Received: 23 June 2021 | Revised: 9 December 2021 | Accepted: 13 December 2021

DOI: 10.1002/hbm.25759

RESEARCH ARTICLE

WILEY

Age-specificity and generalization of behavior-associated structural and functional networks and their relevance to behavioral domains

Junhong Yu[1] | Nastassja Lopes Fischer[2,3]

[1]Psychology, School of Social Sciences, National Technological University, Singapore, Singapore

[2]Centre for Family and Population Research, Faculty of Arts and Social Sciences, National University of Singapore, Singapore, Singapore

[3]Centre for Research and Development in Learning (CRADLE), Nanyang Technological University, Singapore, Singapore

Correspondence
Junhong Yu, Psychology, School of Social Sciences, Nanyang Technological University, Singapore, Singapore.
Email: junhong.yu@ntu.edu.sg

Funding information
Nanyang Assistant Professorship, Grant/Award Number: 021080-00001

Abstract

Behavior-associated structural connectivity (SC) and resting-state functional connectivity (rsFC) networks undergo various changes in aging. To study these changes, we proposed a continuous dimension where at one end networks generalize well across age groups in terms of behavioral predictions (age-general) and at the other end, they predict behaviors well in a specific age group but fare poorly in another age group (age-specific). We examined how age generalizability/specificity of multimodal behavioral associated brain networks varies across behavioral domains and imaging modalities. Prediction models consisting of SC and/or rsFC networks were trained to predict a diverse range of 75 behavioral outcomes in a young adult sample ($N = 92$). These models were then used to predict behavioral outcomes in unseen young ($N = 60$) and old ($N = 60$) subjects. As expected, behavioral prediction models derived from the young age group, produced more accurate predictions in the unseen young than old subjects. These behavioral predictions also differed significantly across behavioral domains, but not imaging modalities. Networks associated with cognitive functions, except for a few mostly relating to semantic knowledge, fell toward the age-specific end of the spectrum (i.e., poor young-to-old generalizability). These findings suggest behavior-associated brain networks are malleable to different degrees in aging; such malleability is partly determined by the nature of the behavior.

KEYWORDS
age-generalization, age-specificity, aging, behavioral predictions, functional connectivity, structural connectivity

1 | INTRODUCTION

In recent years, research using various edge-based prediction modeling approaches has mapped out the brain networks associated with a diverse range of behaviors. These approaches generally involve deriving from a training dataset, the optimal prediction weights assigned to the edges in the connectome, and then applying these weights to unseen subjects' network of edges to predict their behavioral traits. Such weights can exist in the binary form of ones and zeros in the connectome-based prediction model (Shen et al., 2017) or as continuous values in other prediction approaches such as those of partial least squares (Yoo et al., 2018) and ridge regression (Gao, Greene,

Constable, & Scheinost, 2019). These "weighted" connectomes would then constitute a behavior-associated network (BAN). Thus far, most of these networks are predominantly conceived in the resting-state functional connectivity (rsFC) modality. Some research has also shown that BANs can also be constructed using the structural connectivity (SC) modality; importantly, the behavioral predictions obtained from these SC networks are largely comparable to their rsFC counterparts (Yu et al., 2020). Generally, edge-based prediction modeling studies have been mostly carried out among young age groups (Yoo et al., 2018). These studies showed that BANs across a wide spectrum of behavior outcomes, such as personality traits (Jiang et al., 2018), cognition (Yoo et al., 2018), affect and psychopathology (Wang et al., 2021), interpersonal closeness (Hyon et al., 2020), and life satisfaction (Itahashi, Kosibaty, Hashimoto, & Aoki, 2021), generalized fairly well to unseen subjects of the same age group.

Much less research has looked at the across-age generalizability of BANs. In this regard, the limited findings have been mixed. One study showed that the rsFC network associated with processing speed generalized poorly across age groups (Gao et al., 2020), while another showed relatively better generalizability of a sustained-attention-associated functional connectivity (FC) network across age groups (Fountain-Zaragoza, Samimy, Rosenberg, & Prakash, 2019). It is likely that the BANs would undergo changes to varying degrees, in the course of aging, which could consequently affect the network's across-age generalizability. To shed light on this, we propose characterizing networks along a continuum, where age-general and age-specific BANs are located at both extreme ends. Assuming that both age-specific and age-general BANs are derived from the same age group, age-general BANs will generalize well beyond the age group they were derived from—that is they can predict behavioral traits with similar precision across age groups. On the other hand, although age-specific BANs will fare poorly in predicting behavioral traits beyond the age group they were derived from, they will outperform age-general BANs in behavioral predictions in the age group they are trained in.

While it will not be surprising if the age-specificity and generalizability of BANs can be manipulated by varying the age range of the training sample they are derived from, we put forth two other important factors that could determine the age-specificity and generalizability of BANs. The first relates to the nature of the behavior. We expect that stable behavioral traits, such as those relating to personality, are maintained by BANs that would undergo relatively few changes across the adult lifespan. On the other hand, cognitive functions, which typically undergo observable age-related decline and as a result require compensatory activation of auxiliary brain networks (Park & Reuter-Lorenz, 2009) to maintain an appropriate level of functioning, are likely to be underpinned by age-specific BANs.

The second relates to the modality of the BAN. Although whole-brain SC is on a general downward trajectory starting from early adulthood, this SC network appears to be held together by a consistent set of highly connected and central hubs which remain mostly unchanged across the lifespan (Betzel et al., 2014). Therefore, given the age-related stability in the SC network topography, we would expect that the subset of highly weighted SC edges associated with behavioral measures to undergo relatively minor changes across aging. On the other hand, rsFC is highly susceptible to experience-dependent plasticity that could occur in the context of aging. Task-based functional magnetic resonance imaging (fMRI) studies have shown previously untapped functional connections can become behaviorally relevant through the redistribution and reorganization of task-related activity (Kelly & Castellanos, 2014). Given the high correspondence between task-related FC and rsFC networks (Cole, Bassett, Power, Braver, & Petersen, 2014), we would expect experienced-related redistribution and reorganization to influence rsFC networks in a similar manner. In this regard, rsFC networks have been previously shown to undergo an age-related decrease in within-network and increase in between-network connectivity, resulting in less segregated brain networks (Damoiseaux, 2017; Zonneveld et al., 2019). In particular, the longitudinal decline in segregation between the default mode and executive control rsFC networks was associated with decreases in processing speed (Ng, Lo, Lim, Chee, & Zhou, 2016). Consequently, such segregation meant that the subset of highly weighted resting-state functional connections associated with a particular behavior would change drastically with age. To these ends, we postulated, in relative terms, that SC and rsFC BANs are age-general and age-specific, respectively.

