






# Predicting executive functioning from brain networks: modality specificity and age effects

Marisa K. Heckner <sup>1,2,\*</sup>, Edna C. Cieslik <sup>1,2</sup>, Lya K. Paas Oliveros <sup>1,2</sup>, Simon B. Eickhoff <sup>1,2</sup>, Kaustubh R. Patil <sup>1,2</sup>, Robert Langner <sup>1,2</sup>

<sup>1</sup>Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, 52425 Jülich, Germany,

<sup>2</sup>Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, 40204 Düsseldorf, Germany

\*Corresponding author: Institute of Neuroscience and Medicine (INM7), Research Centre Jülich, 52425 Jülich, Germany. Email: m.heckner@fz-juelich.de

Healthy aging is associated with structural and functional network changes in the brain, which have been linked to deterioration in executive functioning (EF), while their neural implementation at the individual level remains unclear. As the biomarker potential of individual resting-state functional connectivity (RSFC) patterns has been questioned, we investigated to what degree individual EF abilities can be predicted from the gray-matter volume (GMV), regional homogeneity, fractional amplitude of low-frequency fluctuations (fALFF), and RSFC within EF-related, perceptuo-motor, and whole-brain networks in young and old adults. We examined whether the differences in out-of-sample prediction accuracy were modality-specific and depended on age or task-demand levels. Both uni- and multivariate analysis frameworks revealed overall low prediction accuracies and moderate-to-weak brain-behavior associations ( $R^2 < 0.07$ ,  $r < 0.28$ ), further challenging the idea of finding meaningful markers for individual EF performance with the metrics used. Regional GMV, well linked to overall atrophy, carried the strongest information about individual EF differences in older adults, whereas fALFF, measuring functional variability, did so for younger adults. Our study calls for future research analyzing more global properties of the brain, different task-states and applying adaptive behavioral testing to result in sensitive predictors for young and older adults, respectively.

**Key words:** mental abilities; cognitive aging; multilevel resting-state fMRI; structural MRI; predictive modeling.

## Introduction

Healthy, cognitive aging is associated with significant structural changes in the brain like cortical thinning, volumetric shrinkage, and decline in white-matter integrity (Park and Reuter-Lorenz 2008; Reuter-Lorenz and Park 2014) as well as changes in the functional network architecture (Spreng and Turner 2019). These brain changes are thought to be accompanied by a decline in cognitive capacities, in which information processing in several cognitive tasks becomes less efficient, especially in demanding tasks that tap into executive functioning (EF) (Park et al. 2002; Park and Reuter-Lorenz 2008). By contrast, performance appears to remain rather stable in tasks taxing semantic abilities (Salthouse 1996; Park et al. 2002), implicit memory, or general knowledge (Park et al. 2002). Therefore, the effects of aging on the brain, and also implications on the behavioral performance, are quite heterogeneous. Overall, previous studies indicate that the cognitive system is highly adaptive and dynamic (Greenwood 2007; Park and Reuter-Lorenz 2008), which implies that modifications of the neural architecture, such as functional reorganization, occur to maintain sufficient levels of cognitive functioning.

EF abilities are relevant for goal-directed thought and adaptive behavior in complex environments and are thus critical for everyday life. Rather than being defined as a single process, EF is a multidimensional construct that involves diverse cognitive abilities. Different lines of research suggest 3 core subcomponents: inhibitory control, working memory, and cognitive flexibility (Lehto 1996; Miyake et al. 2000; Alvarez and Emory 2006; Diamond 2013). Inhibitory control has been linked to controlling one's attention, thoughts, or emotions to attain higher-order or long-term goals. Working memory is associated with holding

content in mind and working with it. For instance, when incorporating new information in plans or considering alternatives. Finally, cognitive flexibility is important in the context of changing one's perspective or adapting to changing rules/demands. At the neural level, EF has been linked to a distributed set of brain regions that have been unified into the so-called multiple-demand network (intraparietal sulcus, inferior frontal sulcus, dorsolateral prefrontal cortex [DLPFC], anterior insula/frontal operculum, pre-supplementary motor area [pre-SMA], and anterior cingulate cortex [ACC]), and also to other brain areas, depending on specific task demands (Teuber 1972; Duncan and Owen 2000; Duncan 2010; Miyake and Friedman 2012; Camilleri et al. 2018).

Earlier studies found age-related differences in EF performance to be partially accounted for by changes in resting-state functional connectivity (RSFC) within brain networks associated with EF (Steffener et al. 2009; Langner et al. 2015; Hausman et al. 2020), and they were able to predict EF abilities of previously unseen individuals from RSFC (Reineberg et al. 2015; He et al. 2021). However, in our companion study (Heckner et al. 2023), investigating the same dataset and applying the same data analysis strategy as in the current study, but focusing on network specificity, we demonstrated overall low prediction accuracies (as indicated by the root mean squared error [RMSE], mean absolute error [MAE], and correlation coefficient [Pearson's  $r$ ]) for individual EF performance levels from within-network RSFC. Furthermore, we did not identify any network specificity, that is, EF performance was not better predicted from an EF-related brain network than from EF-unrelated networks (i.e. a perceptuo-motor network and a whole-brain approach). The overall low prediction accuracies and brain-behavior associations (coefficients of determination  $R^2 \leq 0.04$ )

Received: June 6, 2023. Revised: August 25, 2023. Accepted: August 26, 2023

© The Author(s) 2023. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

raised the question of whether the associations found are indeed meaningful. Together with previous research (Finn 2021; Finn and Bandettini 2021), these findings challenged the notion that biomarkers for individual EF performance can be identified using RSFC patterns.

Since the effects of cognitive aging as well as the behavioral consequences appear quite heterogeneous, magnetic resonance imaging (MRI) metrics capturing different aspects of brain structure and function may need to be applied. A commonly used metric derived from resting-state fMRI, regional homogeneity (ReHo), has been shown to be sensitive to age differences during rest (Wu et al. 2007) and offered a better prediction accuracy of crystallized intelligence compared to RSFC (Larabi et al. 2021). ReHo measures the local similarity of a voxel's time series to its neighboring voxels and is based on the assumption that meaningful brain activity is represented in clusters of neighboring voxels rather than single voxels (Zang et al. 2004). It has been discussed as local connectivity that is necessary to induce global connectivity (Jiang and Zuo 2016). Another metric derived from resting-state fMRI is the fractional amplitude of low-frequency fluctuations (fALFF), which reflects the relative contribution of low-frequency fluctuations within a specific frequency band to the whole frequency range (Zou et al. 2008) and can thus be taken as a measure of functional within-subject brain variability. Previous studies have identified a negative association between age and functional brain variability as measured via fALFF. These changes were associated with cortical atrophy, measured through cortical thickness or gray-matter volume (GMV), and a decline in inhibitory control (Hu et al. 2014; Vieira et al. 2020). GMV is a well-established and widely used method for quantifying regional brain morphology, which has been shown to be linked to cognitive aging, atrophy, and performance in tasks taxing executive functioning (e.g. Good et al. 2001; Gunning-Dixon and Raz 2003; Oh et al. 2014). Its wide use throughout the literature offers the advantage of comparability with other studies and results in the field.

As our recent companion paper questioned RSFC's potential as a biomarker for individual EF abilities, the aim of the current study was to investigate and systematically compare further widely investigated functional and structural brain metrics for their potential as a biomarker for EF performance. For this purpose, we defined an EF network (EFN) in the brain by integrating the results of previous neuroimaging meta-analyses (Rottschy et al. 2012; Langner et al. 2018; Worringer et al. 2019), each encompassing diverse facets of EF. We further included a perceptuo-motor (Heckner et al. 2021) and a whole-brain network (Power et al. 2011) for prediction, as EF-unrelated control networks. Then, we examined to what degree individual abilities in the 3 main EF subcomponents (i.e. inhibitory control, cognitive flexibility, and working memory) could be predicted from GMV, RSFC, ReHo, and fALFF within these networks in young and older adults, respectively. For each to-be-predicted EF performance score, we separately sought to predict performance in a high-demand (HD) task (representative of increased EF demand) and a low-demand (LD) EF control condition. We implemented a linear approach using partial least squares regression (PLSR) for prediction, as previous studies revealed comparable prediction accuracies for a nonlinear approach (random forest) or connectome-based predictive modeling (Finn et al. 2015; Shen et al. 2017; Heckner et al. 2023). Overall, we investigated (i) whether one of the structural or functional metrics (GMV, RSFC, ReHo, and fALFF) outperforms the others in predicting EF; (ii) if this pattern changes depending on the network, task-demand level, or age group; and (iii) if

young and older adults differ in their predictability depending on the structural or functional metric, network, or task-demand level.

