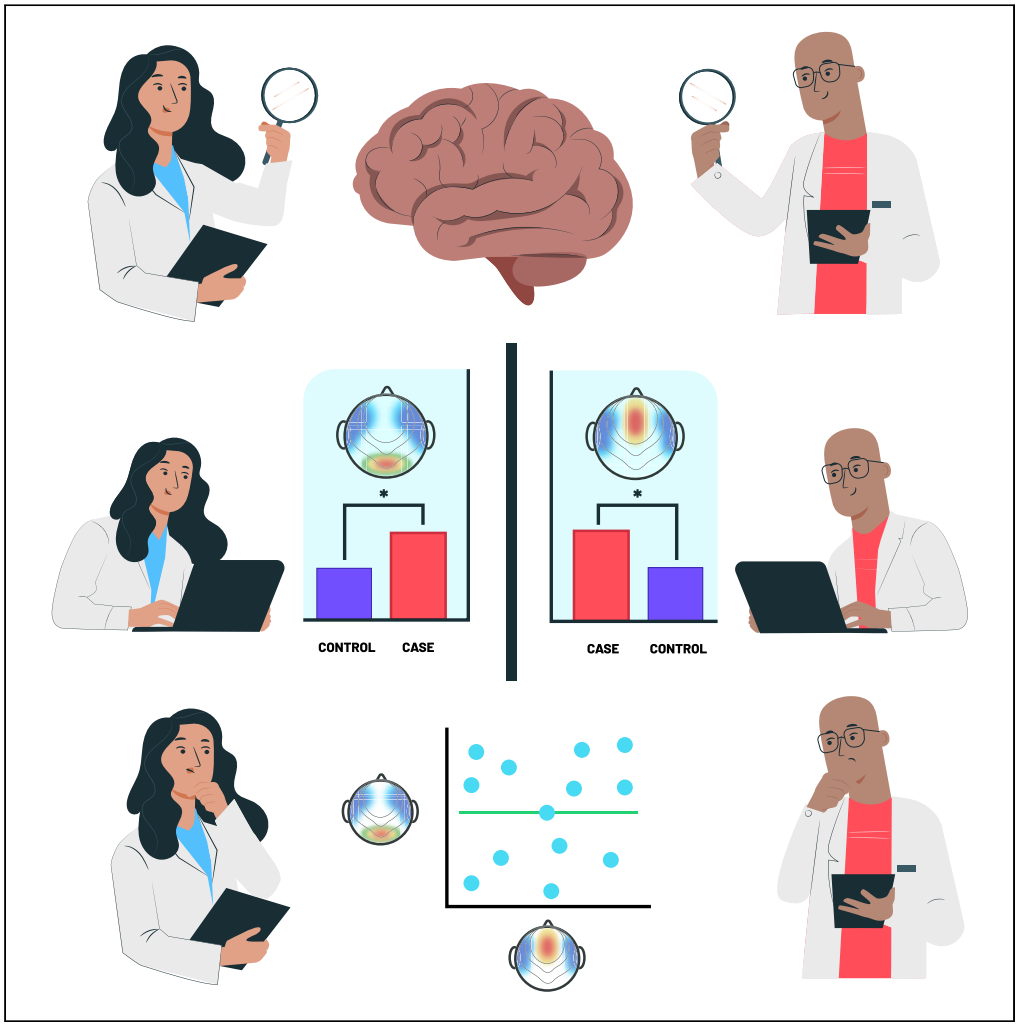
iScience

CellPress
OPEN ACCESS

Article

Do we really measure what we think we are measuring?



Dario Gordillo, Janir Ramos da Cruz, Dana Moreno, Simona Garobbio, Michael H. Herzog

dario.gordillolopez@epfl.ch

Highlights

Many tests in science are assumed to measure the same phenomena

EEG features showing significant effects do not strongly correlate with each other

Cognitive tasks are only poorly predicted by EEG features

A significant result may tell less of a research question than believed

Gordillo et al., iScience *26*, 106017
February 17, 2023
https://doi.org/10.1016/j.isci.2023.106017

iScience

CellPress
OPEN ACCESS

Article

Do we really measure what we think we are measuring?

Dario Gordillo,[1,4,*] Janir Ramos da Cruz,[1,2,3] Dana Moreno,[1] Simona Garobbio,[1] and Michael H. Herzog[1]

SUMMARY

Tests used in the empirical sciences are often (implicitly) assumed to be representative of a given research question in the sense that similar tests should lead to similar results. Here, we show that this assumption is not always valid. We illustrate our argument with the example of resting-state electroencephalogram (EEG). We used multiple analysis methods, contrary to typical EEG studies where one analysis method is used. We found, first, that many EEG features correlated significantly with cognitive tasks. However, these EEG features correlated weakly with each other. Similarly, in a second analysis, we found that many EEG features were significantly different in older compared to younger participants. When we compared these EEG features pairwise, we did not find strong correlations. In addition, EEG features predicted cognitive tasks poorly as shown by cross-validated regression analysis. We discuss several explanations of these results.

INTRODUCTION

Representative paradigms with elaborated tests are crucial in all empirical sciences. In the brain sciences, neuroimaging methods are used to investigate mechanisms underlying cognition and perception. Typically, a link between a neuroimaging feature (e.g., brain volume, connectivity) and a cognitive function of interest is declared if the chosen test delivers a significant result (after accounting for confounding variables). There is often the implicit assumption that the significant neuroimaging feature is representative of the neural mechanism under investigation. Here, we explicitly tested this assumption with the example of resting-state electroencephalogram (EEG).

In resting-state studies, EEG is recorded for around 5 min during which participants do nothing else than rest quietly. Signal processing methods are applied to quantify spatial and/or temporal characteristics of spontaneous brain activity. The outcomes of the analysis methods, i.e., EEG features, are interpreted to reflect brain processes linked to certain aspects of perception and cognition. For example, activity in the alpha and theta bands has been linked to memory and executive functions,[1–3] alpha-band activity to visual perception,[4,5] temporal autocorrelations of alpha-band oscillations and EEG microstates dynamics to reaction times,[6,7] connectivity features and alpha activity to intellectual abilities,[8–10] just to give a few examples. Similarly, EEG features reveal abnormalities in patients with schizophrenia,[11–15] depression,[16–19] and healthy older adults,[20–23] among others.

Each of these findings indicates a significant link between a given EEG feature and an aspect of cognition or a disease. In this respect, this approach has been very successful. Yet, these results provoke the questions of how the different EEG features relate to each other and how representative they are of the underlying mechanisms. For example, one might expect different EEG features, recorded from the same patients, to correlate with each other if they are supposed to point to the same aspect of the disease.

Here, we analyzed data from resting-state EEG recordings and a battery of cognitive tests. To obtain a comprehensive set of neurophysiological features, we applied widely used analysis methods to the same EEG data, including time-domain, frequency-domain, connectivity, and nonlinear dynamical analysis methods both in the electrode and source spaces. We extracted 175 EEG features. From the battery of cognitive tests, we obtained 12 cognitive variables describing several cognitive aspects. We correlated each EEG feature with each cognitive variable using methods that permitted us to examine linear and nonlinear relationships. Next, we correlated the features revealing significant correlations with the same

[1]Laboratory of Psychophysics, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[2]Institute for Systems and Robotics – Lisboa (LARSyS), Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal

[3]Wyss Center for Bio and Neuroengineering, CH-1202 Geneva, Switzerland

[4]Lead contact

*Correspondence: dario.gordillolopez@epfl.ch

https://doi.org/10.1016/j.isci.2023.106017

cognitive variable using univariate and multivariate correlation methods. This comparison allowed us to investigate whether the features, showing a significant correlation with one of the cognitive variables, point to a common mechanism. In a second project, we conducted group comparisons between younger and older adults using each EEG feature. A significant group difference would indicate that the EEG features tap into important age-related changes in brain processing. To test whether the features showing group differences target common age-related aspects, we correlated the EEG features revealing significant group differences. Furthermore, we used principal component analysis to assess the latent dimensions of multiple EEG features showing significant correlations to a cognitive variable and significant group differences between younger and older adults. As a complementary analysis, we evaluated cross-validated regression models using each EEG feature to predict the cognitive variables. Importantly, we did not want to elaborate on any particular relationship between an EEG feature and cognitive ability or an EEG feature and aging. We were interested in how significant results from single analyses relate to each other.

RESULTS

We analyzed data from the publicly available LEMON database.[24] This database includes resting-state EEG recordings and a battery of cognitive tasks. The sample used for the present study consisted of 201 participants, 138 younger adults (mean age = 25.43, SD = 3.39, 42 females) and 63 older adults (mean age = 67.66, SD = 4.79, 31 females). Using multiple analysis methods, we obtained 175 EEG features from the resting-state EEG recordings. The EEG features can be composed either of 61, 80, or 4 variables, corresponding to the number of electrodes, brain regions, or microstate parameters, respectively. From the battery of 6 cognitive tests, we obtained 12 cognitive variables. Details are shown in experimental model and subject details.

Correlations between EEG features and cognitive variables

We computed Spearman and distance correlations between each EEG feature and cognitive variable. Thus, for each age group and correlation type, we performed 2100 (175*12) analyses. With this evaluation, we sought to identify the EEG features reflective of a neural process linked to each cognitive ability. Next, the EEG features showing a significant correlation to a cognitive variable were pairwise correlated either using Spearman or distance correlations and multivariate distance correlations. Strong correlations between EEG features would suggest that the features point to the same mechanism representative of the cognitive aspect under study. In younger adults, 109 analyses were significant using Spearman correlations and 121 using distance correlations (after correction for multiple comparisons for each pair of EEG feature and cognitive variable). For most cognitive variables, we found more than one EEG feature showing a significant correlation (Figure 1A). The correlations between these EEG features were weak in most of the cases. Similar results were found using multivariate distance correlations ($\sqrt{|\mathscr{R}^*_n|}$), which permitted us to correlate EEG features considering all the variables, i.e., electrodes, brain regions, or microstate parameters (Figure 1B). Results for older adults were similar and are presented in Figure S1.

For instance, for younger adults, both the life time statistics of the amplitude envelopes in the theta band (*life time theta*) and the node strength of delta connectivity measured in the electrode space using phase locking value (*node str e-plv delta*) correlated significantly with the working memory variable obtained from the test of attentional performance (*Tap working memory*; $\rho max = -0.28$ and $\rho max = 0.29$, respectively). However, *life time theta* and *node str e-plv delta* did not correlate strongly with each other ($\rho = -0.11$; $\sqrt{|\mathscr{R}^*_n|} = 0.07$; Figure 2). Similarly, using distance correlations, both the occurrence of microstate class C (*microstate C*) and the betweenness centrality of gamma connectivity measured in the electrode space using weighted phase lag index (*betw cen e-wpli gamma*) showed significant correlations with the module A variable of the trail making test (*Tmt-A*; $\mathscr{R}_n$ *max* = 0.25 and $\mathscr{R}_n$ *max* = 0.36, respectively). However, the two EEG features only weakly correlate with each other ($\mathscr{R}_n = 0.15$; $\sqrt{|\mathscr{R}^*_n|} = 0.01$; Figure 3). ρmax and $\mathscr{R}_n$ *max* denote the maximum significant Spearman or distance correlation (among all the electrodes, brain regions, or microstate parameters) of the EEG feature with the cognitive variable.

While the correlations between EEG features were generally low (Figure 1B), some EEG features obtained from different analysis methods were highly correlated with each other. For example, in younger adults, both the Hjorth activity parameter (*hjorth activity*) and the standard deviation of the amplitude envelopes in the beta band (*std ampl beta*) showed significant correlations with the attention span module of the Cvlt

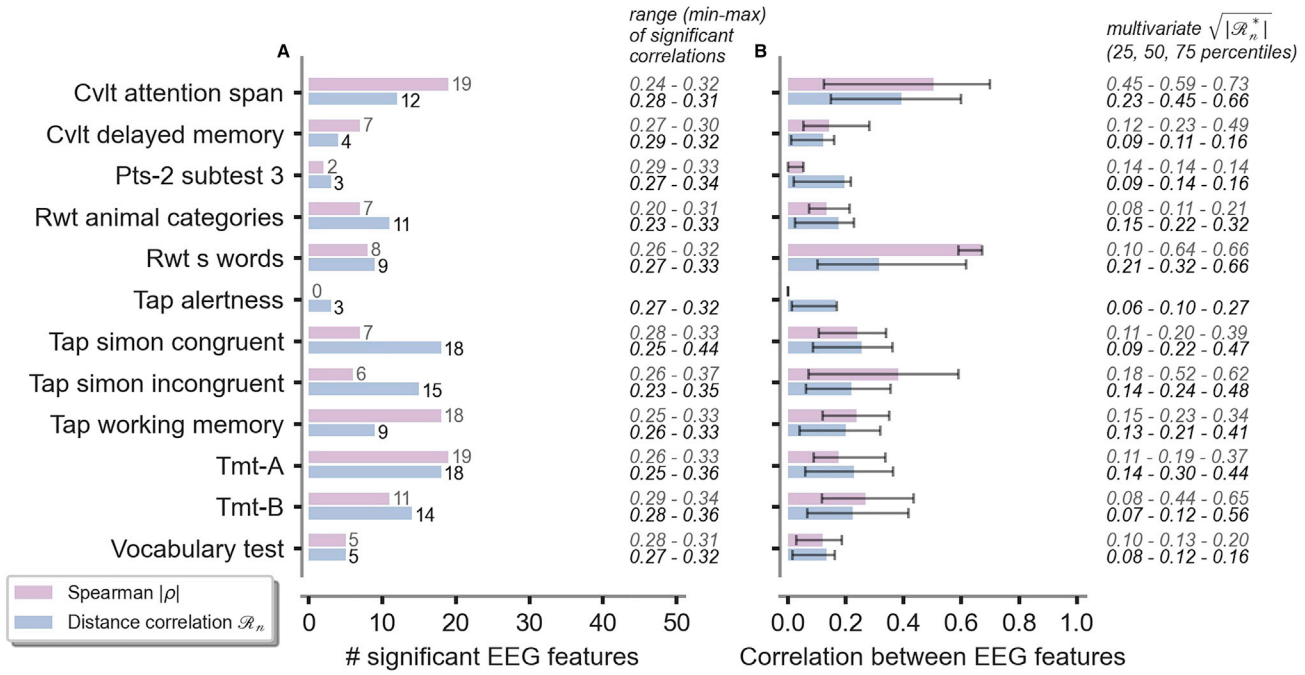


Figure 1. Result of the correlation analysis in younger adults

(A) EEG features with significant correlations to cognitive variables. On the right side of the panel, we indicate the range (min-max) of the magnitudes of the significant correlations (see correlations between EEG features and cognitive variables in quantification and statistical analysis).

