

Decoding the complexity of cognitive impairment in people with REM Sleep Behaviour Disorder

Beatrice Orso

beatrice27orso@gmail.com

Mayo Clinic

Alessio Martino

Luiss University of Rome

Andrea Brugnolo

University of Genoa

Nicola Girtler

University of Genoa

Pietro Mattioli

University of Genoa

Giulia Scalas

Luiss University of Rome

Francesco Famà

University of Genoa

Federico Massa

University of Genoa

Luigi Lorenzini

University of Genoa

Mattia Losa

University of Genoa

Matteo Pardini

University of Genoa

Dario Arnaldi

University of Genoa

Article

Keywords:

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11040184/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. BO: received a research grant from G.E. Healthcare. M.L received honoraria from Eli Lilly S.p.A. FM: received speaker honoraria from Eli Lilly. LL receives funding from the MSCA postdoctoral fellowship (#101204296) and Alzheimer Nederland (WE.03-2025-02). MP: received research support from Novartis and Nutricia, received fees from Novartis, Merck and Biogen. DA: received fees from Idorsia, Bruno, Italfarmaco and PIAM for lectures, consultation and board participation.

Abstract

Backgrounds

Isolated REM sleep behaviour disorder (iRBD) is a prodromal marker of α -synucleinopathies. Although cognitive impairment is consistently reported in iRBD, its assessment varies considerably across studies. This study characterized mild cognitive impairment (MCI) patterns in iRBD using machine learning and identified the minimal neuropsychological test battery needed for accurate diagnosis.

Methods

Comprehensive neuropsychological data from 216 assessments (162 with MCI, 54 without) in 110 iRBD patients were analysed. Four supervised classifiers (logistic regression, support vector machine, random forest, gradient boosting) were trained across three frameworks (raw scores, raw scores plus demographics, binarized scores) to distinguish MCI status. A multi-objective hyperparameter optimization approach then identified the minimal test subset preserving diagnostic accuracy.

Results

Of MCI cases, 30% were single-domain and 70% multi-domain, with attention/executive functions and memory most frequently impaired. Binarized neuropsychological scores yielded the best classification performance, with support vector machine achieving 94.7% accuracy. Feature importance analyses consistently highlighted episodic memory, executive, and visuospatial measures. The optimization framework identified a mean of 8 informative tests spanning these domains, achieving comparable diagnostic accuracy to the full battery.

Conclusions

Cognitive impairment is common in iRBD and predominantly multidomain. Machine learning approaches applied to neuropsychological data can accurately classify MCI status and identify a concise, clinically feasible set of tests with high diagnostic value. These findings support comprehensive, multidomain neuropsychological assessment in iRBD and may inform optimized risk stratification in prodromal α -synucleinopathies.

What is already known on this topic

RBD subjects may experience cognitive decline during the course of the disease. The cognitive profile is usually multidomain, but no consensus exists on the optimal test battery for detecting mild cognitive impairment (MCI) in this population.

What this study adds: Machine learning approaches applied to neuropsychological data can accurately identify MCI and highlight a concise set of cognitive tests with high diagnostic value.

How this study might affect research, practice or policy: These findings have direct implications for clinical neuropsychological assessment, as they suggest that shorter assessments can remain clinically informative if they maintain representation across all four cognitive domains, rather than oversampling a single cognitive system.

Introduction

Isolated REM sleep behaviour disorder (iRBD) is a parasomnia characterized by dream-enactment behaviours associated with the loss of physiological muscle atonia during REM sleep.¹ Compelling longitudinal evidence indicates that iRBD represents one of the most robust prodromal markers of α -synucleinopathies, with more than 80% of patients eventually phenoconverting to Parkinson's disease (PD), dementia with Lewy bodies (DLB), or multiple system atrophy (MSA) over time.^{1,2} As such, iRBD provides a unique clinical window into the earliest stages of neurodegenerative disease, when disease-modifying interventions may be most effective.³

Cognitive impairment in absence of overt parkinsonism or dementia is increasingly recognized in iRBD.⁴ Cross-sectional and longitudinal studies have consistently shown that people with iRBD perform worse than healthy controls across multiple cognitive domains, particularly attention, executive functions, visuospatial abilities, and episodic memory.⁴⁻⁶ Importantly, the presence of mild cognitive impairment (MCI) in iRBD has been associated with a higher risk of phenoconversion, especially toward DLB, suggesting that cognitive dysfunction may represent a critical marker of disease severity and trajectory.^{1,7,8} However, despite increasing recognition of the clinical significance of cognitive impairment in people with iRBD, substantial variability in neuropsychological assessments, MCI diagnostic criteria employed, and analytical methodologies across studies continues to exist.

Cognitive impairment in iRBD populations is often assessed using screening tests (mostly the Mini-Mental State Examination -MMSE-⁹ or the Montreal Cognitive Assessment -MoCA-¹⁰). However, a comprehensive neuropsychological assessment is more sensitive for diagnosing MCI than screening tests alone.¹¹ Indeed, the detection of MCI through comprehensive cognitive batteries serves as a more robust risk factor for phenoconversion, particularly for developing dementia, in iRBD than screening tests for overall cognition.¹ On the contrary, comprehensive neuropsychological batteries are time-consuming and require trained neuropsychologists, and there is no clear consensus yet on the extent of cognitive assessment in people with iRBD. It remains to be determined whether all cognitive domains should always be explored or whether a small subset of tests focusing on specific cognitive functions might be sufficient.¹²

In recent years, machine learning approaches have emerged as powerful tools for identifying complex, multivariate patterns in clinical and neuropsychological data.¹³ Such methods have demonstrated promise for early detection of cognitive impairment and for predicting disease progression in

neurodegenerative disorders.¹⁴ Their application to iRBD populations may help in shedding light on the intrinsic heterogeneity of such disorders, especially regarding cognitive impairment.

The aims of the present study are: (i) to characterize MCI in iRBD patients using multiple supervised machine learning models, and (ii) to identify which neuropsychological domains and tests are most informative for MCI detection. The ultimate goal is to inform the design of comprehensive yet clinically feasible assessments that cover all relevant cognitive domains.