## Methods

### Sample

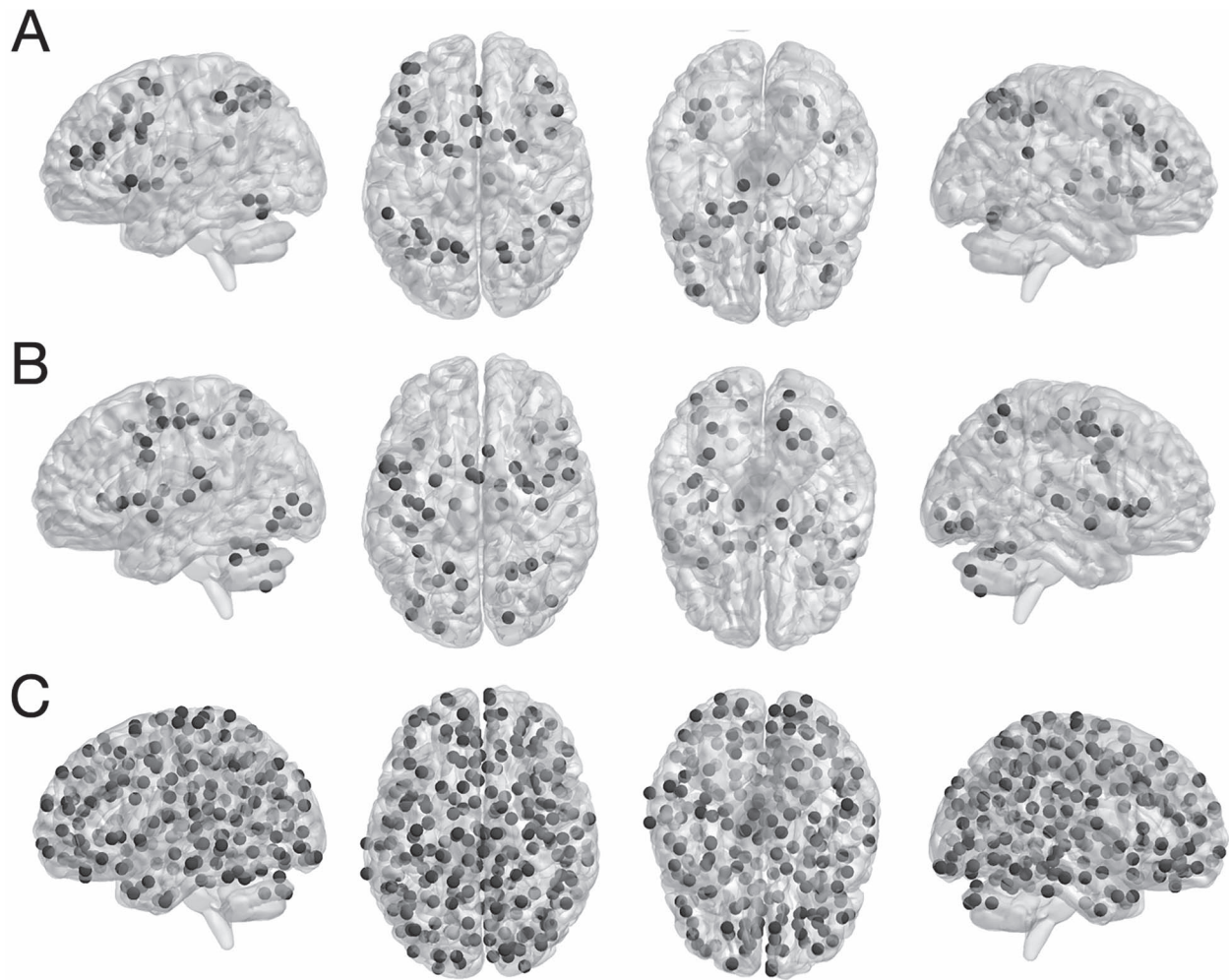
Whole-brain magnetic resonance images of 116 healthy young adults (age range = 20–40 yr, mean age = 26.67, SD = 5.80, 64 females) and 111 old adults (age range = 60–80 yr, mean age = 68.19, SD = 5.66, 72 females) were obtained from the publicly available enhanced Nathan Kline Institute-Rockland Sample (eNKI-RS; Nooner et al. 2012). These age bins were chosen to maximize the age variance for studying age-related differences in the association between brain features and behavioral target variables. We excluded participants with acute and/or severe psychiatric or neurological disorders in the past or when currently taking medication presumably affecting brain activity. The reanalysis of the data was approved by the local ethics committee of the Medical Faculty at the Heinrich Heine University Düsseldorf. All participants underwent the same protocol. The sample used was the same as in our companion paper (Heckner et al. 2023). The specific sample used is available upon request.

### Neuroimaging data acquisition and processing

Brain images were acquired on a Siemens Tim Trio 3T MRI scanner (Siemens Medical Systems, Erlangen, Germany). T1-weighted structural images were obtained using a MPRAGE sequence (TR = 1.9 s, TE = 2.52 ms, flip angle = 8°, in-plane resolution = 1.0 × 1.0 × 1.0 mm<sup>3</sup>) and were further analyzed using SPM12 (Wellcome Trust Centre Neuroimaging, London, <https://www.fil.ion.ucl.ac.uk/spm/>) and the CAT12 toolbox (Gaser et al. 2022). Whole-brain resting-state fMRI data were obtained using blood oxygen level-dependent (BOLD) contrast (gradient-echo echo planar imaging [EPI] pulse sequence, TR = 1.4 s, TE = 30 ms, flip angle = 65°, voxel size = 2.0 × 2.0 × 2.0 mm<sup>3</sup>, 64 slices, 404 volumes). Participants were instructed to keep their eyes open and maintain fixation on a central dot. Physiological and movement artifacts were removed from the RS data by using FIX (FMRIB's ICA-based Xnoiseifier, version 1.061 as implemented in FSL 5.0.9; Griffanti et al. 2014; Salimi-Khorshidi et al. 2014), which decomposes the data into independent components and identifies noise components using a large number of distinct spatial and temporal features via pattern classification. Unique variance related to the identified artifactual components is then regressed from the data. Data were further preprocessed using SPM12 and in-house MATLAB scripts. After removing the first 4 functional dummy volumes, the remaining EPI volumes were corrected for head movement by a 2-pass affine registration procedure. First, images were aligned to the initial volume and, subsequently, to the mean of all volumes. The mean EPI image was then coregistered to the gray-matter probability map provided by SPM12 using normalized mutual information and by keeping all EPI volumes aligned. Next, the mean EPI image of each participant was spatially normalized to Montreal Neurological Institute (MNI)-152 space using the "unified segmentation" approach (Ashburner and Friston 2000). The resulting deformation parameters were then applied to all other EPI volumes.

### Brain networks

We used 3 different networks for our prediction analyses, which were the same as in our companion study (Heckner et al. 2023):



**Fig. 1.** Nodes of meta-analytically defined A) executive function and B) perceptuo-motor networks, and C) Power et al.'s coordinates of putative functional areas. Taken with permission and modified from Heckner et al. (2023), copyright © 2023 Oxford University Press.

(i) An EF-related network (EFN) based on the maximum conjunction of the 3 pertinent meta-analyses investigating working memory (Rottschy et al. 2012), cognitive action regulation (Langner et al. 2018), and multitasking (Worringer et al. 2019). The resulting network comprised 50 nodes (i.e. brain coordinates). (ii) An EF-unrelated, perceptuo-motor network that integrated visual, auditory, and motor processes and comprised 59 nodes (Heckner et al. 2021). (iii) A whole-brain network as a control. We employed Power et al.'s (2011) graph of putative functional areas, which includes 264 nodes. All networks are displayed in Fig. 1.

### Brain metrics

Considering our multimodal approach, we computed metrics for structural (GMV) and functional (RSFC, ReHo, and fALFF) modalities for every node in each network. Network nodes covered a sphere with 6-mm radius around each peak coordinate. Apart from the whole-brain network based on Power et al. (2011), peaks of meta-analytic convergence were extracted using the SPM Anatomy Toolbox version 3 (Eickhoff et al. 2005, 2007) and were manually checked so that they would not overlap with each other or exceed the cortex when spheres were added. A gray-matter mask, including subcortical regions, was used to ascertain which nodes comprised gray matter (<https://zenodo.org/record/6463123#.Y1ltjsjMJ3h>).

### Gray-matter volume

Structural T1-weighted images were preprocessed and analyzed using SPM12 and the CAT12 toolbox. Within a unified segmentation model (Ashburner and Friston 2000), images were corrected for bias-field inhomogeneities; the brain tissue was classified into gray matter, white matter, and cerebrospinal fluid, and the images were spatially normalized to the MNI template using DARTEL (Ashburner and Friston 2011). Then, the segmented images were nonlinearly modulated using the Jacobian determinant derived from the normalization process to adjust them to the amount of expansion and contraction applied during normalization. GMV values were then obtained for each voxel in a given node as computed in CAT12 and were then averaged across the node.

### Resting-state functional connectivity

The variance explained by the mean white matter and cerebrospinal fluid signal was removed from the time series to reduce spurious correlations. Subsequently, the data were band-pass filtered with the cut-off frequencies of 0.01 and 0.1 Hz.

