Supplementary Material

Cognitive Performance and Learning Parameters Predict Response to Working Memory Training in Parkinson's Disease

Supplementary Material 1.

Details on learning parameters and working memory tasks

Learning parameters

We used three different, min-max normalized composite scores which have been observed continuously across the training period and approximate individual training performance with respect to different aspects of working memory. The simple span task score, i.e., simple training score, is calculated based on the sum of the individual performances in the exercises 'Path Finder Reverse', 'Path Finder', and 'Polaroid Picture'. The complex span task score, i.e., complex training score, consists of the sum of the individual performances in the exercises 'Memory Interrupted', 'Memobox', 'Turnabout', 'Shuffler', and 'Parita'. The 'n-back' training score is based solely on the training task 'Memoflow'. A detailed description of the NeuroNation training tasks can be found below (adapted from Ophey et al., 2020 [1]). The aggregation of individual scores to composite scores has been based on the recommendation of the NeuroNation Software (https://www.neuronation.com, Synaptikon GmbH, Berlin, Germany) as well as theoretical considerations and was conducted to reduce the feature space. We used a subset of the available training sessions to include only sessions containing at least one exercise of each composite score (Supplementary Figure 1).

REFERENCE

[1]   Ophey A, Giehl K, Rehberg S, Eggers C, Reker P, van Eimeren T, Kalbe E (2020) Effects of working memory training in patients with Parkinson's Disease with out cognitive impairment: A randomized controlled trial. *Parkinsonism Relat Disord* **72**, 13-22.

Description of the Working Memory Training Tasks Selected from NeuroNation (Synaptikon GmbH)

|	Task	Description
1	Path Finder Forward	A sequence of dots gets connected. The sequence has to be memorized and re-clicked following that order. The sequence lengthens with progressing level of difficulty.
2	Path Finder Backward	A sequence of dots gets connected. The sequence has to be memorized and re-clicked in the reverse order. The sequence lengthens with progressing level of difficulty.
3	Shuffler	Symbols of the face-up cards have to be memorized. The cards will then be shuffled and the location of the memorized cards has to be determined. The number of cards and to be memorized symbols increases with progressing levels of difficulty.
4	Memory interrupted	Simple math equations have to be solved mentally. Afterwards, it has to be stated whether a shown result is correct. Meanwhile, letters and numbers are shown that have to be recalled later. The math equations get more complex and the sequence of letters and numbers lengthens with progressing levels of difficulty.
5	Memoflow	A sequence of symbols is presented. When the current stimulus matches the symbol n-steps back, a button has to be pressed. The load factor n increases with progressing levels of difficulty.
6	Parita	A sequence of symbols is presented visually and a sequence of numbers auditory. When the current symbol matches the symbol n-steps back, a button has to be pressed. The load factor n increases with progressing levels of difficulty. Simultaneously, it has to be determined whether the number heard corresponds to the one memorized in the beginning.
7	Memobox	It has to be observed how many balls leave and enter a box. After each trial, the number of balls of the same color in each box has to be entered. The number of movements increases with progressing levels of difficulty.
8	Turnabout	Symbols on a grid card have to be memorized. After one or more rotations, their locations have to be indicated by clicking on the grid position. The number of symbols and rotations increases with progressing levels of difficulty.
9	Polaroid Picture	A number of symbols appears successively in a grid. The positions of all the briefly shown symbols have to be remembered and indicated by clicking on the grid position. The number of symbols increases with progressing levels of difficulty.

Supplementary Material 2.

Random forest model evaluation: out-of-bag (OOB) error

The OOB error reflects the deviation between predicted and observed value for participants not part of the random forest regression model generation. During random forest regression, decision trees are generated based on random subsets of the data set (= bagging). Therefore, not all participants are included when generating each decision tree. Consequently, one can predict the value of a participant based on the subset of decision trees that did not include this participant during their generation. The OOB error reflects the mean square error between this prediction and the observed values of the participants. Therefore, it represents a measure of 'external' model evaluation. Due to the mathematical properties of random sampling (with replacement) approximately one third (N=10) of the total sample are not part of the decision tree generation and therefore, can be predicted when assessing the OOB error. As this 'external' evaluation is based on a rather small sample, we report the results in the Supplementary Material.