Methods

Patients

Patients were enrolled in IRCCS Azienda Ospedaliera Metropolitana, Ospedale Policlinico San Martino, Genova (Italy), between January 2013 and December 2024. The inclusion criterion was a diagnosis of iRBD, confirmed by video-polysomnography (vPSG).¹⁵ Exclusion criteria for all participants were the presence of other brain diseases, assessed through Brain MRI or CT¹⁶, presence of white matter hyperintensities if the Wahlund scale was > 1 for each brain region,¹⁷ as well as the presence of overt parkinsonism or dementia. At baseline, all patients underwent the Movement Disorder Society Unified Parkinson Disease rating scale, third part (MDS-UPDRS-III) to quantify motor impairment, the MMSE or the MoCA, as a measure of global cognitive functioning.

Video Polysomnography Recording

Sleep scoring was performed following current criteria.¹⁵ PSG derivations were placed according to recommended rules^{15,18,19} to evaluate sleep features, respiratory, cardiac, and limb events. Chin electromyography (EMG) was used to visually quantify REM sleep without atonia (RWA), and the 'any' REM percentage was used.^{20,21} If used, patients were asked to withdraw melatonin, hypnotic medications, and antidepressant drugs for two weeks before the recording.

Neuropsychological assessment

Cognitive status of patients was assessed using a comprehensive neuropsychological battery, including at least two tests for each of the main cognitive domains (memory, attention and executive functions, visuospatial abilities, and language).¹⁶ To evaluate the presence of MCI, impairment must be observed in at least two tests, either within the same cognitive domain (single-domain MCI) or across different domains (multiple-domain MCI).²² Moreover, the cognitive change needs to be endorsed by the person or informant, and it should be considered relevant at clinical assessment by a neurologist with specific expertise in cognition.

Clinical information on activities and instrumental activities of daily living and cognitive assessments was evaluated at baseline and prospectively every two years.

In detail, the following tests for each domain were used:

1. Memory: Corsi Span^{23,24}, Digit Span²⁴, Rey Auditory Verbal Learning Test (RAVLT, immediate and delayed recall),²⁵ and Craft Story.²⁵
2. Attention and Executive functions: Trail Making Test A and B (TMTA-TMTB)²⁶, Symbol Digit task²³, Stroop Colour Word²⁷, Wisconsin Card Sorting Test (WCST).²⁸
3. Visuospatial abilities: Clock Drawing Test²⁹, Constructional apraxia (simple copy and copy with guiding landmarks).³⁰
4. Language: semantic verbal fluency and phonemic verbal fluency.²⁵

The neuropsychological battery was repeated every two years, according to our clinical routine.

Standard protocol approval

All patients provided written consent to the study. The local ethics committee approved the study protocol (CER Liguria: 105/2023-DB id 13027), and all participants signed an informed consent form in compliance with the Helsinki Declaration of 1975.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Study overview and statistical analysis

The neuropsychological test scores were used in the following analysis frameworks as follows: a) the raw, uncorrected scores, which were treated as continuous variables; and b) the binarized 0–1 scores, where 0 indicated tests considered impaired, while 1 indicated tests considered normal, based on currently published cut-off and normative data for each test.

To study demographics and clinical differences between iRBD patients with and without MCI, a Pearson's Chi-squared test was used, while a descriptive Analysis of Variance (ANOVA) was performed to summarize differences in cognitive performance. Furthermore, a MANOVA was used to assess whether MCI status had a global effect on the joint distribution of cognitive variables. The presence of missing values was addressed using a K-Nearest Neighbours (K-NN) imputation strategy.

In the first step of this study, we aimed to characterize MCI in patients with RBD by exploring the combinations of neuropsychological tests that are consistently altered. To do this, four supervised

machine learning models were trained and evaluated for binary classification of patients who were classified as having MCI versus those who did not:

1. Logistic Regression (a statistical model that estimates the probability that a given input belongs to a particular class);
2. Support Vector Machine (SVM, a model capable of handling non-linear decision boundaries using kernel functions);
3. Random Forest (an ensemble of decision trees that aggregates predictions to reduce overfitting and improve generalizability);
4. Gradient Boosting (a boosting technique that builds models sequentially to correct the errors of previous ones, often yielding high predictive accuracy).

These analysis models were applied across three experimental frameworks:

Experiment 1) Considering the raw neuropsychological test scores;

Experiment 2) Considering both demographic information (age, sex and education) and raw neuropsychological test scores;

Experiment 3) Considering the binarized (0–1) neuropsychological test scores.

Each experiment involved a holdout set (20% of the assessments), with the remaining 80% used for training the models. Within the training set, a 5-fold cross-validation and grid search were conducted to estimate hyperparameters. The best-performing model was then evaluated on the holdout set. This process was repeated across five independent runs to simulate real-world prediction scenarios and to assess the robustness of patient classification. To enhance interpretability, feature importance for each neuropsychological test was extracted from the Random Forest and SVM models.

As a second step, we applied a multi-objective optimisation approach to minimise both the number of neuropsychological tests used and the classification error, aiming to identify the minimum set of tests required to identify MCI patients. A hyperparameter optimisation framework (optuna: <https://github.com/optuna/optuna>)³¹ was tested on both continuous and binary neuropsychological variables, using the SVM classifier. This process was repeated across 20 independent runs to identify the smallest subset of features that could achieve optimal performance. All neuropsychological assessments were treated as independent entries in the statistical analyses to increase the number of available assessments and boost statistical power.

Results

Group characteristics and Neuropsychological Features

A total of 110 subjects were enrolled (mean age: 69 ± 7.57 ; 78% males), of which 32 eventually phenocconverted to an overt α -synucleinopathy during a mean follow-up time of 27 ± 24 months, either

PD or DLB based on the clinical syndrome seen at follow-up, according to current criteria^{32,33}, and 78 subjects remained free from overt parkinsonism and/or dementia at the last available follow-up.