There were no significant correlations between the target variables (i.e. task scores) and sex in either subgroup. The correlation between within-scanner movement (derivative of root mean square variance over voxels) and age was significant in the older subgroup. However, we refrained from additionally correcting for movement (i.e. removing movement-related variance that is

partially shared with age). In our companion study, we additionally computed RSFC without global signal regression and with statistically removing the influence of the 6 head movement parameters ( $x, y, z$  translations and  $\alpha, \beta, \gamma$  rotations) derived from realignment, their squared values as well as their derivatives to control for possible age-specific effects of global signal regression and movement. Importantly, both corrections did not alter the results, except for the main effect of age after movement correction, which was, however, qualified by the crossed age  $\times$  task-demand-level interaction and was therefore not interpreted. As such, movement regression may have removed age-related variance from the BOLD signal time series (Heckner et al. 2023) and was therefore not applied here.

In each network, RSFC was computed by first extracting the BOLD signal time courses of all voxels within each network node expressed as the first eigenvariate. Then, pair-wise functional connectivity was computed as Fisher's Z-transformed linear (Pearson) correlation between the first eigenvariate of the time series of each network's nodes.

### Regional homogeneity

ReHo represents the homogeneity of a voxel's time series with respect to its nearest local neighbors' time courses (Zang et al. 2004) and is computed through Kendall's coefficient of concordance (KCC; Kendall and Gibbons 1990). Thus, each voxel was assigned a KCC value based on its time series homogeneity toward its nearest neighbors and was then averaged across the node.

### Fractional amplitude of low-frequency fluctuations

fALFF was computed as the ratio between power spectrum in the frequency range (0.01–0.1 Hz) and spectral power in the entire frequency range. Therefore, the time series of each voxel was transformed to the frequency domain without band-pass filtering. Then, the square root was calculated at each frequency range of the power spectrum. Per voxel, the sum of power in the 0.01–0.1 Hz frequency range was divided by the summed power spectrum across the entire frequency range (Zou et al. 2008). t-tests comparing the age groups with respect to their behavioral performance in all 3 EF tasks and 2 conditions as well as descriptive statistics can be found in our previous paper (Heckner et al. 2023).

### Behavioral measures

Executive function target variables were obtained from the eNKIRS and comprised a HD and LD (i.e. control) condition for each of 3 classical EF tasks (i.e. working memory, inhibitory control, and cognitive flexibility). All tasks used were previously evaluated and were shown to have moderate to high reliability (Delis et al. 2001; Homack et al. 2005; Gur et al. 2010). Performance distributions per age group can be found in the supplement (see Figs. S1–S6).

#### Working memory

Working memory ability was quantified using reaction times (RT) of correct responses of the 1-back (HD) and 0-back (LD) conditions of the Short Letter-N-Back Test, which is part of Penn's Computerized Neurocognitive Battery (Gur et al. 2010). In this test, participants are required to press a button if the letter on the screen is the same as the one presented in  $N$  trials before.

#### Inhibitory control

Inhibition performance was measured using RT of the incongruent (HD) and congruent (LD) conditions of the Color Word Interference (CWI) Test, which is part of the Delis-Kaplan Executive Function System (D-KEFS; Delis et al. 2004). Here,

participants are asked to name the ink color of a written word but inhibit the response to the word naming a color (same or different as the ink) itself.

### Cognitive flexibility

For quantifying cognitive flexibility, we used RT of the number and letter switching and sequencing conditions of the Trail Making Test (TMT), which is part of the D-KEFS. In this test, participants are asked to connect consecutive targets of 1 type (e.g. numbers; LD) or of 2 types (numbers and letters; HD) in an alternating fashion.

Raw performance scores for all tasks were z-transformed, and outliers with  $>3$  times the standard deviation below or above the mean were removed.

### Prediction

Individual z-transformed performance scores were then predicted from within-network GMV, RSFC, ReHo, and fALFF using PLSR (Krishnan et al. 2011). PLSR is similar to a supervised principal component regression (based on eigen-decomposition) and is thus advantageous when dimensionality reduction is beneficial for the analysis. In contrast to principal component regression, dimensionality reduction in PLSR is supervised (i.e. it uses information about the target variables), yielding the advantage that the resulting latent variables are all related to the target variables. After dimensionality reduction, a linear regressor was applied to the transformed data.

For prediction, a 10-fold cross-validation was performed for which the data were split into 10 sets, 9 of which were used for training, while the 10th was held back as a test set and was subsequently used for prediction of the unseen data. This was done with each set being the test set once. In total, 100 repetitions of this 10-fold cross-validation were computed to ensure robustness. Prediction accuracy was assessed via RMSE, MAE, and Pearson's  $r$ .

Prediction accuracy, as indicated by RMSE for the 100 repetitions, was then submitted to a 2 (age group)  $\times$  3 (network)  $\times$  2 (task-demand level)  $\times$  4 (modalities) mixed-measures ANOVA ( $P < 0.00005$ , Bonferroni-adjusted for the  $10 \times 100$  cross-validation scheme) to further assess age differences in prediction accuracy, modality specificity, and the impact of task-demand level. Therefore, prediction results for LD (i.e. 0-back, CWI congruent, and TMT consecutive) and HD (i.e. 1-back, CWI incongruent, and TMT switch) conditions were averaged into LD and HD compound scores, respectively. When Mauchly's test of sphericity was significant, Greenhouse–Geisser corrected results were interpreted.

To further corroborate the ANOVA main effects, machine-learning-adjusted t-tests for significant differences were computed (Nadeau and Bengio 2003). To account for violating the independence assumption in a paired Student's t-test, here, the variance estimate is adjusted by training and sample size.

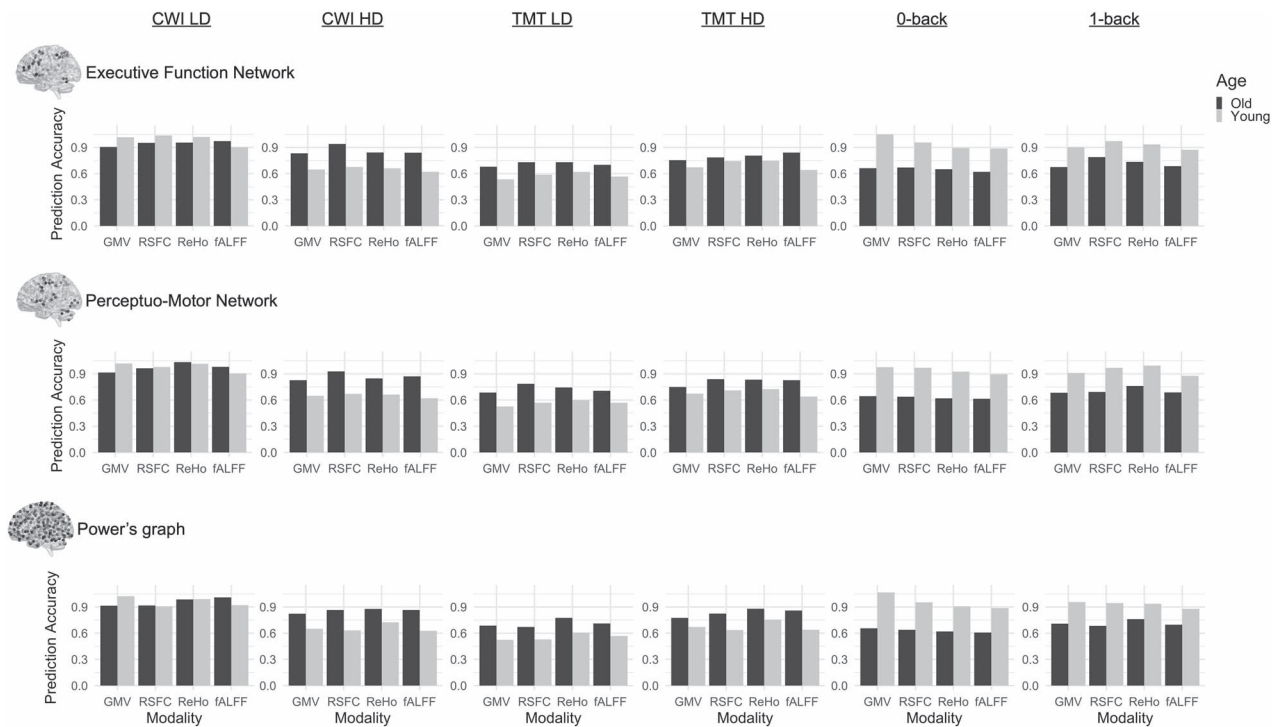
## Results

### Prediction

The averaged prediction results from the test set as indicated by RMSE are displayed in Fig. 2 and as indicated by Pearson's  $r$  are displayed in Table 1. Additional accuracy measures, including RMSE and MAE, can be found in Tables S1–S6.

### Mixed-measures ANOVA

To further assess age differences in prediction accuracy as well as network specificity and the impact of the task-demand level,



**Fig. 2.** Prediction accuracies expressed as RMSE for CWI LD congruent condition, CWI HD incongruent condition, TMT LD consecutive condition, TMT HD switch condition, 0-back, and 1-back for old (dark) and young (light) adults from prediction within the executive function network, perceptuo-motor network, and Power's graph of putative functional areas.

we submitted the prediction accuracies as given by RMSE values for each of the 100 repetitions to a 2 (age group)  $\times$  3 (network)  $\times$  2 (task-demand level)  $\times$  4 (modalities) mixed-measures ANOVA ( $P < 0.00005$ , Bonferroni-adjusted for the 10  $\times$  100 cross-validation scheme).

