The relationship between sleep and cognition in Parkinson’s disease: A meta-analysis

Running head: Sleep and cognition in Parkinson’s disease

Pushpanathan, Maria E.¹,⁴
Loftus, Andrea, M.²,⁴
Thomas, Meghan, G.³,⁴,⁵
Gasson, Natalie,²,⁴ &
Bucks, Romola, S.¹,⁴

1. School of Psychology, University of Western Australia
2. School of Psychology and Speech Pathology, Curtin University
3. Experimental and Regenerative Neuroscience, School of Animal Biology, University of Western Australia
4. ParkC Collaborative
5. Parkinson’s Centre, Vario Health Institute, Edith Cowan University, Australia

The authors have no conflicts of interest to declare

Correspondence to: Romola S. Bucks
M304, School of Psychology
The University of Western Australia
35 Stirling Highway
Crawley 6009,
Western Australia

Tel: +61 8 6488 3232
Fax: +61 8 6488 1006
Email: romola.bucks@uwa.edu.au

Abstract: 195 words
Article: 7846 words
SUMMARY

It is well established that sleep disorders have neuropsychological consequences in otherwise healthy people. Studies of night-time sleep problems and cognition in Parkinson’s disease (PD), however, paint a mixed picture, with many reporting no relationship between sleep problems and neuropsychological performance. This review aims to meta-analyse this research and to examine the factors underlying these mixed results. A literature search was conducted of published and unpublished studies, resulting in 16 papers that met inclusion criteria. Data were analysed in the domains of: global cognitive function; memory (general, long-term verbal recognition, long-term verbal recall); and executive function (general, shifting, updating, inhibition, generativity, fluid reasoning).

There was a significant effect of sleep on global cognitive function, long-term verbal recall, long-term verbal recognition, shifting, updating, generativity, and fluid reasoning.

Although there are effects on memory and executive function associated with poor sleep in PD, the effects were driven by a small number of studies. Numerous methodological issues were identified. Further studies are needed to reliably determine whether disturbed sleep impacts on cognition via mechanisms of hypoxia, hypercapnia, sleep fragmentation, chronic sleep debt or decreased REM and/or slow wave sleep in PD, as this may have important clinical implications.

Keywords: Parkinson’s disease; sleep disorder; REM sleep behaviour disorder; cognition; neuropsychology; executive function; memory; review
Glossary

Generativity - speed and efficiency of access to long-term memory
Inhibition - the ability to override prepotent or automatic responses
Updating - updating and monitoring of working memory representations
Fluid Reasoning - concept formation and novel problem solving tasks
Shifting - switching back and forth between different tasks or mental sets
Type 2 errors - incorrectly concluding that there is no relationship (false negative)

Abbreviations

EF - executive function
ESS - Epworth sleepiness scale
FAB - frontal assessment battery
FTD - frontotemporal dementia
H&Y - Hoehn and Yahr score
ISCS - inappropriate sleep composite score
MCI - mild cognitive impairment
MMSE - mini-mental state examination
MSA - multiple system atrophy
OSA - obstructive sleep apnoea
PD - Parkinson’s disease
PDSS - Parkinson’s disease sleep scale
PSG - polysomnography
PSP - progressive supranuclear palsy
PSQI - Pittsburgh sleep quality inventory
RBD - rapid eye movement sleep behaviour disorder
REM - rapid eye movement
RLS - restless legs syndrome
SCOPA-cog - scales for outcomes in Parkinson’s disease cognitive scale
SCOPA-sleep - scales for outcomes in Parkinson’s disease sleep scale
SRBD - sleep related breathing disorders
SSS - Stanford sleepiness scale
SWS- slow wave sleep

UPDRS- united Parkinson’s disease rating scale
INTRODUCTION

Parkinson’s disease (PD) is classified as a movement disorder, but non-motor symptoms are common and have a profound effect on patient experience and quality of life. [1] Up to a decade before the first motor symptoms emerge, patients often experience significant sleep disruption. [2] Sleep quality is strongly correlated with health related quality of life [3] and the impact of PD treatment (levodopa and dopamine agonists) on the sleep-wake cycle is an area of debate, with treatment effects varying between patients and compounds. [4,5]

Sleep disorders affect up to 98% of PD patients. [6] A range of sleep disturbances are common in PD, including insomnia, sleep fragmentation, sleep-related breathing disorders (SRBD), hallucinations, nightmares, narcolepsy, REM sleep behaviour disorder (RBD) and non-REM parasomnias. [7] Similar to other neurodegenerative diseases (e.g. Alzheimer’s disease), [8] significant sleep problems occur in PD, alongside significant cognitive dysfunction. This is noteworthy as cognitive impairment greatly increases disease burden. [9] Moreover, sleep problems have been shown to contribute to neuropsychological deficits in otherwise healthy people [10–12] The cognitive changes observed in early to moderate PD are primarily deficits in executive function (EF) and memory. [13–15] These are the same domains affected in those with SRBD [16] and in insomnia. [17] SRBD causes sleep fragmentation and hypoxia, leading to daytime tiredness and cognitive deficits. [18] Daytime tiredness has an effect on attention and motivation, affecting neuropsychological test performance, [10] and ability to manage daytime activities. [19] Diminished slow wave sleep (SWS) is thought to interfere with the capacity to learn new information [20] and with specific deficits in spatial learning. [21]

Given that sleep problems are almost universal in PD, and that they are associated with cognitive impairment even when experienced in isolation, this suggests that that the
neuropsychological deficits in mild to moderate PD may be compounded by chronic sleep
debt or consistently fragmented sleep. That is, cognitive impairment may not only arise
directly through the pathology of PD, but also indirectly via the mechanism of chronic sleep
disruption. Studies of night-time sleep problems and cognition in PD, however, paint a mixed
picture, with many reporting no relationship between sleep problems and neuropsychological
performance. Critical appraisal of this literature, taking account of sample size and
methodology is needed, as no such synthesis has been conducted.

This meta-analysis systematically examined the relationship between sleep and
cognition in PD. The objectives of this study were: 1) to identify whether there are specific
cognitive deficits associated with sleep problems (insomnia, sleep fragmentation, SRBD,
hallucinations, nightmares, narcolepsy, RBD, and non-REM parasomnias) in PD and 2)
which neuropsychological tests are sensitive to sleep-associated cognitive impairment in PD.
We analysed, separately, measures of global cognitive function, EF and memory.
Additionally, we analysed the sub-domains of memory: long-term verbal recall and long-term
verbal recognition and the sub-domains of EF: shifting, updating, inhibition, generativity and
fluid reasoning.

METHOD

Search strategy

Electronic search of Medline, PsychInfo, PubMed, Proquest: Theses and Dissertations and
Web of Science: Conference Proceedings databases to 19/12/2013 was conducted restricted
to papers in English, supplemented by hand searches of reference lists from included and
seminal papers.
Figure 1 list search terms which produced a total of 2283 papers. Following exclusion of duplicates and irrelevant reports, judging by title and abstract screening, 43 papers were retained for full-text evaluation. Two reviewers (M.E.P., R.S.B.) independently evaluated all papers retained for full-text screening. Studies were evaluated by a priori inclusion criteria (described below). Disagreements, of which there were very few, were resolved through discussion.

Insert Figure 1

Study Eligibility Criteria

Inclusion criteria:

1) Participants must be adults diagnosed with idiopathic PD
2) Sleep must be measured in a reliable manner: PSG, validated questionnaire, clinical interview, actigraphy (or a combination of these)
3) Cognition must be measured using validated neuropsychological tests
4) Relationship between sleep and cognition must be reported statistically
5) Samples must be independent (for prospective studies, we used baseline data, when multiple studies were published by the same authors, we confirmed independence of samples with the authors or used the study with the largest N).

Exclusion Criteria

1) Atypical PD or Parkinsonian syndromes
2) Sleep measured by a single item score or sub-scale from non-motor symptom scale
3) Cognition measured by subjective report.

Outcomes
For each study, the primary outcome was neuropsychological test scores. For two studies [22,23] the outcome was the proportion of each group (RBD+, RBD-) that fell below pre-specified cut-off scores on the mini-mental state examination (MMSE) and the frontal assessment battery (FAB).

We categorised the neuropsychological tests described in each paper according to memory theory employed in Wallace and Bucks [24] and EF was divided according to the multi-dimensional approach used by Olaithe and Bucks. [25]

Detailed description of the neuropsychological domains is provided in Table S2. The domains were: global cognitive function, memory, executive function, visuospatial/constructional ability, and psychomotor ability. Memory was further subdivided into global memory performance, short-term verbal memory, short-term spatial memory, long-term verbal recall, and long-term verbal recognition. EF was divided into: global EF, shifting, updating, inhibition, generativity, and fluid reasoning.

Cognitive domains assessed in just one study (complex attention; long-term visual recall; short-term spatial memory), could not be included in the analysis.

Data Extraction and Coding

Data extracted and coded from the final articles included: authors, publication status, year of publication, journal, study design, source of participants (clinic or community), inclusion and exclusion criteria, mean age, mean disease duration, sleep measurement, neuropsychological tests, and covariates used.