Of the 110 iRBD patients included with a baseline neuropsychological assessment, of which 56% (61/110) met the MCI criteria, 65 had one cognitive follow-up, 38 had two cognitive follow-ups, and 17 had three cognitive follow-ups. Given that the MCI diagnosis is based on a single cognitive assessment, regardless of prior assessment results, this approach is methodologically appropriate. Seven patients who had developed dementia during the follow-up time were excluded from the analysis, so to focus on the prodromal stage. Thus, a total of 216 neuropsychological assessments were retained in the study, consisting of 162 iRBD patients with MCI and 54 iRBD patients without MCI (Table 1).

Table 1
Demographic and clinical characteristics of iRBD subjects with and without MCI.

Age at assessment	RBD with MCI (n = 162)	RBD without MCI (n = 54)	p-value
	71.9 ± 7	68.5 ± 7.2	0.001
Sex			0.2
Males	132 (81%)	40 (74%)	
Females	30 (19%)	14 (26%)	
Education	9.8 ± 4	11.8 ± 4.1	0.002
MCI Type			
Single-domain	49 (30%)	/	
Multi-domain	113 (70%)	/	
MMSE	27.3 ± 2.6	29 ± 1.1	< 0.001

In the iRBD with MCI group, 49 (30%) assessments were classified as “Single-domain”, while 113 (70%) were classified as “Multi-domain” MCI. Among patients with single-domain MCI, the Attention and Executive Function domain was found impaired in 33 evaluations, the Memory domain in 12, the Visuo-spatial domain in 3 and the Language domain in 1 evaluation.

Of the 113 assessments labelled as Multi-domain MCI, 55 highlighted two impaired cognitive domains (Memory + Attention and Executive functions = 44; Attention and Executive functions + Visuo-spatial functions = 5; Memory + Visuo-spatial functions = 5; Attention and Executive functions + Language = 1), 43 highlighted three impaired cognitive domains (Memory + Attention and Executive functions + Visuo-spatial = 27; Memory + Attention and Executive functions + Language = 13; Attention and Executive functions + Visuo-spatial + Language = 2; Memory + Visuo-spatial + Language = 1). Lastly, fifteen evaluations highlighted impairment across all four cognitive domains (Fig. 1). Of note, Memory and

attention/executive functions were the most frequently affected domains across multi-domain MCI patterns.

Additionally, taking into consideration patients with both single-domain and multi-domain MCI, Attention and Executive functions were rated as impaired 140 times, Memory 117 times, Visuo-spatial 58 times, and Language 33 times (Fig. 1).

Descriptive analysis

When comparing RBD with MCI and RBD without MCI patients across cognitive measures, both ANOVA and Kruskal–Wallis analyses found that all variables reached statistical significance ($p < 0.05$), confirming that the observed differences are not artefacts of distributional assumptions. Moreover, the MANOVA analysis confirmed the presence of widespread cognitive differences between MCI and non-MCI groups, distributed across multiple neuropsychological domains (Wilks' Lambda: $F(19.204) = 44.19$, $p < 0.0001$).

Attention and Executive functions tasks (namely: TMTA, TMTB, Stroop Colour, Stroop Colour-Word) showed the strongest difference, alongside working memory measures (Digit Span, Corsi Span) and episodic memory (Rey Immediate and Delayed Recall). These findings indicate that the earliest cognitive impairments in RBD with MCI patients are distributed across multiple domains rather than confined to a single deficit.

Classification performance across experimental frameworks

For each experiment, the dataset was split into a training set (80%, $n = 174$) and a holdout test set (20%, $n = 43$); classification performance (accuracy, sensitivity, specificity) for all four models across the three experimental frameworks is summarized in Table 2.

Table 2

Classification performance of four machine learning models across the three experimental frameworks. Values are mean $\pm$ standard deviation across 5 independent runs.

	Model	Accuracy	Sensitivity	Specificity
Experiment 1: raw neuropsychological tests scores	Gradient Boosting	85.78 $\pm$ 7.3	89.23 $\pm$ 7.4	81.05 $\pm$ 7.98
	Random Forest	85.33 $\pm$ 2.98	90 $\pm$ 5.83	78.95 $\pm$ 6.45
	Support Vector Machine	84.44 $\pm$ 7.03	89.23 $\pm$ 7.4	77.89 $\pm$ 9.42
	Logistic Regression	81.78 $\pm$ 5.31	85.38 $\pm$ 5.01	76.84 $\pm$ 10.26
Experiment 2: demographics and raw neuropsychological tests scores	Gradient Boosting	84.89 $\pm$ 5.31	86.15 $\pm$ 6.44	83.16 $\pm$ 9.42
	Random Forest	84 $\pm$ 4.82	86.92 $\pm$ 6.99	80 $\pm$ 5.77
	Support Vector Machine	83.56 $\pm$ 6.21	86.92 $\pm$ 6.99	78.95 $\pm$ 8.32
	Logistic Regression	81.78 $\pm$ 6.92	86.15 $\pm$ 7.98	75.79 $\pm$ 7.06
Experiment 3: binarized neuropsychological tests scores	Gradient Boosting	88 $\pm$ 6	86.92 $\pm$ 6.44	89.47 $\pm$ 6.45
	Random Forest	92 $\pm$ 6	92.31 $\pm$ 6.66	91.58 $\pm$ 7.98
	Support Vector Machine	94.67 $\pm$ 2.53	94.62 $\pm$ 2.11	94.74 $\pm$ 6.45
	Logistic Regression	93.33 $\pm$ 4.16	92.31 $\pm$ 7.2	94.74 $\pm$ 5.26

Table 3

Classification performance of the hyperparameter-optimized minimal test subsets, across 20 independent runs, using the **A)** uncorrected, raw neuropsychological test scores, and the **B)** binarized neuropsychological test scores.