In a first step, prediction results for LD (i.e. 0-back, CWI congruent, and TMT consecutive) and HD (i.e. 1-back, CWI incongruent, and TMT switch) conditions were averaged into LD and HD compound scores, respectively. The ANOVA yielded significant main effects for the factors modality, network, task-demand level, and age group. These effects were qualified by 2-way, 3-way, and a 4-way interaction among all factors (see Table 2).

The complementary machine-learning-adjusted t-tests conducted to corroborate the main effects did not confirm the main effects of age and task-demand level. Network differences were only significant for perceptuo-motor versus Power networks, but not for EF versus perceptuo-motor and EF versus Power networks, respectively. All modality differences, but fALFF versus GMV, were significant (see Table S7).

We obtained the following results from the 2-way interactions. While prediction accuracy in older participants was significantly better for LD than HD task conditions, prediction accuracy in younger participants was better for HD than LD conditions (see Fig. 3A and Table 3). Regarding network differences, prediction accuracy for LD conditions was best for the whole-brain approach, as compared to the perceptuo-motor network and the EFN, whereas for HD conditions, it was best for the EFN (see Fig. 3B and Table 3). Furthermore, in older adults, prediction accuracy was best for the EFN, relative to the whole-brain and perceptuo-motor networks. Conversely, for younger adults, prediction accuracy was generally best for the whole-brain network (see Fig. 3C and Table 3).

For LD conditions, prediction accuracy was generally best for fALFF, as compared to RSFC, GMV, and ReHo, while for HD conditions, prediction accuracy was generally best for GMV (see Fig. 3D and Table 3). For ReHo and fALFF, prediction accuracy was generally best for the EFN, while for RSFC, prediction accuracy was best for the whole-brain network and for GMV for the perceptuo-motor network (see Fig. 3E and Table 3). For older adults, prediction accuracy was generally the highest for GMV, as compared to fALFF, RSFC, and ReHo, while for younger adults, prediction accuracy was best for fALFF (see Fig. 3F and Table 3).

Post-hoc pairwise comparisons revealed that prediction accuracy (i.e. RMSE) was better for older than younger participants (see Fig. 4A and Table 4). Prediction accuracy was better for HD as compared to LD (see Fig. 4B and Table 4) conditions across networks, modalities, and age groups. Across the demand level, networks, and age groups, prediction accuracy was best for fALFF as compared to GMV, RSFC, and ReHo (see Fig. 4C and Table 4). Prediction accuracy was best for the Power nodes as compared to the EFN and perceptuo-motor network (see Fig. 4D and Table 4). The EFN and perceptuo-motor network did not differ significantly.

## Discussion

The current study investigated and systematically compared to what extent the regional brain morphology (GMV) and 3 different functional brain metrics (RSFC, ReHo, and fALFF) predict the individual differences in EF abilities and whether these brain-behavior associations are modality- and/or age-specific. For this purpose, we defined 3 brain networks: an EF-related network, a perceptuo-motor network linked to visual, auditory, and motor processing, and a whole-brain network. We predicted individual EF performance scores of 3 critical EF subcomponents (i.e.

**Table 1.** Prediction results as indicated by Pearson's *r* according to brain modality, brain network and target variable for the young and old subgroup.

	CWI_LD	CWI_HD	TMT_LD	TMT_HD	0-back	1-back
GMV						
EFN						
Old	.19 <sup>a</sup>	.19 <sup>a</sup>	.13	.24 <sup>a</sup>	.01	.22 <sup>a</sup>
Young	-.09	.10	.18 <sup>a</sup>	-.01	-.18	.10
PercMot						
Old	.15	.21 <sup>a</sup>	.08	.27 <sup>a</sup>	.08	.19 <sup>a</sup>
Young	-.08	.08	.25 <sup>a</sup>	-.02	-.06	.12
Power						
Old	.18 <sup>a</sup>	.24 <sup>a</sup>	.11	.20 <sup>a</sup>	.07	.12
Young	-.07	.11	.26 <sup>a</sup>	.08	-.16	.02
RSFC						
EFN						
Old	.10	-.12	.04	.22 <sup>a</sup>	.04	-.13
Young	-.04	-.02	.02	-.27	.10	.01
PercMot						
Old	.01	-.06	-.21	-.06	.11	.15
Young	.05	-.01	.06	-.22	-.02	-.04
Power						
Old	.13	.03	.11	-.15	-.05	.10
Young	.12	.05	.19 <sup>a</sup>	.12	-.15	-.07
ReHo						
EFN						
Old	.04	.12	-.07	.07	.05	.02
Young	-.04	.12	-.15	-.26	.12	.12
PercMot						
Old	-.17	.09	-.07	-.10	.16 <sup>a</sup>	-.18
Young	-.07	.10	-.10	-.22	.01	-.11
Power						
Old	.03	.09	-.07	-.12	.16	-.04
Young	-.01	-.14	-.02	-.23	.10	.10
fALFF						
EFN						
Old	-.02	.06	.01	-.13	.14	.08
Young	.06	.09	-.15	.00	.07	.10
PercMot						
Old	-.06	.04	-.14	-.04	.18 <sup>a</sup>	.05
Young	.06	.09	-.18	.03	.04	.06
Power						
Old	-.14	.01	.04	-.15	.18 <sup>a</sup>	.05
Young	.04	.07	-.10	.04	.07	.08

Note. EFN = Executive Function Network, PercMot = Perceptuo-Motor Network, GMV = Gray-Matter Volume, RSFC = Resting-State Functional Connectivity, ReHo = Regional Homogeneity, fALFF = Fractional Amplitude of Low-Frequency Fluctuations, CWI = Color-Word Interference, TMT = Trail Making Test, LD = Low-Demand, HD = High-Demand. <sup>a</sup>*P* < 0.05.

working memory, inhibitory control, and cognitive flexibility) from within-network GMV, RSFC, ReHo, and fALFF. Finally, we submitted the prediction results to a 2 (age group) × 4 (modalities) × 3 (network) × 2 (task-demand level) mixed-measures ANOVA to assess the effects of modality and age. While prediction accuracy was overall rather low to moderate, it was better for HD than LD task conditions. This difference was especially pronounced for fALFF and GMV. However, this effect might be driven by the age × task-demand-level interaction, as prediction accuracy for younger adults was better for HD (vs. LD), whereas for older adults, it was better for LD (vs. HD) conditions. Prediction accuracy for younger adults was best with fALFF, while for older adults, the highest accuracy was achieved with GMV.