Studying the generalizability of BANs across age groups may be useful to understand age-related differences and similarities of behavioral traits across the adult lifespan and how they relate to age-related changes in the brain. For instance, some studies have demonstrated that normal aging may lead to a reduction in rsFC preferentially at the default-mode network and the dorsal attention networks; which are thought to be implicated in cognitive functioning, such as attention and memory processes (van den Heuvel & Hulshoff Pol, 2010). Others have also reported FC decrements in normal elderly subjects in motor and salience networks (Allen et al., 2011; Onoda, Ishihara, & Yamaguchi, 2012). However, in comparison to the relatively well-documented course of the cognitive functioning across the adult's lifespan, age-related changes in the other behavioral domains and their neural correlates, such as those of affective and social skills are still understudied (Knight & Mather, 2006; Mather, 2016). The scant available evidence has pointed to relatively preserved neural mechanisms of emotion across the life span related to both positive and negative affective states (Mather, 2016). Although these results are informative, it remains to be known if brain connectivity patterns associated with behavioral traits could be generalized across age groups.

To these ends, the current study examined how generalizable these BANs are across age groups. While the concept of age generalizability or specificity generally relates to how the models derived from any age group would generalize to other age groups, in the current research context, due to sample size limitations we are only examining how well the young-derived models generalized to the old age group, but not vice-versa. Hence, subsequent references to the age-specificity and generalizability concepts should be narrowly interpreted in this context of generalizing young-derived models to the old age group. In the current study, we constructed multimodal

connectome models of a diverse range of behaviors from a young adult sample, and then assess these models' predictions in unseen subjects in the young and old age groups. The accuracy metrics of these predictions are then used to infer the age-specificity and generalizability of the BANs. In the current study, we hypothesized that (a) BANs derived from the young adult sample would produce more accurate predictions in the young than old unseen subjects; (b) multimodal networks associated with personality traits and cognitive functions are likely to fall at the age-general and age-specific ends of the spectrum, respectively; and (c) SC BANs are significantly more age-generalizable than their rsFC counterparts.

2 | MATERIALS AND METHODS

2.1 | Participants and procedures

We used the publicly accessible "Leipzig Study for Mind-Body-Emotion Interactions" (LEMON; Babayan et al., 2019) dataset in the current study. Participants from this dataset were recruited via public advertisements, leaflets, online advertisements, and information events disseminated at the University of Leipzig. To be eligible, participants should not report any neurological disorders, head injury, alcohol or other substance abuse, hypertension, pregnancy, claustrophobia, chemotherapy, and malignant diseases, current and/or previous psychiatric disease, or any medication affecting the cardiovascular and/or central nervous system in a telephone prescreening. After completing the health report, eligible individuals were invited to the laboratory. Participants earned monetary compensation upon completing the data collection procedures, which comprised of the MRI scanning, a resting-state electroencephalography recording, and a battery of psychological assessments, such as emotion, cognitive and personality tests, and a psychiatric interview. The study protocol conformed to the Declaration of Helsinki and was approved by the ethics committee at the medical faculty of the University of Leipzig (reference number 154/13-ff).

This dataset included 227 adults, which were divided into two age groups—the young age group consisted of 153 individuals (45 females) aged from 20 to 35 years old ($Age_{mean} = 25.1$; $SD = 3.1$); and the old age group consisted of 74 (37 females) participants aged from 59 to 77 years old ($Age_{mean} = 67.6$; $SD = 4.7$). We subsequently excluded 1 and 14 subjects from the young and old age groups, respectively, due to excessive head motion during the MRI scans.

2.2 | Behavioral measures

Participants in the LEMON study were administered a diverse range of behavioral measures. All measures were administered via their German-translated versions. Among the behavioral measures assessed in the LEMON study, we excluded state measures of affect. Given the volatility of such measures, it is likely that participants' affective states would change from the time they were measured to the moment they undergo the MRI scan, thus confounding brain-behavior associations. Additionally, we also excluded cognitive measures involving error counts, as most participants scored 0 in these measures. Consequently, a total of 75 behavioral outcomes were included, they can be broadly grouped in the personality-general, coping, affect, socio-cognitive and cognitive domains (see Table 1). These selected behavioral outcomes are described in detail in the supplementary materials.

2.3 | MRI acquisition

Participants were scanned using a 3 T scanner (MAGNETOM Verio, Siemens Healthcare GmbH, Erlangen, Germany) equipped with a 32-channel head coil. T1-weighted images were acquired using a Magnetization Prepared 2 Rapid Acquisition Gradient Echoes protocol (TE = 2,920 ms; TR = 5,000 ms; TI1 = 700 ms; TI2 = 2,500 ms; FOV = 256 mm; 176 sagittal slices; voxel size = 1 mm isotropic). For the resting-state fMRI, 657 T2*-weighted gradient-echo EPI volumes were acquired (TR = 1,400 ms; TE = 30 ms; 64 axial slices; matrix = 88 × 88; voxel size = 2.3 mm isotropic). During the resting-state fMRI scan, subjects were instructed to remain awake and lie still with their eyes open while looking at a fixation cross. Diffusion-weighted images were acquired using a high angular resolution diffusion imaging protocol, consisting of 7 b0 images, 60 diffusion directions at $b = 1{,}000$ s/mm^2 (TE = 80 ms; TR = 7,000 s; 128 × 128 matrix, 88 axial slices, voxel size = 1.7 mm isotropic). Additionally, gradient echo field maps (TR = 680 ms; TE1 = 5.19 ms; TE2 = 7.65 ms), and spin-echo images with reversed phase encoding (TR = 7,000 ms; TE = 80 ms; echo spacing = 0.78 ms) acquired were acquired for correcting the rsfMRI and DWI distortions, respectively.