	Metric	Accuracy	Sensitivity	Specificity
A	Mean	81 $\pm$ 4.4	80.4 $\pm$ 6.3	81.8 $\pm$ 9.5
	Best run	89	88.5	89.5
B	Mean	89.9 $\pm$ 5.8	87.5 $\pm$ 8.2	93.2 $\pm$ 6.3
	Best run	100	100	100

Using raw neuropsychological scores alone (Experiment 1) or combined with demographic data (Experiment 2), Gradient Boosting slightly outperformed Random Forest and Support Vector Machine, while Logistic Regression showed the lowest performance; when adding demographic variables, the model did not meaningfully improve classification over raw scores alone. Binarizing neuropsychological scores (Experiment 3) improved performance overall: Support Vector Machine achieved the best accuracy, exceeding 94%, followed by Logistic Regression, Random Forest, and Gradient Boosting. Across all three experiments, sensitivity exceeded specificity, indicating a greater ability to detect MCI-positive patients.

Feature importance

Feature importance analysis showed that the Random Forest model ranked TMTB-A, Rey Auditory Verbal Learning Test (immediate recall), Craft Story, TMTB, Symbol Digit, Stroop Colour, and TMTA as the most informative predictors of MCI status in when considering demographics and neuropsychological raw test scores (Experiment 2) (Fig. 2A). Using binarized scores (Experiment 3), the ten most important Random Forest features were TMTB, Craft Story, Rey Auditory Verbal Learning Test (immediate and delayed recall), Corsi Span, Digit Span, Stroop Colour-Word, Constructional apraxia (with guiding landmarks), TMTB-A, and Clock Drawing Test (Fig. 2B).

In Experiment 2, the Support Vector Machine model instead prioritized Craft Story, education, age at assessment, Rey Auditory Verbal Learning Test (immediate recall), Stroop Colour-Word, Clock Drawing Test, TMTB-A, and TMTB (Fig. 2C). In Experiment 3, the Support Vector Machine model selected largely the same tests (Craft Story, Rey Auditory Verbal Learning Test immediate recall, TMTB, Stroop Colour, Corsi Span, Rey Auditory Verbal Learning Test delayed recall, Constructional apraxia, Digit Span, Stroop Colour-Word, Clock Drawing Test), reinforcing memory, executive, and visuospatial measures as the most robust predictors of MCI status (Fig. 2D). Across both models, episodic memory (Story Recall, Rey Auditory Verbal Learning Test immediate recall) and executive-function (Trail Making Test) measures were consistently top-ranked.

Minimal test-battery selection: hyperparameter optimization

The multi-objective optimization framework, applied to the Support Vector Machine model in Experiments 1 and 3, identified the smallest test subset that preserved diagnostic accuracy for both uncorrected and binarized scores (Table 3A-B; Fig. 3A-B).

For uncorrected scores, the most frequently selected tests across 20 runs were Craft Story (16/20), Corsi Span (14/20), Rey Auditory Verbal Learning Test immediate recall (12/20), and Constructional apraxia with guiding landmarks (9/20), with a mean of 6.3 ± 1.8 tests selected per run.

For binarized scores, the top selections were Craft Story (19/20), Clock Drawing Test (16/20), Rey Auditory Verbal Learning Test immediate recall (13/20), and Constructional apraxia with guiding landmarks (13/20), with a mean of 8.1 ± 2.0 tests selected per run. In both frameworks, the optimized

subsets spanned memory, executive, and visuospatial domains, consistent with the feature-importance analyses above.

We performed a post hoc analysis, applying all the aforementioned statistical models only to baseline data. The overall results did not change significantly, showing a slight increase in performance variance, likely due to the reduced sample size and more unstable models.

Discussion

In this study, we showed that cognitive impairment is both highly prevalent and predominantly multidomain, even at the prodromal stage of α -synucleinopathies. More than half of our cohort of iRBD patients met the criteria for MCI at baseline, with more than one-third of this exhibiting impairment across three or four cognitive domains. These findings reinforce the notion that cognitive deficits in iRBD mirror widespread network dysfunction, already present at the prodromal stage. A key clinical implication of these findings is that cognitive assessment in iRBD should systematically cover all four major cognitive domains (namely, memory, attention and executive functions, visuospatial abilities, and language) regardless of the specific tests chosen. Indeed, no single domain alone captures the full complexity of cognitive impairment in this clinical population.

The most frequently observed cognitive profile was characterized by deficits in attention and executive functions, followed by deficits in working memory, episodic memory, and visuospatial abilities. This pattern aligns with previous studies describing a non-amnestic or multi-domain MCI phenotype in iRBD, which differs from the predominantly amnestic profile, typically associated with prodromal Alzheimer's disease.^{6,34,35} The high frequency of executive–attentional and visuospatial impairments is particularly noteworthy, as these domains are closely linked to fronto-striatal and posterior cortical networks known to be affected early in α -synucleinopathies.^{36,37}

Using supervised machine learning, we demonstrate that neuropsychological tests reliably distinguishing iRBD patients based on their MCI status span the four main clinical domains. Notably, no single domain or a fixed combination of domains characterizes MCI in people with iRBD; rather, impaired domains co-occur in highly variable patterns across patients. This heterogeneity is clinically meaningful: while attention and executive functions were the most frequently impaired domains, memory, visuospatial, and, less commonly, language deficits also contributed substantially to MCI classification. These findings are in line with prior studies describing the neuropsychological profile of iRBD.^{6,38}

Among the four domains examined, language emerged as the least frequently impaired, both in overall prevalence and in feature selection analyses. However, a minority of patients exhibited isolated or predominant language-domain deficits, and the hyperparameter optimization analysis confirmed that excluding language tests reduced overall diagnostic accuracy. This observation underscores that, while language impairment is uncommon in iRBD, its presence may carry specific diagnostic or prognostic significance and therefore warrant systematic assessment as part of any comprehensive