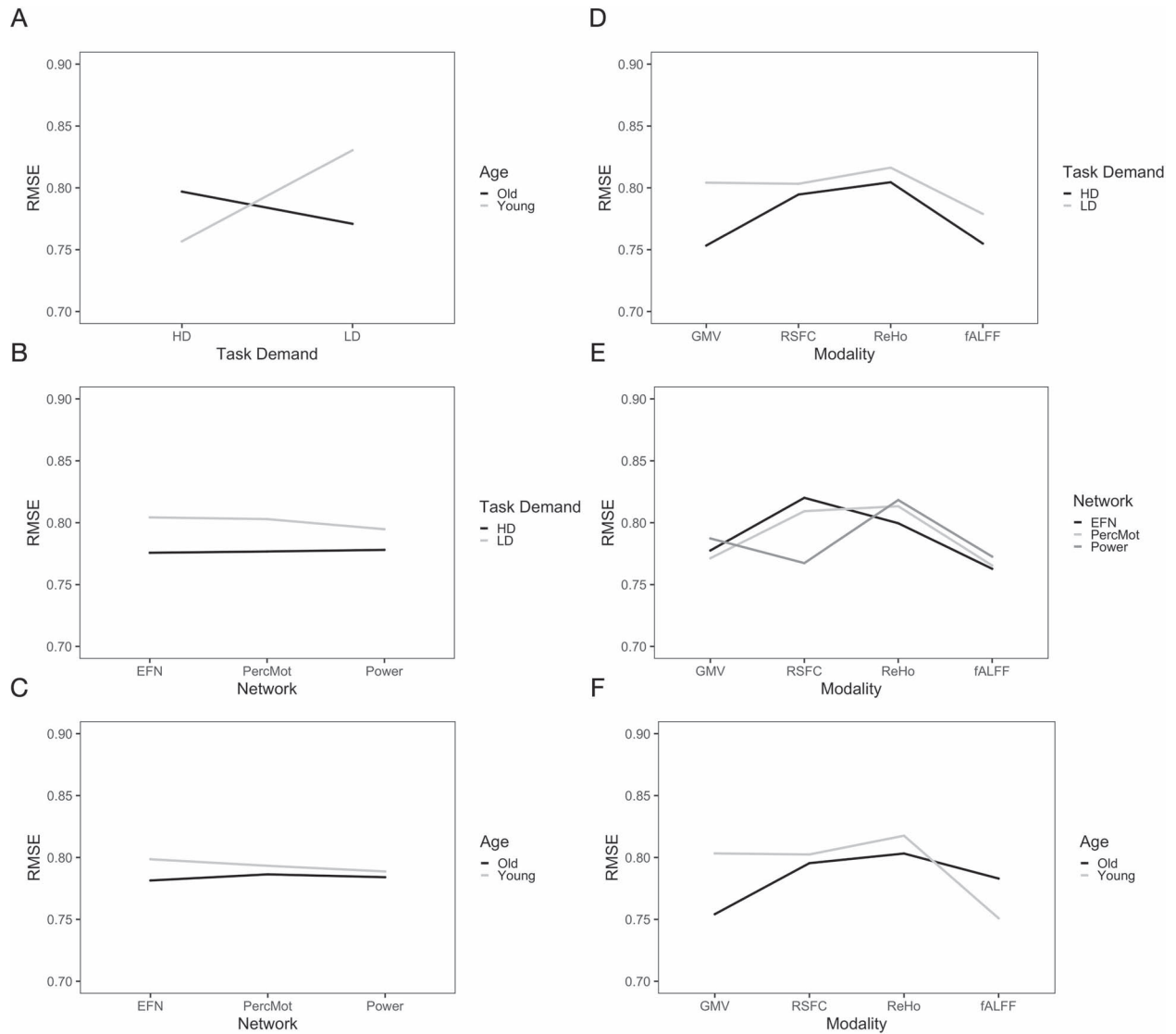
### Prediction of EF abilities

In line with our companion study (Heckner et al. 2023), as well as published guidelines for prediction analyses (Scheinost et al. 2019; Poldrack et al. 2020), we assessed our prediction results with RMSE, MAE, as well as Pearson's correlation coefficient (*r*)

as these scores offer different, yet complementary, information about the accuracy of predictive models and the association between brain metrics and behavioral target variables. Here, we will discuss prediction accuracy as measured with RMSE (<0.8) and the respective Pearson's *r* correlation coefficient (note that cognitive performance was z-scored). As mentioned above, prediction accuracies and brain-behavior associations were moderate at best, and this was also evident from the explained variance of the prediction models, as measured with the coefficient of determination *R*<sup>2</sup> (Scheinost et al. 2019). In our study, the explained variance did not exceed 6%. Whole-brain RSFC was, for example, only able to explain 3.7% of the variance in the TMT LD condition for younger adults (*r*=0.19). GMV within the EFN was able to explain 3.7% of the variance in the TMT HD condition for older adults (*r*=0.24). Similarly, but using the perceptuo-motor network, 5.2% variance was explained for the same target for older adults (*r*=0.27) and only 2.1% targeting TMT LD condition in the young (*r*=0.25) was explained. Additionally, using GMV from whole-brain features,

**Table 2.** ANOVA results.

ANOVA results	F	df	P	$\eta^2_p$
<b>Main effects</b>				
Modality	5,694.22	2,88, 569.45	0	.966
Network	79.76	2, 396	<0.001	.287
Demand level	8,387.54	1, 198	<0.001	.977
Age	1,373.24	1, 198	<0.001	.874
<b>Two-way interactions</b>				
Age × modality	4,192.85	2,88, 569.45	0	.955
Age × network	231.43	2, 396	<0.001	.539
Age × demand level	36,476.45	1, 198	<0.001	.995
Modality × network	1,701.77	5.64, 1116.63	0	.896
Modality × demand level	1,514.61	2.80, 554.22	<0.001	.884
Network × demand level	173.28	2, 396	<0.001	.467
<b>3-way interactions</b>				
Modality × network × age	54.16	6, 1188	<0.001	.215
Modality × network × demand level	115.27	5.46, 1080.37	<0.001	.368
Age × modality × demand level	368.65	2.80, 554.22	<0.001	.651
Age × network × demand level	117.33	2, 396	<0.001	.372
<b>4-way interactions</b>				
Age × network × demand level × modality	32.16	6, 1188	<0.001	.140

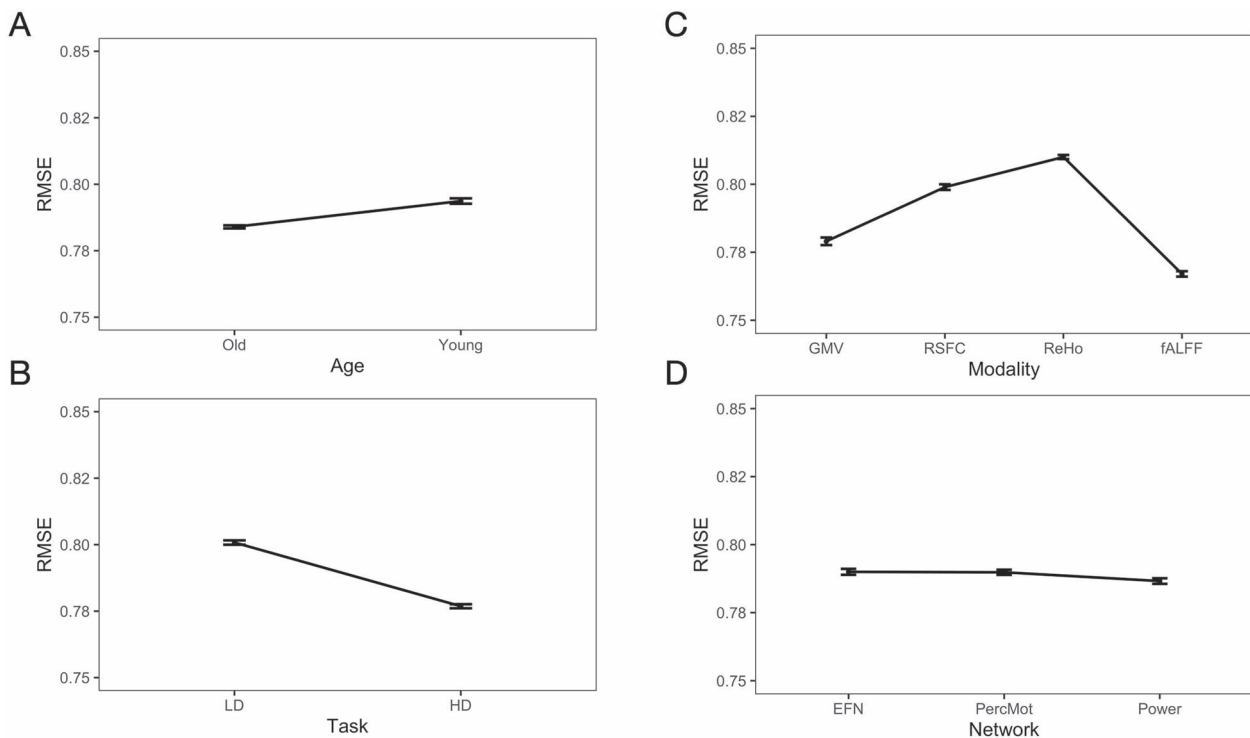


**Fig. 3.** Interaction effects for A) age × demand level, B) demand level × network, C) age × network, D) demand level × modality, E) network × modality, and F) age × modality.

**Table 3.** ANOVA interaction effects displayed as mean and standard error.

		Old	Young	LD	HD	
Age × modality	GMV	.754 (.0004)	.804 (.0005)	.805 (.0004)	.753 (.0003)	Demand × modality
	RSFC	.795 (.0003)	.802 (.0006)	.803 (.0003)	.795 (.0003)	
	ReHo	.803 (.0004)	.818 (.0004)	.816 (.0004)	.805 (.0005)	
	fALFF	.783 (.0003)	.751 (.0004)	.779 (.0003)	.755 (.0003)	
Age × network	EFN	.781 (.0003)	.799 (.0003)	.804 (.0003)	.776 (.0003)	Demand × network
	PercMot	.786 (.0003)	.793 (.0003)	.803 (.0003)	.777 (.0003)	
	Power	.784 (.0003)	.789 (.0003)	.795 (.0003)	.778 (.0004)	
Age × demand	LD	.771 (.0003)	.831 (.0003)			
	HD	.797 (.0003)	.757 (.0003)			
Modality × network	GMV		RSFC	ReHo	fALFF	
	EFN	.778 (.0003)	.820 (.0005)	.800 (.0005)	.763 (.0003)	
	PercMot	.771 (.0004)	.809 (.0004)	.713 (.0005)	.765 (.0004)	
	Power	.788 (.0005)	.767 (.0003)	.818 (.0006)	.773 (.0005)	

Note. EFN = Executive-Function-related Network, PercMot = Perceptuo-Motor-related Network, Power = Power et al.'s (2011) graph of putative functional areas, HD = High-Demand, LD = Low-Demand, GMV = Gray-Matter Volume, RSFC = Resting-State Functional Connectivity, ReHo = Regional Homogeneity, fALFF = Fractional Amplitude of Low-Frequency Fluctuations.

**Fig. 4.** Main effects for age (A), task (B), modality (C), and network (D) (mean ± standard error).

1.5% variance in TMT HD condition was explained for older adults ( $r = 0.20$ ) and 2.7% in the TMT LD condition for the young ( $r = 0.26$ ) was explained.