The OOB results are visualized in Supplementary Figure 3. The OOB prediction error of the 'all' model was 0.192 (95%-CI = [0.191; 0.194]) at POST and 0.264 (95%-CI = [0.262; 0.266]) at FU. OOB prediction error indicates comparable model performance of the 'all' model and the 'cog/learning' model at both POST and FU ('all' vs. 'cog/learning': $p > 0.050$). Whereas the model ranking regarding RMSE and OOB is similar for FU, the ranking varies at POST. In terms of OOB, the 'cog' model significantly outperformed all other models at POST ($p < 0.001$); however, in terms of RMSE, it was the third best model only.

Comparison of model performance (OOB prediction error) by pairwise permutation tests

| | Permutation Test | | | | Effect Size | |
| | POST | | 3-month FU | | POST | 3-month FU |
	statistic	*p**	statistic	*p**	Cohen's *d*	Cohen's *d*
'all' vs. 'cog/learning'	1.65	0.592	2.52	0.071	0.07	0.11
'all' vs. 'cog'	15.78	<0.001	-11.9	<0.001	0.75	0.55
'all' vs. 'learning'	-30.28	<0.001	-11.26	<0.001	1.84	0.52
'cog/learning' vs. 'cog'	-13.84	<0.001	-14.04	<0.001	0.65	0.66
'cog/learning' vs. 'learning'	-30.41	<0.001	-13.48	<0.001	1.85	0.63
'cog' vs. 'learning'	-35.33	<0.001	0.32	1.000	2.58	0.01