Two of the studies [26,27] examined visual hallucinations in addition to RBD dividing participants into 4 or 3 groups, respectively. Participants were categorised both by their endorsement of RBD symptoms and their experience of visual hallucinations. As visual hallucinations were not relevant to this review, groups were collapsed into RBD+ and RBD-.
by calculating the pooled means and standard deviations from statistics provided in each paper. The effect sizes reported for these studies are for between-groups differences for these groups.

There were anomalies in the published data in two studies. [26,28] We found a significant difference in the Meral et al. [26] Stroop data, in the opposite direction to that expected (RBD+ better than RBD-). This pattern is inconsistent both with previously published literature and the other cognitive test results published in Meral’s paper. We were unable to make contact with the authors, so excluded these Stroop data from the inhibition analysis.

Likewise, concerns about the z-score data from Nardone et al. [28] led us to exclude them. The study examined people with concomitant PD and RBD using a neuropsychological battery which assessed cognition in a number of domains. Aside from the MMSE and the dementia rating scale, all results were reported in the form of z-scores. No information was provided about which normative data were used or how z-scores were calculated. Again, we were unsuccessful in contacting the authors.

The majority of studies examined the effect of RBD status on cognitive performance. The exceptions to this were Cochen de Cock et al. (2010), [29] who examined obstructive sleep apnoea (OSA); Verbaan, van Rooden, van Hilten and Rjisman (2010) [30] who compared those with and without restless legs syndrome (RLS); Stavitsky, Neargarder, Bogdanova, McNamara & Cronin-Golomb (2012) [31] who looked at overall sleep quality using actigraphy, and Goldman et al., (2013) [32] who explored overall sleep quality using the Pittsburgh sleep quality index. Naismith, Terpening, Shine and Lewis, (2011) [33] examined the neuropsychological performance of both ‘good’ and ‘poor’ sleepers in terms of overall sleep quality and RBD status. For this study, we included data pertaining to overall...
sleep quality, as both sleep quality and RBD status were gauged by questionnaire response. We suggest that PD patients are able more accurately to report their overall sleep quality than their RBD status. Sensitivity to detect RBD in patients with PD using self-report is only 33% [34] as RBD in PD manifests in a less dramatic fashion than in idiopathic RBD. As RBD is less obvious in PD and may remain undetected by both the patient and their bed-partner, diagnostic classification based on self-report measures is unlikely to reflect actual RBD status and may produce a confused picture of group differences.

Data Analysis

Our initial objective was to meta-analyse these data to determine whether specific, dissociable sleep problems in PD are associated with distinct types of neuropsychological deficits. However, examination of the data extracted revealed that this approach would not be possible given the heterogeneity in the 16 studies found. We therefore operationalised sleep problems by comparing PD patients who reported good sleep (or no symptoms of RBD or RLS) with those who reported poor sleep (or symptoms of RBD or RLS). Comprehensive Meta-Analysis version 2.2.064 was used to synthesise data, calculate effect sizes, run moderator and post hoc analyses and to create forest plots. A random effects model was used to calculate effect sizes. Effect sizes are displayed in Hedges’ g.

RESULTS

Description of Studies

From the 2283 articles initially identified, 16 studies met inclusion criteria. Study details, including methodology and methodological issues, can be found in Table S1.

Across these 16 studies, there were 1882 people with PD. The average age was 66.28±8.65 years, and the average time since diagnosis was 7.17±6.04 years. All studies were conducted in a clinical setting. Mean united Parkinson’s disease rating scale (UPDRS)
scores were provided for 10 studies, while either mean or median Hoehn and Yahr (H&Y) scores were reported in 8. One study did not provide either measure of motor disease stage for the entire sample (mean H&Y was provided only for RBD subgroup). [35] Three studies accepted only people in the early stages of PD (disease duration < 5 years), [23,36,37] with two of these also requiring medication naïve participants. [36,37]

PSG was used in six studies. One study [31] used actigraphy to quantify markers of sleep quality while the remainder used sleep questionnaires, clinical interviews or both to identify sleep disorder for the purposes of categorising participants or measure sleep quality. Table S3 details the sleep measures used in each included study.

Nine of the studies assessed cognition in a manner that partially met the criteria for a level II neurological assessment capable of detecting mild cognitive impairment (MCI) in PD as defined by the movement disorders society. [38] A level II neuropsychological assessment must contain at least 2 tests in each of 5 domains; i) attention and working memory ii) EF iii) language iv) memory and v) visuospatial. The remaining seven used assessments of global cognition, which are classed as a level I (abbreviated) neuropsychological assessment (e.g. SCOPA-COG, Mattis dementia rating scale) or used the MMSE or FAB which were not recommended for use in PD.

Effect Sizes

The average effect size estimates for each cognitive domain are displayed in Table 1. See Figures 2 to 13 for forest plots by domain. Significant, overall effects were found in global cognitive function, long-term verbal recall, long-term verbal recognition, shifting, updating, generativity, fluid reasoning and visuospatial/constructional ability. These were all moderate, positive effects with the exception of global cognitive function, which generated a small
positive effect \( (g = 0.33) \) and long-term verbal recognition, which approached a large positive effect \( (g = 0.78) \).

\[
\begin{align*}
\text{Insert Forest Plots}
\end{align*}
\]

**Heterogeneity**

Heterogeneity was inspected visually using forest plots and Cochrane’s \( Q \), which indicates whether the observed variability in effect sizes is greater than would be expected by chance. The \( Q \) statistic was significant for global cognitive function, inhibition, generativity, fluid reasoning, and visuospatial/constructional ability. \( \hat{I}^2 \) values ranged from 65.73 to 90.25, suggesting that at least 65.73% of the variance was generated by real differences between studies and may be explained by covariates. Moderator analysis was used to explore this unexplained variance.

**Moderator Analysis**

According to Borenstein et al. (2009), \cite{39} meta-regression should only be conducted for outcomes in which there are at least 10 samples to 1 covariate. Global cognitive functioning was the only domain to meet this criterion. Average disease severity for good and poor sleepers (UPDRS or H & Y score) was not reported consistently across studies (H & Y score was reported in 8 of the 16 studies, UPDRS was reported in 10 studies, and no measure of motor symptom severity by group was reported in 3 studies), thus motor symptom severity could not be explored as a covariate. Paired samples t-tests revealed no differences between ‘good’ and ‘poor’ sleepers in terms of disease duration (\( t(11)=-1.65, \ p=.13 \)) or levodopa equivalent dose (\( \text{LED}; \ t(11)=-1.34, \ p=.19 \)). However, ‘poor’ sleepers were significantly older than good sleepers (\( t(12)=-2.90, \ p=0.13 \)).

Given that there was a significant difference in age between the 2 groups, we calculated a difference score (Age of good sleepers (years) - Age of poor sleepers (years)) for
each study that reported this information, and used this as a moderator variable in meta-
regression. Age-difference between good and poor sleepers was not a moderator of the effect
of poor-sleep on general cognition (Q= 0.02, df=1, p=.90), nor was mean age (standardised
mean of good and poor sleepers within each study) (Q=0.01, df=1, p=.93).

As there were no sleep-group differences in disease duration or LED, we did not
examine difference scores for these variables, but examined average disease duration and
LED by study as moderators of sleep-group differences in general cognition. Neither disease
duration (Q=0.68, df=1, p=.41) nor LED (Q=0.03, df=1, p=.87) moderated the effects.

Categorical moderator analysis was conducted to examine whether test type (MMSE
or other) impacted on the effects found for general cognitive functioning. There were no
differences between good and poor sleepers in MMSE scores (g= 0.22, p = .05). However,
differences remained on other tests of global cognitive ability (SCOPA-Cog, Mattis dementia
rating scale, short test of mental status) (g=0.47, p< .05), for which heterogeneity was not a
problem (Q= 5.85,df = 2, p> .05).

Publication Bias

Funnel plots and publication bias statistics were examined in all cognitive domains in which
there was a significant effect. Generativity was the only domain which generated a
symmetrical funnel plot and robust statistics, indicating little risk of publication bias. In all
other domains, Orwin’s fail-safe N and Duval and Tweedie’s trim and fill indicated that
effect sizes had been inflated by publication bias.

While it appears that publication bias may be inflating effect sizes in some domains,
these statistics are prone to distortion in small, heterogeneous samples, [40] and it is difficult
to tease out whether publication bias is a systemic issue in this field or whether it is an
artefact of sample characteristics. [41] Although the effects were driven by a small number of
studies that used sensitive tests, the fact that two-thirds of included studies concluded no
effect of sleep on cognition in PD indicates that non-significant results are being accepted for
publication.

DISCUSSION

This meta-analysis indicates that, relative to those without sleep problems, people
with PD who experience sleep problems demonstrate poorer cognitive performance. Poor
sleepers had poorer performance on tests of global cognitive ability such as the Mattis
dementia rating scale, and the SCOPA-Cog. When cognition was measured using brief or
compound measures (such as the MMSE to measure global cognitive ability, the FAB to
measure EF and memory or EF summary scores), no significant effects of sleep on cognition
were evident. Conversely, targeted neuropsychological tests revealed significant effects for
all sub-domains (inhibition excepted), including: long-term verbal recall, long-term verbal
recognition, shifting, updating, generativity and fluid reasoning.

