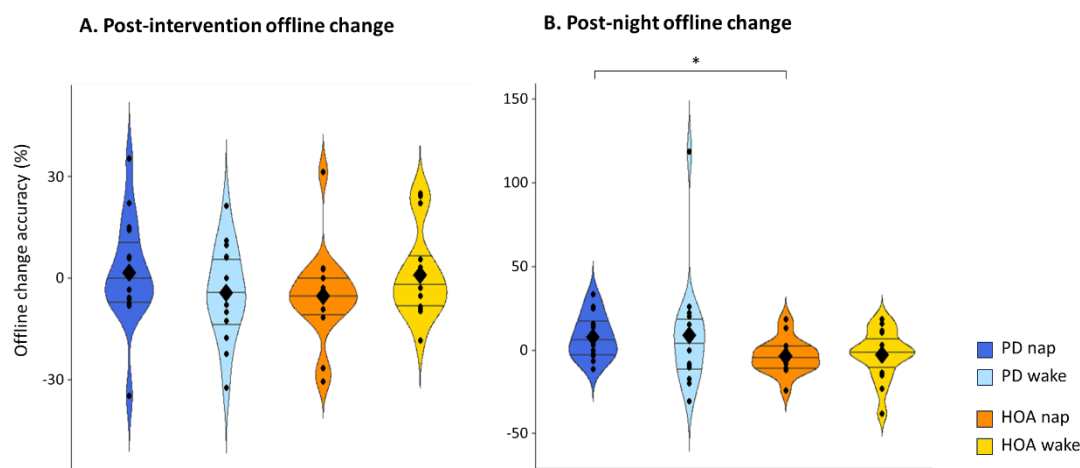


Figure S5. Offline changes of single-task accuracy of the MSL sequence in the PD group (blue shades) and HOA (orange shades) (A) post-intervention and (B) post-night.

Violin plot: mean (diamond) median (central horizontal bar), and 25th (lower bar) and 75th (higher bar) percentiles.

“*” = moderate evidence for a difference ($3 \leq BF < 10$); “***” = strong evidence for a difference ($10 \leq BF < 100$); “****” = decisive evidence for a difference (> 100)

Post-night offline changes

Weak evidence for a difference ($BF_{10} = 0.37$, $W = 140$, $p = 0.66$, $r_{rb} = -0.08$, $p = 0.66$; Figure S5B, in cold colours) was observed between nap and wake in the PD group on post-night offline changes measured with accuracy. The same comparison in the HOA group also showed weak evidence for no difference between the nap and wake interventions ($BF_{10} = 0.34$, $t_{(30)} = 0.22$, 95% CI [-7.49–9.30], $p = 0.83$, Cohen's $d = 0.08$; Figure S5B, in warm colours).

The offline change in performance showed moderate evidence for a difference between PD and HOA allocated to nap ($BF_{10} = 4.06$, $t_{(30)} = 2.63$, 95% CI [2.51–20.01], $p = 0.01$, Cohen's $d = -0.93$; Figure S5B), but it was weakly in favour of no difference in the wake groups ($BF_{10} = 0.74$, $W = 159.5$, $p = 0.24$, $r_{rb} = 0.21$, $p = 0.24$; Figure S5B).

Dual-task cost

The findings on automaticity measured with accuracy suggested weak evidence for no difference in dual-task cost between people with PD allocated to nap or wake intervention, neither at post-intervention ($BF_{10} = 0.35$, $W = 115.5$, $p = 0.65$, $r_{rb} = 0.08$, $p = 0.65$; Figure S6A, in cold colours) nor at post-night ($BF_{10} = 0.36$, $W = 125.5$, $p = 0.94$, $r_{rb} = 0.02$, $p = 0.93$; Figure S6B, in cold colours). Between the HOA allocated to the nap and to the wake interventions there was weak evidence in favour of a difference post-intervention ($BF_{10} = 2.63$, $t_{(30)} = 2.37$, 95% CI [1.50–19.96], $p = 0.02$, Cohen's $d = 0.84$; Figure S6A, in warm colours) and post-night ($BF_{10} = 0.81$, $W = 167.5$, $p = 0.14$, $r_{rb} = -0.27$, $p = 0.14$; Figure S6B, in warm colours).