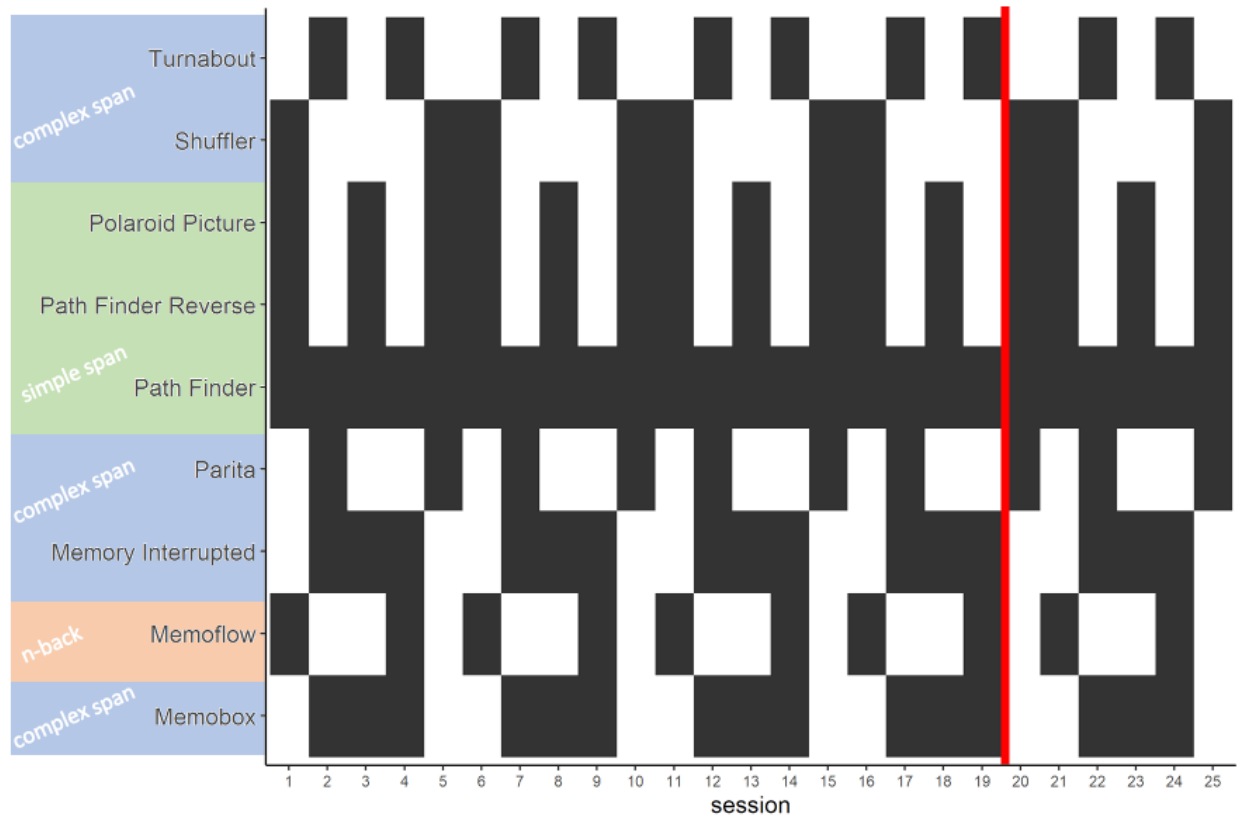
Supplementary Figure 1. Computerized WMT schedule for each participant. The training schedule was compiled based on the online multi-domain cognitive training program NeuroNation (https://www.neuronation.com, Synaptikon GmbH, Berlin, Germany) and consisted of 9 different working memory tasks. Black rectangles indicate that a given task was trained at a given day. Each training session consisted of a selection of 5 tasks. We included each training session in which at least one task per composite score (simple, complex, nback) was trained until session 19 (red bar).