Methodological Problems

Sleep Measurement

A fundamental problem with most papers was the reliance on self-report measures for
assessing sleep. While questionnaires are a pragmatic measurement method, they have
significant limitations. Sleep problems are one of the earliest manifestations of PD. [2]
Symptoms begin so early that patients may not even realise that their sleep symptoms are
attributable to PD. [1,42] People also habituate to sleep symptoms, as exemplified by under-
reporting of symptoms in OSA, [43] and patients’ perceptions of their sleep are frequently
inconsistent with objective indices. [44] A response to a sleep-rating scale may reflect the
patient’s state of mind, personality, trait anxiety or other psychological factors rather than
how well they actually sleep. [45,46]
Högl et al., (2010) [47] articulate more specific difficulties of using rating scales to measure sleep in PD. Firstly, sleep in PD is multifactorial and sleep scales that screen for a particular disorder potentially miss other facets. This is pertinent to this review, as 5 of 16 studies used screening questionnaires to identify a sleep disorder. The questionnaires that Högl et al. [47] recommended to measure sleep in PD were: the SCOPA-Sleep, the Pittsburgh sleep quality index (PSQI), the Parkinson’s disease sleep scale (PDSS), the Epworth sleepiness scale (ESS), the inappropriate sleep composite score (ISCS) and the Stanford sleepiness scale (SSS). The sleep scales used in studies included in this review are included in Table S3. Of five studies that used questionnaires to measure sleep, only two used measures recommended by Högl et al. [47] Goldman et al. [32] used the PSQI and Naismith et al. [33] used the SCOPA-Sleep, both of which measure overall sleep quality. Several of the clinical interviews used were based on validated questionnaires, which were not recommended by the MDS taskforce. Moreover, these papers lacked information about how the interviews had been changed from the original and did not report psychometrics.

Structured clinical interviews offer an opportunity for the researcher to expand on questionnaire responses, for the patient to raise important issues, and to clarify any points of confusion. Several of the studies [22,48,49] obtained more information by conducting interviews with both the patient and a family member or bed partner. This is a significant methodological improvement, as people may be unaware of their own sleeping behaviours (e.g. many snorers are unaware that they do so). [44] Most clinical interviews were conducted using criteria defined by the international classification of sleep disorders and were administered by an experienced clinician. In several studies, [23,35,49] clinical interviews were administered alongside questionnaires, providing a good example of a mixed-methods approach to sleep measurement when PSG was, presumably, unavailable or impractical.
Actigraphy is often used as an adjunct to questionnaires and interviews, as it provides objective data with which to compare subjective accounts. Actigraphy provides measures of sleep latency, wake after sleep onset, total sleep time and sleep efficiency. However, actigraphy is problematic in PD. Although, Stavisky et al. (2012), [31] followed the standard protocol for conducting actigraphy, Sadeh (2011) [50] argues that it is inappropriate to use standard actigraphy algorithms in special populations as specificity (ability to detect wakefulness) is unacceptably low. As actigraphy uses movement to infer sleep, it is questionable to assume that it is accurate in PD: a movement disorder. First, validation studies against PSG, and development of an algorithm that takes the idiosyncratic sleep movements of PD into account must be conducted to demonstrate reliability and validity.

Although tremor is attenuated during sleep, [51] there are unusual patterns of movement during sleep in PD. During REM, skeletal muscle movement is ordinarily inhibited; but in PD the mechanism responsible for skeletal muscle atonia during REM often fails. [52] Further, dyskinesias and other purposeful movements are often observed during sleep in PD. [53] Recently, Sixel Doring et al., (2014) [37] recorded movement during REM sleep in 51% of people with early PD. If these results are typical, then actigraphy will confuse REM and wake in approximately half of people with PD. At this stage, the necessary groundwork to validate actigraphy in PD has not been done; therefore, its use is premature.

Focus on RBD

RBD has inspired much research interest, as it appears very early in the disease course and heralds a more severe form of PD. [54,55] However, the focus on RBD is at the expense of research into other sleep disorders in PD. Eleven of 16 papers reviewed focussed on RBD, with one additional paper examining both RBD status and overall sleep quality. In contrast, there is a lack of accurate data regarding the prevalence of SRBDs at different disease stages and how these might affect cognition. A recent review [56] found relatively mild SRBD in
PD both in terms of apnoea-hypopnoea index and hypoxia and proposed that relatively low BMI, a lack of atonia during RBD in this group and selection bias (mild PD) could have contributed to these findings. Further investigation is warranted, as both PD and SRBDs are associated with deficits in executive function and memory. [15,16] PD may cause SRBD which may cause memory and executive deficits via the mechanisms of sleep fragmentation and hypoxia. [18] Alternatively, the dual injury of SRBD in addition to PD may compound memory and executive deficits.

Different sleep disorders impact sleep in different ways. For example, insomnia can produce delayed sleep onset or early morning wakening which reduce total sleep time/sleep efficiency. [57] By contrast, RLS and SRBD cause frequent arousals, which disrupt sleep architecture. [18,58] RBD seems more strongly associated with REM sleep disruption. [59] Thus, the nature of the sleep disorder experienced in PD may be predictive of the type and severity of cognitive deficits produced. Sleep problems that disturb SWS impact new learning and maintenance of brain plasticity; [60,61] reductions in slow-wave sleep (SWS) in PD have been documented since the 1970s. [62] Insomnia, in otherwise healthy individuals, produces deficits in tasks that measure attention and memory. [10] and is also common in PD [6]. Further, SRBD and RBD have both been associated with significant cognitive deficits, [16,63] and are common in PD. However, few of these have been studied in detail in PD using appropriate methods, such as PSG. Taken together, this suggests that focussed, well-designed research is needed to investigate the prevalence and aetiology of the full range of sleep problems in PD; any number of which could contribute to the cognitive problems frequently observed.

Measurement of Cognition
Many of the studies that concluded that there was no relationship between sleep problems and cognition in PD used neuropsychological tests which may have been insensitive to sleep-related deficits in PD. For example, the MMSE was not recommended for use to detect cognitive impairment in PD by the 2010 movement disorders task force, as it is insensitive to the cognitive changes that occur in PD. Being designed to detect frank dementia, it has pronounced ceiling effects. In a large study (N=873) where the MMSE was used to screen for dementia in people with PD, the MMSE was shown to have ‘strikingly low sensitivity’ at only 50%. Yet the MMSE was used in 10 of the 16 included studies, used the MMSE alone or in combination with the FAB instead of a targeted neuropsychological battery. These studies must, therefore, be treated with caution.

Similarly, the FAB was designed to detect frontal lobe dysfunction and has been validated in PD, Multiple System Atrophy (MSA), corticobasal degeneration, frontotemporal dementia (FTD) and Progressive Supranuclear Palsy (PSP). Indeed, regression analysis indicates that 69.7% of people with PSP and FTD were correctly classified using the FAB, suggesting that it successfully captures deficits associated with predominantly medial-prefrontal dysfunction. However, dysfunction in PD primarily involves the dorsolateral and ventrolateral prefrontal cortices and the FAB is not sensitive to these changes. A 2012 study which sought to validate the FAB in PD found sensitivity (66.3%) and specificity (72.3%) to detect dementia was worse than the MMSE (sensitivity 79.9%; specificity 74%) at a cut-off of 26 points. Even though dementia in PD is characterised primarily by deficits of EF, the FAB was poorer at detecting dementia in PD than the MMSE, which does not test EF at all.

**Summary Data**

Two reviewed studies conducted neuropsychological tests measuring sub-domains of memory and EF. They then combined these scores for analysis. In calculating summary
scores in this manner, sensitivity to subtle, dissociable cognitive impairments in PD may be compromised. Although both memory and EF may be impaired with poor sleep in PD, the size of any such effects, and the mechanisms by which sleep impacts on memory and EF could be radically different. Summing scores may remove effects altogether, obscure such possible differences, or make interpretation of any results problematic. Indeed, meta-analysis of summary score data found no significant effects.

**Recommendations for Future Research**

In light of the present review, future research must consider the findings from both the sleep literature and from research into cognitive dysfunction in PD. Neuropsychological tests should be selected based on their demonstrated sensitivity to the cognitive changes in PD and their association with the sleep disorder of interest. In early to moderate PD, deficits will typically be related either to EF or memory and may be very subtle. Consequently, sensitive tests are required. The dorsolateral prefrontal cortex is a locus of disruption in PD. [69] Tasks that ‘tap’ the dorsolateral prefrontal cortex include EF tasks across a number of sub-domains such as the Wisconsin card sorting task (set-shifting), the tower of London task (fluid reasoning), digit span backwards (updating), n-back tasks (updating), the Stroop test (inhibition) and verbal fluency tasks (generativity). Specific memory deficits that have been linked with damage to the dorsolateral prefrontal cortex have been detected using delayed memory tasks and word list learning tasks with interference trials, such as the Rey auditory verbal learning test or the California verbal learning test. [71–73] When test selection is theoretically driven, the risk of type 2 errors is diminished.