2.4 | Image processing

Distortion correction was carried out on the EPI volumes using FSL FUGUE with the subject's magnitude and phase difference images. Following which, these corrected volumes were preprocessed using data processing assistant for resting-state fMRI (DPARSF), advance version (Yan, Wang, Zuo, & Zang, 2016). Briefly, the first 10 volumes were removed to allow for T1 equilibrations effects. Next, the middle slice was used as the reference slice for slice time correction and motion correction realignment. Following this, the corrected images were coregistered to their respective T1-weighted image, and segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) tissue maps using the DARTEL algorithm. Then, nuisance covariates regression, consisting of Friston 24 head-motion parameters, mean WM and CSF, and linear detrending was carried out. Global signal regression was also carried out to maximize the associations between FC and behavioral variables (Li et al., 2019). To further remove the effects of head motion, scrubbing was carried out on volumes with framewise displacement ($FD_{Jenkinson}$) > 0.2 mm (Yan et al., 2013). Additionally, participants with excessive head motion, as

TABLE 1 List of behavioral measures

Personality-general	Coping	Affect	Socio-cognitive	Cognitive
LOT-R optimism	COPE seeking emotional support	Hamilton depression rating scale	TAS identifying feelings	CVLT Total learning
LOT-R pessimism	COPE behavioral disengagement	NEO-FFI neuroticism	TAS describing feelings	CVLT short delay
LOT-R overall score	COPE positive reframing	NEO-FFI extraversion	TAS externally-oriented thinking	CVLT long delay
NEO-FFI openness	COPE humor	STAI trait anxiety	TAS overall score	CVLT delayed recognition
NEO-FFI agreeableness	COPE substance use	STAXI trait anger	TeiQueSF global trait emotional intelligence	LPS-2 fluid intelligence
NEO-FFI conscientiousness	COPE use of informational support	STAXI trait anger temperament	TeiQueSF self-control	RWT category fluency
UPPS urgency	COPE venting	STAXI trait anger response	TeiQueSF emotionality	RWT letter fluency
UPPS lack of premeditation	COPE planning	STAXI anger-in	TeiQueSF sociability	TAP alertness (with audio RT)
UPPS lack of perseverance	COPE acceptance	STAXI anger-out	TeiQueSF well-being	TAP alertness (without audio RT)
UPPS sensation seeking	COPE self-blame	STAXI anger-control		TAP incompatibility (congruent RT)
BIS/BAS drive fun seeking	COPE religion			TAP incompatibility (incongruent RT)
BIS/BAS drive reward responsiveness	COPE denial			TAP incompatibility (*F* value)
BIS/BAS drive avoidance behavior	COPE active coping			TAP incompatibility (*F*%)
	CERQ self-blame			TAP working memory
	CERQ acceptance			TMT part A
	CERQ rumination			TMT part B
	CERQ positive refocusing			WST verbal intelligence
	CERQ refocusing on planning			
	CERQ positive reappraisal			
	CERQ putting into perspective			
	CERQ catastrophizing			
	CERQ blaming others			
	CERQ self-distraction			
	ERQ reappraisal			
	ERQ suppression			

Abbreviations: BIS/BAS, Behavioral Inhibition and Approach System; CERQ, Cognitive Emotion Regulation Questionnaire; COPE, Coping Orientations to Problems Experienced; CVLT, California Verbal Learning Task; ERQ, Emotion Regulation Questionnaire; LOT-R, Optimism Pessimism Questionnaire-Revised; LPS, Performance Testing System; NEO-FFI, Big-Five of Personality; RWT, Regensburger Word Fluency Test; STAI, State–Trait Anxiety Inventory (Short version); STAXI, State–Trait Anger Expression Inventory; TAP, Test of Attentional Performance; TAS, Toronto-Alexithymia Scale; TEIQue-SF, Emotional Intelligence Questionnaire; TMT, Trail Making Test; UPPS, Impulsivity Questionnaires; WST, Vocabulary Test.

operationalized by having a mean $FD_{Jenkinson}$ > 0.2 mm, were excluded. Next, the images were normalized by DARTEL to MNI space with a voxel size of $3 \times 3 \times 3\ mm^3$ and smoothed using a 4 mm FWHM kernel. A band-pass filter of 0.01–0.1 Hz was applied to the signal to remove the high-frequency physiological noise and low-frequency drift.

For the network construction, the brainnetome atlas (Fan et al., 2016) was used to parcellate the whole brain into 246 anatomical regions corresponding to the nodes of the network. For each participant, the time course of each node was extracted, and Fisher's z transformed correlations between each pair of nodes (i.e., edge) were computed, resulting in a 246 × 246 connectivity matrix.

For the preprocessing of diffusion-weighted images, inhomogeneity field maps were first generated from the subject's pair of spin-echo images using FSL topup (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). Subsequent preprocessing of the diffusion images, tractography, and construction of the structural connectome were carried out using MRtrix3 (Tournier et al., 2019). Briefly, MP-PCA denoising (Veraart et al., 2016) and removal of Gibbs ringing artifacts (Kellner, Dhital, Kiselev, & Reisert, 2016) were carried out on the raw images. Subsequently, they were corrected for motion, eddy currents, and susceptibility-induced distortion using the inhomogeneity field maps obtained previously. Following this, bias field correction (Tustison et al., 2010) was carried out. Next, the GM), WM, and CSF response functions were obtained using the Dhollander (Dhollander & Connelly, 2016) algorithm which are in turn used for estimating fiber orientation distributions (FOD) in the WM, GM, and CSF tissues from diffusion data using spherical deconvolution. Then, anatomically constrained tractography (Smith, Tournier, Calamante, & Connelly, 2012) was carried out using the WM FOD. This involved the prior preparation of a GM mask from the segmentation of the subject's T1 structural image using FSL FAST, and then using this mask to seed streamlines.