Supplementary material

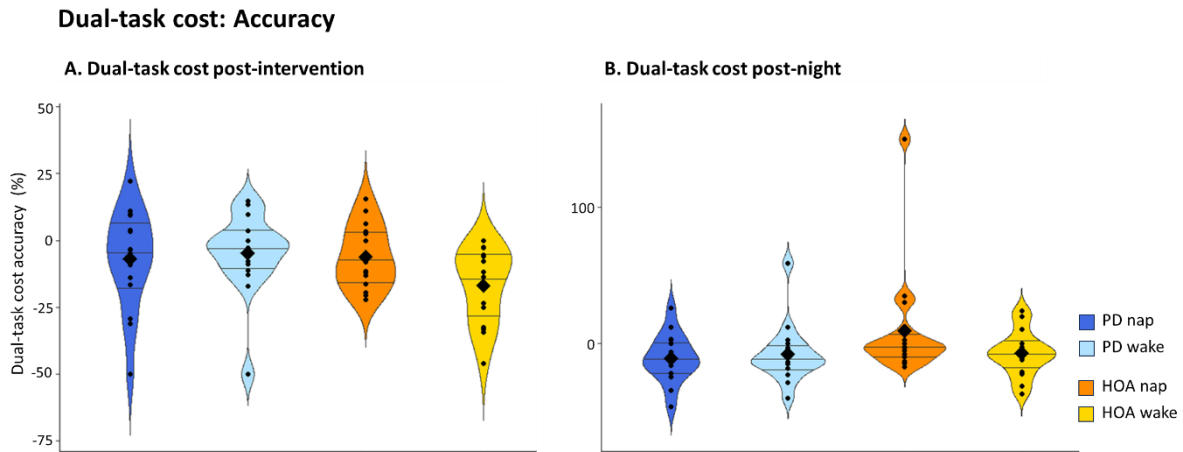


Figure S6. (A) Post-intervention dual-task cost of accuracy performance of the MSL sequence in the PD and HOA groups. (B) Post-night dual-task cost of accuracy performance of the MSL sequence in the PD and HOA groups. Violin plot: mean (diamond), median (central horizontal bar), and 25th (lower bar) and 75th (higher bar) percentiles. “*” = moderate evidence for a difference ($3 \leq BF < 10$); “**” = strong evidence for a difference ($10 \leq BF < 100$); “***” = decisive evidence for a difference (> 100)

Comparisons of the dual-task costs at post-intervention between people with PD and HOA allocated to the nap intervention showed weak evidence suggesting no difference between the two populations ($BF_{10} = 0.34$, $t_{(30)} = -0.14$, 95% CI [-11.83–10.33], $p = 0.89$, Cohen’s $d = 0.05$), but this was strongly in favour of a difference when comparing the groups allocated to the wake intervention ($BF_{10} = 2.65$, $W = 195.5$, $p = 0.01$, $r_{rb} = 0.46$, $p < 0.01$). At post-night, analyses revealed weak evidence for a difference between people with PD and HOA allocated to the nap ($BF_{10} = 1.17$, $W = 77.5$, $p = 0.06$, $r_{rb} = -0.28$, $p = 0.11$), and those allocated to the wake intervention ($BF_{10} = 0.34$, $W = 110$, $p = 0.51$, $r_{rb} = -0.12$, $p = 0.51$).

Extended practice post-intervention and post-night

When comparing people with PD allocated to the nap intervention, and those allocated to the wake intervention, we found weak evidence in favour of no difference ($BF_{10} = 0.34$, $W = 113.5$, $p = 0.60$, $r_{rb} = 0.10$, $p = 0.59$), and similar findings were evidenced in the HOA group ($BF_{10} = 0.38$, $W = 117$, $p = 0.69$, $r_{rb} = 0.07$, $p = 0.68$).

The analyses comparing people with PD and HOA allocated to the nap intervention and to the wake intervention showed weak evidence for no difference between the populations of study (nap: $BF_{10} = 0.36$, $t_{(30)} = 0.44$, 95% CI [-0.62–0.96], $p = 0.51$, Cohen’s $d = 0.16$; wake: $BF_{10} = 0.42$, $W = 145$, $p = 0.53$, $r_{rb} = -0.12$, $p = 0.53$).

At post-night, we found weak evidence for a difference between people with PD allocated to the two interventions ($BF_{10} = 0.55$, $W = 149.5$, $p = 0.43$, $r_{rb} = -0.15$, $p = 0.43$), while for the HOA we found moderate evidence for a difference between nap and wake interventions ($BF_{10} = 4.32$, $t_{(30)} = -2.66$, 95% CI [-2.15– -0.28], $p = 0.01$, Cohen’s $d = -0.94$). Comparing people with PD and HOA, we found weak evidence for those allocated to the nap intervention ($BF_{10} = 1.09$, $t_{(30)} = -1.77$, 95% CI [-1.65–0.11], $p = 0.09$, Cohen’s $d = -0.63$), and moderate evidence for no difference for the participants allocated to the wake intervention ($BF_{10} = 0.04$, $W = 182.5$, $p = 0.04$, $r_{rb} = -0.37$, $p = 0.04$).