Whilst dividing participants in the reviewed studies into those with good vs. poor sleep was a practical approach to the heterogeneity of methods used, the blunt nature of this distinction may limit power to find significant cognitive effects. Accordingly, greater
consistency of more specific and powerful methods of assessing sleep is required. Studies that analyse sleep architecture in PD ideally require PSG and a sleep technician. Where PSG is not accessible, a mixed-methods approach is recommended to circumvent some of the problems with self-report measures of sleep. A sleep diary, which requires the patient to record sleep habits immediately on waking may ameliorate response bias inherent in questionnaires (which often ask patients to rate their sleep habits over the previous two weeks or more). [74]

While mindful of the limitations of self-report measures, some sleep questionnaires are more suitable for use in PD than others. Several sleep questionnaires have been developed for PD, including the SCOPA-Sleep [75] which measures indices related to sleep quality and the PDSS which screens for a range of sleep disturbances common in PD. [76] A combination of sleep questionnaires may be needed to capture both sleep quality and to screen for specific sleep disorders, e.g. the Berlin questionnaire for OSA risk or the international restless legs syndrome study group screening questionnaire for RLS. A clinical interview could then be used to probe further into the sleep symptomatology experienced by the patient.

Even though this field of research is in its infancy, it is important to delineate the relationships between sleep and cognition in PD as the results may have critical clinical implications. For instance, the standard treatment for RBD is low dose Clonazepam. [77] But if it is common for people to have concomitant RBD and OSA in PD, then treatment for RBD could potentially exacerbate OSA leading to downstream cognitive problems. [78,79] Cognitive problems interfere with effective management of motor symptoms and lead to poorer quality of life. A better strategy may be to focus research efforts on long-acting dopaminergic treatments which will control both conditions throughout the night-time hours. [80]
Conclusions

Despite a sparse literature characterised by methodological problems, the results of this meta-analysis reveal significant effects of sleep problems on neuropsychological performance in PD across several cognitive domains. Most affected are the sub-domains of executive function, that is: shifting, updating, generativity and fluid reasoning and long-term verbal recall and long-term verbal recognition, both subdomains of memory.

What is needed is a proper examination of the relationship between sleep and cognition in PD, informed by both the sleep and neuropsychological literatures.

Practice Points

1. Sleep problems cause neuropsychological deficits in otherwise healthy people. People with Parkinson’s disease typically experience chronically disrupted sleep and may be vulnerable to these same neuropsychological deficits.
2. If sleep problems are a factor underlying cognitive dysfunction in Parkinson’s disease, this has major clinical implications (e.g., treatment for REM sleep behaviour disorder can exacerbate concomitant sleep apnoea, increasing downstream cognitive dysfunction)
Research Agenda

1. Research design in this population must be informed by both the sleep and neuropsychological literatures: neuropsychological deficits in both disordered sleep and in early Parkinson’s disease are likely to pertain to memory and executive dysfunction. Neither the mini-mental state examination nor the frontal assessment battery will detect subtle deficits in these domains.

2. If polysomnography is not available, a mixed-methods approach is recommended for sleep measurement. Questionnaires may be supplemented by a clinical interview (preferably also including a family member) to circumvent unawareness of sleep behaviours and to probe for more detailed information. Sleep diaries can capture information on a day-to-day basis and thus ameliorate some response bias.

3. The prevalence and aetiology of sleep disorders, other than REM sleep behaviour disorder, are not well understood in PD, but the existence of sleep-related neuropsychological deficits in otherwise healthy people is well-established. There is no reason to believe that people with Parkinson’s disease are not equally susceptible to these same deficits and might, therefore, experience a double-jeopardy of cognitive impairment due to PD and sleep dysfunction.
References


**Search terms:** Parkinson's disease AND sleep OR sleep disorders OR sleep problems OR REM sleep behaviour disorder OR obstructive sleep apnoea OR Restless Legs Syndrome AND attention OR cognition OR cognitive processes OR episodic memory OR executive functions OR long term memory OR memory OR memory disorders OR neuropsychology OR prefrontal cortex OR selective attention OR short term memory OR spatial memory OR sustained attention OR verbal memory OR visual attention OR visual memory OR visuospatial memory

**Databases searched:**
- Medline (n=253)
- PsychInfo (n=26)
- PUBMED (n=828)
- Proquest Theses and Dissertations (n=1130)
- Web of Science Conference Proceedings (n=18)

**Manual Searching:**
- Hand-searching of reference lists checking citations of seminal papers in the field (n=26)
- Papers recommended by contacted authors (n=2)
- Total n = 2283 papers

**Irrelevant and duplicate titles removed**
(n=2136)

**Abstracts Excluded**
(n=104)
- Reasons:
  1. Sleep not measured
  2. Cognition not measured
  3. Not a peer reviewed study
  4. Review paper
  5. Conference presentation later released as a paper

**Studies Excluded**
(n=27)
- Reasons:
  1. Sleep measured in an unreliable manner (4)
  2. Cognition measured in an unreliable manner (3)
  3. Measured neither sleep nor cognition in an acceptable manner (1)
  4. Relationship between sleep and cognition cannot be evaluated from data provided (12)
  5. Participants were included in multiple studies (4)
  6. Participants were a non-representative sample (1)
  7. Motor learning/consolidation task (1)

**Full text copies retrieved for evaluation using quality assessment criteria**
(n=43)

**Abstracts Screened**
(n=147)

**Data extracted for review**
(n=16)

**Figure 1**
*Flow Chart of search strategy, retrieval and selection process*
## Meta Analysis

**Study Name** | **Outcome** | **_d_** | **Standard error** | **Variance** | **Lower bound** | **Upper bound** | **Z-value** | **p-value** | **Std. diff. in means and 95% CI.**
---|---|---|---|---|---|---|---|---|---
Bugaio, 2011  | Global Cognitive Function (WMSE) | 0.03 | 0.43 | 0.18 | -0.80 | 0.66 | 0.51 | 0.94 |
Cochrane de Cock, 2016 | Global Cognitive Function (WMSE) | 0.00 | 0.23 | 0.05 | -0.44 | 0.29 | 0.70 | 1.00 |
Lanfors, 2010 | Global Cognitive Function (WMSE) | 0.23 | 0.27 | 0.27 | -0.20 | 0.65 | 0.21 | 0.23 |
Mared, 2007 | Global Cognitive Function (Short Test of Mental Status) | 0.84 | 0.24 | 0.96 | 0.58 | 1.05 | 0.56 | 0.00 |
Husseini, 2011 | Global Cognitive Ability (WMSE) | 0.43 | 0.21 | 0.64 | 0.02 | 0.85 | 0.27 | 0.94 |
Mehmet, 2013 | Combined | 0.27 | 0.39 | 0.06 | -0.30 | 0.84 | 0.92 | 0.36 |
Schwerdt, 2006 | Global Cognitive Ability (WMSE) | 0.71 | 0.23 | 0.05 | 0.26 | 1.16 | 0.06 | 0.60 |
Schroder, 2011 | Global Cognitive Ability (WMSE) | 0.14 | 0.09 | 0.51 | -0.32 | 0.05 | 1.47 | 0.14 |
Stein-Doring, 2011 | Global Cognitive Ability (WMSE) | 0.24 | 0.16 | 0.03 | -0.07 | 0.56 | 1.53 | 0.13 |
Stein-Doring, 2013 | Global Cognitive Ability (WMSE) | 0.10 | 0.20 | 0.04 | -0.28 | 0.49 | 0.91 | 0.60 |
Vanbeek, 2010 | Global Cognitive Ability (SCOPA-COG) | 0.50 | 0.24 | 0.06 | 0.03 | 0.97 | 0.09 | 0.94 |
Yoshida, 2009 | Global Cognitive Ability (WMSE <C) | 1.22 | 0.46 | 0.21 | -0.33 | 2.12 | 2.67 | 0.01 |
Yoshida, 2013 | Combined | 0.23 | 0.11 | 0.01 | 0.11 | 0.55 | 2.36 | 0.00 |

Figure 2: Forest plot for global cognitive ability displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Std diff in means</th>
<th>Std error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goldman, 2013</td>
<td>Memory (summary score)</td>
<td></td>
<td>-0.26</td>
<td>0.21</td>
<td>0.05</td>
<td>-0.68</td>
<td>0.15</td>
<td>-1.23</td>
<td>0.22</td>
</tr>
<tr>
<td>Stavisky, 2012</td>
<td>Combined</td>
<td></td>
<td>-0.17</td>
<td>0.36</td>
<td>0.13</td>
<td>-0.88</td>
<td>0.55</td>
<td>-0.46</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.24</td>
<td>0.18</td>
<td>0.03</td>
<td>-0.60</td>
<td>0.12</td>
<td>-1.30</td>
<td>0.19</td>
</tr>
</tbody>
</table>

![Forest plot for memory (general) displaying effect size (Hedges' g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers](image)

Figure 3: Forest plot for memory (general) displaying effect size (Hedges' g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
**Meta Analysis**

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ougron, 2009</td>
<td>Combined</td>
<td>Std diff in means</td>
</tr>
<tr>
<td>Meral, 2007</td>
<td>Combined</td>
<td>0.82</td>
</tr>
<tr>
<td>Navigli, 2011</td>
<td>Combined</td>
<td>0.72</td>
</tr>
<tr>
<td>Sironi, 2008</td>
<td>Long term verbal memory (Logical Memory)</td>
<td>0.45</td>
</tr>
<tr>
<td>Silk-Croping, 2013</td>
<td>Combined</td>
<td>0.70</td>
</tr>
</tbody>
</table>