The AAL-90 atlas (Tzourio-Mazoyer et al., 2002) was used to parcellate the whole brain. For each subject, the AAL-90 template was first warped to the subject's native DTI space to obtain the transformations. Then, the warped template was overlaid onto the subject's diffusion tensors for visual inspection of the alignment. The transformations obtained previously were then applied to warp the AAL-90 atlas into the subject's native diffusion space. For the network construction in the subject's native diffusion space, the nodes i and j were thought to be connected by an edge ($e_{ij} = [i, j]$), if at least one reconstructed streamline was found with its two endpoints located within the two nodes, respectively. The edges in the connectivity matrix for each participant were operationalized as the number of streamlines connecting between each pair of regions. Finally, the thresholded matrices were normalized using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010) within MATLAB.

2.5 | Statistical analysis

The partial least squares regression (PLSR; Yoo et al., 2018) approach was used to construct connectome models for behavioral predictions. Briefly, PLSR finds a linear regression solution by projecting predictors onto a new dimensional space in relation to the outcome variables. Such dimensional reduction meant the PLSR approach is well-suited in situations where the predictors are highly correlated among themselves, such as in the case of brain connectomes.

PLSR was carried out separately on the SC and rsFC modalities to construct the prediction models associated with each behavioral measure in a training dataset, which were subsequently applied to the testing dataset. First, the SC and FC matrices from the young sample were assigned to a training ($N = 92$; young-train) and testing dataset ($N = 60$; young-test). We decided on this ratio of train-test assignment, such that the testing datasets for the young and old equally are matched in terms of sample size. In order to minimize overfitting, we determined the optimal component number (k) via a fivefold cross-validation procedure within the young-train sample; we tested PLSR models with k values of 1–10, and selected the k-component solution corresponding to the lowest mean square error of prediction. Having identified the optimal k value, we carried out the PLSR again on the full young-train dataset, this time setting the number of components to k, to obtain the beta coefficients, which are then applied to the young-test dataset, as well as the old age group (i.e., old-test; $N = 60$) to obtain two sets of modality-specific predicted behavioral scores, in these two unseen samples.

Next, we created another array of models that combined the SC and rsFC features. First, we calculated the network strength—the dot product of the model coefficients (derived from the young-test sample) and the subject's connectivity matrix, separately in the SC and rsFC modalities. Then, within the young-train sample, we regressed the behavioral scores onto the SC and rsFC network strengths, to obtain their regression coefficients which were subsequently applied to the network strengths in the young-test and old-test samples, to generate their predicted behavioral scores. Hence, for each behavioral measure in each subject in the young-test and old-test datasets, there were three predicted scores from the SC-only, rsFC-only, and the combined SC and rsFC models.

To assess the accuracy of predictions we computed the normalized root mean square error (NRMSE) between the predicted and observed scores for each of the three models. The NRMSE is the root mean square error (RMSE) divided by the mean observed score, and is calculated separately for the young-test and old-test samples (i.e., $\mathrm{NRMSE_{old}} = \mathrm{RMSE_{old}}/\text{mean observed score}_{\mathrm{old}}$; $\mathrm{NRMSE_{young}} = \mathrm{RMSE_{young}}/\text{mean observed score}_{\mathrm{young}}$). The NRMSE can be compared across measures and samples; lower NRMSE values correspond to more accurate predictions. To assess the degree of young-to-old age generalizability, we calculate the ratio of $\mathrm{RMSE_{young}}/\mathrm{RMSE_{old.}}$ These RMSE ratios can similarly be compared across measures. Ratios close to 1 meant that the model predicted scores at similar levels of precision in both the young-test and old-test samples, alluding to greater generalizability of the young-train model to the old-test sample. On the other hand, lower RMSE ratios indicate that predictions are more accurate in the young-test than the old-test samples. Finally, we also calculated the standardized regression coefficients of the SC and rsFC network strengths, and take their ratio as a rough indicator of the SC and rsFC networks' share of contribution in predicting the behavioral scores in the young sample.

Given that the way the young subjects were shuffled into the young-train and young-test samples would affect the optimal k value and subsequently the beta coefficients and prediction metrics, the above-mentioned procedures were repeated 1,000 times, with participants randomly assigned into the young-train and young-test samples each time. After which, the prediction metrics and beta coefficients obtained at the end of each iteration were averaged across the 1,000 iterations. These analyses are carried out in MATLAB (R2019b). An overview of the training/testing paradigm is illustrated in Figure 1.

We do not expect the PLSR model derived from the young-train sample to produce meaningful predictions in the young-test sample for all behavioral measures. If the model fared poorly in predicting