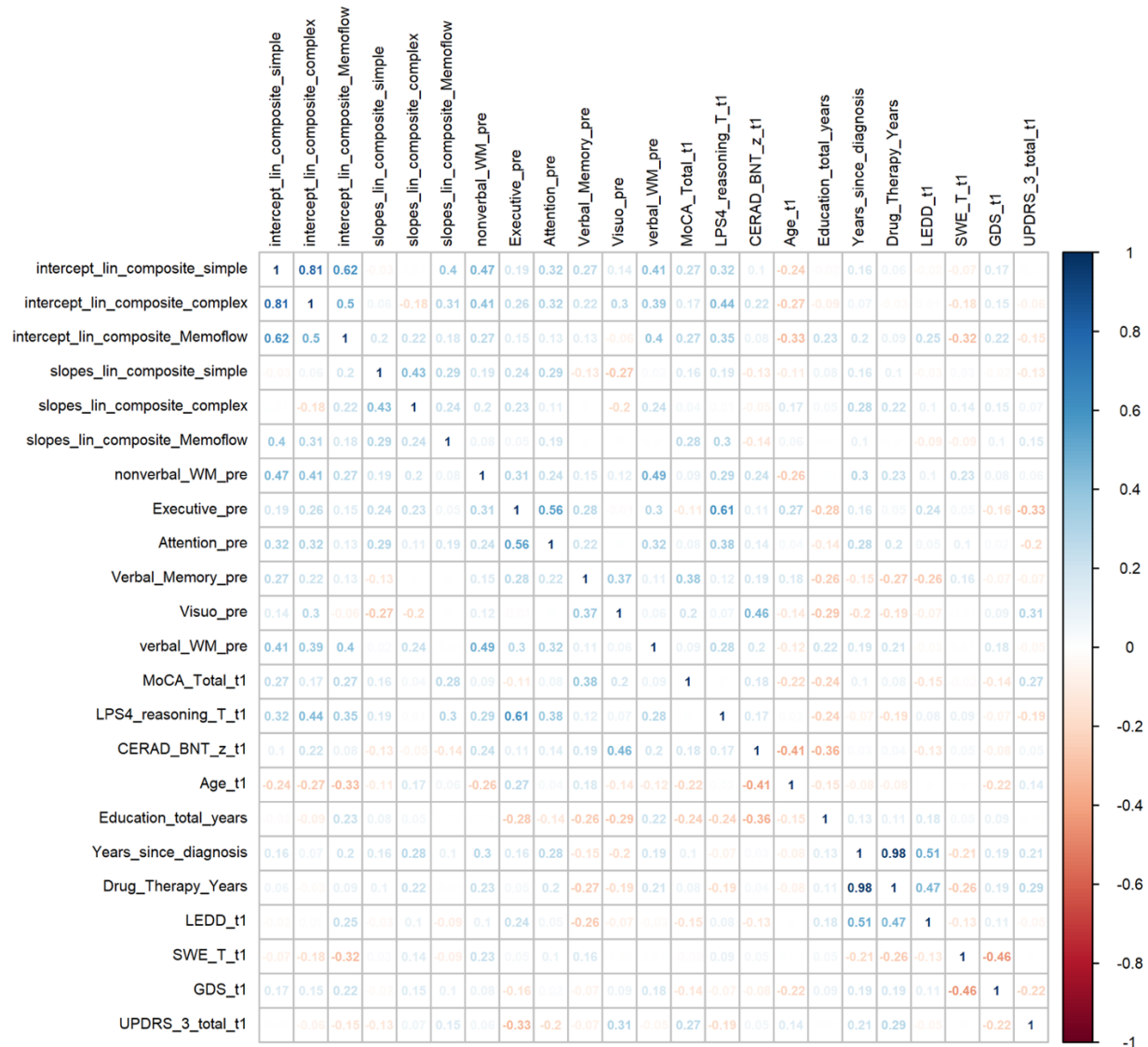
Supplementary Figure 2. Pairwise spearman correlations between predictor variables. Predictor variables showing pairwise correlation coefficients higher than 0.8 were excluded prior to random forest regression analysis to reduce the feature space and avoid interpretational biases. We excluded one learning parameter ('intercept_lin_composite_simple' = intercept parameter for composite score simple) and one clinical variable ('Years_since_diagnosis') due to high collinearity to other variables.