**Figure 4:** Forest plot for long-term verbal recall displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers, Favours B: Favours good sleepers
Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon, 2009</td>
<td>Long term verbal recognition (RAVLT, Recognition)</td>
<td>0.57</td>
<td>0.32</td>
<td>0.11</td>
<td>-0.06</td>
<td>1.21</td>
<td>1.77</td>
<td>0.08</td>
</tr>
<tr>
<td>Metal, 2007</td>
<td>Long term verbal recognition (SBST recognition)</td>
<td>0.91</td>
<td>0.24</td>
<td>0.06</td>
<td>0.45</td>
<td>1.38</td>
<td>3.84</td>
<td>0.00</td>
</tr>
</tbody>
</table>

-1.00 -0.50 0.00 0.50 1.00

Favours A

Favours B

Figure 5: Forest plot for long-term verbal recognition displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
### Meta Analysis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bugnho, 2011 Executive Function (FAB)</td>
<td>0.01</td>
<td>0.27</td>
<td>0.08</td>
<td>-0.53</td>
<td>0.55</td>
<td>0.03</td>
<td>0.96</td>
</tr>
<tr>
<td>Goldman, 2011 Executive Function (summary score)</td>
<td>0.06</td>
<td>0.21</td>
<td>0.04</td>
<td>-0.35</td>
<td>0.47</td>
<td>0.28</td>
<td>0.78</td>
</tr>
<tr>
<td>Lavault, 2010 Executive Function (FAB)</td>
<td>0.20</td>
<td>0.27</td>
<td>0.07</td>
<td>-0.28</td>
<td>0.77</td>
<td>0.93</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Figure 6: Forest plot for executive function (general) displaying effect size (Hedges' g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers.
Figure 7: Forest plot for shifting displaying effect size (Hedges' g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
## Meta Analysis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon, 2000: Updating (Digit Span)</td>
<td>0.76 0.33 0.11 1.41 2.31 0.02</td>
<td>0.60 0.18 0.03 0.94 3.36 0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naismith, 2011: Updating (Digit Span, backwards)</td>
<td>0.53 0.21 0.04 0.94 2.50 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 8: Forest plot for updating displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon, 2009</td>
<td>Combined</td>
<td>0.86</td>
<td>0.33</td>
<td>0.11</td>
<td>0.93</td>
<td>1.53</td>
<td>2.64</td>
<td>0.01</td>
</tr>
<tr>
<td>Sixel-Doring, 2018</td>
<td>Inhibition (Stroop, interference time)</td>
<td>0.20</td>
<td>0.17</td>
<td>0.03</td>
<td>-0.13</td>
<td>0.53</td>
<td>1.19</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.48</td>
<td>0.34</td>
<td>0.11</td>
<td>-0.14</td>
<td>1.14</td>
<td>1.43</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**Std diff in means and 95% CI**

-0.86 to 2.64

**Favours A**

**Favours B**

---

**Figure 9:** Forest plot for inhibition showing effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers.
### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon, 2009</td>
<td>Combined</td>
<td>1.05</td>
<td>0.34</td>
<td>0.12</td>
<td>0.38</td>
<td>1.72</td>
<td>3.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Madon, 2006</td>
<td>Generativity (Verbal Fluency)</td>
<td>0.19</td>
<td>0.26</td>
<td>0.07</td>
<td>-0.31</td>
<td>0.70</td>
<td>0.75</td>
<td>0.45</td>
</tr>
<tr>
<td>Meral, 2007</td>
<td>Generativity (Verbal fluency, semantic)</td>
<td>2.32</td>
<td>0.29</td>
<td>0.09</td>
<td>1.75</td>
<td>2.69</td>
<td>7.95</td>
<td>0.00</td>
</tr>
<tr>
<td>Naismith, 2011</td>
<td>Generativity (Verbal Fluency, z score)</td>
<td>0.09</td>
<td>0.21</td>
<td>0.04</td>
<td>-0.32</td>
<td>0.49</td>
<td>0.41</td>
<td>0.68</td>
</tr>
<tr>
<td>Slifstein, 2006</td>
<td>Generativity (verbal fluency, phonemic)</td>
<td>0.34</td>
<td>0.23</td>
<td>0.05</td>
<td>-0.11</td>
<td>0.76</td>
<td>1.48</td>
<td>0.14</td>
</tr>
<tr>
<td>Sixel-Doring, 2018</td>
<td>Combined</td>
<td>0.17</td>
<td>0.16</td>
<td>0.03</td>
<td>-0.15</td>
<td>0.49</td>
<td>1.06</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.67</td>
<td>0.31</td>
<td>0.09</td>
<td>0.06</td>
<td>1.27</td>
<td>2.17</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Figure 10: Forest plot for generativity displaying effect size (Hedges' g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers.
### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meral, 2007</td>
<td>Fluid Reasoning (Clock drawing)</td>
<td>0.93</td>
<td>0.24</td>
<td>0.06</td>
<td>0.46</td>
<td>1.39</td>
<td>3.90</td>
<td>0.00</td>
</tr>
<tr>
<td>Sinforni, 2006</td>
<td>Fluid Reasoning (Raven’s Matrices)</td>
<td>0.44</td>
<td>0.23</td>
<td>0.05</td>
<td>-0.01</td>
<td>0.89</td>
<td>1.93</td>
<td>0.05</td>
</tr>
<tr>
<td>Sixel-Doring, 2011</td>
<td>Combined</td>
<td>0.18</td>
<td>0.16</td>
<td>0.03</td>
<td>-0.14</td>
<td>0.49</td>
<td>1.09</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.49</td>
<td>0.22</td>
<td>0.05</td>
<td>0.05</td>
<td>0.92</td>
<td>2.20</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Figure 11: Forest plot for fluid reasoning displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
## Meta Analysis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marion, 2008 Psychomotor (Symbol Digit Modalities)</td>
<td>0.04</td>
<td>0.26</td>
<td>0.07</td>
<td>-0.46</td>
<td>0.55</td>
<td>0.17</td>
<td>0.87</td>
</tr>
<tr>
<td>Stavitsky, 2008 Combined</td>
<td>0.07</td>
<td>0.37</td>
<td>0.13</td>
<td>-0.65</td>
<td>0.78</td>
<td>0.19</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.21</td>
<td>0.04</td>
<td>-0.36</td>
<td>0.46</td>
<td>0.25</td>
<td>0.81</td>
</tr>
</tbody>
</table>

-1.00 -0.50 0.00 0.50 1.00

Favours A

Favours B

---

**Figure 12**: Forest plot showing effect size (Hedges' g) for psychomotor ability calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers
### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon, 2009</td>
<td>Combined</td>
<td>0.63</td>
<td>0.34</td>
<td>0.11</td>
<td>0.17</td>
<td>1.49</td>
<td>2.47</td>
<td>0.01</td>
</tr>
<tr>
<td>Goldman, 2013</td>
<td>Visuospatial Abilities (summary score)</td>
<td>-0.01</td>
<td>0.21</td>
<td>0.04</td>
<td>-0.42</td>
<td>0.41</td>
<td>-0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>Meral, 2007</td>
<td>Combined</td>
<td>0.94</td>
<td>0.24</td>
<td>0.06</td>
<td>0.47</td>
<td>1.41</td>
<td>3.90</td>
<td>0.00</td>
</tr>
<tr>
<td>Sixel-Droring, 2019</td>
<td>Combined</td>
<td>0.21</td>
<td>0.16</td>
<td>0.03</td>
<td>-0.11</td>
<td>0.53</td>
<td>1.27</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.45</td>
<td>0.22</td>
<td>0.05</td>
<td>0.01</td>
<td>0.89</td>
<td>2.02</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Ifigure 13: Forest plot for visuospatial/ constructional ability displaying effect size (Hedges’ g) calculated using a random effects model. Favours A: Favours poor sleepers; Favours B: Favours good sleepers.
Table 1
Mean overall effect sizes, confidence intervals and homogeneity statistics