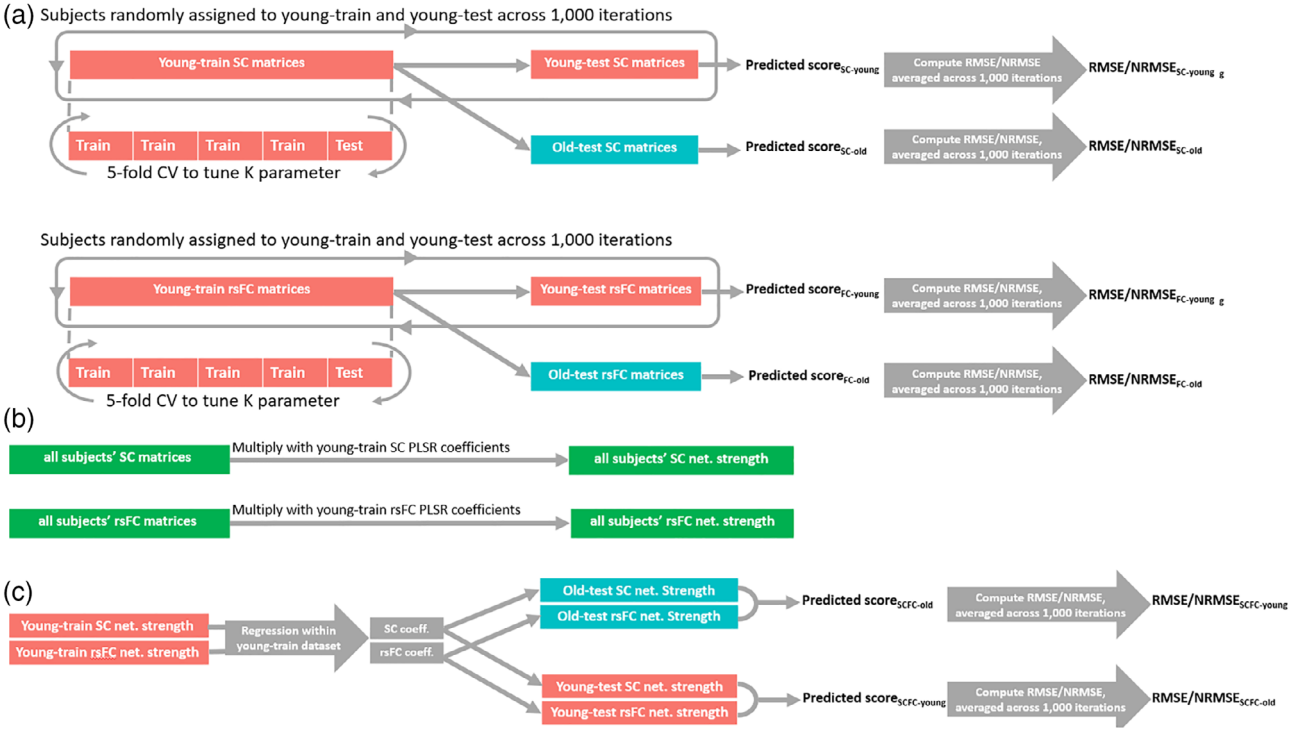

FIGURE 1 Overview of training/testing paradigm. (a) Within the SC and rsFC modalities, PLSR models were trained on the young-train dataset to predict scores in the young-test and old-test datasets. The optimal component number was tuned within the young-train dataset via fivefold cross-validation. This whole process was repeated 1,000 times, with young subjects randomly assigned to the young-train and young-test datasets each time. (b) SC and functional connectivity (FC) network strengths were computed as the dot products of the network matrices and their respective PLSR coefficients. (c) The young-train SC and FC network strengths were entered into a regression to obtain their respective coefficients, which are subsequently applied to the young-test and old-test network strengths to obtain predicted scores in the combined SC and rsFC model. NRMSE, normalized root mean square error; PLSR, partial least squares regression; RMSE, root mean square error; rsFC, resting-state functional connectivity; SC, structural connectivity

behavioral scores among unseen subjects of the same age group, it would be meaningless to examine the generalizability of the model to unseen subjects of another age group. To this end, for the reporting of the results, we excluded behavioral measures that are poorly predicted in the young-test sample, as defined by having an NRMSE ≥0.5 for the combined SC and rsFC model.

For the comparison of the model characteristics and prediction metrics between modalities or age groups, we use paired samples *t* tests and repeated measures analysis of variance (ANOVA) with the modality as the within-subject factor. One-way between-subject ANOVAs were used to compare prediction metrics between behavioral domains. Statistical significance was set at $p < .05$. The MATLAB and R codes for the analyses and generating the figures are available at https://osf.io/u6zck/.

3 | RESULTS

3.1 | Comparing behavioral outcomes across age groups

Figure 2 illustrates the differences in behavioral outcomes, as quantified by Cohen's *d*s, across both age groups. In general, relatively large differences (i.e., Cohen's $d > 0.5$) exist across most outcomes in the cognitive domains. On the contrary, scores of affective outcomes tend to be largely similar across age groups.

3.2 | Characteristics of trained PLSR models and validation in the young-test sample

PLSR models were trained using the young-train dataset, across the behavioral spectrum and modalities. Figure 3a shows the average *k* values in these models. These *k* values were generally low (≤5). In particular, this mean was consistently lower in the SC than the rsFC modality as suggested by a paired samples *t* test ($t(66) = 6.16$, $p < .001$, Cohen's $d = 0.75$). Figure 3b shows the ratio of the SC and rsFC regression coefficients across the behavioral spectrum. In general, the rsFC modality explained a larger proportion of variance (mean SC:rsFC ratio = 0.62) across the behavioral spectrum in the young-train sample.

Next, we validated these models in the young-test sample. As shown in Figure 3c, most of the $NRMSE_{young}$ values in the combined SC and rsFC modality were well below 0.5. In particular, those of eight behavioral outcomes (TAP_I_Fval, Hamilton_Scale, CERQ_Catastrophizing, CERQ_BlamingOthers, CERQ_positiveRefocusing, LOT_Pessimism.