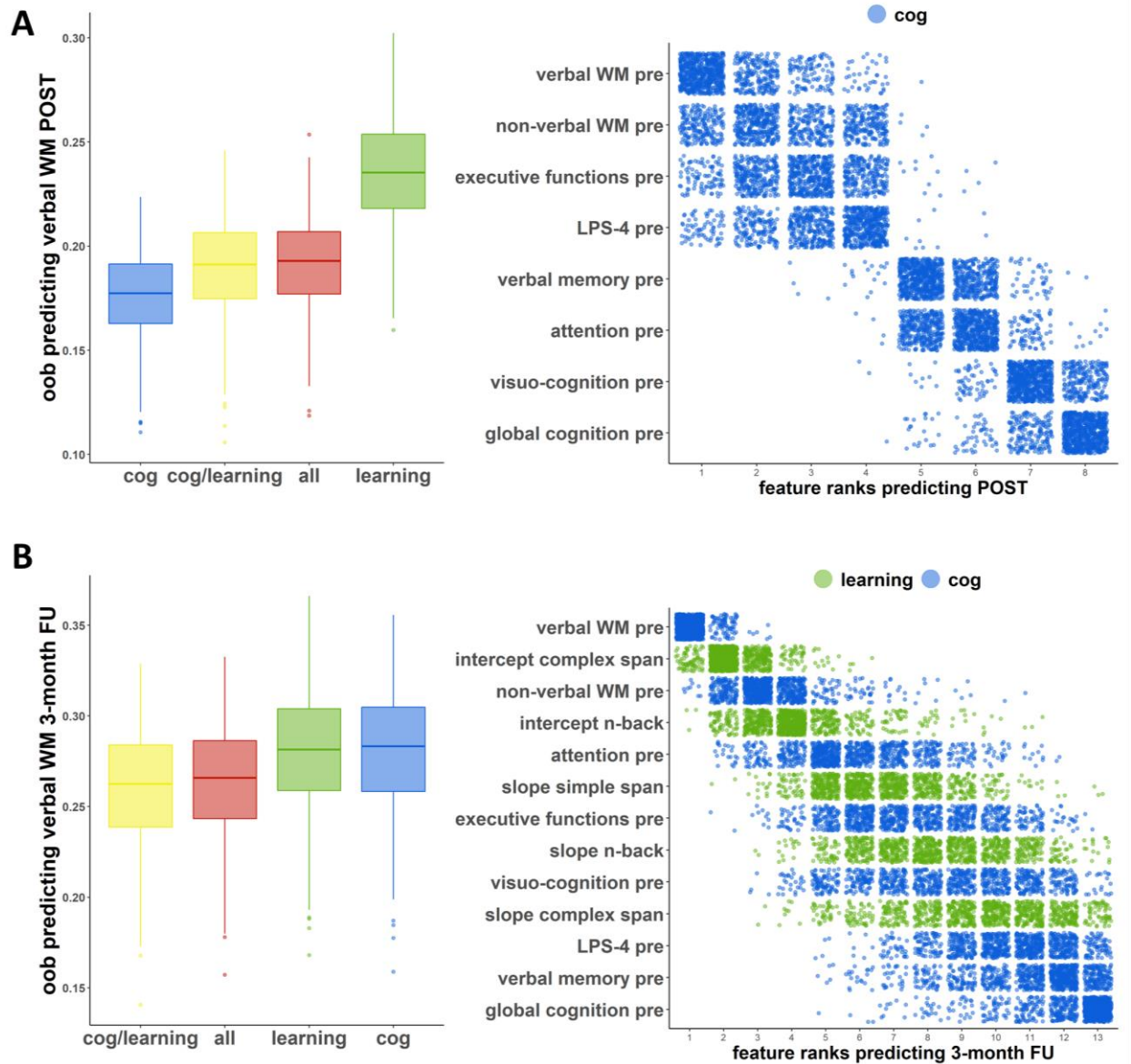
Supplementary Figure 3. Random forest regression model prediction for verbal WM POST and 3-month FU for out-of-bag (OOB) error estimate. We used random forest regression to evaluate the predictive performance of different subsets of predictor variables: cognitive (cog), learning, clinical and demographic (all). The graph shows the performance of the models generated through resampling (N=1000) measured by OOB error and the feature importance ('impurity') of the 'cog' model for the prediction of verbal working memory at timepoints POST (A) and the feature importance of the 'cog/learning' model at timepoint 3-month FU (B).