<table>
<thead>
<tr>
<th>Domain</th>
<th>Hedges' g</th>
<th>95% CI</th>
<th>Z</th>
<th>P</th>
<th>Q (df)</th>
<th>P</th>
<th>Tau</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LL</td>
<td>UL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Cognitive Function</td>
<td>0.33</td>
<td>0.11</td>
<td>0.55</td>
<td>2.95</td>
<td>&lt;0.01</td>
<td>34.42(11)</td>
<td>0.00</td>
<td>0.30</td>
</tr>
<tr>
<td>Memory (General)</td>
<td>0.23</td>
<td>-0.59</td>
<td>0.12</td>
<td>-1.30</td>
<td>0.20</td>
<td>0.06(1)</td>
<td>0.82</td>
<td>0.00</td>
</tr>
<tr>
<td>Long Term Verbal Recall</td>
<td>0.51</td>
<td>0.24</td>
<td>0.77</td>
<td>3.79</td>
<td>&lt;0.01</td>
<td>7.09(4)</td>
<td>0.13</td>
<td>0.19</td>
</tr>
<tr>
<td>Long Term Verbal Recognition</td>
<td>0.78</td>
<td>0.41</td>
<td>1.15</td>
<td>4.14</td>
<td>&lt;0.01</td>
<td>0.75(1)</td>
<td>0.39</td>
<td>0.00</td>
</tr>
<tr>
<td>Executive Function (General)</td>
<td>0.10</td>
<td>-0.18</td>
<td>0.37</td>
<td>0.70</td>
<td>0.49</td>
<td>0.45(2)</td>
<td>0.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Executive Function- Shifting</td>
<td>0.27</td>
<td>0.04</td>
<td>0.49</td>
<td>2.35</td>
<td>0.02</td>
<td>5.19(4)</td>
<td>0.27</td>
<td>0.12</td>
</tr>
<tr>
<td>Executive Function- Updating</td>
<td>0.59</td>
<td>0.25</td>
<td>0.93</td>
<td>3.36</td>
<td>&lt;0.01</td>
<td>0.33(1)</td>
<td>0.56</td>
<td>0.00</td>
</tr>
<tr>
<td>Executive Function- Inhibition</td>
<td>0.47</td>
<td>-0.17</td>
<td>1.11</td>
<td>1.44</td>
<td>0.15</td>
<td>3.29(1)</td>
<td>0.07</td>
<td>0.39</td>
</tr>
<tr>
<td>Executive Function- Generativity</td>
<td>0.66</td>
<td>0.06</td>
<td>1.26</td>
<td>2.16</td>
<td>0.03</td>
<td>51.29(5)</td>
<td>0.00</td>
<td>0.70</td>
</tr>
<tr>
<td>Executive Function- Fluid Reasoning</td>
<td>0.49</td>
<td>0.05</td>
<td>0.92</td>
<td>2.21</td>
<td>0.03</td>
<td>6.86(2)</td>
<td>0.03</td>
<td>0.32</td>
</tr>
<tr>
<td>Visuospatial/Constructualional</td>
<td>0.45</td>
<td>0.01</td>
<td>0.88</td>
<td>2.02</td>
<td>0.04</td>
<td>11.63(3)</td>
<td>0.01</td>
<td>0.38</td>
</tr>
<tr>
<td>Psychomotor Ability</td>
<td>0.05</td>
<td>-0.35</td>
<td>0.46</td>
<td>0.25</td>
<td>0.81</td>
<td>0.00(1)</td>
<td>0.96</td>
<td>0.00</td>
</tr>
</tbody>
</table>
All contributing authors should sign this form and submit it with their manuscript

CONFLICT OF INTEREST DISCLOSURE

All funding sources supporting this work are fully acknowledged. All of the authors will disclose to the Editor(s) any pertinent personal financial interests associated with the development, testing, manufacture or marketing of any drug or product described in this manuscript entitled:

Romola Bucks has no pertinent personal financial interests associated with this manuscript.

Andrea Loftus has no pertinent personal financial interests associated with this manuscript.

Maria Pushpanathan has no pertinent financial interests associated with this manuscript.

Natalie Gasson has no pertinent financial interests associated with this manuscript.

Meghan Thomas has no pertinent financial interests associated with this manuscript.

Signed by the author(s)  Date

Romola Bucks  5th May, 2014

Andrea Loftus  6th May, 2014

Maria Pushpanathan  6th May, 2014

Natalie Gasson  6th May, 2014

Meghan Thomas  6th May 2014
6th May 2014

Editorial Office Sleep Medicine Reviews
Gio Bakker
P.O. Box 993
1000 AZ Amsterdam
The Netherlands
ysmrv@elsevier.com;
fax: ++31204853378
<table>
<thead>
<tr>
<th>Reference</th>
<th>Methodology and Main Findings</th>
</tr>
</thead>
</table>
| Sinforiani et al. 2006 [27]| **Methodology**: Compared the cognitive performance of PD patients who were RBD+ and RBD- (as identified by clinical interview with the patient and their bed partner) on neuropsychological assessment  
- **Findings**: People with RBD performed more poorly on the Wisconsin card sorting task than those without RBD. Those who also experienced visual hallucinations in addition to RBD also demonstrated poorer performance in Corsi’s block tapping test, logical memory and Raven’s matrices than those without RBD and concurrent visual hallucinations.  
- **Methodological Issues**: 1, 2, 8                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Meral et al. 2007 [26]     | **Methodology**: Compared the cognitive performance of PD patients who were RBD+ and RBD- (as identified by clinical interview) on comprehensive neuropsychological assessment  
- **Findings**: Found no statistically significant differences between performance of groups on any cognitive domain  
- **Methodological Issues**: 1, 2, 6, 7                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Marion et al. 2008 [35]    | **Methodology**: Compared the cognitive performance of PD patients who were RBD+ and RBD- (as identified by Mayo RBD Screening Questionnaire and clinical interview) on neuropsychological assessment  
- **Findings**: No statistically significant differences in any of the tests but...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Mean Age</th>
<th>Mean Disease Duration</th>
<th>Exclusion Criteria</th>
<th>Methodology</th>
<th>Findings</th>
<th>Methodological Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon et al. 2009 [76]</td>
<td>N= 48</td>
<td>65.89</td>
<td>5.27</td>
<td>MMSE&lt;26, &lt;7 years education, atypical PD, major depression, dementia, obstructive sleep apnoea</td>
<td>Compared the cognitive performance of PD patients who were RBD+ and RBD- (as identified by PSG) on neuropsychological assessment</td>
<td>Compared to people with PD only, people with PD and concomitant RBD demonstrated poorer performance on the majority of cognitive measures including digit span, trail making test- part b, the Stroop task, RAVLT (Trials 1-5, immediate recall, delayed recall), Rey-Osterreith complex figure and Block Design.</td>
<td>1, 2, 3, 5, 8, 10</td>
</tr>
<tr>
<td>Yoritaka et al. 2009 [22]</td>
<td>N= 150</td>
<td>68.5</td>
<td>6.4</td>
<td>Hoehn &amp; Yahr of 5, people with secondary parkinsonism, Lewy Body Dementia, single patients (needed bed partner to verify symptoms)</td>
<td>Compared the cognitive status of PD patients who were RBD+ and RBD- (as identified by clinical interview with participant and their partner).</td>
<td>27.2% of RBD+ group and 10.1% of RBD- group fell below pre-specified cut-off on the MMSE (23 points). This was a significant difference.</td>
<td>1, 2, 3, 4, 5, 8</td>
</tr>
<tr>
<td>Cochen de Cock et al. 2010  [29]</td>
<td>N= 100</td>
<td>62.4</td>
<td>7.5</td>
<td>MMSE&lt;21</td>
<td>Compared 50 unselected PD patients and 50 PD patients selected for daytime sleepiness. Examined both groups using PSG to determine respective rates of obstructive sleep apnoea (OSA). They then</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
looked for evidence of cognitive dysfunction associated with OSA by comparing the MMSE scores of those with and without OSA.

- **Findings** There were no group differences in MMSE scores.
- **Methodological Issues** 2, 3, 4, 5, 8, 9