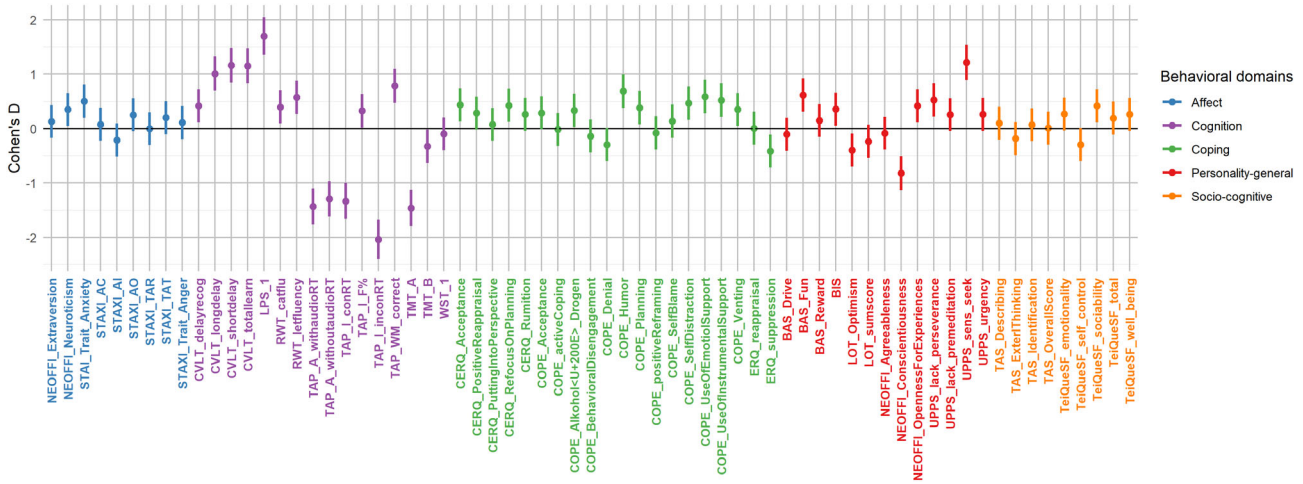

FIGURE 2 Forest plot showing the effect size (i.e., Cohen's *d*) associated with the between age-group differences across all behavioral outcomes. Positive Cohen's *d* values suggest higher scores in the old than in the young age groups. The error bars represent 95% confidence intervals

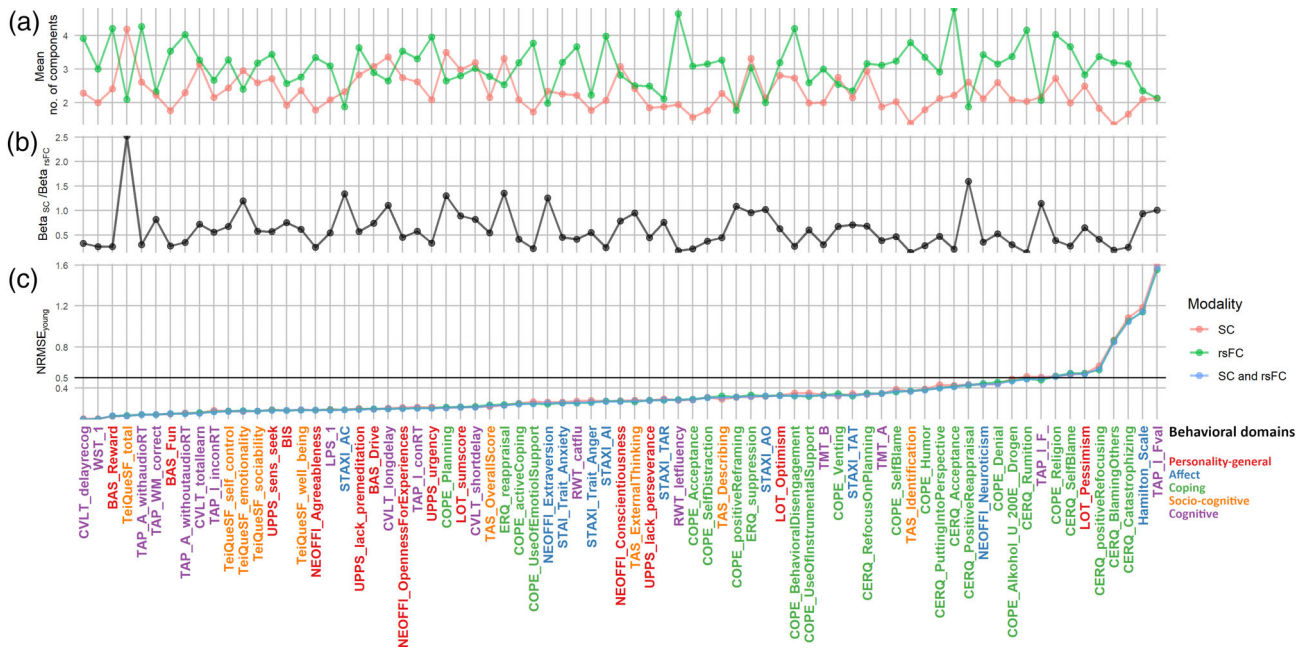

FIGURE 3 (a) Mean number of components selected in 1,000 fivefold cross-validation iterations of the PLSR model. (b) Ratio of SC and functional connectivity (FC) network strength coefficients in the regression model predicting behaviors in the young sample. Higher values correspond to larger SC contributions in behavioral prediction. (c) NRMSE of PLSR models (trained from the young-train group) predictions in the young-test dataset. Behavioral outcomes are arranged in order of increasing NRMSE in the combined SC and resting-state functional connectivity (rsFC) modality. NRMSE, normalized root mean square error; PLSR, partial least squares regression; RMSE, root mean square error; rsFC, resting-state functional connectivity; SC, structural connectivity

CERQ_SelfBlame, and COPE_Religion) exceeded 0.5. These behavioral outcomes are thus excluded in the subsequent reporting of results.

It was observed that the *k* values for the SC models correlated significantly and negatively with $NRMSE_{young}$ ($r = -.29$, $p = .012$), suggesting that sparse SC models produced more accurate predictions. On the other hand, the *k* values for the rsFC models were not significantly correlated with $NRMSE_{young}$ ($r = -.11$, $p = .333$). Finally, one-way ANOVAs suggests that the *k* values in the SC ($F(4,70) = 1.09$, $p = .367$) and rsFC ($F(4,70) = 2.29$, $p = .069$) models were not significantly different across behavioral domains.

3.3 | Comparing predictions across age groups

Using the remaining 67 behavioral outcomes, we compared the performance (i.e., NRMSE) of the young-train PLSR models in predicting behavioral scores in the young-test and old-test samples. Paired samples *t* test suggests that the NRMSE values, collapsed across behavioral outcomes and modalities, were significantly different ($t(100) = 2.96$, $p = .003$, Cohen's $d = .21$; see Figure 4a) in the young-test and old-test samples. As expected, behavioral predictions derived from the young-train PLSR models were more precise in the young-test than in the old-test samples.