(B) Median (confidence interval: 25th and 75th percentiles) Spearman and distance correlations between the EEG features showing a significant correlation with the same cognitive variable. On the right side of the panel, we indicate the 25th, 50th, and 75th percentiles of the multivariate distance correlations ($\sqrt{|\mathscr{R}_n^*|}$; ranging from 0 to 1) between the EEG features (with all its variables) showing a significant correlation with the same cognitive variables.

(*Cvlt attention span*; ρmax = 0.28 and ρmax = 0.30, respectively). The two EEG features correlated strongly with each other (ρ = 0.84; $\sqrt{|\mathscr{R}_n^*|}$ = 0.77). Similarly, using distance correlations, the animal categories variable of the Rwt test (*Rwt animal categories*) correlated significantly with the occurrence of microstate class B (*microstate B*; $\mathscr{R}_n$ *max* = 0.27) and with the clustering coefficient of alpha connectivity measured in the electrode space using weighted phase lag index (*clust coef e-wpli alpha*; $\mathscr{R}_n$ *max* = 0.28). The EEG features *microstate B* and *clust coef e-wpli alpha* exhibit a moderate correlation with each other ($\mathscr{R}_n$ = 0.41; $\sqrt{|\mathscr{R}_n^*|}$ = 0.35). All pairs of EEG feature and cognitive variable that showed significant results are presented in Figure 4.

Next, we used a principal component analysis (PCA) to examine whether EEG features, showing a significant correlation with a cognitive variable, can be grouped into a set of latent variables. Then, we used the EEG latent variables in a multiple regression model to predict the cognitive scores (see dimensionality reduction and multiple regression in quantification and statistical analysis). We found that a small number of EEG latent variables tended to explain a considerable amount of the variance of the EEG features that had a significant correlation (Spearman or distance correlation) with a cognitive variable (see Figures S5–S10). For instance, in younger adults, we applied PCA on the 18 EEG features that showed a significant Spearman correlation to the *Tap working memory* scores in younger adults (Figure 2). The first principal component explained 31.99% of the variance of the 18 EEG features. The first three principal components explained 57.61% of the variance of the 18 EEG features (Figure S6). Results are similar across cognitive variables. The proportion of variance explained by the first principal components of the EEG features showing a Spearman correlation to a cognitive variable ranged from 29.32% (*Rwt animal categories*; 7 EEG features) to 62.83% (*Rwt s words*; 8 EEG features; median across cognitive variables: 44.15%). For EEG features showing a significant distance correlation to a cognitive variable, the variance explained by the first principal components ranged from 27.81% (*Rwt animal categories*; 11 EEG features) to 58.64% (*Rwt s words*; 9 EEG features; median across cognitive variables: 33.90%).

Figure 2. Spearman correlations between the EEG features that correlated significantly with the *Tap working memory* variable in younger adults

The main diagonal has the *ρmax* (maximum *Spearman rho*) of the electrode or brain region showing the largest significant correlation to the cognitive variable. On the right side, we show the pairwise multivariate distance correlations between the EEG features (with all its variables).

Finally, we investigated whether a combination of EEG latent variables explains the cognitive variables better than a single latent variable. Thus, we asked whether various uncorrelated EEG features carry complementary information of the cognitive variables. To this end, we computed a multiple regression model. First, we used only the first PC as the predictor variable and then added, one by one, more PCs to the model (up to the third PC, i.e., three predictors). To compare models with different numbers of predictor variables, we used adjusted-R^2 to measure the goodness of fit. The analysis was performed for younger and older adults separately. For the EEG features showing a significant Spearman correlation with a cognitive variable in younger adults, the adjusted-R^2 values ranged from 0.00 (*Pts 2 subtest 3*) to 0.22 (*Tmt-B*) using only the first principal component in the regression model (median adjusted-R^2 across cognitive variables: 0.15), and from 0.11 (*Cvlt attention span*) to 0.32 (*Cvlt delayed memory*) using the first three principal components (median adjusted-R^2 across cognitive variables: 0.21; Table S2). For EEG features showing a significant distance correlation, in younger adults, the adjusted-R^2 values ranged from 0.07 (*Tap alertness*) to 0.20 (*Tmt-B*) using only the first principal component in the regression model (median adjusted-R^2 across cognitive variables: 0.12), and from 0.09 (*Tap alertness*) to 0.22 (*Rwt animal categories*) using the first three principal components (median adjusted-R^2 across cognitive variables: 0.14; Table S4). Results for older adults are presented in Tables S3 and S5. Importantly, the estimates of predictive performance were not obtained using cross-validation. As such, results should be taken with caution.

Interim conclusions: There are significant correlations between cognitive variables and EEG features obtained with different analysis methods, including connectivity, spectral power, and microstate methods. Classically, studies in the field investigate the relationship between one EEG feature and one cognitive variable in great detail with the tacit assumption that the EEG feature is representative of a proposed brain mechanism. However, we found that even though various EEG features show a significant correlation with a cognitive variable, these EEG features usually do not strongly correlate with each other. We found that a set of latent dimensions composed of multiple EEG features may explain some cognitive variables better than a single latent dimension. Yet, this was not the case for most cognitive variables. Hence, one cannot take it for granted that an EEG feature is representative of the research question at hand just because there is a significant correlation between the feature and a cognitive or another variable. We

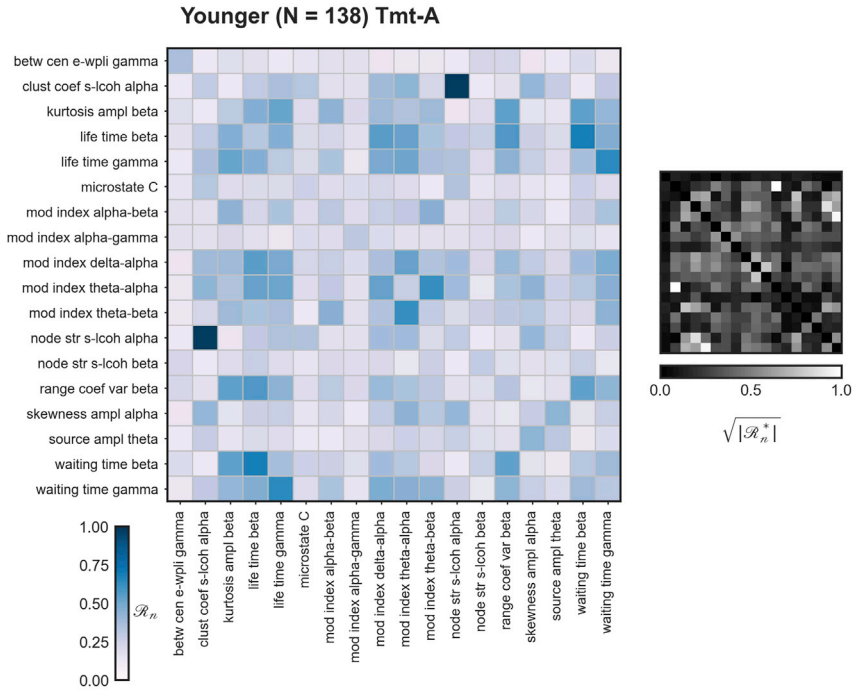


Figure 3. Distance correlations between the EEG features that correlated significantly with the *Tmt-A* variable in younger adults

The main diagonal contains the $\mathscr{R}_n$ *max* (maximum distance correlation) of the electrode, brain region, or microstate parameter showing the largest significant correlation to the cognitive variable. On the right side, we show the pairwise multivariate distance correlations between the EEG features.

are not claiming that studies based on single correlations cannot provide meaningful information about brain mechanisms. We are just pointing out that a significant result does not guarantee it. We examine this notion further in the general discussion.

Prediction of cognitive variables using EEG features

The features studied in EEG research are hypothesized to reflect neurophysiological processes involved in cognitive function. Therefore, we would expect that EEG features predict cognitive scores adequately. In this section, to test for predictive ability, we go beyond correlations and used cross-validated machine learning (ML) models. Thus, we tested the ability of single EEG features to predict each cognitive variable in an out-of-sample manner. This approach has several advantages. First, ML models handle multivariate predictors very well. Neuroimaging data are often composed of information from several brain regions or electrodes, and thus, multivariate methods provide a compact way to use all the information. Second, models can be evaluated using cross-validation, where different parts of the data are used to train and test the models, providing a rigorous test for the generalizability of results. We tested two models, namely ridge models, which are sensitive to linear relationships between predictors and predicted variables, and random forest models, which detect nonlinear relationships.

In total, 2100 (175*12) models were built using one EEG feature and one cognitive variable for each ML model (ridge or random forest) and age group (younger and older). Predictive performance was estimated using the coefficient of determination (R^2). Models were trained using cross-validation on 67% of the data and tested on the left-out 33%. We repeated the entire procedure 50 times, with different allocations of participants in the train and test sets, and obtained the median predictive performance (see cross-validated prediction of cognitive variables using EEG features in quantification and statistical analysis). Note that R^2 calculations (using the sums of squares formula and not the squared correlation) can result in negative values when the model prediction on data not used in model training is less accurate than it would be by just predicting the mean value of the data.[25,26] For younger adults, the 25th, 50th, and 75th percentiles of the 2100 (175*12) R^2 values obtained using ridge regression were 0.00, 0.00, and 0.03 for the

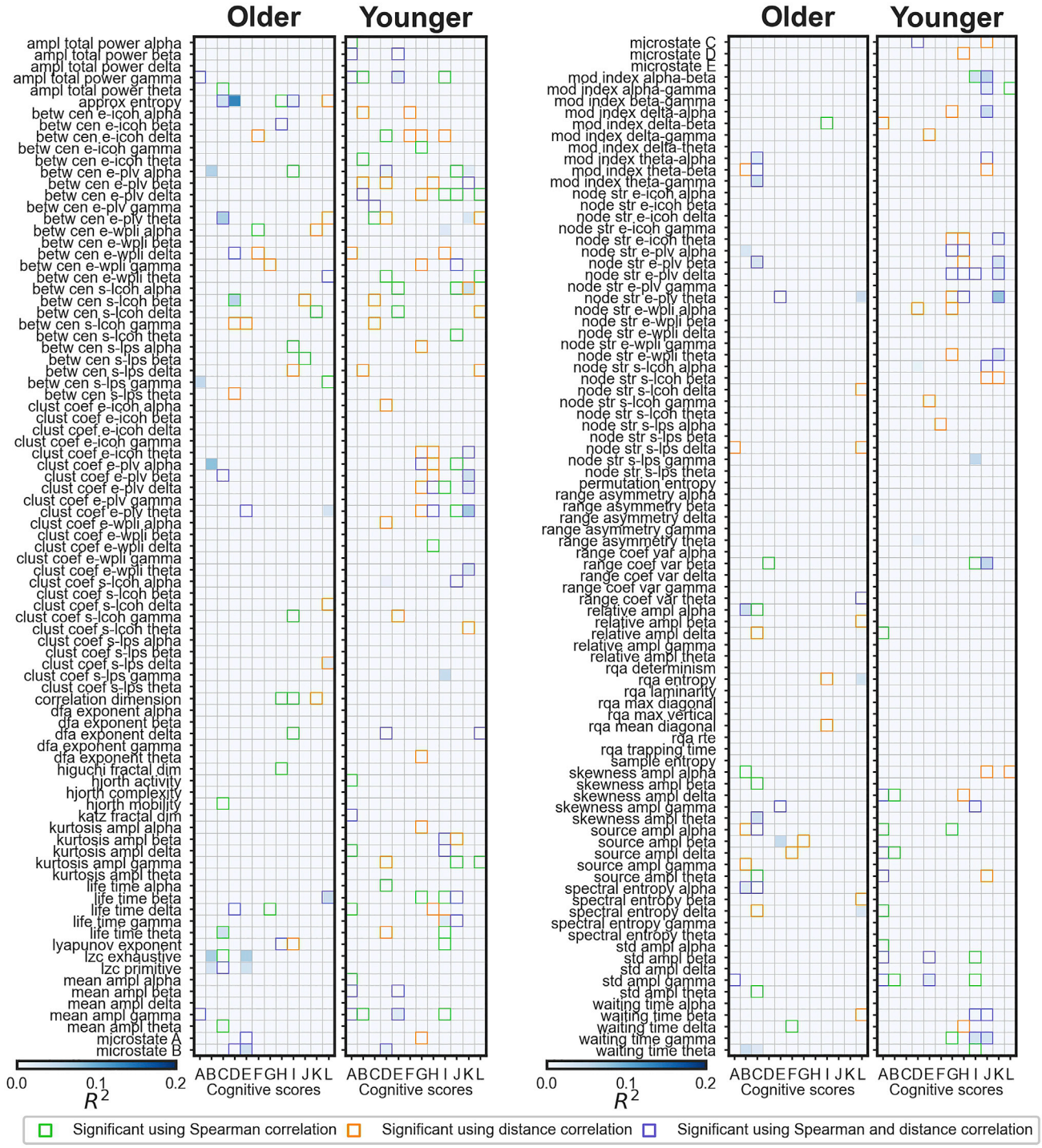


Figure 4. Prediction of cognitive variables using EEG features

Cross-validated R^2 is shown (median R^2 across 50 iterations). A ridge regression model was built for each pair of EEG feature and cognitive variable. Abbreviations: A = *Cvlt attention span*, B = *Cvlt delayed memory*, C = *Pts-2 subtest 3*, D = *Rwt animal categories*, E = *Rwt s words*, F = *Tap alertness*, G = *Tap simon congruent*, H = *Tap simon incongruent*, I = *Tap working memory*, J = *Tmt-A*, K = *Tmt-B*, L = *Vocabulary test*. Green and orange squares indicate that Spearman and distance correlation analyses were significant, respectively. Purple squares indicate that both Spearman and distance correlations were significant for the same EEG and cognitive variable pair. Colormap limits are set between 0 and 0.2. Negative R^2 values are shown as zero.