Predictions based on ReHo were not able to explain any variance in the target variables. Within-EFN fALFF, on the other hand, explained 2.1% of variance in the working memory LD condition for older adults ( $r = 0.14$ ) and explained 2.6% when predicting from the whole-brain ( $r = 0.18$ ). From these results, it is surprising to note that, while prediction from within-network fALFF resulted in the best overall prediction accuracy, only very little variance could be eventually explained. GMV was able to explain variance in more tasks and conditions. Nevertheless, the amount of explained variance was overall still quite low ( $R^2 < 0.06$ ). Thus, the question arises how to harmonize the findings of – on the one hand –

better prediction accuracies but – on the other hand – quite low brain–behavior associations for features extracted from fALFF. One possibility would be that within-subject functional brain variability may be highly important, even necessary, for EF but that interindividual differences in this variability, as reflected by fALFF, do not scale with individual EF abilities, at least not in the normal range of performance. More work is needed to understand the neural mechanisms and functional meaning of fALFF.

Summing up, age group, as well as task and demand levels, show moderate modality specificity. Brain–behavior associations were generally rather low but were more pronounced when predicting from structural as compared to functional features. These findings bring into question whether it is feasible to predict



**Table 4.** Post-hoc pairwise comparisons of ANOVA effects.

Post-hoc pairwise comparisons			
Factor	Mean (SE) <sub>group1</sub>	Mean (SE) <sub>group2</sub>	P
Age	.784 (.0002) <sub>old</sub>	.794 (.0002) <sub>young</sub>	$5.30 \times 10^{-91}$
Modality	.779 (.0002) <sub>GMV</sub>	.767 (.0002) <sub>fALFF</sub>	$3.27 \times 10^{-87}$
	.779 (.0002) <sub>GMV</sub>	.799 (.0002) <sub>RSFC</sub>	$4.33 \times 10^{-129}$
	.779 (.0002) <sub>GMV</sub>	.810 (.0003) <sub>ReHo</sub>	$7.67 \times 10^{-153}$
	.799 (.0002) <sub>RSFC</sub>	.767 (.0002) <sub>fALFF</sub>	$3.75 \times 10^{-166}$
	.799 (.0002) <sub>RSFC</sub>	.810 (.0003) <sub>ReHo</sub>	$3.27 \times 10^{-74}$
	.767 (.0002) <sub>fALFF</sub>	.819 (.0003) <sub>ReHo</sub>	$4.48 \times 10^{-161}$
Network	.790 (.0002) <sub>EFN</sub>	.790 (.0002) <sub>PercMot</sub>	.583
	.790 (.0002) <sub>EFN</sub>	.787 (.0002) <sub>Power</sub>	$5.01 \times 10^{-25}$
	.790 (.0002) <sub>PercMot</sub>	.787 (.0002) <sub>Power</sub>	$1.24 \times 10^{-25}$
Demand level	.777 (.0002) <sub>HD</sub>	.801 (.0002) <sub>LD</sub>	$4.84 \times 10^{-180}$

Note. SE = standard error.

individual EF abilities from functional metrics at rest as well as from a priori brain networks defined via group-level analyses. Together with recent evidence from RSFC-based predictive modeling (Heckner et al. 2023), the present results argue against network specificity, but this time, even across different brain feature modalities. Importantly, the results across both of our studies stress the need for brain measures that are not just somewhat associated with EF but can actually explain variance in individual EF abilities. While brain-behavior associations found in the present study are generally low, they are comparable to other results in the field (Ferguson et al. 2017; Greene et al. 2018; He et al. 2021), calling for more informative measures or methods and for a critical reevaluation of the predictive and explanatory values of the models examined so far.

### Modality specificity and age effects

The ANOVA yielded a main effect of modality on prediction accuracy, and post hoc pairwise comparisons revealed significant differences between all modalities. Overall, the best prediction accuracy was achieved when predicting from fALFF, followed, in descending order, by GMV, RSFC, and ReHo. However, this main effect may be explained by the age  $\times$  modality interaction. While for younger adults, prediction was best from fALFF, followed, in descending order, by GMV, RSFC, and ReHo; for older, it was best from GMV, followed by RSFC, fALFF, and ReHo. RSFC's better prediction accuracy, however, may be explained by the modality  $\times$  network interaction. This interaction revealed that prediction from RSFC was best for the whole-brain approach. One possible explanation for this finding might be the greater feature space gain for the whole-brain approach with 34,716 connections as compared to 1,225 connections for the EFN. However, in our companion study (Heckner et al. 2023), 10 random networks of the same size as the EFN still resulted in a significantly better prediction accuracy than the EFN and the perceptuo-motor network. Therefore, it does not seem to be the sheer number of features that is responsible for the prediction outcome. Lastly, all metrics were better at predicting HD (vs. LD) task conditions. This effect was especially pronounced for features extracted from GMV and fALFF. Again, this main effect is qualified by the age  $\times$  demand-level interaction, which revealed the best prediction accuracies for older adults for LD conditions, whereas for younger adults, best prediction accuracies were achieved for HD conditions. As such, the main effect of task demand appears to be driven by older adults and should not be interpreted because of the underlying crossed interaction.

Our results revealed that GMV and fALFF contained more information on the individual EF performance than did the other modalities and that this effect was age dependent. Regional GMV is very well linked to global atrophy observed in advanced age. Previous research, however, has shown that the age-related decline in GMV is especially pronounced in brain regions associated with EF, such as fronto-parietal areas (Taki et al. 2004; Chee et al. 2006; Hu et al. 2014). Furthermore, the pattern of global decline is thought to be rather consistent across older adults (bilateral pre-SMA, supplementary motor area [SMA], insula, ACC, DLPFC, inferior parietal lobule, and caudate; Bergfield et al. 2010; Giorgio et al. 2010; Taki et al. 2011). In line with these findings, the current results suggest that GMV may be a possible marker for individual EF ability levels in older adults—although one must keep in mind that prediction accuracies were still only small to moderate.

For younger adults, the best prediction accuracy was achieved with fALFF. Previous research suggested that low-frequency fluctuations of the BOLD signal are spontaneous and reflect the intrinsic connectivity of the brain (Biswal et al. 1995; Fox and Raichle 2007). Thus, fALFF is understood as a measure of functional within-subject variability that—similar to other measures of brain variability—possibly reflects cognitive adaptability (i.e. the ease of mental set reconfiguration) to task demands (Bolt et al. 2018; Uddin 2020). In young adults, the fALFF pattern appears to be linked to behavior more closely, while in older adults, fALFF does not seem to contain relevant information about individual EF performance. Previous research has shown that an age-related decrease in fALFF and GMV overlapped in prefrontal/frontal brain regions, including pre-SMA, SMA, and DLPFC (Hu et al. 2014). It was concluded that prefrontal brain regions, critical for EF, show concurrent age-related changes in structure and function. Earlier, it had been suggested that younger, faster participants show a higher variability in brain activity (as assessed by SD) across tasks and greater regional dedifferentiation of signal variability than older adults (Garrett et al. 2011). As variability may provide the kinetic energy for brain networks to explore possible functional architectures (McIntosh et al. 2010; Deco et al. 2011), an intrinsically more variable brain might be able to configure optimal networks for processing a given input toward a particular behavioral goal more flexibly and efficiently (Garrett et al. 2011).

One reason for the rather low brain-behavior associations achieved from within-network RSFC might be its unconstrained nature, as discussed in detail in our companion paper (Heckner et al. 2023). Several recent studies have shown that behavioral prediction from brain connectivity during tasks (or

movie watching) may work somewhat better than from rest (Greene et al. 2018; Sripada et al. 2020; Finn and Bandettini 2021; Kraljević et al. 2023). Tasks modulate functional brain states and may thus offer important information about the individual differences in the brain functional organization and their association with behavior (Greene et al. 2018). ReHo, a measure of local synchronicity/connectivity, is thought to induce global connectivity (i.e. RSFC; Jiang and Zuo 2016). Therefore, it would not be surprising if both local and global connectivity measures may be affected by, for example, mind wandering or thinking about a task during rest (Gregory et al. 2016). fALFF, on the other hand, as a measure of local variability or adaptability, reflecting the spontaneous, intrinsic connectivity of the brain, may therefore be influenced to a lesser degree by unconstrained thoughts during rest. Such a differential susceptibility to state effects might possibly explain the lower prediction accuracies observed for RSFC and ReHo and the superiority of fALFF even though all metrics are based on brain activity during “rest” (i.e. in a state without an externally driven task).

Differential susceptibility to state effects might have also contributed to the relatively better prediction performance observed for the within-network GMV, which was best for older adults and second best for younger adults, as compared to the other brain metrics. In particular, functional metrics might be overall more susceptible to state effects and, thus, have a reliability disadvantage, relative to structural metrics. Hence, lower prediction accuracies achieved for the functional metrics, especially RSFC and ReHo, should not prematurely be marked as weak markers for any individual performance differences but rather as a product of state–trait interactions—which would make these metrics less suitable for capturing stable, trait-level aspects of brain activity that are thought to share variance with stable cognitive traits, in particular when the amount of brain activity data available per participant is relatively limited. For more conclusive answers, future studies are needed that investigate the comparative reliability of different functional and structural brain metrics as well as the impact of age on this issue (see, e.g. Song et al. 2012, for reporting on age-related RSFC reliability differences).