Hence, accuracy performance change rate during training seemed not to be different between people with PD and HOA, but it appears that a 2-hour nap may have disrupted post-night performance in the HOA.

Supplementary material

Aggregated effect across sessions

Analysing the differences among the mean accuracy of the first four blocks of the training of each session, weak to moderate evidence for no effect was found for all the factors and their interactions (group: $BF_{10} = 0.39$, $F_{(1,60)} = 1.04$, $p = 0.31$, partial $\eta^2 = 0.02$; intervention: $BF_{10} = 0.47$, $F_{(1,60)} = 0.65$, $p = 0.42$, partial $\eta^2 = 0.01$; session: $BF_{10} = 0.42$, $F_{(2,120)} = 1.02$, $p = 0.36$, partial $\eta^2 = 0.02$; group by intervention interaction: $BF_{10} = 0.44$, $F_{(1,60)} = 0.35$, $p = 0.56$, partial $\eta^2 < 0.01$; group by session interaction: $BF_{10} = 0.09$, $F_{(2,120)} = 0.75$, $p = 0.48$, partial $\eta^2 = 0.01$; intervention by session interaction: $BF_{10} = 0.18$, $F_{(2,120)} = 0.70$, $p = 0.50$, partial $\eta^2 = 0.01$; group by intervention by session interaction: $BF_{10} = 0.21$, $F_{(2,120)} = 0.10$, $p = 0.91$, partial $\eta^2 < 0.01$).

Correlation analysis: accuracy change and sleep micro-architecture

Correlation parameters	Correlation test, r, p-value	BF_{10}
Accuracy– NREM2 (%) + NREM3 (%)	PD: $S = 650.98$, $r = 0.04$, $p = 0.88$ HOA: $t_{(13)} = 2.26$, $r = 0.53$, $p = 0.04$	$BF_{10} = 0.55$ $BF_{10} = 2.44$
Accuracy– Spindle density	PD: $S = 583.93$, $r = 0.14$, $p = 0.60$ HOA: $t_{(14)} = -0.32$, $r = -0.09$, $p = 0.75$	$BF_{10} = 0.54$ $BF_{10} = 0.54$
Accuracy– Spindle amplitude	PD: $S = 330.74$, $r = 0.51$, $p = 0.04$ HOA: $t_{(14)} = 0.51$, $r = 0.13$, $p = 0.62$	$BF_{10} = 2.38$ $BF_{10} = 0.57$
Accuracy– Spindle frequency	PD: $S = 772.07$, $r = -0.14$, $p = 0.62$ HOA: $t_{(14)} = -2.32$, $r = -0.53$, $p = 0.04$	$BF_{10} = 0.92$ $BF_{10} = 2.66$
Accuracy– Slow wave density	PD: $S = 504.87$, $r = 0.26$, $p = 0.34$ HOA: $t_{(13)} = 1.63$, $r = 0.41$, $p = 0.13$	$BF_{10} = 49.24$ $BF_{10} = 1.26$
Accuracy– Slow wave amplitude	PD: $S = 380.78$, $r = 0.44$, $p = 0.09$ HOA: $S = 459$, $r = -0.009$, $p = 0.98$	$BF_{10} = 1.81$ $BF_{10} = 0.55$
Accuracy– Slow wave slope	PD: $S = 446.83$, $r = 0.34$, $p = 0.19$ HOA: $t_{(12)} = -0.28$, $r = -0.08$, $p = 0.79$	$BF_{10} = 1.29$ $BF_{10} = 0.56$
Accuracy– ndPAC	PD: $S = 1087.3$, $r = -0.60$, $p = 0.01$ HOA: $t_{(12)} = 0.07$, $r = 0.07$, $p = 0.81$	$BF_{10} = 2.50$ $BF_{10} = 0.56$