training data and −0.04, −0.03, and −0.02 for the testing data (Figure 4). For the random forest regression models, the 25th, 50th, and 75th percentiles of the R^2 values were 0.65, 0.74, and 0.78 for the training data, and −0.14, −0.10, and −0.06 for the testing data. For older adults, the 25th, 50th, and 75th percentiles of the R^2 values obtained using ridge regression were 0.00, 0.00, and 0.06 for the training data and −0.09, −0.06, and −0.04 for the testing data. For random forest regression models, the 25th, 50th, and 75th percentiles of the R^2 values were 0.80, 0.82, and 0.83 for the training data, and −0.24, −0.16, and −0.10 for the testing data. See Data S2 to S5 for detailed results.

Interim conclusions: Predictions play a crucial role in science. With this analysis, we set out to assess the ability of EEG features to predict cognitive variables. We used cross-validated prediction models. There is a hypothesized relationship between neurophysiological features at rest and cognitive performance. Thus, if EEG features truly reflect core aspects of cognitive functioning, one might expect EEG to predict cognitive performance well. Surprisingly, we found generally weak predictive performance using two different regression models. Hence, there is the possibility that we might need to rethink to what extent neurophysiological features obtained from resting-state recordings truly have a clear-cut link to behavioral measures. Another option is that the relationships might be less strong than often implicitly thought. We like to stress that our results provide only a general overview of brain-behavior predictive success and are only related to resting-state EEG features.

Group comparisons of the EEG features between younger and older adults

Classically, case-control studies using EEG are set out to identify neurophysiological processes differing in two groups (e.g., patients and controls, younger and older adults). The tacit assumption is that a significant result shows that an EEG feature under study points, for example, to a cause of a disease. In this analysis, we examined differences between older and younger adults. Each of the 175 EEG features was subjected to group comparisons between older and younger adults. 108 out of the 175 EEG features (61.71%) contained at least one variable showing significant group differences between older and younger adults, indicating that important age-related effects are detected. The absolute effect sizes (*r;* ranging from 0 to 1) of the representative variables ranged between 0.18 (*microstate E*) and 0.58 (*spectral entropy beta*), corresponding to small to large effect sizes.[27] The 25th, 50th, and 75th percentiles of the absolute significant effect sizes (one value per significant EEG feature) were 0.26, 0.31, and 0.42, respectively. For 56 out of the 108 EEG features showing significant group differences, the effects were positive, namely, older adults showed higher values than younger adults and the opposite was the case for the remaining 52 EEG features (Figure 5).

Older adults showed significantly decreased node strength in theta connectivity measured in the source space using lagged phase synchronization (*node str s-lps theta;* r = -0.31), increased long-range temporal correlations in the delta band (*dfa exponent delta;* r = 0.29), as well as longer mean duration of the microstate class A (*microstate A;* r = 0.50), to name a few. Group differences were also observed in EEG features in the different frequency bands, for instance, older adults showed reduced spectral entropy in the delta band (*spectral entropy delta;* r = -0.41), reduced spectral amplitudes in the theta band in the source space (*source ampl theta;* r = -0.40), reduced node strength in alpha connectivity measured in the source space using lagged phase synchronization (*node str s-lps alpha;* r = -0.25), increased waiting time statistics of the amplitude envelopes in the beta band (*waiting time beta;* r = 0.33), and increased node strength in gamma connectivity measured in the electrode space using phase locking value (*node str e-plv gamma;* r = 0.26).

Interim conclusions: We identified various EEG features that showed group differences in older compared to younger participants. The effect sizes of the group differences ranged from small to large, with a median significant effect size of r = 0.31. Hence, these features point to age-related changes in brain processing. There is the question of whether the EEG features, showing clear-cut group differences, point to the same neurophysiological mechanism differing in older participants.

Correlations between EEG features showing age-related differences

In the previous analysis, we found 108 EEG features showing differences in older compared to younger participants. In this section, by pairwise correlating these EEG features, we ask whether the targeted brain processes point to a common mechanism underlying age-related differences. We calculated Spearman correlations between the representative variables (i.e., showing the largest group effect) of the 108 EEG features showing group differences (Figure 6). We found that 41.74% of the correlation values were

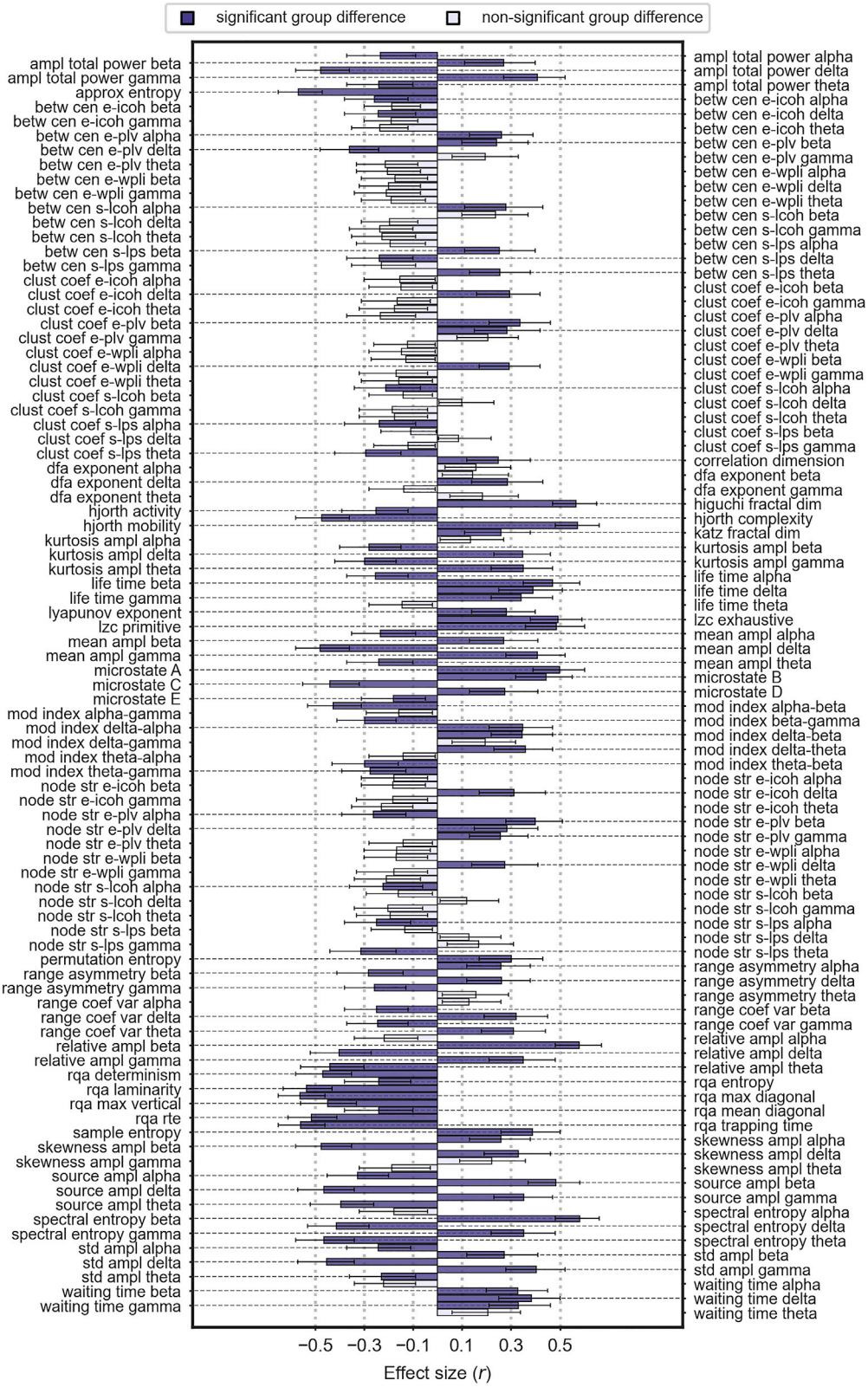


Figure 5. Effect size and confidence intervals of the group differences between younger and older adults for each of the 175 EEG features

Negative effect sizes indicate that older adults had significantly reduced values compared to younger adults. Black dotted horizontal lines serve as a guide to the labels of the EEG features showing significant group differences.

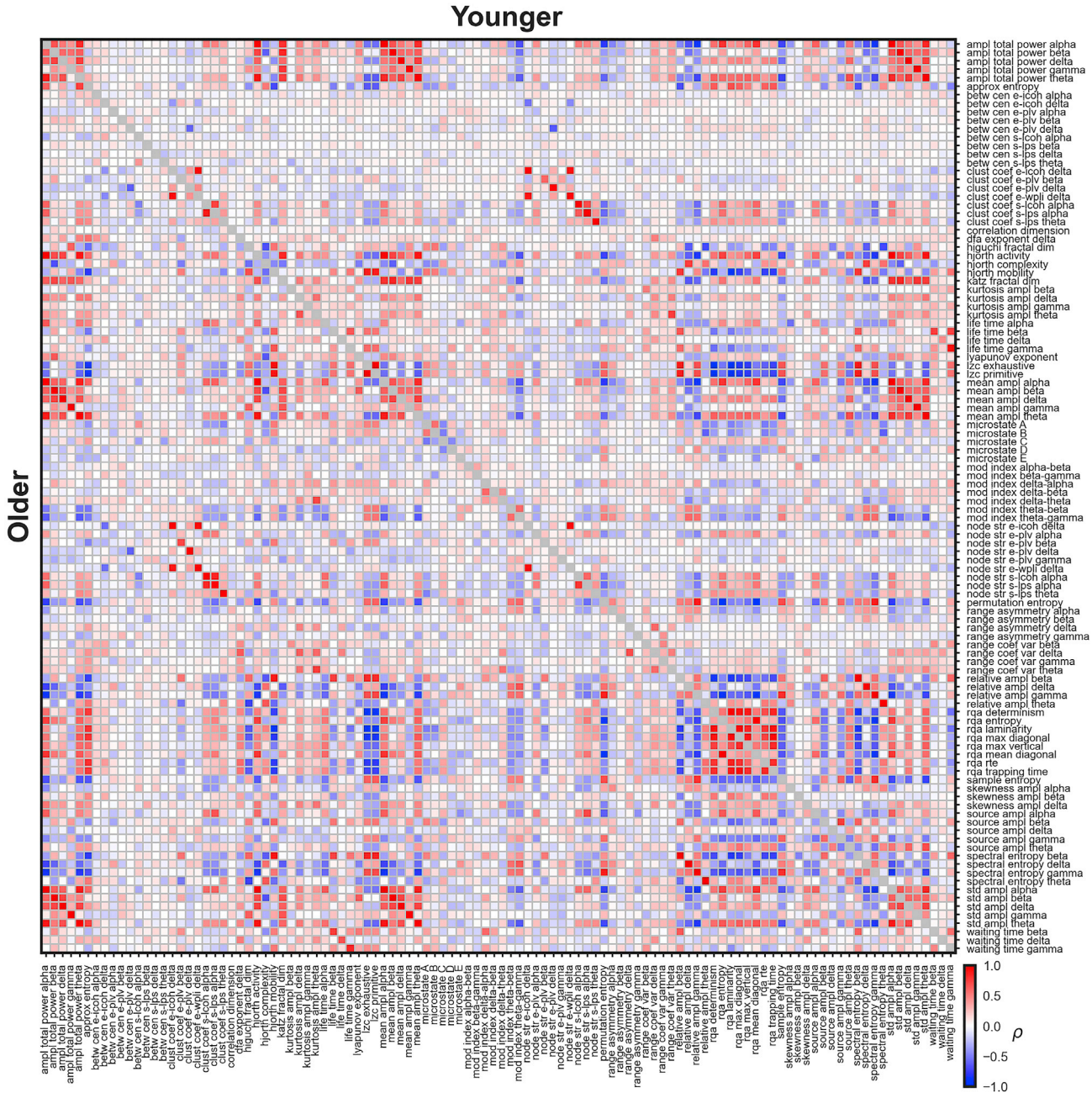


Figure 6. Spearman correlations between the 108 EEG features that showed a significant group difference between younger and older adults

The ρ correlations belonging to younger and older adults are presented in the upper and lower triangular parts of the matrix, respectively. To calculate the correlations, from each EEG feature, we selected the electrode, brain region, or microstate parameter showing the biggest effect size in the group comparisons between older and younger adults. See Figure S2 for the results using multivariate distance correlations.

significant for younger adults and 33.77% for older adults (without correction for multiple comparisons). Since significance depends on the sample size, we focus on the magnitudes of the pairwise correlations. The 25th, 50th, and 75th percentiles of the magnitudes of the 5778 (108*107/2) correlation values were 0.06, 0.13, and 0.29, for younger adults and 0.08, 0.17, and 0.31, for older adults.