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Methodology</th>
<th>Exclusion Criteria</th>
<th>Findings</th>
<th>Methodological Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lavault et al. 2010 [48]</td>
<td>N= 61 participants with PD (39 men, 22 women), mean age 64.33 years (in 2005), mean disease duration 6.89 years (in 2005)</td>
<td>- Methodology Compares the cognitive performance of people with PD who are RBD+ and RBD- (as identified by questionnaire and clinical interview with patient and bed partner).</td>
<td>MMSE &lt;18</td>
<td>- Findings There were no group differences in either global cognitive ability or executive function.</td>
<td>1, 2, 3, 5, 10</td>
</tr>
<tr>
<td>Verbaan et al. 2010 [30]</td>
<td>N= 269 participants with PD, 172 men, 97 women. Mean age 61.5 years (SD 9.8 years). Mean disease duration 13.3 years (SD 6 years).</td>
<td>- Methodology Compared the cognitive performance of PD patients with RLS (assessed by the IRLSSG) to those without RLS using the SCOPA-COG assessment.</td>
<td>MMSE &lt;24, non-Caucasian ethnicity</td>
<td>- Findings There were no differences in cognitive performance between groups</td>
<td>1, 2, 3, 4, 5, 9</td>
</tr>
<tr>
<td>Bugalho et al. 2011 [23]</td>
<td>N= 75 participants with PD 32 men, 43 women. Mean age 72.56 years (SD 7.20 years). Mean disease duration 2.76 years (SD 1.37 years).</td>
<td>- Methodology Compared the cognitive performance of PD patients who were RBD+ and RBD- (as identified by the RBD Screening Questionnaire and clinical interview) and neuropsychological assessment of basic global cognitive function and executive function.</td>
<td>Neurological conditions other than PD, Hoehn &amp; Yahr &gt; 2.5, Disease duration &gt; 5 years</td>
<td>- Findings No difference in cognitive performance between those who had a history of RBD symptoms and those with no history of RBD symptoms</td>
<td>1, 2, 3, 4, 5, 9, 10</td>
</tr>
<tr>
<td>Study</td>
<td>Methodological Issues</td>
<td>Participants</td>
<td>Exclusion Criteria</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Naismith et al. 2011 [33]</td>
<td>1, 2, 3, 5, 8</td>
<td>N=101</td>
<td>MMSE&lt;24, dementia, depression</td>
<td>Compared the cognitive performance of people with PD who reported disturbed night-time sleep (on the SCOPA-night questionnaire) with those who reported undisturbed sleep. Compared the cognitive performance of those who reported RBD symptoms (on the REM Sleep Behaviour Disorder Screening Questionnaire) with those who reported no RBD symptoms.</td>
<td>Disturbed sleep was associated with poorer performance on the Digit Span backwards task as well as in the Logical Memory I and II tasks. People who reported symptoms of RBD demonstrated poorer performance on the Digit Span backwards task and the Verbal Fluency task.</td>
</tr>
<tr>
<td>Sixel-Doring et al. 2011 [77]</td>
<td>1</td>
<td>N=457</td>
<td>Severe dementia (MMSE&lt; 10), acute psychosis, patients not able to co-operate in sleep lab</td>
<td>Compared the MMSE scores of people with and without RBD (identified by PSG)</td>
<td>There were no group differences in MMSE scores.</td>
</tr>
<tr>
<td>Stavitsky et al. 2012 [31]</td>
<td>3, 5, 8, 9</td>
<td>N= 35</td>
<td>MMSE &lt;25, history of substance abuse, head injury, neurological disorders other than PD</td>
<td>This study correlated measures of sleep quality such as sleep latency, wake after sleep onset, total sleep time, sleep efficiency and sleep fragmentation (as measured by actigraphy) with measures of cognitive performance.</td>
<td>They found a significant correlation between sleep efficiency and measures of attention and executive function.</td>
</tr>
<tr>
<td>Study</td>
<td>Methodological Issues</td>
<td>Participants</td>
<td>Exclusion Criteria</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Goldman et al. 2013 [32]</td>
<td>1, 3, 9</td>
<td>N= 93 participants with PD, 69 men, 24 women. Mean age 73.62 years (SD 6.33 years). Mean disease duration 10.73 years (SD 4.12 years).</td>
<td>- Disease duration &lt;4 years, severe or unstable depression, anticholinergic medications, neurological problems other than PD, known untreated or unstable sleep disorders (e.g. OSA), participant must be excluded if cognitive symptoms preceded motor symptoms and they developed dementia.</td>
<td>Correlated neuropsychological performance over a range of tasks with sleep quality (as measured by the Pittsburgh Sleep Quality Index: PSQI scores).</td>
<td>There were no correlations between any of the cognitive test scores and PSQI scores.</td>
</tr>
<tr>
<td>Plomhause et al. 2013 [36]</td>
<td>1, 3, 8, 9</td>
<td>N= 57 participants with PD (all participants were newly diagnosed and drug naïve), 35 men, 22 women. Mean age 61.49 years (SD 10.81 years). Mean disease duration 13.81 months (SD 9.61 months).</td>
<td>- Dementia, neurologic comorbidities, antiparkinsonian medications, irregular use of psychotropic medications in last 3 months</td>
<td>Compared PD patients who were RBD+ or RBD- (identified by PSG) on a range of neuropsychological measures</td>
<td>There were no significant differences between the groups on any of the cognitive measures.</td>
</tr>
<tr>
<td>Nardone et al. 2013 [28]</td>
<td>2, 3, 8, 9</td>
<td>N= 23 participants with PD, 17 men, 6 women. Mean age 64.66 years (SD 6.44 years). Mean disease duration 5.57 years (SD 2.58 years).</td>
<td>- Other neurological, psychiatric or cardiovascular disease. Beck Depression Inventory Score &gt;14, Vigilance impairment (as measured by a score &gt;1 second on the four choice reaction time test). Participants treated with anticholinergic medications, those currently using benzodiazepines.</td>
<td>Participants with PD who were RBD+ or RBD- (identified by PSG) were compared across a comprehensive neuropsychological measures</td>
<td></td>
</tr>
</tbody>
</table>
Findings 90% of participants with RBD were found to have MCI compared to 38% of participants without RBD. This represents a significant difference between groups. This study used transcranial magnetic stimulation (TMS) to investigate the cholinergic networks implicated in MCI in PD.

Methodological Issues 2, 6, 9, 10

<table>
<thead>
<tr>
<th>Study</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sixel-Doring et al. 2014 [37]</td>
<td>N= 159 participants with PD (all participants were drug naïve) 104 men, 55 women. Mean age 65 years (SD 10 years). All participants were newly diagnosed.</td>
</tr>
<tr>
<td></td>
<td>- Exclusion Criteria MMSE &lt;26, current or past treatment with antipsychotics, severe vascular encephalopathy or normal pressure hydrocephalus on magnetic resonance imaging (MRI), signs or symptoms suggestive of multiple system atrophy or progressive supranuclear palsy or medication induced PD</td>
</tr>
<tr>
<td></td>
<td>- Methodology Compared PD patients with events related to RBD (RBEs; movements observed on video PSG during REM sleep) to PD patients who did not demonstrate RBEs on a range on neuropsychological measures.</td>
</tr>
<tr>
<td></td>
<td>- Findings There were no significant differences between the RBE and non-RBE groups on any of the cognitive measures.</td>
</tr>
<tr>
<td></td>
<td>- Methodological Issues 2, 8, 9</td>
</tr>
</tbody>
</table>

**Methodological Issues:**

1. Sleep measure may be insensitive to detect sleep problems common in PD
2. Study focused on one dimension of sleep
3. Cognitive measure may be insensitive to types of cognitive dysfunction associated with stage of PD assessed or summary scores used which will hide effects
4. Cognitive measures not targeted at sleep related changes
5. Inappropriate statistical methods used
6. Very high number of participants excluded, potentially reducing variance
7. Failed to exclude depression or to control for depressive symptomatology
8. Failed to control for years of education or premorbid IQ
9. Failed to control for age/gender
References for Table S1 (not included in main article)


### Table S2

**Descriptions of cognitive domains examined in this review, neuropsychological tests used in included studies and the deficits observed in these domains in the PD literature**

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Tests that tap this cognitive skill</th>
<th>Pattern of deficits observed in PD</th>
</tr>
</thead>
</table>
| **Global Cognitive Function**  | Refers to a composite score that summarises a person’s overall mental abilities. Tests that measure global cognitive function will tap a number of cognitive domains typically memory, executive functions, working memory and visuospatial/constructional ability. | - Mini Mental State Exam [23; 26; 27; 28; 29; 33; 36; 37; 48; 77]  
  - Mattis Dementia Rating Scale [24; 28]  
  - SCOPA-COG [30]  
  - Short Test of Mental Status [26] | At diagnosis, global cognitive function in PD is typically close to ceiling when measured by a common test like the MMSE. [1] MMSE scores, however, change at a rate of approximately +1.4 to -6.8 points per year (mean -0.3 points, SD 0.1 points; Williams-Grey et al., 2007). [2] Risk factors for rapid decline in global cognitive function are age (>70 years), a non-tremor dominant subtype of PD, a high UPDRS motor score and poor baseline performance on the pentagon copying (visuospatial) component of the MMSE at baseline. |
| **Memory**                    | Memory refers to the way that information is encoded, stored and retrieved. There are many sub-categories of memory (verbal vs. visual; short term vs. delayed; recall vs. recognition) and a person who experiences difficulties in one sub-domain may demonstrate intact performance in other sub-domains. | - Memory Summary Score [31; 32] | In early PD, memory tasks are often unaffected compared to healthy aged-matched controls (Farina, 2000). [3] In moderate PD, amnestic deficits are the second most common form of mild cognitive impairment (MCI) in PD according to a 2007 study by Caviness et al. [4] And memory problems can be predictive of eventual dementia in PD (Levy et al., 2002) [5]. |
| * Short Term Verbal Recall    | The ability to recall (non-cued) verbal information within 2 minutes of presentation                                                                                                                                 | - Digit Span Forwards [27] | Cooper et al., 1991 [6] found that people with early PD, who were not yet on medication had indistinguishable results on measures of short term verbal recall. |
| * Short Term Spatial Recall   | The ability to recall (non-cued) spatial information within 2 minutes of presentation                                                                                                                        | - Corsi’s Block Tapping Test [27] | Lange et al (1992) [7] noted that people with PD performed worse than controls on Corsi’s block tapping task. This disparity was exacerbated when LDopa therapy was withdrawn. |
| Long Term Verbal Recall | The ability to recall (non-cued) verbal information after a period >2 minutes from presentation | - SBST Immediate Verbal Recall [26]  
- RAVLT Immediate Verbal Recall [24; 76]  
- Logical Memory I [33]  
- SBST Delayed Verbal Recall [26]  
- Logical Memory Test [27]  
- RAVLT Delayed Verbal Recall [36; 76]  
- Logical Memory II [33] | Aarsland et al (2009) [8] examined a community sample of newly diagnosed, unmedicated PD patients and found that PD patients performed slightly worse on the California Verbal Learning Test Immediate Recall Test (partial Eta Square= 0.027). When examining newly diagnosed patients with PD, Muslimovic et al. (2005) [9] found that 11% of patients scored significantly below normative data for delayed verbal learning. |
| Long Term Verbal Recognition | The ability to recognise previously presented verbal information amongst distractors after a period >10 minutes after presentation | - SBST Delayed Verbal Recognition [26]  
- RAVLT Recognition [36; 76] | Previous studies have found that delayed verbal recognition memory typically remains intact as PD develops. [10; 11] |
| *Long Term Visual Recall | The ability to recall visual information after a period >2 minutes following presentation | - Visual Memory Subtest of Weschler Memory Scale [26] | In a small study of highly functional PD patients Mohr et al., (1990) [12] demonstrated specific deficits in visual immediate recall, despite most neuropsychological function remaining intact. They demonstrated poorer performance on the delayed visual reproduction sub-section of the Weschler Memory Scale in a group of highly functional Parkinson’s patients with largely intact neuropsychological function. |
| * Language | Language skills require a person to access words stored in their lexicon and produce them at an appropriate time. Language skills are one of the first skills to develop and correspondingly, are one of the most robust skills when a patient is facing cognitive decline. | - Boston Naming Test [32] | Muslimovic et al., 2005 [9] demonstrated in a recent study that newly diagnosed people with PD demonstrated only a relatively small deficit in language tests relative to a healthy age-matched sample. |
| Executive Function (EF-General) | Executive or frontal functions collectively refer to the ‘higher order’ mental processes involved in planning, reasoning, set-shifting, inhibition of prepotent responses and updating mental sets. | - Frontal Assessment Battery [23; 48]  
- Executive Function Summary Score [31; 32] | Consistently, research has shown that one of the earliest domains to be affected in PD is executive function (Muslimovic et al., 2005 [9]; Lewis et al. 2003 [13]; Levin et al., 1995[14]). People who experience executive dysfunction early in the course of PD are at greater risk of developing significant cognitive decline as PD advances [2; 5]. |
**EF- Shifting**  
Shifting back and forth between multiple tasks, operations or mental sets. Requires the disengagement of an irrelevant task set and subsequent engagement of a relevant task set when a new operation must be performed on a set of stimuli, necessary to overcome proactive interference or negative priming due to having recently performed a different operation.  
- Wisconsin Card Sorting Test [26; 27; 37]  
- Trails B [26; 33; 37; 76]  
Lees et al., (1983) [15] notes that PD patients, even in the very early stages of the disease course experience difficulty with set-shifting. They demonstrate this principle in this study using the Wisconsin Card Sorting Test with participants who have early stage PD.