3.4 | Comparing predictions and young-to-old age generalizability across modalities and behavioral domains

Then, we compared the NRMSE values (in the young-test and old-test samples) and RMSE ratios, collapsed across behavioral outcomes, between the different modalities. Results of repeated measures ANOVAs suggest that the NRMSE values and RMSE ratios were highly similar across modalities (see Figure 4b) (*p*s ≥ .997).

Next, we compared the NRMSE values and RMSE ratios, collapsed across modalities, between the different behavioral domains. Results of a one-way ANOVA suggested significant differences in all three metrics across the behavioral domains (see Figure 4c). Post hoc Tukey tests revealed significant differences between several pairs of behavioral domains in all three metrics. Generally, the $NRMSE_{young}$ values were significantly higher in the affect and coping domains relative to the other domains (adjusted *p*s < .001), suggesting that the BANs produced relatively less accurate predictions in these measures. Of major relevance to our hypothesis, the RMSE ratios in the cognition domain were significantly lower compared to the other four domains (adjusted *p*s < .001).

As we narrow down further into the individual behavioral outcomes (see Figure 5), we observed that most of the behavioral outcomes in the cognition domains were associated with relatively low RMSE ratios, apart from some interesting exceptions such as the LWT_catflu, LWT_letfluency, TAP_I_conRT, and WST_1. We also observed a handful of behavioral outcomes, especially those measured by the STAXI, which were consistently associated with RMSE ratios >1. This meant that the young-train PLSR models actually performed better in the old-test than in the young-test datasets for these behavioral outcomes. Interestingly, the RMSE ratios in all three conditions were highly and negatively correlated (*r*s ranging from −.68 to −.69) with the magnitude of the Cohen's *d*s obtained earlier (differences in behavioral outcomes across age groups). This suggests that behavioral outcomes with larger age-group differences are associated with less age-general BANs. The chord diagrams representing the BANs of three behavioral measures with the highest and lowest RMSE ratios are shown in Figure 6.

3.5 | Effect of functional network parcellation scheme

While we had intended to use different parcellation schemes across modalities to optimize predictions within their respective modalities,

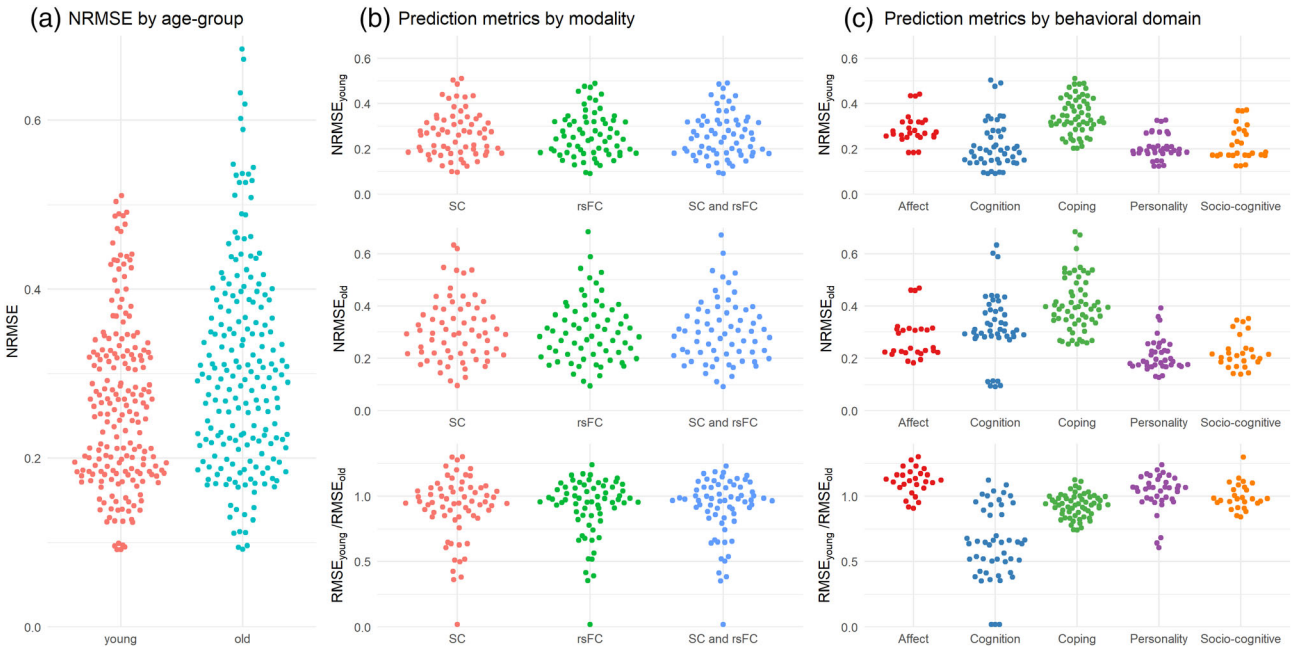

FIGURE 4 Beehive plots illustrating the comparison of (a) normalized root mean square error (NRMSE) values across age groups, and NRMSE and root mean square error (RMSE) ratios across (b) modality and (c) behavioral domain. Note that $NRMSE_{old}$ values of TMT_B are not shown in the above plots as they are extreme outliers (structural connectivity [SC] = 4.43; functional connectivity [FC] = 4.42; combined SC and resting-state FC [rsFC] = 4.42)