Importantly, since the choice of EEG reference may influence the results,[28] we obtained zero-referenced EEG features and compared them to those with average reference and current source density (CSD; for

connectivity features in the electrode space), which were the ones used in the previous analyses. We found a very good agreement between average/CSD and zero-referenced EEG features, as quantified by intra-class correlations (ICC) and distance correlations (see comparison between EEG reference choices in quantification and statistical analysis). For younger adults, the 25th, 50th, and 75th percentiles of the ICC values between average/CSD and zero-referenced EEG features were 0.57, 0.92, and 0.97 (ICC ranges from 0 to 1), while for older adults, the percentile values were 0.65, 0.93, and 0.97, suggesting that the choice of reference does not affect the results. For younger adults, the 140 distance correlation values between average/CSD and zero-referenced EEG features were 0.79, 0.98, and 0.99, while for older adults, the distance correlations were 0.81, 0.99, and 1, for the 25th, 50th, and 75th percentiles. While for most EEG features the ICC and distance correlation values were high, network EEG features, in particular, betweenness centrality features showed rather low ICCs (see Data S6). However, this did not change the relationship between EEG features (see: Figures S3 and S4). For zero-referenced EEG features, the pairwise correlations between the 140 EEG features (see Data S6) in younger adults were 0.09, 0.19, and 0.36, whereas in older adults the pairwise correlations were 0.11, 0.17, and 0.33, for the 25th, 50th, and 75th percentiles. For average/CSD EEG features, the correlations were 0.10, 0.20, and 0.38 for younger adults, and 0.12, 0.20, and 0.35 for older adults (25th, 50th, and 75th percentiles of the multivariate distance correlation).

EEG features may be more adequately summarized considering the whole set of electrodes, brain regions, or microstate parameters. Hence, in addition to the previous univariate correlation assessment, we calculated multivariate distance correlations, which allowed us to compare EEG features using all their variables (Figure S2). The results are similar. For younger adults, the magnitudes of the multivariate distance correlations ($\sqrt{|\mathscr{R}^*_n|}$) were 0.12, 0.23, and 0.41, whereas for older adults the magnitudes were 0.12, 0.21, and 0.38, for the 25th, 50th, and 75th percentiles. In younger adults, 58.54% of the multivariate distance correlations were significant, whereas 53.01% were significant for older adults (without correction for multiple comparisons).

Furthermore, to investigate whether the EEG features showing a group difference between older and younger participants can be grouped into a set of latent variables, we used PCA on the representative variables (i.e., the variables showing the largest group difference between older and younger participants) of the 108 EEG features. For older adults, the first PC explained 24.01% of the variance of the EEG features. The second and third PCs explained 13.47% and 7.54%, respectively (Figure S11). The first PC consists essentially of EEG features obtained from nonlinear dynamical and entropy analysis methods (e.g., *rqa trapping time*, *lzc primitive*, and *permutation entropy*). The second PC is mainly composed of beta and delta band features (e.g., *mean ampl beta*, *node str e-plv delta*, and *relative ampl delta*) and also nonlinear dynamical features (e.g., *katz fractal dim* and *lyapunov exponent*). The third PC contains mainly temporal and connectivity EEG features (e.g., *microstate D*, *node str e-plv alpha*, and *life time delta*). These latent dimensions could be interesting for future investigations of age-related differences in neurophysiology.

Most of the strong correlations were found between EEG features obtained from very similar methods. For example, the node strength and the clustering coefficient of delta connectivity estimated in the electrode space using the imaginary part of coherence (*node str e-icoh delta* and *clust coef e-icoh delta*, respectively) showed strong correlations with each other (ρ = 0.97, $\sqrt{|\mathscr{R}^*_n|}$ = 0.99 and ρ = 0.97, $\sqrt{|\mathscr{R}^*_n|}$ = 1 in younger and older adults, respectively). Similarly, EEG features obtained using recurrence quantification analysis, including determinism (*rqa determinism*) and laminarity (*rqa laminarity*), were strongly correlated (ρ = 0.99, $\sqrt{|\mathscr{R}^*_n|}$ = 0.98 and ρ = 0.99, $\sqrt{|\mathscr{R}^*_n|}$ = 1 in younger and older adults, respectively). The life and waiting time statistics of the amplitude envelopes in the gamma band (*life time gamma* and *waiting time gamma*) showed also strong positive correlations (ρ = 0.91, $\sqrt{|\mathscr{R}^*_n|}$ = 0.97 and ρ = 0.83, $\sqrt{|\mathscr{R}^*_n|}$ = 0.98 in younger and older adults, respectively), to name a few examples.

Interim conclusions: Identifying brain mechanisms underlying cognition or perception is crucial in the brain sciences. Our analysis showed that various EEG features, all showing group differences in older compared to younger adults, mainly correlate weakly with each other. While there are strong correlations between similar methods, for example, between entropy and nonlinear measures, these features did not correlate, for instance, with EEG microstates, connectivity, or autocorrelation features. Hence, while the EEG features point to meaningful brain processes showing clear-cut group differences, they do not point to a general

neurophysiological deficit in older compared to younger adults. The main conclusion from this analysis, as well as the previous ones, is that statistically significant effects might explain much less of the research question than it is often implicitly assumed.

DISCUSSION

Science relies on tests targeting the crucial aspects of a research field. Classically, these tests show a significant group difference between an intervention and a control condition, between a case and a control group, or strong correlations to real-world or experimental outcomes. Then, in-depth studies are carried out to describe the test in great detail in order to understand the causes of the observed effects. It usually is assumed that these tests are representative of the interrogated mechanisms in the sense that other tests supposed to target the same mechanism should strongly correlate with each other. Here, we have shown that this rationale might not always hold. To exemplify our argument, we analyzed a publicly available database containing resting-state EEG recordings and performance scores from a battery of cognitive tests of older and younger adults.[24]

We extracted 175 EEG features from the EEG recordings and 12 cognitive variables from the battery of cognitive tests. To identify the associations between EEG features and cognitive variables, we correlated each EEG feature with each cognitive variable using Spearman and distance correlations. For younger adults, Spearman correlations were significant for 109 analyses, while 121 analyses were significant using distance correlations. For older adults, Spearman correlations were significant for 57 analyses and distance correlations for 60 (Figure S1). Then, we correlated the EEG features that showed a significant correlation with a cognitive variable (Figure 1). Surprisingly, the correlations were weak in most cases, suggesting that not all the EEG features are representative of the investigated cognitive variable. Using PCA, we found that a set of latent dimensions composed of multiple EEG features may explain some cognitive variables better than a single latent dimension. Next, we found 108 EEG features revealing a significant group difference between older and younger participants. These features also did not show strong correlations, even though they showed clear-cut group effects. Using cross-validated regression analysis, we found very weak evidence that EEG features are adequate predictors of the cognitive variables suggesting that it may be possible that the link between a cognitive aspect and an EEG feature is less informative than believed.

How can these results be explained? There are at least five possibilities. First, EEG features might have low test-retest reliability. Even though there are significant correlations with the cognitive tasks, the low test-retest reliability may have led to weak correlations between EEG features (cf.[29]). While here we do not have a measure of test-retest, certain EEG features have shown adequate reliabilities in previous studies.[30–35] Second, there might be clear-cut group differences but the variance in the groups is low. In this case, one would not expect high correlations, the so-called *reliability paradox*.[36] However, variance, for example, was rather high in the older participants in our study. Third, the EEG features do not reflect the intended brain aspects well. For instance, the EEG features may point to nonlinear brain mechanisms. Yet, some of the EEG analysis methods are linear and as such, they might miss nonlinear patterns. Hence, the correlations between EEG features might be misled by the linear methods. Fourth, the EEG features target a common mechanism and, to a substantial amount, also target-unspecific, i.e., idiosyncratic, aspects, which have little to do with the targeted mechanism but contribute to the large inter-individual variability. For example, some EEG features might be sensitive to fatigue,[37] whereas others are not. Hence, if substantial parts of the variance correspond to such target-unspecific aspects and different EEG features tap into different target-unspecific aspects, the correlations may be low. Fifth, each EEG feature targets a different aspect of a highly heterogeneous, multifactorial mechanism. Previous research indicates that this scenario might be true in certain instances. For example, by combining different resting-state or evoked EEG features, previous studies have improved classification between groups or experimental conditions.[38–43] Similarly, previous studies have reported that a combination of EEG features allowed a better characterization of certain brain processes.[44–46] Our results using PCA also show evidence that, for some cognitive variables, latent dimensions of EEG features might be more informative than single features. All these studies suggest that combining measures and features from the same paradigms might offer new insights into complex processes.

In our analysis, significant results came with *small* to *medium* effect sizes in the range from 0.20 to 0.37 ($|\rho|$) in the correlation analysis, and with *small* to *large* effect sizes in the range of 0.18–0.58 ($|r|$) in the case-control analyses according to Cohen[27] and from *typical* (0.2–0.3) to *relatively large* (>0.3) according to Gignac and Szodorai.[47] However, even for "large" effect sizes, there is a large proportion of unexplained variance. For

example, for $r = 0.5$, the unexplained variance is 75%, and a Cohen's d of 0.8 corresponds to a discriminability of 65% only (for the optimal decision criterion, i.e., hits = correct rejections). A good discriminability of 90% corresponds to a $d = 2.5$. Hence, the question is where does all the dominant noise come from? In an optimistic scenario, it comes from measurement noise. For example, EEG is a relatively noisy technique (e.g., electrode misplacement and volume conduction). Thus, the true effects might be larger. The pessimistic scenario is that noise is low and inter-individual variability is high, i.e., there are multiple factors and each paradigm taps into one, or a nonlinear combination of all of them plus a large amount of target-unspecific variance. This large amount of unexplained variance could account for why one may obtain both significant group differences and low correlations. Hence, even if clear-cut effects are found, this does not guarantee that a paradigm represents the intended aspects well. Therefore, when a test leads to a significant result, one needs to ask how representative the test is for the research question at hand. Particularly in complex systems where everything is correlated with everything to some degree, there can be many tests, which show significant but negligible effects.

Our impression is that the above considerations have been overlooked and hold true in many other research areas. As an example, in schizophrenia research, several studies have found atypical patterns in several resting-state EEG features, which are thereon studied in detail and linked to the crucial aspects of the disease. Here, again, the tacit assumption is that the EEG feature under study taps into the common and representative aspects of the disease, and for that reason, they should correlate with similar features. However, we have shown that this is not always the case. In a previous EEG study, we extracted 194 EEG features from the resting-state recordings of 121 patients with schizophrenia and 75 healthy controls. We found that 69 out of the 194 EEG features showed a significant group difference between patients with schizophrenia and healthy controls. However, the features showed mainly weak correlations with each other, questioning to what extent a single EEG feature is representative of the disease.[48] In another example, in vision research, weak correlations have been found between performances in various visual tasks in older and younger adults.[49–52] Visual illusions also correlate weakly,[53,54] suggesting that the underlying visual functions cannot be explained using only one visual paradigm. Similarly, four visual tasks assumed to capture visual magnocellular stream function showed weak relationships with each other.[55] The authors concluded that none of these tasks is a general measure of magnocellular function, as was assumed. Eisenberg et al.[56] showed that several questionnaires and cognitive tasks thought to point to the same psychological construct correlate weakly with each other and predict real-world outcomes poorly. The authors concluded that the construct lacks coherence.

Most natural sciences face severe crises. The brain sciences are among the fields hard-hit. First, many studies are underpowered and/or subject to questionable research practices. False positives are the consequence.[57–59] Second, even if the very same data are used, different analysis tools can lead to different results. This problem becomes more severe with complex analysis pipelines and more degrees of freedom.[60–62] Third, the uncontrolled use of open data for hypothesis testing can increase false positives.[63–65] Fourth, here, we have shown that there is one additional problem. Studies may have been conducted perfectly with clear-cut, significant results. Still, the studies may not target the mechanism assumed or they are less representative of the research question than believed. Overall, our results show that single measurements, even with "large" effect sizes, may be less meaningful than thought. To what extent the above-mentioned scenarios hold must be shown for each study individually.