neuropsychological battery. Across multiple experimental frameworks, ensemble-based methods (Random Forest and Gradient Boosting) and Support Vector Machines consistently outperformed logistic regression, highlighting the importance of capturing non-linear relationships and complex interactions among cognitive measures. Across models, the tests most consistently identified as informative were the Craft Story and Rey Auditory Verbal Learning Test (episodic memory), Trail Making Test B and Stroop tasks (executive functions), Digit and Corsi Span (attention and working memory), Clock Drawing Test and constructional apraxia (visuospatial abilities). The convergence of these tests across multiple frameworks and feature selection methods lends robustness to their role as cognitive markers in prodromal synucleinopathies. Importantly, the consistent selection of tests from multiple cognitive domains across all analytical frameworks reinforces a core clinical message: a neuropsychological battery restricted to one or two domains' risks missing an important proportion of MCI patients. The cognitive profile of iRBD is inherently multidomain, and any cognitive assessment protocol should reflect this complexity. With the aim of striking a pragmatic balance between comprehensive assessment and clinical feasibility, we conducted a hyperparameter optimization analysis to identify the minimal yet informative subset of neuropsychological tests that accurately classify MCI in iRBD patients. Using a multi-objective optimization framework, we show that good diagnostic performance can be achieved with a mean of 8 tests across the different analysis runs. Notably, the selected tests span attention, executive function, memory, and visuospatial function. The prominence of episodic memory and visuospatial measures in the optimized test sets is of particular interest, as these domains are often associated with posterior cortical dysfunction and may signal an increased risk for DLB rather than PD.^{8,38–40} These findings support the use of detailed, multidomain neuropsychological assessments for the accurate diagnosis of MCI in people with iRBD. Whether the specific pattern of domain involvement can additionally serve as a marker of phenoconversion risk, and toward which clinical syndrome, remains an important question for future longitudinal analyses. Conversely, pure language tests were not selected in this analysis. However, we found that a minority of patients had specific language-domain impairment. Indeed, the hyperparameter optimization achieved excellent, though not perfect, accuracy in identifying MCI among patients with iRBD. This result suggests that a neuropsychological assessment not including language tests might diagnose the vast majority but not all patients with MCI.

The integration of machine learning techniques provides novel insights into test selection and diagnostic optimization.

Several limitations should be acknowledged. First, external validation in independent cohorts is needed to confirm the generalizability of the machine learning models. Then, including biomarker data (e.g., dopamine transporter imaging, seed amplification assay methods) may further enhance predictive accuracy and should be explored in future studies.

In conclusion, cognitive impairment in iRBD is common, multidomain, and detectable well before the onset of overt neurodegenerative disease. Machine learning approaches applied to neuropsychological data can accurately identify MCI and highlight a concise set of cognitive tests with high diagnostic value.

These findings have direct implications for clinical neuropsychological assessment. Routine practice often encourages abbreviated cognitive batteries due to time constraints, leading to a focus on selected cognitive domains. Our results suggest that shorter assessments can remain clinically informative if they maintain representation across all four cognitive domains, rather than oversampling a single cognitive system, as cognitive impairment in people with iRBD does not follow a fixed pattern. These findings have important implications for early diagnosis, risk stratification, and the design of targeted interventions in prodromal α -synucleinopathies.

Declarations

Competing Interests

BO: received a research grant from G.E. Healthcare. M.L received honoraria from Eli Lilly S.p.A. FM: received speaker honoraria from Eli Lilly. LL receives funding from the MSCA postdoctoral fellowship (#101204296) and Alzheimer Nederland (WE.03-2025-02). MP: received research support from Novartis and Nutricia, received fees from Novartis, Merck and Biogen. DA: received fees from Idorsia, Bruno, Italfarmaco and PIAM for lectures, consultation and board participation.

Author Contribution

BO: Execution, analyses, drafting of manuscript, reviewing and editing the final version of the manuscript, AM: Execution, analyses, reviewing and editing the final version of the manuscript, AB: reviewing and editing the final version of the manuscript, NG: reviewing and editing the final version of the manuscript, PM: reviewing and editing the final version of the manuscript, GS: execution and analyses, FF: reviewing and editing the final version of the manuscript, FM: reviewing and editing the final version of the manuscript, LL: reviewing and editing the final version of the manuscript, ML: reviewing and editing the final version of the manuscript, MP: reviewing and editing the final version of the manuscript, DA: Execution, design, drafting of manuscript, reviewing and editing the final version of the manuscript.

Acknowledgement

This work was supported by: NEXTGENERATIONEU (NGEU), funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022) for the activities of Federico Massa. #NEXTGENERATIONEU (NGEU), funded by the European Union, NRRP M6C2 – Investment 2.1 Enhancement and strengthening of biomedical research in the NHS (PNRR-MCNT2-2023-12377527 - CUP C13C23001070006) for the activities of Matteo Pardini. Bando Ricerca Finalizzata RF-2021-12374240, funded by the Italian Ministry of Health, for the activities of Dario Arnaldi. The views and opinions expressed are those of the authors alone and do not

necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

1. Postuma, R. B. *et al.* Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: A multicentre study. *Brain* 142, 744–759 (2019).
2. Galbiati, A., Verga, L., Giora, E., Zucconi, M. & Ferini-Strambi, L. The risk of neurodegeneration in REM sleep behavior disorder: A systematic review and meta-analysis of longitudinal studies. *Sleep Med. Rev.* 43, 37–46 (2019).
3. Videnovic, A. *et al.* Clinical trials in REM sleep behavioural disorder: challenges and opportunities. *J. Neurol. Neurosurg. Psychiatry* 91, 740–749 (2020).
4. Leitner, C. *et al.* Neuropsychological Changes in Isolated REM Sleep Behavior Disorder: A Systematic Review and Meta-analysis of Cross-sectional and Longitudinal Studies. *Neuropsychol. Rev.* <https://doi.org/10.1007/s11065-022-09572-1> (2023) doi:10.1007/s11065-022-09572-1.
5. Yoon, E. J. *et al.* Brain metabolism related to mild cognitive impairment and phenocconversion in patients with isolated REM sleep behavior disorder. *Neurology* 98, e2413–e2424 (2022).
6. Gagnon, J.-F., Bertrand, J.-A. & Génier Marchand, D. Cognition in rapid eye movement sleep behavior disorder. *Front. Neurol.* 3, 82 (2012).
7. Joza, S. *et al.* Prodromal dementia with Lewy bodies in REM sleep behavior disorder: A multicenter study. *Alzheimer's & Dementia* (2023).
8. Joza, S. *et al.* Progression of clinical markers in prodromal Parkinson's disease and dementia with Lewy bodies: a multicentre study. *Brain* awad072 (2023) doi:10.1093/brain/awad072.
9. Folstein, M. F., Robins, L. N. & Helzer, J. E. The mini-mental state examination. *Arch. Gen. Psychiatry* 40, 812 (1983).
10. Nasreddine, Z. S. *et al.* The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699 (2005).
11. Diniz, B. S. D. O. *et al.* Mild cognitive impairment: cognitive screening or neuropsychological assessment? *Revista Brasileira de Psiquiatria* 30, 316–321 (2008).
12. Fiamingo, G. *et al.* Neuropsychological evaluation of phenocconversion risk in REM sleep behaviour disorder: A scoping review. *J. Sleep Res.* 32, (2023).
13. Shah, M. *et al.* Neuropsychological detection and prediction using machine learning algorithms: a comprehensive review. *Intelligent Medicine* 4, 177–187 (2024).
14. Pellegrini, E. *et al.* Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 10, 519–535 (2018).
15. Berry, R. B. *et al.* AASM scoring manual updates for 2017 (version 2.4). *Journal of Clinical Sleep Medicine* vol. 13 665–666 Preprint at (2017).