ndPAC = normalized direct phase-amplitude coupling

Supplementary material

5. Correlations of behavioural outcomes with levodopa equivalent daily dose

To test whether levodopa had an influence on the post-intervention and post-night offline changes, and dual-task costs, we performed a correlation analysis between the levodopa equivalent daily dose and each of these behavioural performance measures, using a Bayesian correlation test. For these control analyses we found weak evidence for a correlation between LEDD and performance measured with PI changes (post-intervention nap: $BF_{10} = 1.36$, $t_{(14)} = 1.71$, $r = 0.42$, $p = 0.11$; post-intervention wake: $BF_{10} = 0.52$, $t_{(14)} = 0.09$, $r = 0.03$, $p = 0.93$; post-night nap: $BF_{10} = 0.52$, $S = 534.57$, $r = 0.21$, $p = 0.43$; post-night wake: $BF_{10} = 1.03$, $t_{(13)} = -1.41$, $r = -0.37$, $p = 0.18$; post-intervention dual-task cost nap: $BF_{10} = 0.61$, $t_{(14)} = 0.67$, $r = 0.18$, $p = 0.51$; post-intervention dual-task cost wake: $BF_{10} = 0.59$, $t_{(13)} = 0.54$, $r = 0.15$, $p = 0.60$; post-night dual-task cost nap: $BF_{10} = 0.56$, $t_{(14)} = 0.43$, $r = 0.12$, $p = 0.67$; post-night dual-task cost wake: $BF_{10} = 0.57$, $t_{(14)} = -0.51$, $r = -0.13$, $p = 0.62$).

Similarly, we found weak evidence for no correlation with sequence duration (post-intervention offline changes nap: $BF_{10} = 0.63$, $S = 768.36$, $r = -0.13$, $p = 0.63$; post-intervention offline changes wake: $BF_{10} = 0.52$, $t_{(14)} = -0.10$, $r = -0.03$, $p = 0.93$; post-night offline changes nap: $BF_{10} = 0.67$, $S = 801.36$, $r = -0.18$, $p = 0.51$; post-night offline changes wake: $BF_{10} = 0.90$, $S = 454.81$, $r = 0.19$, $p = 0.50$; post-intervention dual-task cost nap: $BF_{10} = 0.65$, $S = 620.83$, $r = 0.09$, $p = 0.75$; post-intervention dual-task cost wake: $BF_{10} = 0.64$, $t_{(13)} = 0.71$, $r = 0.19$, $p = 0.49$; post-night dual-task cost nap: $BF_{10} = 0.60$, $S = 608.79$, $r = 0.10$, $p = 0.70$; post-night dual-task cost wake: $BF_{10} = 0.65$, $S = 637.94$, $r = 0.06$, $p = 0.82$).

For correlations with accuracy we found weak evidence for a correlation with offline changes at any time point (post-intervention offline changes nap: $BF_{10} = 0.88$, $S = 499.94$, $r = 0.26$, $p = 0.32$; post-intervention offline changes wake: $BF_{10} = 1.24$, $t_{(14)} = 1.62$, $r = 0.40$, $p = 0.13$; post-night offline changes nap: $BF_{10} = 0.59$, $S = 499.47$, $r = 0.27$, $p = 0.32$; post-night offline changes wake: $BF_{10} = 0.63$, $S = 690.01$, $r = -0.01$, $p = 0.96$; post-intervention dual-task cost nap: $BF_{10} = 0.64$, $S = 542.6$, $r = 0.20$, $p = 0.45$; post-intervention dual-task cost wake: $BF_{10} = 0.60$, $S = 65.98$, $r = 0.006$, $p = 0.98$; post-night dual-task cost nap: $BF_{10} = 0.53$, $S = 780.95$, $r = -0.15$, $p = 0.58$; post-night dual-task cost wake: $BF_{10} = 0.73$, $S = 780.72$, $r = -0.15$, $p = 0.58$).

Given that the evidence was weak overall, we decided not to account for LEDD in our analyses.

6. Additional findings on electrophysiological markers of plasticity during sleep and demographics

In the absence of a group effect, we next explored the effect of age and gender on sleep micro-architecture. Decisive evidence in favour of an effect of age was found for slow wave amplitude ($BF_{10} = 964.25$, Spearman $S = 7220.6$, $r = -0.47$, $p < 0.01$) and for slow wave slope ($BF_{10} = 155.71$, Spearman $S = 7497.1$, $r = -0.51$, $p < 0.01$), suggesting a decrease with age. Strong evidence for an effect of gender was found for spindle amplitude ($BF_{10} = 14.76$, $t_{(24)} = -3.45$, 95% CI [-17.60–4.45], $p < 0.01$, Cohen's $d = 1.24$), with females showing greater spindle amplitude, but weak evidence of an effect of gender was found for slow wave density ($BF_{10} = 1.76$). No effects of age or gender were found for the phase-amplitude coupling ($BF_{10} > 1$). Finally, no effect of AHI on sleep micro-architecture metrics was found during the experimental nap (spindle amplitude: $BF_{10} = 0.57$, spindle frequency: $BF_{10} = 0.34$, spindle density: $BF_{10} = 0.69$, slow wave amplitude: $BF_{10} = 0.37$, slow wave slope: $BF_{10} = 0.37$, slow wave density: $BF_{10} = 0.68$, ndPAC: $BF_{10} = 0.40$).