Similar to previous studies (Pläschke et al. 2020; Heckner et al. 2023), prediction accuracy across modalities, networks, and task-demand levels was better for older (vs. younger) subjects, suggesting that brain–behavior associations become tighter with advancing age. Possibly, such increased associations might be due to overall age-related neural decline, such as brain atrophy or white-matter degeneration, influencing network integrity (Cabeza et al. 2016) and reorganization that is linked to EF performance. However, this main effect is qualified by the crossed age  $\times$  demand-level interaction and should thus only be interpreted with great caution.

Interestingly, we replicated this age  $\times$  demand-level interaction already reported in Heckner et al. (2023) across all structural and functional modalities such that the prediction accuracy for younger adults was consistently better when predicting HD (vs. LD) conditions, while for older adults, the reverse pattern was observed. One possible explanation is that age-related effects on the network level might still be compensated for in LD conditions, but they might not be compensated for in HD conditions taxing EF abilities. This is in line with the compensation-related utilization of neural circuits hypothesis of cognitive aging (Reuter-Lorenz and Cappell 2008) and previous research showing that age-related neurobiological decline comprises BOLD responsivity during task state. Older adults might be able to compensate LD task conditions but reach a ceiling at a certain level of processing demands

such that compensatory activation cannot be further increased in HD conditions (Nagel et al. 2011). Hence, our results emphasize the relevance of behavioral testing procedures that are more adaptive to performance differences (e.g. because of compensatory efforts) in order to have tests that are sensitive enough to capture meaningful brain–behavior associations across ability levels (cf. Heckner et al. 2023). Interestingly, our results revealed a trend toward better prediction accuracies as well as stronger brain–behavior associations for TMT scores, as compared to CWI and n-back performance, across demand level, networks, and age groups for features extracted from GMV and RSFC. Although we did not specifically aim to investigate prediction performance depending on the individual EF tasks, the TMT might be a more sensitive target variable in capturing the individual differences in EF performance as compared to the other tests used. Possibly, because the TMT measures different facets and stages of cognitive processing and might therefore be sensitive to several changes in cognition. This might, especially, be the case, when targeting LD and HD conditions separately and not subtracted (i.e. HD – LD).

Overall, despite generally low prediction accuracies, our results point out the superiority of GMV and fALFF in predicting individual differences in EF performance and indicate that this effect is age dependent. While older adults' EF performance was predicted best by features extracted from GMV, younger adults' EF performance was predicted best by fALFF features. For older adults, overall patterns, like global atrophy, seem to be most predictive of EF performance in comparison to the other metrics applied. For younger adults, fALFF appears to be most predictive of EF performance. As a measure of neural variability, fALFF may reflect the ease of exploring the best-suited network constellation for a given task. Additionally, our results stress the importance of adaptive testing in order to find meaningful brain–behavior associations.

## Conclusion and outlook

The current study investigated and systematically compared to what extent individual differences in EF performance can be predicted from structural (GMV) as well as resting-state functional (RSFC, ReHo, and fALFF) modalities. In addition, we examined whether the pattern in prediction changes with age, brain network, or task-demand level. Our results revealed overall rather low-to-moderate prediction accuracies and brain–behavior associations. Explained variance in the target variables did not exceed 6%. These findings generally question the utility of the brain metrics examined here for predicting individual differences in EF abilities. However, our results did point out the superiority of GMV and fALFF as compared to ReHo and RSFC in predicting individual EF performance. One possibility could be that individual differences in EF abilities are more strongly driven by global brain characteristics that can be better assessed with metrics for global atrophy or variability. Furthermore, our results revealed an age-related modality specificity: For older adults, structural measures of overall atrophy might be more informative, while for younger adults, functional measures of brain variability seem to contain more information about individual EF abilities.

The overall low-to-moderate prediction accuracy, as well as the missing network specificity, questions the potential of the single metrics to be applied as biomarkers for individual differences in EF performance. Rather, our findings suggest that future research may need to analyze more global properties of the brain, possibly combining different structural and functional metrics, to result in more sensitive predictors for young and older adults,

respectively. This also applies to adaptive behavioral testing as our results revealed better prediction accuracies in LD (vs. HD) task conditions for older adults, while for younger adults, prediction accuracies were better for HD (vs. LD) task conditions. Furthermore, it is important to consider the possible impact of the feature space size, especially when comparing different metrics (e.g. edge-level RSFC vs. node-level ReHo). A replication with a different, larger sample as well as different cognitive states (i.e. task performance, movie watching) and a continuous age distribution might prove useful for revealing more information contained in the brain about individual mental abilities. Lastly, considering the complexity of machine learning outputs and the increased use and relevance of these approaches in the field of cognitive neuroscience, developing appropriate methods for comparing the outcomes of different models accounting for intrinsic dependencies in the cross-validation scheme is strongly warranted.

## Author contributions

Marisa Heckner (Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing - original draft), Edna Cieslik (Conceptualization, Supervision, Writing - review & editing), Lya Paas Oliveros (Writing - review & editing), Simon B Eickhoff (Conceptualization, Funding acquisition, Supervision, Writing - review & editing), Kaustubh R. Patil (Funding acquisition, Methodology, Resources, Supervision, Writing - review & editing), Robert Langner (Conceptualization, Methodology, Supervision, Writing - review & editing)

## Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

## Funding

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, EI 816/11-1, PA 3634/1-1, EI 816/21-1, SPP2041, CRC), the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme “Supercomputing and Modeling for the Human Brain,” the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1), 785907 (HBP SGA2), and 945539 (HBP SGA3).

**Conflict of interest statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Alvarez JA, Emory E. Executive function and the frontal lobes: a meta-analytic review. *Neuropsychol Rev*. 2006;16(1):17–42.
- Ashburner J, Friston KJ. Voxel-based morphometry—the methods. *NeuroImage*. 2000;11(6):805–821.
- Ashburner J, Friston KJ. Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. *NeuroImage*. 2011;55(3):954–967.
- Bergfield KL, Hanson KD, Chen K, Teipel SJ, Hampel H, Rapoport SI, Moeller JR, Alexander GE. Age-related networks of regional covariance in MRI gray matter: reproducible multivariate patterns in healthy aging. *NeuroImage*. 2010;49(2):1750–1759.
- Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 1995;34(4):537–541.
- Bolt T, Anderson ML, Uddin LQ. Beyond the evoked/intrinsic neural process dichotomy. *Netw Neurosci*. 2018;2(1):1–22.
- Cabeza R, Nyberg L, Park DC. *Cognitive neuroscience of aging: linking cognitive and cerebral aging*. Oxford, United Kingdom: Oxford University Press; 2016.
- Camilleri JA, Müller VI, Fox P, Laird AR, Hoffstaedter F, Kalenscher T, Eickhoff SB. Definition and characterization of an extended multiple-demand network. *NeuroImage*. 2018;165:138–147.
- Chee MWL, Goh JOS, Venkatraman V, Tan JC, Gutchess A, Sutton B, Hebrank A, Leshikar E, Park D. Age-related changes in object processing and contextual binding revealed using fMR adaptation. *J Cogn Neurosci*. 2006;18(4):495–507.
- Deco G, Jirsa VK, McIntosh AR. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat Rev Neurosci*. 2011;12(1):43–56.
- Delis DC, Kaplan E, Kramer JH. Delis-Kaplan executive function system. 2001.
- Delis DC, Kramer JH, Kaplan E, Holdnack J. Reliability and validity of the Delis-Kaplan executive function system: an update. *J Int Neuropsychol Soc*. 2004;10(2):301–303.
- Diamond A. Executive functions. *Annu Rev Psychol*. 2013;64(1):135–168.
- Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci*. 2010;14(4):172–179.
- Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci*. 2000;23(10):475–483.
- Eickhoff SB, Klaas SE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*. 2005;25(4):1325–1335.
- Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage*. 2007;36(3):511–521.
- Ferguson MA, Anderson JS, Spreng RN. Fluid and flexible minds: intelligence reflects synchrony in the brain’s intrinsic network architecture. *Netw Neurosci*. 2017;1(2):192–207.
- Finn ES. Is it time to put rest to rest? *Trends Cogn Sci*. 2021;25(12):1021–1032.
- Finn ES, Bandettini PA. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage*. 2021;235:117963.
- Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci*. 2015;18(11):1664–1671.
- Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. 2007;8(9):700–711.
- Garrett DD, Kovacevic N, McIntosh AR, Grady CL. The importance of being variable. *J Neurosci*. 2011;31(12):4496–4503.
- Gaser C, Dahnke R, Thompson PM, Kurth F, Luders E, Alzheimer’s Disease Neuroimaging Initiative. CAT – a computational anatomy toolbox for the analysis of structural MRI data. 2022: bioRxiv:2022 June 11.495736.
- Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, De Stefano N, Johansen-Berg H. Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*. 2010;51(3):943–951.
- Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*. 2001;14(1):21–36.