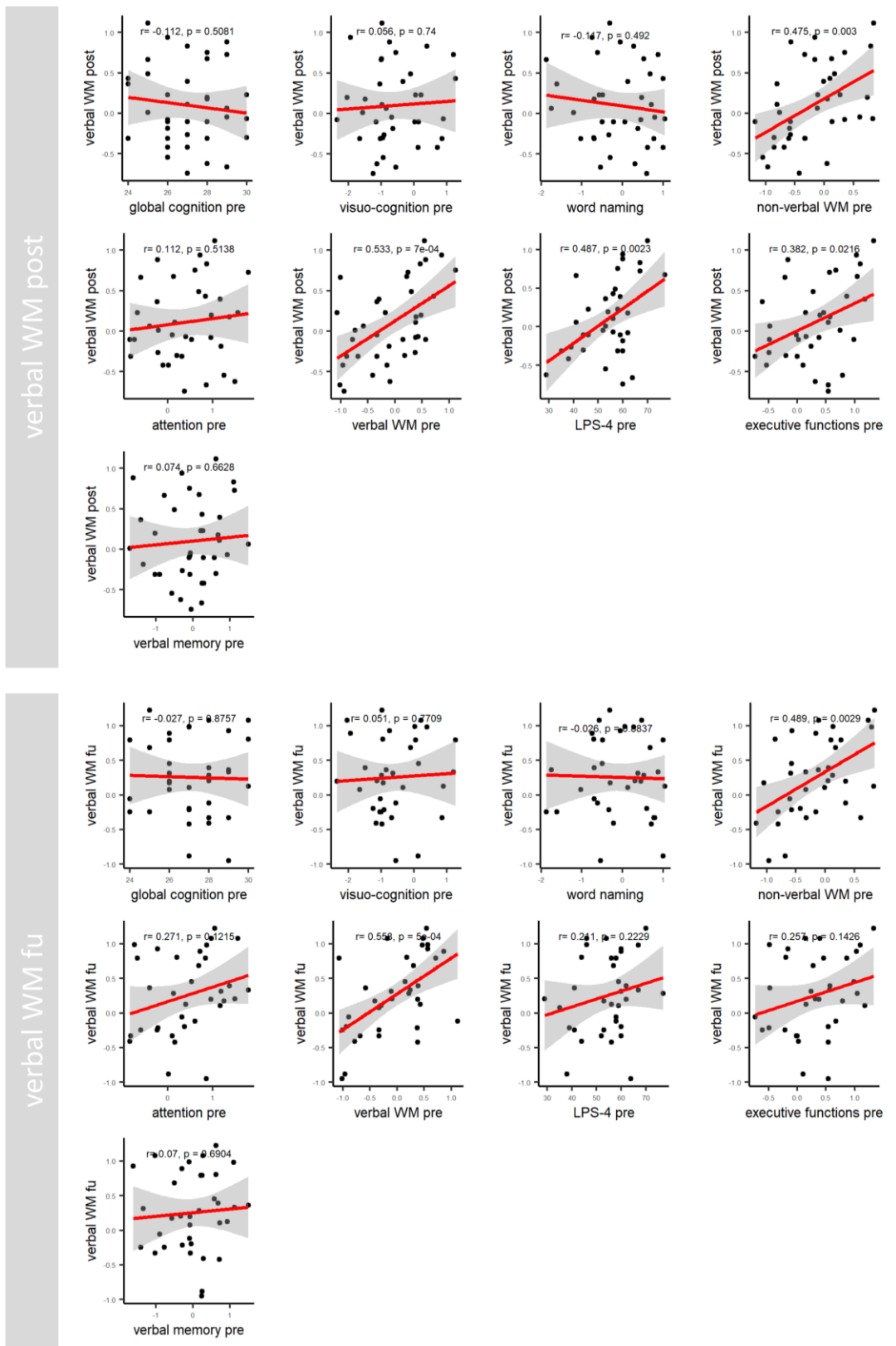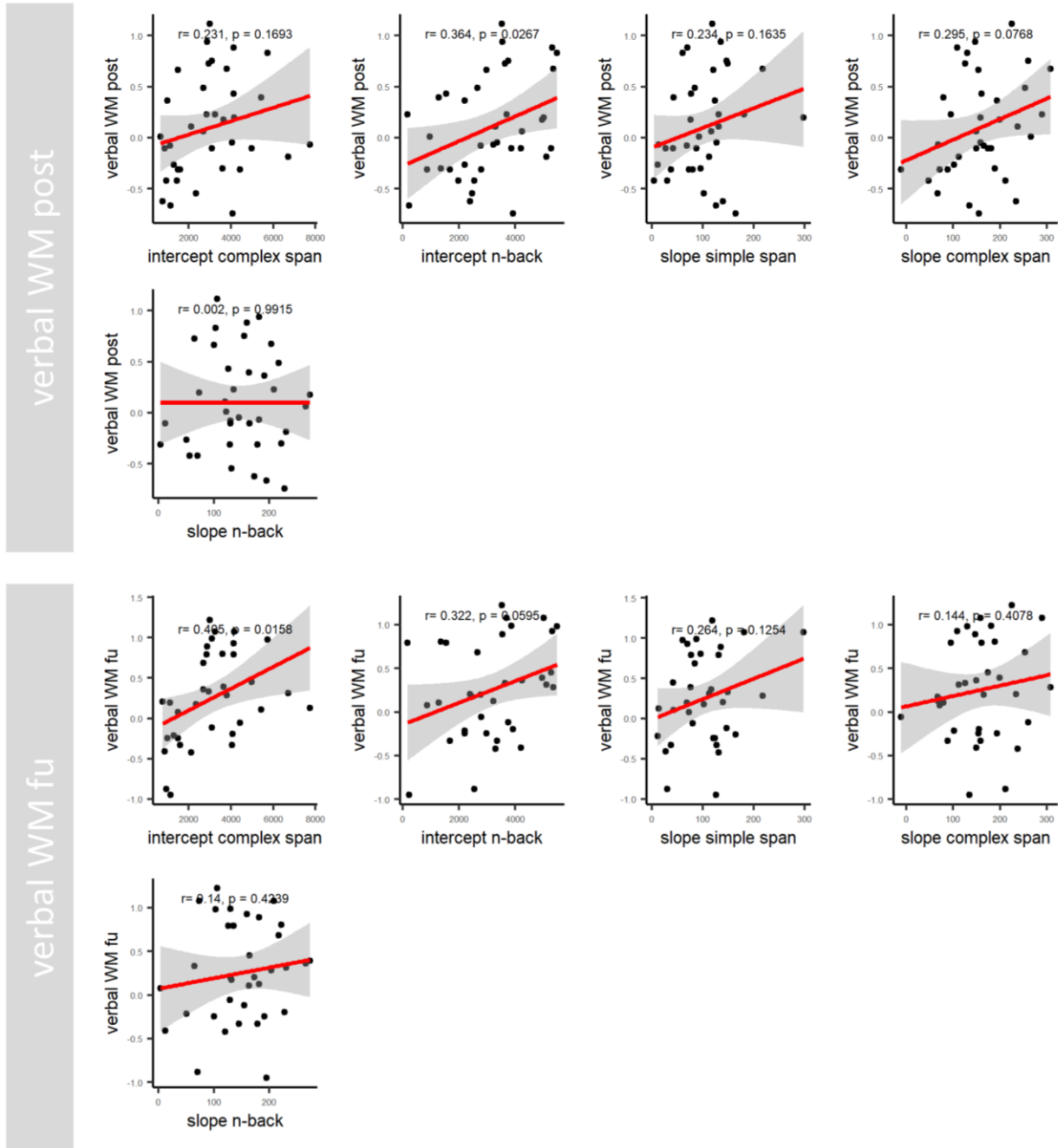
Supplementary Figure 4. Bivariate correlations between verbal working memory (at POST and 3-month FU) and cognitive baseline variables.
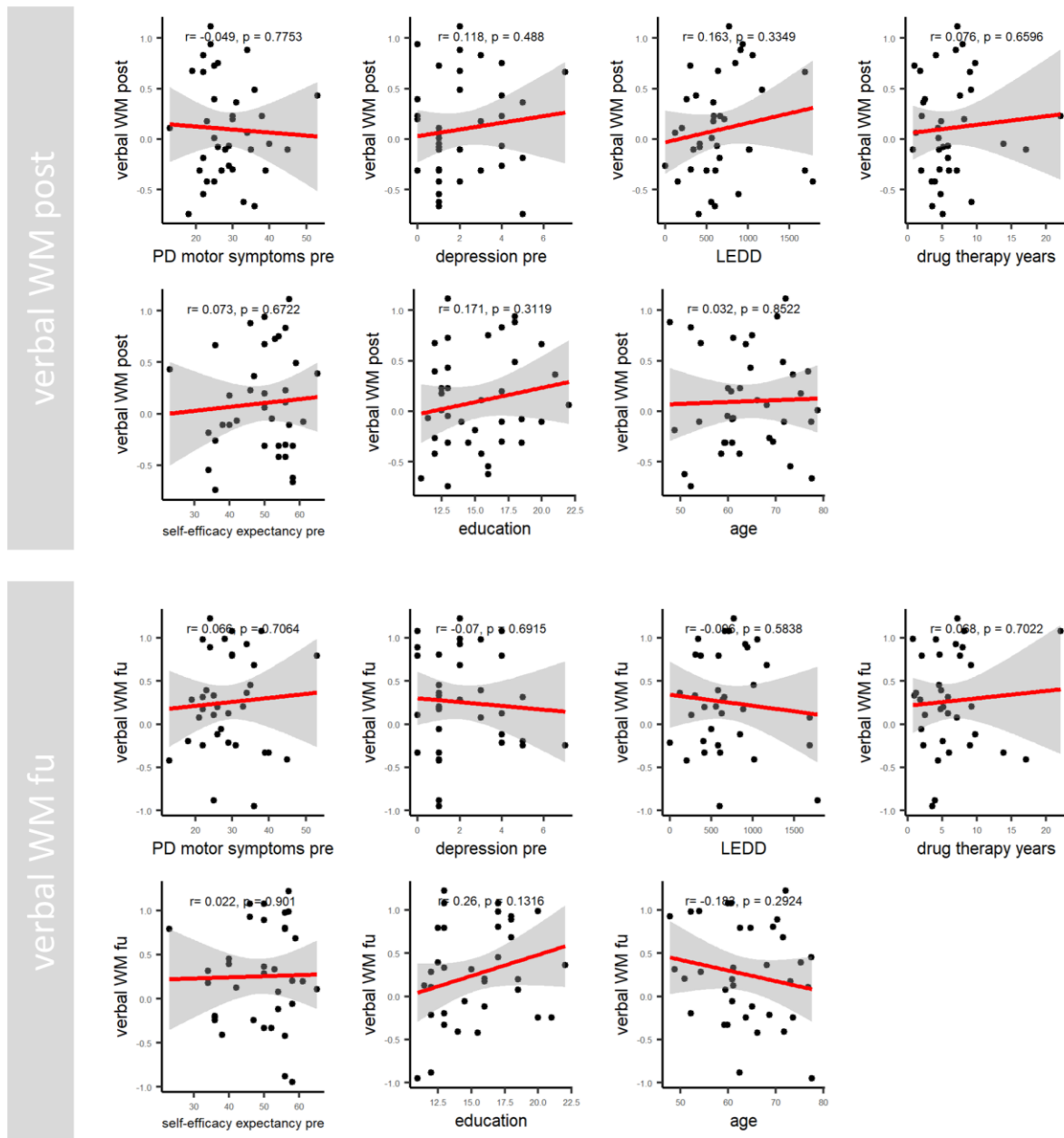
Supplementary Figure 5. Bivariate correlations between verbal working memory (at POST and 3-month FU) and demographic and clinical variables.

Supplementary Figure 6. Bivariate correlations between verbal working memory (at POST and 3-month FU) and learning parameters.