Limitations of the study

One of the main limitations of our study is that resting-state EEG and behavior were not measured simultaneously. Hence, whether there are causal links between EEG features and cognitive variables is not clear-cut. The evoked EEG features show a more direct link to the temporal aspects of cognitive processing, and thus stronger brain-behavior relationships might be expected. As such, our results and interpretations of the weak correlations between EEG and cognition do not concern evoked EEG paradigms. Furthermore, there are at least two limitations that should be considered when interpreting the results from our prediction analysis. First, our sample size is small for a reliable assessment of predictive ability.[66] Second, we do not have an independent dataset to test our predictive models. We tried to account for these limitations using a repeated train-test split procedure, which produces unbiased estimates of predictive ability for small sample sizes. Nonetheless, larger sample sizes and independent datasets are needed. Finally, we do not have a measure of the test-retest reliability of the EEG features. Hence, EEG features with poor reliabilities may mislead the correlations. While our results show that similar EEG features correlate strongly

with each other (resembling test-retest), measuring and accounting for reliability, e.g., using disattenuated correlations, will tell about the "true" correlations between EEG features.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - EEG collection and preprocessing
  - Cognitive assessment
  - EEG features extraction
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Correlations between EEG features and cognitive variables
  - Dimensionality reduction and multiple regression
  - Cross-validated prediction of cognitive variables using EEG features
  - Group comparisons of EEG features between older and younger adults
  - Correlations between EEG features showing age-related differences
  - Comparison between EEG reference choices

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106017.

ACKNOWLEDGMENTS

This work was funded by the National Centre of Competence in Research (NCCR) Synapsy financed by the Swiss National Science Foundation under grant 51NF40-185897. We acknowledge all the contributors to the Leipzig Study for Mind-Body-Emotion Interactions.

AUTHOR CONTRIBUTIONS

Conceptualization, M.H.H., D.G., and J.R.C.; Methodology, D.G., J.R.C., and D.M.; Investigation, D.G. and S.G.; Writing – Original Draft, D.G., M.H.H., J.R.C., and S.G.; Writing- Review & Editing, D.G. and M.H.H.; Supervision, M.H.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 25, 2022
Revised: December 18, 2022
Accepted: January 16, 2023
Published: January 20, 2023

REFERENCES

1. Finnigan, S., and Robertson, I.H. (2011). Resting EEG theta power correlates with cognitive performance in healthy older adults. Psychophysiology *48*, 1083–1087. https://doi.org/10.1111/j.1469-8986.2010.01173.x.

2. Klimesch, W., Vogt, F., and Doppelmayr, M. (1999). Interindividual differences in alpha and theta power reflect memory performance. Intelligence *27*, 347–362. https://doi.org/10.1016/S0160-2896(99)00027-6.

3. Richard Clark, C., Veltmeyer, M.D., Hamilton, R.J., Simms, E., Paul, R., Hermens, D., and Gordon, E. (2004). Spontaneous alpha peak frequency predicts working memory performance across the age span. Int. J. Psychophysiol. *53*, 1–9. https://doi.org/10.1016/j.ijpsycho.2003.12.011.

4. Chen, L., Wu, B., Qiao, C., and Liu, D.-Q. (2020). Resting EEG in alpha band predicts individual differences in visual size perception. Brain Cogn. *145*, 105625. https://doi.org/10.1016/j.bandc.2020.105625.

5. Rogala, J., Kublik, E., Krauz, R., and Wróbel, A. (2020). Resting-state EEG activity predicts frontoparietal network reconfiguration and improved attentional performance. Sci. Rep. *10*, 5064. https://doi.org/10.1038/s41598-020-61866-7.

6. Irrmischer, M., Poil, S.-S., Mansvelder, H.D., Intra, F.S., and Linkenkaer-Hansen, K. (2018).

Strong long-range temporal correlations of beta/gamma oscillations are associated with poor sustained visual attention performance. Eur. J. Neurosci. *48*, 2674–2683. https://doi.org/10.1111/ejn.13672.

7. Zanesco, A.P., King, B.G., Skwara, A.C., and Saron, C.D. (2020). Within and between-person correlates of the temporal dynamics of resting EEG microstates. Neuroimage *211*, 116631. https://doi.org/10.1016/j.neuroimage.2020.116631.

8. Doppelmayr, M., Klimesch, W., Stadler, W., Pöllhuber, D., and Heine, C. (2002). EEG alpha power and intelligence. Intelligence *30*, 289–302. https://doi.org/10.1016/S0160-2896(01)00101-5.

9. Thatcher, R.W., North, D., and Biver, C. (2005). EEG and intelligence: relations between EEG coherence, EEG phase delay and power. Clin. Neurophysiol. *116*, 2129–2141. https://doi.org/10.1016/j.clinph.2005.04.026.

10. Zakharov, I., Tabueva, A., Adamovich, T., Kovas, Y., and Malykh, S. (2020). Alpha band resting-state EEG connectivity is associated with non-verbal intelligence. Front. Hum. Neurosci. *14*, 10. https://doi.org/10.3389/fnhum.2020.00010.

11. da Cruz, J.R., Favrod, O., Roinishvili, M., Chkonia, E., Brand, A., Mohr, C., Figueiredo, P., and Herzog, M.H. (2020). EEG microstates are a candidate endophenotype for schizophrenia. Nat. Commun. *11*, 3089. https://doi.org/10.1038/s41467-020-16914-1.

12. Di Lorenzo, G., Daverio, A., Ferrentino, F., Santarnecchi, E., Ciabattini, F., Monaco, L., Lisi, G., Barone, Y., Di Lorenzo, C., Niolu, C., et al. (2015). Altered resting-state EEG source functional connectivity in schizophrenia: the effect of illness duration. Front. Hum. Neurosci. *9*, 234. https://doi.org/10.3389/fnhum.2015.00234.

13. Nikulin, V.V., Jönsson, E.G., and Brismar, T. (2012). Attenuation of long-range temporal correlations in the amplitude dynamics of alpha and beta neuronal oscillations in patients with schizophrenia. Neuroimage *61*, 162–169. https://doi.org/10.1016/j.neuroimage.2012.03.008.

14. Rieger, K., Diaz Hernandez, L., Baenninger, A., and Koenig, T. (2016). 15 Years of microstate research in schizophrenia – where are we? A meta-analysis. Front. Psychiatr. *7*, 22. https://doi.org/10.3389/fpsyt.2016.00022.

15. Sponheim, S.R., Clementz, B.A., Iacono, W.G., and Beiser, M. (2000). Clinical and biological concomitants of resting state EEG power abnormalities in schizophrenia. Biol. Psychiatr. *48*, 1088–1097. https://doi.org/10.1016/S0006-3223(00)00907-0.

16. Fingelkurts, A.A., Fingelkurts, A.A., Rytsälä, H., Suominen, K., Isometsä, E., and Kähkönen, S. (2007). Impaired functional connectivity at EEG alpha and theta frequency bands in major depression. Hum. Brain Mapp. *28*, 247–261. https://doi.org/10.1002/hbm.20275.

17. Jaworska, N., Blier, P., Fusee, W., and Knott, V. (2012). Alpha power, alpha asymmetry and anterior cingulate cortex activity in depressed males and females. J. Psychiatr. Res. *46*, 1483–1491. https://doi.org/10.1016/j.jpsychires.2012.08.003.

18. Murphy, M., Whitton, A.E., Deccy, S., Ironside, M.L., Rutherford, A., Beltzer, M., Sacchet, M., and Pizzagalli, D.A. (2020). Abnormalities in electroencephalographic microstates are state and trait markers of major depressive disorder. Neuropsychopharmacology *45*, 2030–2037. https://doi.org/10.1038/s41386-020-0749-1.

19. Shim, M., Im, C.-H., Kim, Y.-W., and Lee, S.-H. (2018). Altered cortical functional network in major depressive disorder: a resting-state electroencephalogram study. Neuroimage Clin. *19*, 1000–1007. https://doi.org/10.1016/j.nicl.2018.06.012.

20. Babiloni, C., Binetti, G., Cassarino, A., Dal Forno, G., Del Percio, C., Ferreri, F., Ferri, R., Frisoni, G., Galderisi, S., Hirata, K., et al. (2006). Sources of cortical rhythms in adults during physiological aging: a multicentric EEG study. Hum. Brain Mapp. *27*, 162–172. https://doi.org/10.1002/hbm.20175.

21. Kumral, D., Şansal, F., Cesnaite, E., Mahjoory, K., Al, E., Gaebler, M., Nikulin, V.V., and Villringer, A. (2020). BOLD and EEG signal variability at rest differently relate to aging in the human brain. Neuroimage *207*, 116373. https://doi.org/10.1016/j.neuroimage.2019.116373.

22. Scally, B., Burke, M.R., Bunce, D., and Delvenne, J.-F. (2018). Resting-state EEG power and connectivity are associated with alpha peak frequency slowing in healthy aging. Neurobiol. Aging *71*, 149–155. https://doi.org/10.1016/j.neurobiolaging.2018.07.004.

23. Zappasodi, F., Marzetti, L., Olejarczyk, E., Tecchio, F., and Pizzella, V. (2015). Age-related changes in electroencephalographic signal complexity. PLoS One *10*, e0141995. https://doi.org/10.1371/journal.pone.0141995.

24. Babayan, A., Erbey, M., Kumral, D., Reinelt, J.D., Reiter, A.M.F., Röbbig, J., Schaare, H.L., Uhlig, M., Anwander, A., Bazin, P.-L., et al. (2019). A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. Sci. Data *6*, 180308. https://doi.org/10.1038/sdata.2018.308.

25. Alexander, D.L.J., Tropsha, A., and Winkler, D.A. (2015). Beware of R^2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. J. Chem. Inf. Model. *55*, 1316–1322. https://doi.org/10.1021/acs.jcim.5b00206.

26. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

27. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd ed. (L. Erlbaum Associates).

28. Qin, Y., Xu, P., and Yao, D. (2010). A comparative study of different references for EEG default mode network: the use of the infinity reference. Clin. Neurophysiol. *121*, 1981–1991. https://doi.org/10.1016/j.clinph.2010.03.056.

29. Loken, E., and Gelman, A. (2017). Measurement error and the replication crisis. Science *355*, 584–585. https://doi.org/10.1126/science.aal3618.

30. Gudmundsson, S., Runarsson, T.P., Sigurdsson, S., Eiriksdottir, G., and Johnsen, K. (2007). Reliability of quantitative EEG features. Clin. Neurophysiol. *118*, 2162–2171. https://doi.org/10.1016/j.clinph.2007.06.018.

31. Haartsen, R., van der Velde, B., Jones, E.J.H., Johnson, M.H., and Kemner, C. (2020). Using multiple short epochs optimises the stability of infant EEG connectivity parameters. Sci. Rep. *10*, 12703. https://doi.org/10.1038/s41598-020-68981-5.

32. Khanna, A., Pascual-Leone, A., and Farzan, F. (2014). Reliability of resting-state microstate features in electroencephalography. PLoS One *9*, e114163. https://doi.org/10.1371/journal.pone.0114163.

33. Kondacs, A., and Szabó, M. (1999). Long-term intra-individual variability of the background EEG in normals. Clin. Neurophysiol. *110*, 1708–1716. https://doi.org/10.1016/S1388-2457(99)00122-4.

34. Nikulin, V.V., and Brismar, T. (2004). Long-range temporal correlations in alpha and beta oscillations: effect of arousal level and test–retest reliability. Clin. Neurophysiol. *115*, 1896–1908. https://doi.org/10.1016/j.clinph.2004.03.019.

35. van der Velde, B., Haartsen, R., and Kemner, C. (2019). Test-retest reliability of EEG network characteristics in infants. Brain Behav. *9*, e01269. https://doi.org/10.1002/brb3.1269.

36. Hedge, C., Powell, G., and Sumner, P. (2018). The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav. Res. Methods *50*, 1166–1186. https://doi.org/10.3758/s13428-017-0935-1.

37. Meisel, C., Bailey, K., Achermann, P., and Plenz, D. (2017). Decline of long-range temporal correlations in the human brain during sustained wakefulness. Sci. Rep. *7*, 11825. https://doi.org/10.1038/s41598-017-12140-w.

38. Abel, J.H., Badgeley, M.A., Meschede-Krasa, B., Schamberg, G., Garwood, I.C., Lecamwasam, K., Chakravarty, S., Zhou, D.W., Keating, M., Purdon, P.L., and Brown, E.N. (2021). Machine learning of EEG spectra classifies unconsciousness during GABAergic anesthesia. PLoS One *16*, e0246165. https://doi.org/10.1371/journal.pone.0246165.

39. Al Zoubi, O., Ki Wong, C., Kuplicki, R.T., Yeh, H.W., Mayeli, A., Refai, H., Paulus, M., and Bodurka, J. (2018). Predicting age from brain EEG signals—a machine learning approach.

Front. Aging Neurosci. *10*, 184. https://doi.org/10.3389/fnagi.2018.00184.

40. Imperatori, L.S., Cataldi, J., Betta, M., Ricciardi, E., Ince, R.A.A., Siclari, F., and Bernardi, G. (2021). Cross-participant prediction of vigilance stages through the combined use of wPLI and wSMI EEG functional connectivity metrics. Sleep *44*, zsaa247. https://doi.org/10.1093/sleep/zsaa247.

41. Price, G.W., Michie, P.T., Johnston, J., Innes-Brown, H., Kent, A., Clissa, P., and Jablensky, A.V. (2006). A multivariate electrophysiological endophenotype, from a unitary cohort, shows greater research utility than any single feature in the western Australian family study of schizophrenia. Biol. Psychiatr. *60*, 1–10. https://doi.org/10.1016/j.biopsych.2005.09.010.

42. Sitt, J.D., King, J.-R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., and Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain *137*, 2258–2270. https://doi.org/10.1093/brain/awu141.

43. Wolff, A., Di Giovanni, D.A., Gómez-Pilar, J., Nakao, T., Huang, Z., Longtin, A., and Northoff, G. (2019). The temporal signature of self: temporal measures of resting-state EEG predict self-consciousness. Hum. Brain Mapp. *40*, 789–803. https://doi.org/10.1002/hbm.24412.