16. Mattioli, P. et al. Cuneus/precuneus as a central hub for brain functional connectivity of mild cognitive impairment in idiopathic REM sleep behavior patients. *Eur. J. Nucl. Med. Mol. Imaging* <https://doi.org/10.1007/s00259-021-05205-6> (2021) doi:10.1007/s00259-021-05205-6.
17. Wahlund, L.-O. et al. A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke* 32, 1318–1322 (2001).
18. Yoon, E. J. et al. Brain metabolism related to mild cognitive impairment and phenoconversion in patients with isolated REM sleep behavior disorder. *Neurology* 98, e2413–e2424 (2022).
19. Lee, J.-Y. et al. Nonmotor and dopamine transporter change in REM sleep behavior disorder by olfactory impairment. *J. Mov. Disord.* 12, 103 (2019).
20. Frauscher, B. et al. Normative EMG values during REM sleep for the diagnosis of REM sleep behavior disorder. *Sleep* 35, 835–847 (2012).
21. Orso, B. et al. Polysomnographic correlates of sleep disturbances in de novo, drug naïve Parkinson's Disease. *Neurological Sciences* 1–6 (2021).
22. Litvan, I. et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Movement disorders* 27, 349–356 (2012).
23. Spinnler, H. & Tognoni, G. Italian Group on the Neuropsychological Study of Ageing: Italian standardization and classification of neuropsychological tests. *Ital J Neurol Sci* 6, 1–120 (1987).
24. Orsini, A. et al. Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *The italian journal of neurological sciences* 8, 537–548 (1987).
25. Novelli, G., Papagno, C., Capitani, E. & Laiacona, M. Tre test clinici di ricerca e produzione lessicale. Taratura su soggetti normali. *Arch. Psicol. Neurol. Psichiatr.* (1986).
26. Amodio, P. et al. Variability of trail making test, symbol digit test and line trait test in normal people. A normative study taking into account age-dependent decline and sociobiological variables. *Aging Clin. Exp. Res.* 14, 117–131 (2002).
27. Brugnolo, A. et al. An updated Italian normative dataset for the Stroop color word test (SCWT). *Neurological Sciences* 37, 365–372 (2016).
28. Laiacona, M., Inzaghi, M. G., De Tanti, A. & Capitani, E. Wisconsin card sorting test: a new global score, with Italian norms, and its relationship with the Weigl sorting test. *Neurological Sciences* 21, 279–291 (2000).
29. Caffarra, P. et al. Italian norms for the Freedman version of the Clock Drawing Test. *J. Clin. Exp. Neuropsychol.* 33, 982–988 (2011).
30. Carlesimo, G. A. et al. The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. *Eur. Neurol.* 36, 378–384 (1996).
31. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2623–2631 (2019) doi:10.1145/3292500.3330701.

32. Postuma, R. B. *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Movement disorders* 30, 1591–1601 (2015).
33. McKeith, I. G. *et al.* Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 89, 88–100 (2017).
34. Gagnon, J. F., Bourgouin, P. A., De Roy, J. & Marchand, D. G. Neuropsychological aspects: Cognition in RBD. *Rapid-eye-movement sleep behavior disorder* 491–507 (2018) doi:10.1007/978-3-319-90152-7_34.
35. Goldman, J. G. *et al.* Cognitive impairment in Parkinson's disease: a report from a multidisciplinary symposium on unmet needs and future directions to maintain cognitive health. *NPJ Parkinsons Dis.* 4, 1–11 (2018).
36. Boccalini, C. *et al.* Metabolic connectivity of resting-state networks in alpha synucleinopathies, from prodromal to dementia phase. *Front. Neurosci.* 16, 930735 (2022).
37. Rahayel, S. *et al.* Brain atrophy in prodromal synucleinopathy is shaped by structural connectivity and gene expression. *Brain* 145, 3162–3178 (2022).
38. Postuma, R. B., Gagnon, J. F., Vendette, M. & Montplaisir, J. Y. Markers of neurodegeneration in idiopathic rapid eye movement sleep behaviour disorder and Parkinson's disease. *Brain* 132, 3298–3307 (2009).
39. Marchand, D. G., Montplaisir, J., Postuma, R. B., Rahayel, S. & Gagnon, J.-F. Detecting the cognitive prodrome of dementia with Lewy bodies: a prospective study of REM sleep behavior disorder. *Sleep* 40, (2017).
40. Génier Marchand, D. *et al.* How does dementia with Lewy bodies start? prodromal cognitive changes in REM sleep behavior disorder. *Ann. Neurol.* 83, 1016–1026 (2018).