- Greene AS, Gao S, Scheinost D, Constable RT. Task-induced brain state manipulation improves prediction of individual traits. *Nat Commun*. 2018;9(1):2807.
- Greenwood PM. Functional plasticity in cognitive aging: review and hypothesis. *Neuropsychology*. 2007;21(6):657–673.
- Gregory MD, Robertson EM, Manoach DS, Stickgold R. Thinking about a task is associated with increased connectivity in regions activated by task performance. *Brain Connect*. 2016;6(2):164–168.
- Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, et al. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*. 2014;95:232–247.
- Gunning-Dixon FM, Raz N. Neuroanatomical correlates of selected executive functions in middle-aged and older adults: a prospective MRI study. *Neuropsychologia*. 2003;41(14):1929–1941.
- Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, Brensinger C, Gur RE. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods*. 2010;187(2):254–262.
- Hausman HK, O'Shea A, Kraft JN, Boutzoukas EM, Evangelista ND, Van Etten EJ, Bharadwaj PK, Smith SG, Porges E, Hishaw GA, et al. The role of resting-state network functional connectivity in cognitive aging. *Front Aging Neurosci*. 2020;12:177.
- He L, Liu W, Zhuang K, Meng J, Qiu J. Executive function-related functional connectomes predict intellectual abilities. *Intelligence*. 2021;85:101527.
- Heckner MK, Cieslik EC, Küppers V, Fox PT, Eickhoff SB, Langner R. Delineating visual, auditory and motor regions in the human brain with functional neuroimaging: a BrainMap-based meta-analytic synthesis. *Sci Rep*. 2021;11(1):9942.
- Heckner MK, Cieslik EC, Patil KR, Gell M, Eickhoff SB, Hoffstädter F, Langner R. Predicting executive functioning from functional brain connectivity: network specificity and age effects. *Cerebral Cortex*. 2023;33(11):6495–6507.
- Homack S, Lee D, Riccio CA. Test review: Delis-Kaplan executive function system. *null*. 2005;27:599–609.
- Hu S, Chao HH-A, Zhang S, Ide JS, Li C-SR. Changes in cerebral morphometry and amplitude of low-frequency fluctuations of BOLD signals during healthy aging: correlation with inhibitory control. *Brain Struct Funct*. 2014;219(3):983–994.
- Jiang L, Zuo X-N. Regional homogeneity: a multimodal, multiscale neuroimaging marker of the human connectome. *Neuroscientist*. 2016;22(5):486–505.
- Kendall M, Gibbons J. *Rank correlation methods*. Oxford, United Kingdom: Oxford University Press; 1990.
- Kraljević N, Langner R, Küppers V, Raimondo F, Patil KR, Eickhoff SB, Müller VI. Network and state specificity in connectivity-based predictions of individual behavior. 2023. bioRxiv:20230511540387.
- Krishnan A, Williams LJ, McIntosh AR, Abdi H. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. *NeuroImage*. 2011;56(2):455–475.
- Langner R, Cieslik EC, Behrwind SD, Roski C, Caspers S, Amunts K, Eickhoff SB. Aging and response conflict solution: behavioural and functional connectivity changes. *Brain Struct Funct*. 2015;220(3):1739–1757.
- Langner R, Leiberg S, Hoffstaedter F, Eickhoff SB. Towards a human self-regulation system: common and distinct neural signatures of emotional and behavioural control. *Neurosci Biobehav Rev*. 2018;90:400–410.
- Larabi DI, Gell M, Amico E, Eickhoff SB, Patil KR. Highly accurate local functional fingerprints and their stability. 2021. bioRxiv:2021August.3.454862.
- Lehto J. Are executive function tests dependent on working memory capacity? *Q J Exp Psychol A*. 1996;49(1):29–50.
- McIntosh A, Kovacevic N, Lippe S, Garrett D, Grady C, Jirsa V. The development of a noisy brain. *Arch Ital Biol*. 2010;148(3):323–337.
- Miyake A, Friedman NP. The nature and organization of individual differences in executive functions: four general conclusions. *Curr Dir Psychol Sci*. 2012;21(1):8–14.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn Psychol*. 2000;41(1):49–100.
- Nadeau C, Bengio Y. Inference for the generalization error. *Mach Learn*. 2003;52(3):239–281.
- Nagel IE, Preuschhof C, Li S-C, Nyberg L, Bäckman L, Lindenberger U, Heekeren HR. Load modulation of BOLD response and connectivity predicts working memory performance in younger and older adults. *J Cogn Neurosci*. 2011;23(8):2030–2045.
- Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, Panek LJ, Brown S, Suavity ST, Li Q, et al. The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front Neurosci*. 2012;6:152.
- Oh H, Madison C, Villeneuve S, Markley C, Jagust WJ. Association of gray matter atrophy with age,  $\beta$ -amyloid, and cognition in aging. *Cereb Cortex*. 2014;24(6):1609–1618.
- Park DC, Reuter-Lorenz P. The adaptive brain: aging and neurocognitive scaffolding. *Annu Rev Psychol*. 2008;60(1):173–196.
- Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, Smith PK. Models of visuospatial and verbal memory across the adult life span. *Psychol Aging*. 2002;17(2):299–320.
- Pläschke RN, Patil KR, Cieslik EC, Nostro AD, Varikuti DP, Plächti A, Lösche P, Hoffstaedter F, Kalenscher T, Langner R, et al. Age differences in predicting working memory performance from network-based functional connectivity. *Cortex*. 2020;132:441–459.
- Poldrack RA, Huckins G, Varoquaux G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiat*. 2020;77(5):534–540.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. Functional network organization of the human brain. *Neuron*. 2011;72(4):665–678.
- Reineberg AE, Andrews-Hanna JR, Depue BE, Friedman NP, Banich MT. Resting-state networks predict individual differences in common and specific aspects of executive function. *NeuroImage*. 2015;104:69–78.
- Reuter-Lorenz PA, Cappell KA. Neurocognitive aging and the compensation hypothesis. *Curr Dir Psychol Sci*. 2008;17(3):177–182.
- Reuter-Lorenz PA, Park DC. How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychol Rev*. 2014;24(3):355–370.
- Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB. Modelling neural correlates of working memory: a coordinate-based meta-analysis. *NeuroImage*. 2012;60(1):830–846.
- Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*. 2014;90:449–468.

- Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychol Rev.* 1996;103(3):403–428.
- Scheinost D, Noble S, Horien C, Greene AS, Lake EMR, Salehi M, Gao S, Shen X, O'Connor D, Barron DS, et al. Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage.* 2019;193:35–45.
- Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, Constable RT. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc.* 2017;12(3):506–518.
- Song J, Desphande AS, Meier TB, Tudorascu DL, Vergun S, Nair VA, Biswal BB, Meyerand ME, Birn RM, Bellec P, et al. Age-related differences in test-retest reliability in resting-state brain functional connectivity. *PLoS One.* 2012;7(12):e49847.
- Spreng RN, Turner GR. The shifting architecture of cognition and brain function in older adulthood. *Perspect Psychol Sci.* 2019;14(4):523–542.
- Sripada C, Angstadt M, Rutherford S, Taxali A, Shedden K. Toward a “treadmill test” for cognition: improved prediction of general cognitive ability from the task activated brain. *Hum Brain Mapp.* 2020;41(12):3186–3197.
- Steffener J, Brickman AM, Rakitin BC, Gazes Y, Stern Y. The impact of age-related changes on working memory functional activity. *Brain Imaging Behav.* 2009;3(2):142–153.
- Taki Y, Goto R, Evans A, Zijdenbos A, Neelin P, Lerch J, Sato K, Ono S, Kinomura S, Nakagawa M, et al. Voxel-based morphometry of human brain with age and cerebrovascular risk factors. *Neurobiol Aging.* 2004;25(4):455–463.
- Taki Y, Thyreau B, Kinomura S, Sato K, Goto R, Kawashima R, Fukuda H. Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals. *PLoS One.* 2011;6(7):e22734.
- Teuber H. Unity and diversity of frontal lobe functions. *Acta Neurobiol Exp.* 1972;32(2):615–656.
- Uddin LQ. Bring the noise: reconceptualizing spontaneous neural activity. *Trends Cogn Sci.* 2020;24(9):734–746.
- Vieira BH, Rondinoni C, Garrido Salmon CE. Evidence of regional associations between age-related inter-individual differences in resting-state functional connectivity and cortical thinning revealed through a multi-level analysis. *NeuroImage.* 2020;211:116662.
- Worringer B, Langner R, Koch I, Eickhoff SB, Eickhoff CR, Binkofski FC. Common and distinct neural correlates of dual-tasking and task-switching: a meta-analytic review and a neuro-cognitive processing model of human multitasking. *Brain Struct Funct.* 2019;224(5):1845–1869.
- Wu T, Zang Y, Wang L, Long X, Li K, Chan P. Normal aging decreases regional homogeneity of the motor areas in the resting state. *Neurosci Lett.* 2007;423(3):189–193.
- Zang Y, Jiang T, Lu Y, He Y, Tian L. Regional homogeneity approach to fMRI data analysis. *NeuroImage.* 2004;22(1):394–400.
- Zou Q-H, Zhu C-Z, Yang Y, Zuo X-N, Long X-Y, Cao Q-J, Wang Y-F, Zang Y-F. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J Neurosci Methods.* 2008;172(1):137–141.