**EF- Updating**  
Updating and monitoring of working memory representations. Requires the monitoring and coding of incoming information for relevance to the task, and then appropriately revising items held in working memory by replacing old, no longer relevant information with new more relevant information. Dynamically manipulate the contents of working memory.  
- Digit Span Backwards [33; 76]  
A study by Gabrieli et al. (1996) [16] found that untreated PD patients had impaired working memory performance compared to healthy, age-matched controls. Cooper et al., (1991) [6] found that untreated, early stage PD patients performed more poorly on the digit span backward test than did healthy age-matched controls.

**EF- Inhibition**  
Inhibition of prepotent, dominant or automatic responses when necessary. An internally generated act of control.  
- Stroop Task [26; 36; 76]  
Hsieh et al., 2008 [17] demonstrated that people with PD lack the ability to effectively inhibit a prepotent response as demonstrated by their significantly poorer performance on the interference condition of the Stroop task when compared to healthy age-matched controls.

**EF- Generativity**  
Speed and efficiency of access to long-term memory. An independent ability to create, generate or produce content without any input from what or whom?  
- Verbal Fluency [26; 27; 33; 35; 37; 76]  

**EF- Fluid reasoning**  
Concept formation/abstraction & problem solving tasks. An intentional cognitive process that does not occur automatically, but rather involves the use of deliberate and controlled mental actions to solve novel problems.  
- Raven’s Progressive Matrices [27]  
- WAIS Similarities [25]  
- Clock Drawing Test [26]  
When profiling cognitive dysfunction in early PD Muslimovic et al (2005) [9], found that 11% of people scored at least 1.5 SDs below the mean on the Clock Drawing Test. On the other hand, in 1988, Hietanen et al., demonstrated similar fluid reasoning skills in young onset PD patients, older onset PD patients and age matched healthy controls using the WAIS similarities test. [18]
<table>
<thead>
<tr>
<th><em>Processing Speed/Complex Attention</em></th>
<th>A measure of the time it takes an individual to complete a mental task that requires processing several relevant variables.</th>
<th><em>Choice Reaction Time Task</em> [33]</th>
<th>In contrast to simple reaction time tasks Pullman <em>et al.</em>, 1988 [19] have demonstrated that PD patients appropriately treated with L-Dopa perform similarly to healthy age-matched controls on choice reaction time tasks. When levodopa ceases to have good effect, however, people with PD become much slower at these tasks.</th>
</tr>
</thead>
</table>
| **Visuospatial/Constructional Ability** | Skills that allow an individual to reconstruct a mental representation of a visual image and to accurately estimate distance, depth and angles. | - Benton’s Face Recognition [26]  
- Judgement Line Orientation [26]  
- Rey Osterrieth Complex Figure [36; 76]  
- Block Design [36; 76]  
- Bells Test [36; 76]  
- Visuospatial/Constructional Skill Summary Score [31] | Significant visuospatial/constructional deficits are often observed in PD patients’ performance. Dubois and Pillon (1997) [20] argue that this is due to the relatively high frontal load of many of these tasks, and that in more simple visuospatial tasks (e.g. Judgement line orientation), people with PD typically score in a similar range to healthy age-matched controls. |
| **Psychomotor Ability** | Psychomotor tests measure the speed and accuracy at which one can make purposeful movements in response to verbal instructions. | - Psychomotor Summary Score [31; 32]  
- Symbol Digit Modalities [35] | Psychomotor impairments can be evident even in the early to moderate stages of PD (Heikkila, 1998), and tasks like the Purdue Pegboard correlate highly with the motor severity scores in PD (Stavitsky, 2012). |

* Denotes that domain is not included in meta-analysis due to insufficient data
References (Table S2)


Table S3
Description of each sleep measure used in included studies

<table>
<thead>
<tr>
<th>Sleep Measure</th>
<th>Metrics</th>
<th>Studies That Used This Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polysomnography (PSG)</td>
<td>- The following measures are collected during PSG including; brain activity (Electroencephalogram) to determine sleep stages and arousals, eye movements (Electroocculogram) to determine sleep stages, muscle activation (Electromyogram) to determine wake periods, arousals or limb movements, breathing related effort (chest expansion sensor) to determine respiratory patterns, and blood oxygen level (pulse oximetry) to determine desaturation. Time-locked infrared video records movement that is synchronised to sleep stage data. Taken together, these measures are used to construct a complete picture of the sleep period - Definitively identifies sleep disorders such as REM sleep behaviour disorder (RBD) and Obstructive Sleep Apnoea (OSA) - Can be used to identify sub-clinical sleep problems like REM sleep without atonia and behavioural events during REM sleep (RBEs).</td>
<td>28; 29; 36; 37; 76; 77</td>
</tr>
<tr>
<td>Actigraphy</td>
<td>- Measures sleep onset latency, sleep fragmentation, time spent awake after sleep onset, sleep efficiency and total sleep time.</td>
<td>31</td>
</tr>
<tr>
<td>Clinical Interview</td>
<td>Administered by an experienced clinician or researcher, a clinical interview addresses international classification of sleep disorder guidelines for sleep disorders or a structured questionnaire. Where possible, the bed partner is included as the patient may not be aware of their night-time behaviour.</td>
<td>22; 23; 26; 27; 35; 48</td>
</tr>
<tr>
<td>Mayo Screening Questionnaire (Boeve et al., 2002)</td>
<td>A 16 item, clinician administered scale, that focuses on features of RBD such as ‘appearing to act out dreams’, ‘flailing arms and legs’ and whether a bed partner has ever been injured by this behaviour.</td>
<td>26</td>
</tr>
<tr>
<td>International Restless Legs Syndrome Study</td>
<td>A 10-item scale which measures the presence and severity of facets of RLS including: level of RLS discomfort in limbs, relief on moving limbs, severity of</td>
<td>30</td>
</tr>
<tr>
<td>Instrument</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Group Screening Questionnaire (IRLSSGS)</td>
<td>sleep disturbance from RLS symptoms and effects of RLS symptoms on mood.</td>
<td>The International Restless Legs Syndrome Study Group (2003)</td>
</tr>
<tr>
<td>RBD Screening Questionnaire (RBDSQ) Stiasny-Kolster et al., 2007</td>
<td>A ten item questionnaire with a binary response format which measures the presence of phenomena associated with RBD including vivid dreams, aggressive or action packed dreams, movements that awaken the individual, movements that injure the bed-partner and vocalisations during sleep.</td>
<td>23; 33</td>
</tr>
<tr>
<td>Scale for Outcomes in Parkinson’s disease (SCOPA-night) Marinus et al., (2003)</td>
<td>SCOPA-night consists of 5 items that measure aspects of sleep quality including difficulty falling asleep, been awake too often, lying awake too long, waking too early and had too little sleep.</td>
<td>33</td>
</tr>
<tr>
<td>Pittsburgh Sleep Quality Index (PSQI) Buysse et al., 1989</td>
<td>A questionnaire designed to measure sleep quality over the last month. It yields 7 component scores (sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication and daytime dysfunction) and a global sleep quality score. The scale consists of 19 self-rated items and 5 questions that are completed by the bed-partner or room-mate.</td>
<td>32</td>
</tr>
</tbody>
</table>