Figures

UpSet Plot: Overall MCI Cognitive Profiles (Single & Multi-domain)

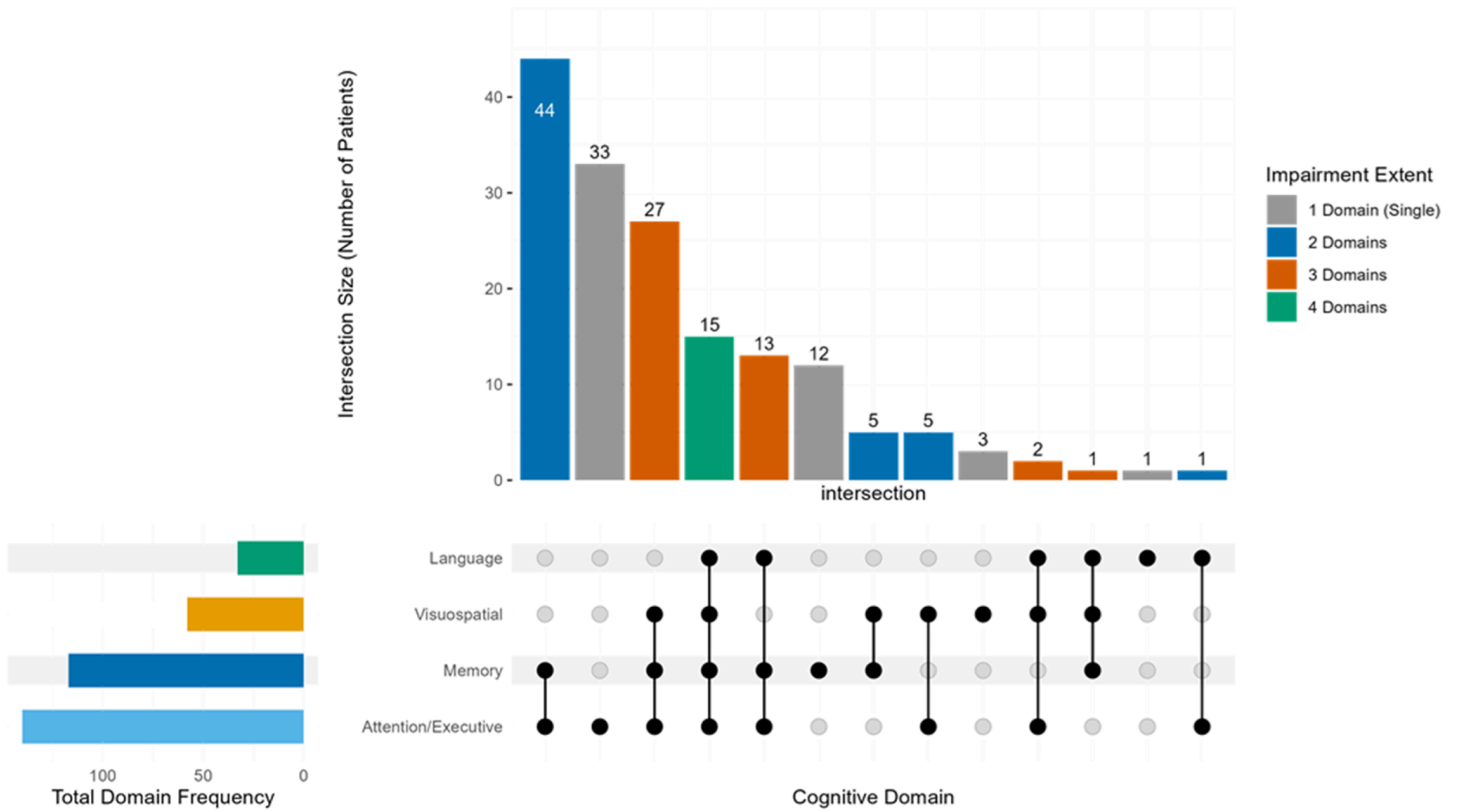


Figure 1

Cognitive domain impairment patterns in iRBD patients with MCI (n = 162). UpSet plot showing the number of assessments per combination of impaired cognitive domains (main panel), colour-coded by the number of domains involved (1-4). The left-hand bars show the total frequency of impairment per cognitive domain across all MCI assessments.