44. Hatz, F., Hardmeier, M., Bousleiman, H., Rüegg, S., Schindler, C., and Fuhr, P. (2016). Reliability of functional connectivity of electroencephalography applying microstate-segmented versus classical calculation of phase lag index. Brain Connect. *6*, 461–469. https://doi.org/10.1089/brain.2015.0368.

45. Hülsemann, M.J., Naumann, E., and Rasch, B. (2019). Quantification of phase-amplitude coupling in neuronal oscillations: comparison of phase-locking value, mean vector length, modulation index, and generalized-linear-modeling-cross-frequency-coupling. Front. Neurosci. *13*, 573. https://doi.org/10.3389/fnins.2019.00573.

46. Imperatori, L.S., Betta, M., Cecchetti, L., Canales-Johnson, A., Ricciardi, E., Siclari, F., Pietrini, P., Chennu, S., and Bernardi, G. (2019). EEG functional connectivity metrics wPLI and wSMI account for distinct types of brain functional interactions. Sci. Rep. *9*, 8894. https://doi.org/10.1038/s41598-019-45289-7.

47. Gignac, G.E., and Szodorai, E.T. (2016). Effect size guidelines for individual differences researchers. Pers. Individ. Dif. *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069.

48. Gordillo, D., da Cruz, J.R., Chkonia, E., Lin, W.-H., Favrod, O., Brand, A., Figueiredo, P., Roinishvili, M., and Herzog, M.H. (2022). The EEG multiverse of schizophrenia. Cereb. Cortex, bhac309. https://doi.org/10.1093/cercor/bhac309.

49. Bosten, J.M., Goodbourn, P.T., Bargary, G., Verhallen, R.J., Lawrance-Owen, A.J., Hogg, R.E., and Mollon, J.D. (2017). An exploratory factor analysis of visual performance in a large population. Vis. Res. *141*, 303–316. https://doi.org/10.1016/j.visres.2017.02.005.

50. Cappe, C., Clarke, A., Mohr, C., and Herzog, M.H. (2014). Is there a common factor for vision? J. Vis. *14*, 4. https://doi.org/10.1167/14.8.4.

51. Garobbio, S., Pilz, K.S., Kunchulia, M., and Herzog, M.H. (2022). No common factor underlying decline of visual abilities in mild cognitive impairment. Exp. Aging Res. *1*, 1–18. https://doi.org/10.1080/0361073X.2022.2094660.

52. Shaqiri, A., Pilz, K.S., Cretenoud, A.F., Neumann, K., Clarke, A., Kunchulia, M., and Herzog, M.H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. Dev. Psychol. *55*, 1775–1787. https://doi.org/10.1037/dev0000740.

53. Cretenoud, A.F., Karimpur, H., Grzeczkowski, L., Francis, G., Hamburger, K., and Herzog, M.H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. J. Vis. *19*, 12. https://doi.org/10.1167/19.14.12.

54. Grzeczkowski, L., Clarke, A.M., Francis, G., Mast, F.W., and Herzog, M.H. (2017). About individual differences in vision. Vis. Res. *141*, 282–292. https://doi.org/10.1016/j.visres.2016.10.006.

55. Goodbourn, P.T., Bosten, J.M., Hogg, R.E., Bargary, G., Lawrance-Owen, A.J., and Mollon, J.D. (2012). Do different 'magnocellular tasks' probe the same neural substrate? Proc. Biol. Sci. *279*, 4263–4271. https://doi.org/10.1098/rspb.2012.1430.

56. Eisenberg, I.W., Bissett, P.G., Zeynep Enkavi, A., Li, J., MacKinnon, D.P., Marsch, L.A., and Poldrack, R.A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. Nat. Commun. *10*, 2319. https://doi.org/10.1038/s41467-019-10301-1.

57. Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. Psychon. Bull. Rev. *19*, 975–991. https://doi.org/10.3758/s13423-012-0322-y.

58. Simonsohn, U., Nelson, L.D., and Simmons, J.P. (2014). P-curve: a key to the file-drawer. J. Exp. Psychol. Gen. *143*, 534–547. https://doi.org/10.1037/a0033242.

59. Yarkoni, T. (2009). Big Correlations in Little Studies: inflated fMRI Correlations Reflect Low Statistical Power—commentary on Vul et al. Perspect. Psychol. Sci. *4*, 294–298. https://doi.org/10.1111/j.1745-6924.2009.01127.x.

60. Botvinik-Nezer, R., Holzmeister, F., Camerer, C.F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J.A., Adcock, R.A., et al. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. Nature *582*, 84–88. https://doi.org/10.1038/s41586-020-2314-9.

61. Carp, J. (2012). On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments. Front. Neurosci. *6*, 149. https://doi.org/10.3389/fnins.2012.00149.

62. Silberzahn, R., Uhlmann, E.L., Martin, D.P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., et al. (2018). Many analysts, one data set: making transparent how variations in analytic choices affect results. Adv. Methods Pract. Psychol. Sci. *1*, 337–356. https://doi.org/10.1177/2515245917747646.

63. Aharoni, E., and Rosset, S. (2014). Generalized α -investing: definitions, optimality results and application to public databases. J. R. Stat. Soc. B *76*, 771–794. https://doi.org/10.1111/rssb.12048.

64. Thompson, W.H., Wright, J., Bissett, P.G., and Poldrack, R.A. (2020). Dataset decay and the problem of sequential analyses on open datasets. Elife *9*, e53498. https://doi.org/10.7554/eLife.53498.

65. Weston, S.J., Ritchie, S.J., Rohrer, J.M., and Przybylski, A.K. (2019). Recommendations for increasing the transparency of analysis of preexisting data sets. Adv. Methods Pract. Psychol. Sci. *2*, 214–227. https://doi.org/10.1177/2515245919848684.

66. Poldrack, R.A., Huckins, G., and Varoquaux, G. (2020). Establishment of best practices for evidence for prediction: a Review. JAMA Psychiatr. *77*, 534–540. https://doi.org/10.1001/jamapsychiatry.2019.3671.

67. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

68. Niemann, H., Sturm, W., Thöne-Otto, A.I., and Willmes, K. (2008). CVLT California Verbal Learning Test. German Adaptation. Manual (Pearson).

69. Donders, J. (2008). A confirmatory factor analysis of the California verbal learning test—second edition (CVLT-II) in the standardization sample. Assessment *15*, 123–131. https://doi.org/10.1177/1073191107310926.

70. Mahjoory, K., Cesnaite, E., Hohlefeld, F.U., Villringer, A., and Nikulin, V.V. (2019). Power and temporal dynamics of alpha oscillations at rest differentiate cognitive performance involving sustained and phasic cognitive control. Neuroimage *188*, 135–144. https://doi.org/10.1016/j.neuroimage.2018.12.001.

71. Zimmermann, P., and Fimm, B. (2012). Testbatterie zur Aufmerksamkeitsprüfung Version 2.3.1. (Psytest).

72. Vandierendonck, A. (2018). Further tests of the utility of integrated speed-accuracy measures in task switching. J. Cogn. *1*, 8. https://doi.org/10.5334/joc.6.

73. Reitan, R.M., and Wolfson, D. (1993). The Halstead-Reitan Neuropsychological Test

Battery Theory and Clinical Interpretation (Neuropsychology Press).

74. Schmidt, K.-H., and Metzler, P. (1992). Wortschatztest : WST (Belt).

75. Kreuzpointner, L., Lukesch, H., and Horn, W. (2013). Leistungsprüfsystem 2 LPS-2 (Hogrefe).

76. Aschenbrenner, S., Lange, K.W., and Tucha, O. (2000). RWT : Regensburger Wortflüssigkeits-Test (Hogrefe).

77. Pascual-Marqui, R.D., Lehmann, D., Koukkou, M., Kochi, K., Anderer, P., Saletu, B., Tanaka, H., Hirata, K., John, E.R., Prichep, L., et al. (2011). Assessing interactions in the brain with exact low-resolution electromagnetic tomography. Philos. Trans. A Math. Phys. Eng. Sci. *369*, 3768–3784. https://doi.org/10.1098/rsta.2011.0081.

78. Hardstone, R., Poil, S.-S., Schiavone, G., Jansen, R., Nikulin, V.V., Mansvelder, H.D., and Linkenkaer-Hansen, K. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. Front. Physiol. *3*, 450. https://doi.org/10.3389/fphys.2012.00450.

79. Montez, T., Poil, S.-S., Jones, B.F., Manshanden, I., Verbunt, J.P.A., van Dijk, B.W., Brussaard, A.B., van Ooyen, A., Stam, C.J., Scheltens, P., and Linkenkaer-Hansen, K. (2009). Altered temporal correlations in parietal alpha and prefrontal theta oscillations in early-stage Alzheimer disease. Proc. Natl. Acad. Sci. USA *106*, 1614–1619. https://doi.org/10.1073/pnas.0811699106.

80. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869. https://doi.org/10.1155/2011/156869.

81. Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. Neuroimage *52*, 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003.

82. Koenig, T., Prichep, L., Lehmann, D., Sosa, P.V., Braeker, E., Kleinlogel, H., Isenhart, R., and John, E.R. (2002). Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. Neuroimage *16*, 41–48. https://doi.org/10.1006/nimg.2002.1070.

83. Custo, A., Van De Ville, D., Wells, W.M., Tomescu, M.I., Brunet, D., and Michel, C.M. (2017). Electroencephalographic resting-state networks: source localization of microstates. Brain Connect. *7*, 671–682. https://doi.org/10.1089/brain.2016.0476.

84. Pascual-Marqui, R.D., Michel, C.M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. *42*, 658–665. https://doi.org/10.1109/10.391164.

85. Brunet, D., Murray, M.M., and Michel, C.M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput. Intell. Neurosci. *2011*, 813870. https://doi.org/10.1155/2011/813870.

86. Pincus, S.M. (1991). Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. USA *88*, 2297–2301. https://doi.org/10.1073/pnas.88.6.2297.

87. Bandt, C., and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. *88*, 174102. https://doi.org/10.1103/PhysRevLett.88.174102.

88. Richman, J.S., and Moorman, J.R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol. *278*, H2039–H2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039.

89. Unakafova, V., and Keller, K. (2013). Efficiently measuring complexity on the basis of real-world data. Entropy *15*, 4392–4415. https://doi.org/10.3390/e15104392.

90. Martínez-Cagigal, V. (2015). Sample Entropy. https://ch.mathworks.com/matlabcentral/fileexchange/69381-sample-entropy.

91. Lempel, A., and Ziv, J. (1976). On the complexity of finite sequences. IEEE Trans. Inf. Theor. *22*, 75–81. https://doi.org/10.1109/TIT.1976.1055501.

92. Thai, Q. (2019). calc_lz_complexity. https://ch.mathworks.com/matlabcentral/fileexchange/38211-calc_lz_complexity.

93. Marwan, N., Carmenromano, M., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Phys. Rep. *438*, 237–329. https://doi.org/10.1016/j.physrep.2006.11.001.

94. Rosenstein, M.T., Collins, J.J., and De Luca, C.J. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. Phys. Nonlinear Phenom. *65*, 117–134. https://doi.org/10.1016/0167-2789(93)90009-P.

95. Theiler, J. (1987). Efficient algorithm for estimating the correlation dimension from a set of discrete points. Phys. Rev. A Gen. Phys. *36*, 4456–4462. https://doi.org/10.1103/PhysRevA.36.4456.

96. Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. Phys. Nonlinear Phenom. *31*, 277–283. https://doi.org/10.1016/0167-2789(88)90081-4.

97. Katz, M.J. (1988). Fractals and the analysis of waveforms. Comput. Biol. Med. *18*, 145–156. https://doi.org/10.1016/0010-4825(88)90041-8.

98. Monge-Álvarez, J. (2015). Higuchi and Katz fractal dimension measures. https://ch.mathworks.com/matlabcentral/fileexchange/50290-higuchi-and-katz-fractal-dimension-measures.

99. Tort, A.B.L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. J. Neurophysiol. *104*, 1195–1210. https://doi.org/10.1152/jn.00106.2010.

100. O'Reilly, D., Navakatikyan, M.A., Filip, M., Greene, D., and Van Marter, L.J. (2012). Peak-to-peak amplitude in neonatal brain monitoring of premature infants. Clin. Neurophysiol. *123*, 2139–2153. https://doi.org/10.1016/j.clinph.2012.02.087.

101. Toole, J.M.O., and Boylan, G.B. (2017). NEURAL: quantitative features for newborn EEG using Matlab. Preprint at ArXiv. https://doi.org/10.48550/arXiv.1704.05694.

102. Hjorth, B. (1970). EEG analysis based on time domain properties. Electroencephalogr. Clin. Neurophysiol. *29*, 306–310. https://doi.org/10.1016/0013-4694(70)90143-4.

103. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

104. Vallat, R. (2018). Pingouin: statistics in Python. J. Open Source Softw. *3*, 1026. https://doi.org/10.21105/joss.01026.

105. Székely, G.J., and Rizzo, M.L. (2013). The distance correlation t -test of independence in high dimension. J. Multivar. Anal. *117*, 193–213. https://doi.org/10.1016/j.jmva.2013.02.012.

106. Seabold, S., and Perktold, J. (2010). Statsmodels: econometric and statistical modeling with python. In 9th Python in Science Conference.

107. RStudio Team (2020). RStudio: Integrated Development Environment for R.

108. Hothorn, T., Hornik, K., Wiel, M.A.v.d., and Zeileis, A. (2008). Implementing a class of permutation tests: the coin package. J. Stat. Softw. *28*. https://doi.org/10.18637/jss.v028.i08.

109. Kassambara, A. (2021). rstatix: Pipe-Friendly Framework for Basic Statistical Tests.

110. Rizzo, M., and Székely, G. (2022). Energy: E-Statistics: Multivariate Inference via the Energy of Data.

111. Dong, L., Li, F., Liu, Q., Wen, X., Lai, Y., Xu, P., and Yao, D. (2017). MATLAB toolboxes for reference electrode standardization technique (REST) of scalp EEG. Front. Neurosci. *11*, 601. https://doi.org/10.3389/fnins.2017.00601.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
EEG and cognitive data	Leipzig Study for Mind-Body-Emotion Interactions	https://doi.org/10.1038/sdata.2018.308
Software and algorithms		
MATLAB	https://matlab.mathworks.com/	Version R2020b
Python	https://www.python.org/	Version 3.6
R Studio	https://www.r-project.org/	Version 4.0.1
Cartool	https://sites.google.com/site/cartoolcommunity	Version 3.8
LORETA	https://www.uzh.ch/keyinst/loreta	Version v20200414
Other		
Scripts for data processing	Github repository	https://github.com/dgl59311/stats_eegfeatures

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact: Dario Gordillo.

Materials availability

The study did not generate new materials.

Data and code availability

- This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- All original code has been deposited at Github and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Data from 227 participants were collected in Leipzig, Germany, as part of the Leipzig Study for Mind-Brain-Body Interactions (LEMON[24]). The sample comprises data from two age groups, 153 younger adults (between 20 and 35 years old) and 74 older adults (between 59 and 77 years old). Participants underwent a physiological and psychological screening at the Day Clinic for Cognitive Neurology of the University Clinic Leipzig and the Max Planck Institute for Human Cognitive and Brain Sciences. Written informed consent was provided by all the participants before data collection. Study protocols were in accordance with the Declaration of Helsinki and were approved by the ethics committee of the University of Leipzig (reference number 154/13-ff).

The data were made publicly available. In the present study, only data from participants that had resting-state EEG recordings were analyzed. Preprocessed resting-state EEG recordings were available for 203 participants (https://ftp.gwdg.de/pub/misc/MPI-Leipzig_Mind-Brain-Body-LEMON/EEG_MPILMBB_LEMON/). We excluded two participants (sub-010276, and sub-010277) due to differences in the sampling rate. The final

sample used for the present study consisted of 201 participants, 138 younger adults (mean age = 25.43, SD = 3.39, 42 females), and 63 older adults (mean age = 67.66, SD = 4.79, 31 females).

METHOD DETAILS

EEG collection and preprocessing

EEG was recorded using a 62-channel active electrode ActiCAP system (Brain Products GmbH, Germany) placed according to a 10-10 system arrangement with the FCz electrode as the reference. The ground was placed on the sternum. Impedances of the electrodes were kept below 5 kΩ. EEG signals were band-pass filtered online between 0.015 Hz and 1 kHz. Data were digitized with a sampling rate of 2500 Hz. During recording, participants alternated between eyes-closed and eyes-open conditions, after 1 min. A 16-min recording was obtained for each participant. Following data acquisition, signals were band-pass filtered between 1 and 45 Hz (eighth order, Butterworth filter). Data were downsampled to 250 Hz.

Further offline preprocessing consisted of the rejection of artifactual channels and segments following a visual inspection. The dimensionality of the data was reduced using principal component analysis (PCA). The PCs ($N \geq 30$) allowing to explain 95% of the total variance were retained. Other physiological artifacts, such as eye movements, blinks, or heartbeats, were identified using independent component analysis and removed. Finally, the retained components were back-projected to the electrode space. Further details on data acquisition and preprocessing are available in.[24]

In the present study, we used 8-min blocks corresponding to the eyes-closed condition segments. Missing electrodes were interpolated using spherical spline interpolation in EEGLAB[67] to fit the same 61-channel montage for all participants. Then, the recordings were re-referenced to the average and down-sampled to 125 Hz.

Cognitive assessment

In total, participants performed six cognitive tests. The data were made publicly available by the LEMON study (https://ftp.gwdg.de/pub/misc/MPI-Leipzig_Mind-Brain-Body-LEMON/Behavioural_Data_MPILMBB_LEMON/Cognitive_Test_Battery_LEMON/). From the six cognitive tests, we extracted 12 variables. The tests and the extracted variables are described below. More details can be found in the documentation of the database.

California verbal learning task

The California Verbal Learning Task (CVLT) measures memory processes and verbal learning capacity.[68] Participants listened to a 16-word list (list A) over five trials. The words belonged to four different semantic categories. After each trial, participants were asked to recall as many words from list A as they could. Then, another 16-word list (list B) was presented as an interference list, which had to be recalled right after its presentation, and had to be followed by a recall of list A. After a delay of 20 min, participants were asked to recall the words from list A, with or without semantic category cues. Based on Donders,[69] and Mahjoory and colleagues,[70] we extracted two scores: attention and delayed-memory scores. The attention score was calculated by adding up the number of words that were correctly recalled after hearing list A for the first time and the number of correctly recalled words from list B. The delayed memory score was calculated by adding the number of correctly recalled words from list A after listening to list B, and the number of correctly recalled words from list A after the 20-min delay, with and without cues.

Test of attentional performance

The Test of Attentional Performance (TAP) consists of three modules that assess different aspects of attention.[71] In the first module, participants had to press a button as soon as a cross appeared on the screen. Two conditions were tested: with and without a pre-stimulus audio signal. The alertness score was estimated from this module as the reaction time averaged across the two conditions. The second module corresponded to the Simon task. In the Simon task, participants had to press a left or right button to indicate the direction of an arrow appearing on the left or the right side of the screen. Congruent (i.e., the direction of the arrow matched its location) and incongruent (e.g., left-pointing arrow on the right side of the screen) trials were presented. The average reaction times (RT) and the percentage correct (PC) were recorded and combined into a rate correct score (i.e., RT/PC.[72] We extracted two scores from this module, given by the rate correct score for the congruent and incongruent trials of the Simon task. In the third module,

participants were presented (serially) with numbers from 1 to 9, and they had to press a button whenever the current number was the same as the second to last number (2-back task). We extracted a working memory score from this module, given by the percentage of correct matches.

Trail making test

The Trail Making Test (TMT) measures cognitive flexibility.[73] In module A, participants had to connect digits from 1 to 25 in ascending order. In module B, 13 numbers and 12 letters had to be alternately connected in their numerical and alphabetical order (e.g., 1-A-2-B-3-C- …). We extracted two scores from the TMT given by the inverse efficiency score (i.e., Task completion time/PC) for modules A and B.

Vocabulary test

The Vocabulary Test (VT) measures verbal intelligence and language comprehension.[74] Participants had to identify a target word among five distracting words. There were 42 trials. We extracted one score from the VT, given by the number of correctly identified target words over all the trials (VT-score).

Performance testing system-2 Subtest 3

The Subtest 3 of the Performance Testing System-2 (Pts-2) assesses logical deductive thinking.[75] For 3 min, 40 rows of eight symbols were presented to the participants. For each row, participants had to identify the symbol that did not follow the logical rule. We extracted one score from the Pts-2, given by the total number of correctly identified symbols.

Regensburger word fluency test

The Regensburger Word Fluency Test (Rwt) measures verbal fluency.[76] The Rwt test consisted of two modules. In the first module, for 2 min, participants had to list as many words starting with the letter "S" as they could. In the second module, participants had to list words representing animals, for 2 min. We extracted two scores from the Rwt, given the total number of correct words in each module.

EEG features extraction

Using time-domain, frequency-domain, nonlinear, and connectivity analysis methods, we extracted 175 features from the resting-state EEG signals. For some analysis methods, we filtered the EEG signals into five frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz). The dimensionality of the analysis outcomes (i.e., EEG features) depended on the analysis method. For instance, for each participant, we obtained either 61 or 80 variables if the analyses were conducted in the electrode or source space, respectively, or 4 variables for EEG features extracted using microstates analysis. Hence, each EEG feature is always composed of more than one variable. We described the analysis methods below. The list of all the EEG features extracted is available in Data S1.

Statistics of amplitude envelopes

For each frequency band, we calculated five statistical descriptors of the distribution of the signal. Amplitude envelopes were extracted using Hilbert transform. The descriptors were: mean and standard deviation of the amplitude envelopes, kurtosis, skewness, and total power of the signals. First, EEG signals were divided into non-overlapping 4-s segments and filtered into five frequency bands. Then, the statistical descriptors were calculated for the amplitude envelope values of each electrode, at each time segment. The average across time segments was used for further analyses. The analyses were conducted in the electrode space. We obtained 25 EEG features from these analyses (Features: 1–5, 68–72, 81–85, 151–155, 166–170; in Data S1).

Spectral amplitudes

To estimate spectral amplitudes, first, EEG signals were divided into non-overlapping 4-s segments. For each electrode time series, at each time segment, we used Fourier analysis to obtain the frequency amplitudes. Relative spectral amplitudes were defined for each frequency band, as the ratio between the sum of the squared Fourier coefficients within the bounds of the frequency band of interest (e.g., within 1–4 Hz, for the delta band), and the squared Fourier coefficients of the full-band signal. The average across time segments was used for further analyses. This analysis was conducted in the electrode space. We obtained five EEG features from this analysis (Features: 137–141; in Data S1).

In addition, we estimated the current source amplitudes for each frequency band using the software LORETA.[77] We defined 80 brain regions of interest (80; 40 per hemisphere) according to the AAL atlas (see Table S1). ROIs included gray matter voxels within a 10-mm radius of the seed. This analysis was conducted in the source space. We obtained five features from this analysis (Features: 156–160; in Data S1).

Temporal correlations

For each frequency band we calculated long-range and short-range temporal correlations of EEG oscillations. Long-range (>1 s) temporal correlations (LRTC) were calculated using detrended fluctuation analysis (DFA[78]). Short (∼< 1 s) temporal correlations were calculated using life – and waiting-time statistics.[79] For DFA, first, we extracted the amplitude envelopes from the EEG time series using Hilbert transform. The amplitude envelopes were integrated. Then, we defined 30 window sizes, varying from 3 to 50 s, distributed evenly on a logarithmic scale. The integrated signal was divided into 50% overlapping segments for each previously defined window-size. At each of these segments, the integrated signal was detrended, and the fluctuation function (i.e., variance) was obtained. The average fluctuation function across segments of each window size was calculated. The average fluctuation functions were plotted in logarithmic axes, and a line was fit. The slope of the line indicated the scaling exponent and this value was used for further analyses. To estimate short-range temporal correlations, we extracted the amplitude envelopes of the EEG time series using Hilbert transform. The median amplitude envelope was used as a threshold, which defined the onset and end of the oscillation bursts. The distributions for life and waiting times were built using the durations of all the burst events that occurred above or below the threshold, respectively. The 95th percentile of the distributions of life and waiting times were used for further analyses. The analyses were performed in the electrode space. We obtained 15 EEG features from these analyses (Features: 58–62, 73–77, 171–175; in Data S1).

Network and connectivity measures

First, EEG functional connectivity was calculated using five connectivity algorithms. Three of these algorithms (i.e., phase locking value, imaginary part of coherence, and weighted phase lag index) were defined in the electrode space, and two (i.e., lagged coherence and lagged phase synchronization) in the source space. EEG Functional connectivity was calculated at each of the five frequency bands. Electrode connectivity analyses were performed using FieldTrip.[80] Before estimating electrode connectivity, scalp current densities were obtained from the EEG time series using the FieldTrip function *ft_scalpcurrentdensity* with the spline method. For source connectivity analyses, first, cortical activity was estimated with the exact low-resolution electromagnetic tomography (eLORETA) algorithm using the software LORETA. We defined 80 brain regions (40 per hemisphere) according to the AAL atlas (see Table S1). ROIs included gray matter voxels within a 10-mm radius of the seed. Then, using the Brain Connectivity Toolbox (BCT [81]), three network statistics (i.e., betweenness centrality, clustering coefficient, and node strength) were calculated from each of the connectivity matrices. The BCT functions employed were *betweenness_wei*, *clustering_coef_wu*, and *strength_und*. The extracted EEG features consisted of the network analysis outcomes. We obtained 45 features in the electrode space (Features: 7–21, 32–46, 101–115, in Data S1) and 30 features in the source space (Features: 22–31, 47–56, 116–125, in Data S1).

Microstates

First, for a given participant, the voltage maps at the peaks of the global field power (GFP) signal were extracted. Maps at GFP peaks have been indicated to have a higher signal-to-noise ratio, providing a more stable representation of the EEG topographies.[82] Then, a *k*-means clustering procedure was performed on these maps with *k* (i.e., the number of cluster centroids) equal to 5. GFP peak maps were then assigned to the cluster centroid to which they showed the highest spatial correlation, as long as the correlation value was above 0.5, otherwise, maps were left unassigned.