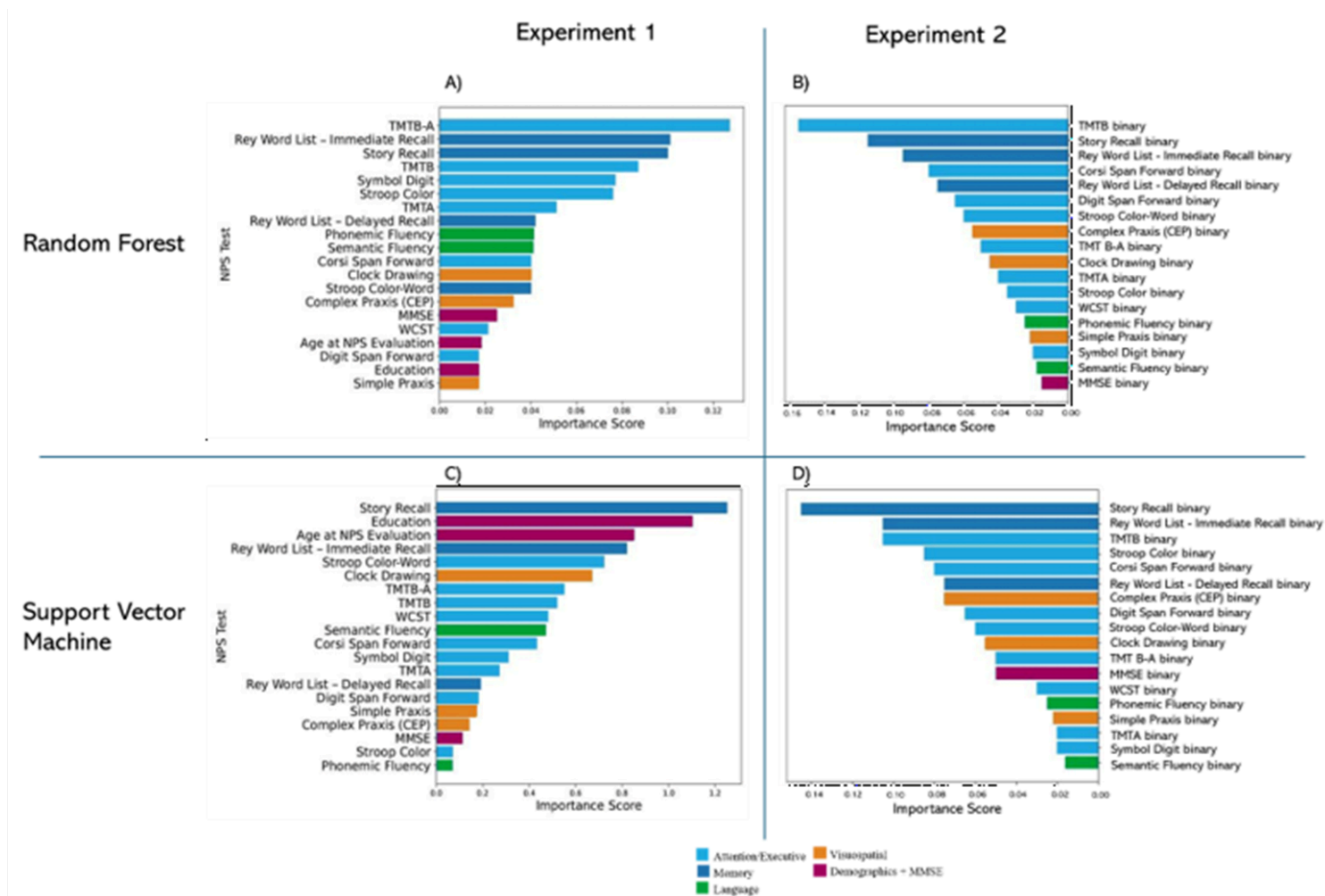


Figure 2

Feature importance for MCI classification, Experiment 1 (demographics + raw neuropsychological scores) and Experiment 2 (binarized neuropsychological scores). **A)** Random Forest in Experiment 1. **B)** Random Forest in Experiment 2. **C)** Support Vector Machine in Experiment 1. **D)** Support Vector Machine in Experiment 2.

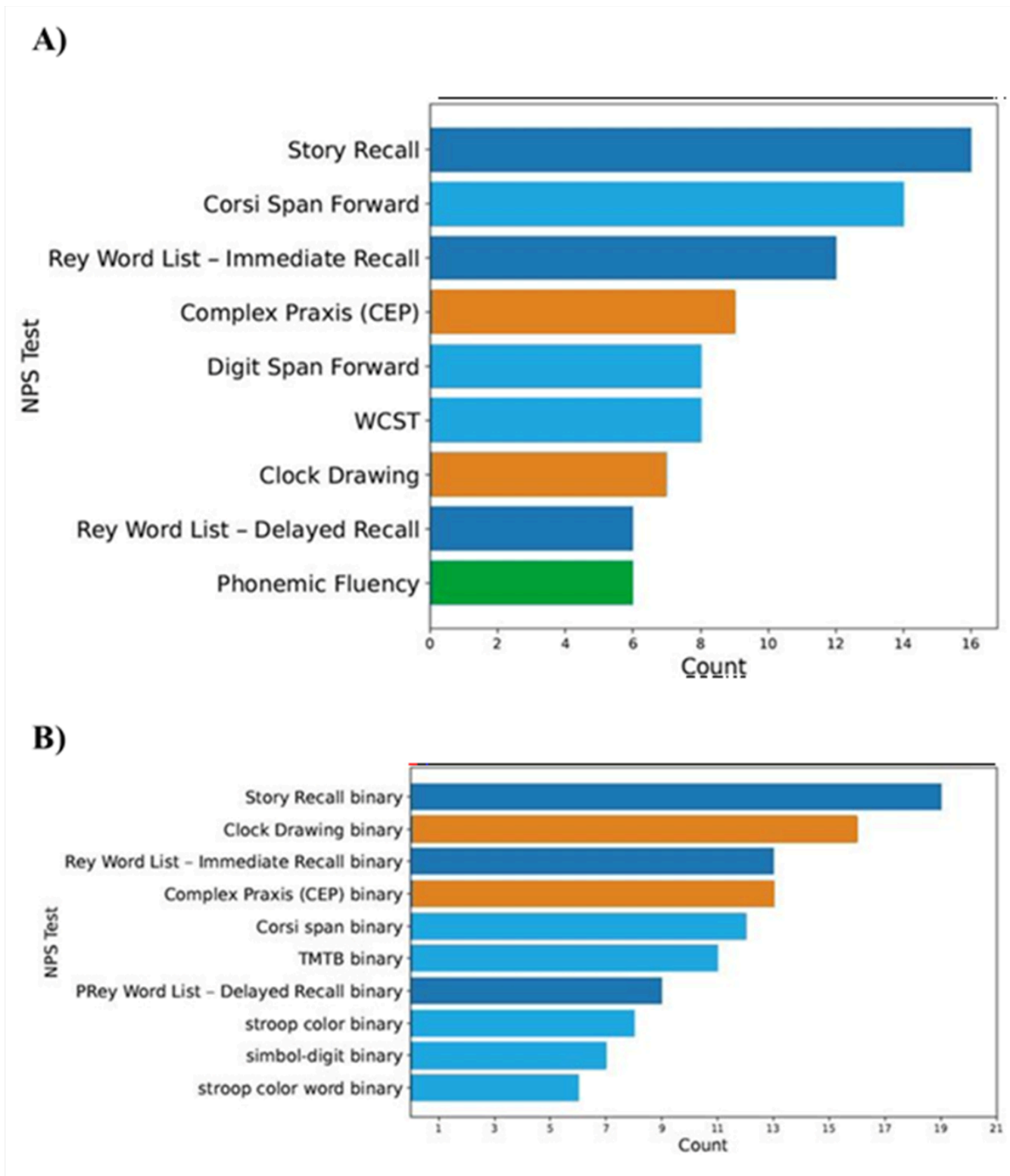


Figure 3

Test-selection frequency from the hyperparameter optimization analysis (20 runs), identifying the minimal informative neuropsychological battery. **A)** Uncorrected/raw scores. **B)** Binarized scores.