Second, a *k*-means clustering procedure was performed on the concatenated cluster centroids from all the participants, obtained in the previous step. The algorithm was initialized 200 times for each value of *k*, with *k* varying from 1 to 15. The optimal number of subject-level cluster centroids was selected according to a metacriterion based on seven independent indicators for optimal cluster solution.[83] Out of the 200 initializations, the cluster solution showing the highest fraction of explained variance was the one retained. The polarity of the voltage maps was always ignored. Maps were only assigned to a given microstate class if they showed a spatial correlation larger than 0.5. Finally, the cluster centroids obtained from the subject-level analysis were assigned to the EEG data of each participant, this time, not only considering the GFP peaks but all the data. Voltage maps showing a correlation below 0.5 to any of the cluster centroids were left

unassigned. Temporal smoothing (*Besag factor* = 10 and *window half size* = 2) was applied to avoid the interruption of quasi-stable segments.[84] Segments equal to or smaller than three samples were rejected. For each microstate class, we extracted four temporal statistics namely the global explained variance, mean duration, time coverage, and frequency of occurrence. Microstates analysis was performed using the software Cartool version 3.8.[85] We obtained five EEG features from this analysis (Features: 86–90, in Data S1).

Entropy and complexity measures

We quantified the complexity of EEG signals using five different methods: approximate entropy, sample entropy, spectral entropy, permutation entropy, and Lempel-Ziv complexity. First, we divided the EEG signals into non-overlapping 4-s segments. Approximate,[86] permutation,[87] and sample[88] entropies were calculated for the full-band EEG signals, using an embedding dimension value of three. Approximate entropy was computed using the function *approximateEntropy* from the Predictive Maintenance Toolbox for MATLAB. Permutation entropy was calculated based on Unakafova.[89] Sample entropy was computed using the code provided by Martínez-Cagigal.[90] The time delay to estimate approximate and permutation entropies was set to one. Lempel-Ziv complexity was calculated from the full-band EEG time series,[91] using the code provided by Thai.[92] Spectral entropy was calculated for each frequency band, and it was defined as the Shannon's entropy of the ratio between the normalized power spectral density (PSD) within the frequency band bounds (e.g., within 1–4 Hz, for the delta band), and the full-band EEG signal. All calculations were performed at each time segment, and the average across segments was used for further analyses. The analyses were conducted in the electrode space. We obtained 10 EEG features from these analyses (Features: 6, 79, 80, 126, 150, 161–165, in Data S1).

Nonlinear dynamical measures

Before obtaining nonlinear dynamical features, the EEG signals were divided into non-overlapping 4-s segments. To obtain recurrence quantification analysis (RQA) features, first, for each electrode time series at each segment, we built recurrence plots and extracted eight RQA features using the CRP toolbox for MATLAB.[93] The RQA features were: determinism, entropy, laminarity, maximal diagonal line length, maximal vertical line length, mean diagonal line length, recurrence times entropy, and trapping time. Recurrence plots were constructed using a fixed radius allowing a 10% recurrence rate. The Lyapunov exponent[94] and correlation dimension[95] were obtained using the functions *lyapunovExponent*, and *correlationDimension*, available in the Predictive Maintenance Toolbox for MATLAB. We also calculated the Higuchi's,[96] and Katz's[97] fractal dimensions using the code provided by Monge-Álvarez.[98] The *kmax* parameter in Higuchi's fractal dimension calculation was set to 25. All nonlinear dynamical features were obtained from the full-band EEG signals. For RQA measures, correlation dimension, and Lyapunov exponent, the embedding parameters (time delay and embedding dimension) were calculated using the function *phaseSpaceReconstruction* from the Predictive Maintenance Toolbox in MATLAB. The analyses were conducted in the electrode space. We obtained 12 EEG features from these analyses (Features: 57, 63, 67, 78, 142–149, in Data S1)

Phase-amplitude coupling

EEG features describing cross-frequency (CF) interactions via phase-amplitude coupling were obtained using the modulation index.[99] First, EEG time series were divided into non-overlapping 4-s segments and were filtered into the five frequency bands. For each band-pass filtered electrode time series, at each segment, the amplitude envelope and the instantaneous phase were extracted using the Hilbert transform. Then, we defined ten different CF interactions, these were between delta phase-theta amplitude, delta phase-alpha amplitude, delta phase-beta amplitude, delta phase-gamma amplitude, theta phase-alpha amplitude, theta phase-beta amplitude, theta phase-gamma amplitude, alpha phase-beta amplitude, alpha phase-gamma amplitude, and beta phase-gamma amplitude. For each CF interaction, we obtained the corresponding phase and amplitude time series (e.g., phase time series in the theta band and amplitude time series in the gamma band for theta phase-gamma amplitude). Then, the phase values were binned into 18 values (from −180 to 180°) and the mean amplitude value (of the modulated frequency) over each bin was calculated. Hence, we obtained the mean amplitude value of the modulated frequency, for each phase value of the phase-modulating frequency. The Kullback-Leibler (KL) divergence indicated whether the amplitude values are uniformly distributed according to the phase values (i.e., no phase-amplitude coupling). The KL divergence was calculated for each of the ten cross-frequency interactions, and the

average value across time segments was used for further analyses. The analysis was conducted in the electrode space. We obtained ten EEG features from this analysis (Features: 91–100, in Data S1).

Time-domain amplitude features

Peak-to-peak amplitude asymmetry and coefficient of variation were calculated for each frequency band using range-EEG analysis.[100] Range features were obtained using the code provided by Toole and Boylan.[101] We also obtained the Hjorth parameters activity, mobility, and complexity from the full band EEG time series.[102] First, we divided the EEG time series into non-overlapping 4-s segments. Then, we calculated the time-domain amplitude features for each time segment and electrode signal. The average across segments was used for further analyses. The analyses were conducted in the electrode space. We obtained 13 EEG features from these analyses (Features: 64–66, 127–136, in Data S1).

QUANTIFICATION AND STATISTICAL ANALYSIS

Correlations between EEG features and cognitive variables

This section describes the analysis behind the results presented in the correlations between EEG features and cognitive variables subsection in results. To investigate the associations between the 175 EEG features and the 12 cognitive variables, we calculated Spearman correlations and distance correlations. Hence, we conducted 2100 correlation analyses (175*12; for each correlation metric) for older and younger adults separately, because of the large age differences between the samples (see experimental model and subject details). Four older adults had missing values in some cognitive variables (sub-010044 in *Tap alertness*, sub-010047 in *Tmt-A*, sub-010050 in *Tap working memory* and *Tap simon incongruent*, and sub-010099 in *Rwt s words* and *Rwt animal categories*). These participants were only excluded in the analyses where the EEG features were correlated with the variables in which they had missing values. These participants were included for the rest of the correlation analyses.

For each pair of EEG feature (with all its electrodes, brain regions, or microstate parameters) and cognitive variable (e.g., *approximate entropy* and *Tmt-A*), we calculated Spearman correlation coefficients and distance correlations. EEG features can have a different number of variables, depending on whether the analysis method was conducted in the electrode space, source space, or using microstates analysis (see method details). Thus, for EEG features obtained in the electrode space (e.g., *approximate entropy*) we obtained 61 correlations to one cognitive variable (each electrode was correlated with the cognitive variable). For EEG features in the source space (e.g., source spectral amplitude in the alpha band, *source ampl alpha*) we obtained 80 correlations and 4 correlations for each microstate class. The p values of the correlations were corrected using False Discovery Rate (FDR[103]) with an error rate of 5%. The p values of the distance correlations were obtained using 1000 repetitions of a bootstrapping procedure and corrected with FDR. We used the *distance_corr* function implemented in the *pingouin* 0.5.1 package for Python.[104]

Next, to evaluate whether EEG features showing significant correlations with the same cognitive variable relate to each other, we calculated Spearman or distance correlations between these EEG features, depending on the method used in the previous step (i.e., either Spearman or distance correlation). From each EEG feature showing a significant correlation to a cognitive variable, we took the electrode, brain region, or microstate parameter showing the highest correlation to be the representative variable for that EEG feature in the analysis. The EEG features revealing a significant correlation to the cognitive variables were also compared using multivariate distance correlations,[105] which considered all the variables of the EEG features (see correlations between EEG features showing age-related differences section below).

Dimensionality reduction and multiple regression

This section describes the analysis behind the results presented in the correlations between EEG features and cognitive variables subsection in results. We analyzed the EEG features that showed a significant correlation to a cognitive variable using principal component analysis (PCA). Importantly, EEG and cognitive variables were log-transformed to improve normality using the function *PowerTransformer* from *scikit-learn*. For example, for the *Tap working memory* variable, we found 18 EEG features showing a significant correlation. We took one variable from each of these 18 EEG features (the one showing the largest correlation to the *Tap working memory* variable) and used a PCA. We calculated the proportion of explained variance. This analysis was performed for each cognitive variable and age group separately.

Then, to understand whether the set of latent variables obtained using PCA can explain the cognitive variables, we used multiple regression. We used the function *OLS* from the *statsmodels* 0.13.2 package for Python.[106] We obtained the latent variables from the EEG features that showed a significant correlation to a cognitive variable. Each EEG feature contributed with one variable (i.e., electrode, brain region, or microstate parameter), which was the one showing the highest correlation to the corresponding cognitive variable. We generated a regression model, first, using the first PC of the EEG data, and then, we added PCs one by one, up to the third PC (i.e., model 1: PC1, model 2: PC1 and PC2, model 3: PC1, PC2, and PC3). To quantify predictive performance, we used adjusted-R^2. In this analysis, we did not use cross-validation to calculate adjusted-R^2 values.

Cross-validated prediction of cognitive variables using EEG features

This section describes the analysis behind the results presented in the prediction of cognitive variables using EEG features subsection in results. We investigated how well the EEG features predicted the cognitive variables using ridge regression and nonlinear random forest regression. We used the *Scikit-learn* 1.0.2 package for Python.[26] For each pair of EEG feature (with all its electrodes, brain regions, or microstate parameters) and cognitive variable, prediction performance was calculated using 50 repetitions of a train-test split procedure. The predictive performance was calculated using the coefficient of determination R^2. First, 33% of the data were left out for validation (testing set) and the remaining 67% of the data (training set) were used for model optimization. For linear ridge models, before training the model we applied a power transform to the data to improve normality using the *PowerTransformer* function. The amount of penalization λ (100 values from 10^{-3} to 10^5 on an evenly spaced logarithmic scale) was selected using cross-validation with the efficient leave-one-out method implemented in the function *RidgeCV*. For random forest models, we used the function *RandomForestRegressor*. The parameters of the random forest models were selected using a grid search procedure with 3-fold cross-validation as implemented in the *GridSearchCV* function. We used 100 estimators and adjusted the tree-depth considering the values 4, 6, 8, or no constraint, and also adjusted the maximum number of features using *log2*, *sqrt*, or *auto* as options. The model with the parameters giving the best performance in the training set was applied to the testing set. The prediction performance R^2 was calculated using the function *r2_score*. The R^2 values were aggregated for each of the 50 repetitions of the procedure and the median predictive performance was reported. The analysis was conducted for the sample of older and younger adults separately.

Group comparisons of EEG features between older and younger adults

This section describes the analysis behind the results presented in the *Group comparisons of the EEG features between younger and older adults* subsection in results. We conducted group comparisons between older and younger adults using each of the 175 EEG features. For each variable (61 electrodes, 80 brain regions, or 4 microstate parameters) of a given EEG feature, we conducted a Mann-Whitney test using the function *wilcox_test* from R Studio version 4.0.1 [107] and the package *coin* 1.4_2.[108] The p values and effect size *r*-values (with bootstrap confidence intervals) were obtained using the R Studio package *rstatix* 0.7.0.[109] The p values were corrected for multiple comparisons using FDR with an error rate of 5%, within each EEG feature.

Correlations between EEG features showing age-related differences

This section describes the analysis behind the results presented in the correlations between EEG features showing age-related differences subsection in results. First, for each EEG feature that contained at least one variable showing a significant difference between younger and older adults (after correcting for multiple comparisons), we selected the variable (i.e., the electrode, brain region, or microstate parameter) with the largest effect size to be the representative variable for that feature for the correlation analysis. For each age group separately, we computed pairwise Spearman correlations between these variables.

Second, to consider not only one but all variables of the EEG features revealing significant group differences, we used an unbiased multivariate distance correlation test for independence in high dimensions.[105] In high dimensions, the original distance correlation statistic (used in the correlation between EEG features and cognitive variables section) increases even under independence. The absolute unbiased estimate accounts for the bias in high dimensions and thus can provide an effect size of the relationship between two EEG features with all its variables (electrodes, brain regions, or microstate parameters) ranging from 0 to 1. We reported the square root of the absolute unbiased distance correlation ($\sqrt{|\mathscr{R}^*_n|}$) because the unbiased

distance correlation approximates the population squared distance correlation. The effect sizes and p values were obtained using the function *dcorT.test* from the R Studio package *energy* 1.7_10.[110]

Comparison between EEG reference choices

To investigate the effect of the EEG reference choice on our results, we re-analyzed the EEG data using zero-reference as implemented in the REST toolbox.[111] We obtained 140 EEG features (not considering source space EEG features) with this reference choice. These zero-referenced EEG features were compared to average or current source density CSD (for electrode connectivity features) referenced EEG features, which were the ones used in the main analyses, using intraclass correlations and distance correlations. For the comparison, we used the intraclass correlation with an absolute agreement (ICC2; to account for mean differences across references) as implemented in the function *intraclass_corr* from the *pingouin* 0.5.1 package for Python. Since EEG features have different numbers of variables (i.e., electrodes or microstate parameters), we calculated the ICC2 values for each variable of each EEG feature. For example, for the EEG feature *microstate A*, we obtained 4 ICC2 values (for global explained variance, mean duration, time coverage, and frequency of occurrence).

Furthermore, to examine whether the choice of reference affects the correlations between EEG features, we pairwise correlated the 140 EEG features obtained with zero reference using multivariate distance correlations, for younger and older adults separately.