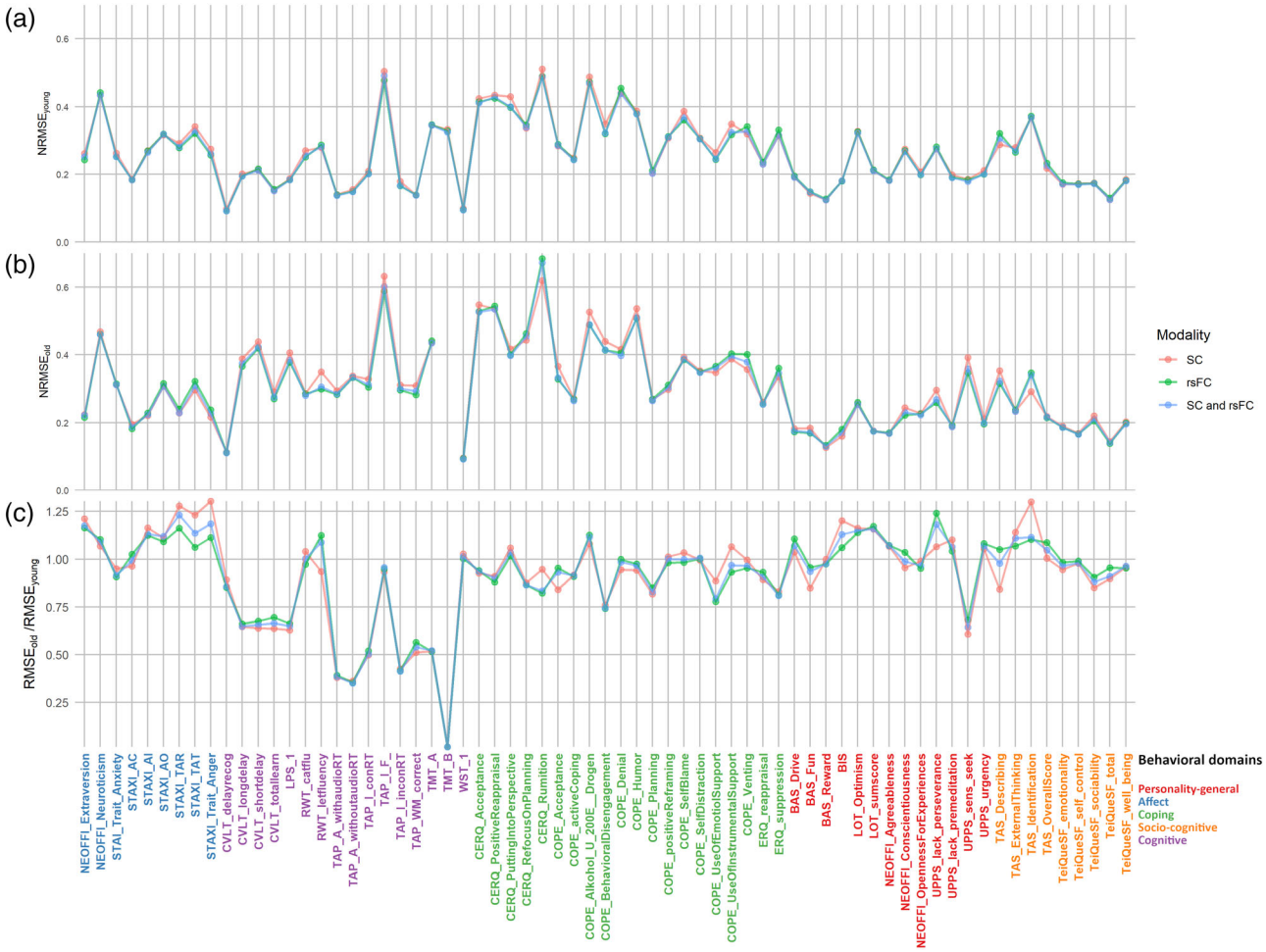

FIGURE 5 NRMSE of PLSR models (trained from the young-train group) predictions in the (a) young-test dataset and (b) old-test dataset. (c) Ratio of RMSE of predictions in the old and young. Larger values correspond to more similar predictions between the young and old groups. NRMSE, normalized root mean square error; PLSR, partial least squares regression; RMSE, root mean square error; rsFC, resting-state functional connectivity; SC, structural connectivity. Note that $NRMSE_{old}$ values of TMT_B are not shown in the above plots as they are extreme outliers (SC = 4.43; functional connectivity [FC] = 4.42; combined SC and rsFC = 4.42)

previous research has shown that the use of different parcellation schemes may affect certain characteristics of brain networks (Wang et al., 2009; Zalesky et al., 2010). Consequently, the use of different parcellation schemes in the SC and rsFC networks may preclude a fair comparison across modalities and may have a confounding influence on the generalizability of behavioral predictions from the young-train model to the old-test sample. Thus, we repeated the above analyses using rsFC connectomes constructed from the AAL-90 atlas.

Paired *t* tests indicated significant differences in the NRMSE values between both parcellation schemes in the young-test ($t(74) = 5.81$, $p < .001$, Cohen's $d = .67$) and old-test ($t(74) = 4.43$, $p < .001$, Cohen's $d = .59$) samples. As expected, rsFC networks constructed from the more granular brainnetome parcellation scheme produced more accurate predictions. However, RMSE ratios were not significantly different between parcellation schemes ($t(74) = 0.77$, $p = .446$, Cohen's $d = .03$). Finally, paired *t* tests comparing the $NRMSE_{young}$ ($t(74) = 1.50$, $p = .139$, Cohen's $d = .17$), $NRMSE_{old}$ ($t(74) = 0.83$, $p = .411$, Cohen's $d = .10$), and RMSE ratios ($t(74) = .03$, $p = .979$, Cohen's $d = .003$) derived from the rsFC and SC networks constructed from the same AAL-90 atlas, did not reveal any significant differences. While these results suggest that different parcellation schemes can systematically influence the accuracy of behavioral predictions, importantly, they do not confound the generalizability of behavioral predictions from the young to the old.

4 | DISCUSSION

In the present study, we constructed SC and rsFC connectome models from young adults and assessed their predictions in unseen young and old subjects, to examine the young-to-old generalizability of BANs. As expected, we showed that these BANs predicted behavioral scores more accurately in unseen subjects of the same young age group, suggesting a general age-specificity effect across the BANs. Next, this age-specificity effect was observed to be greatest among BANs in the cognition domains. That is, BANs in the cognition domains, relative to

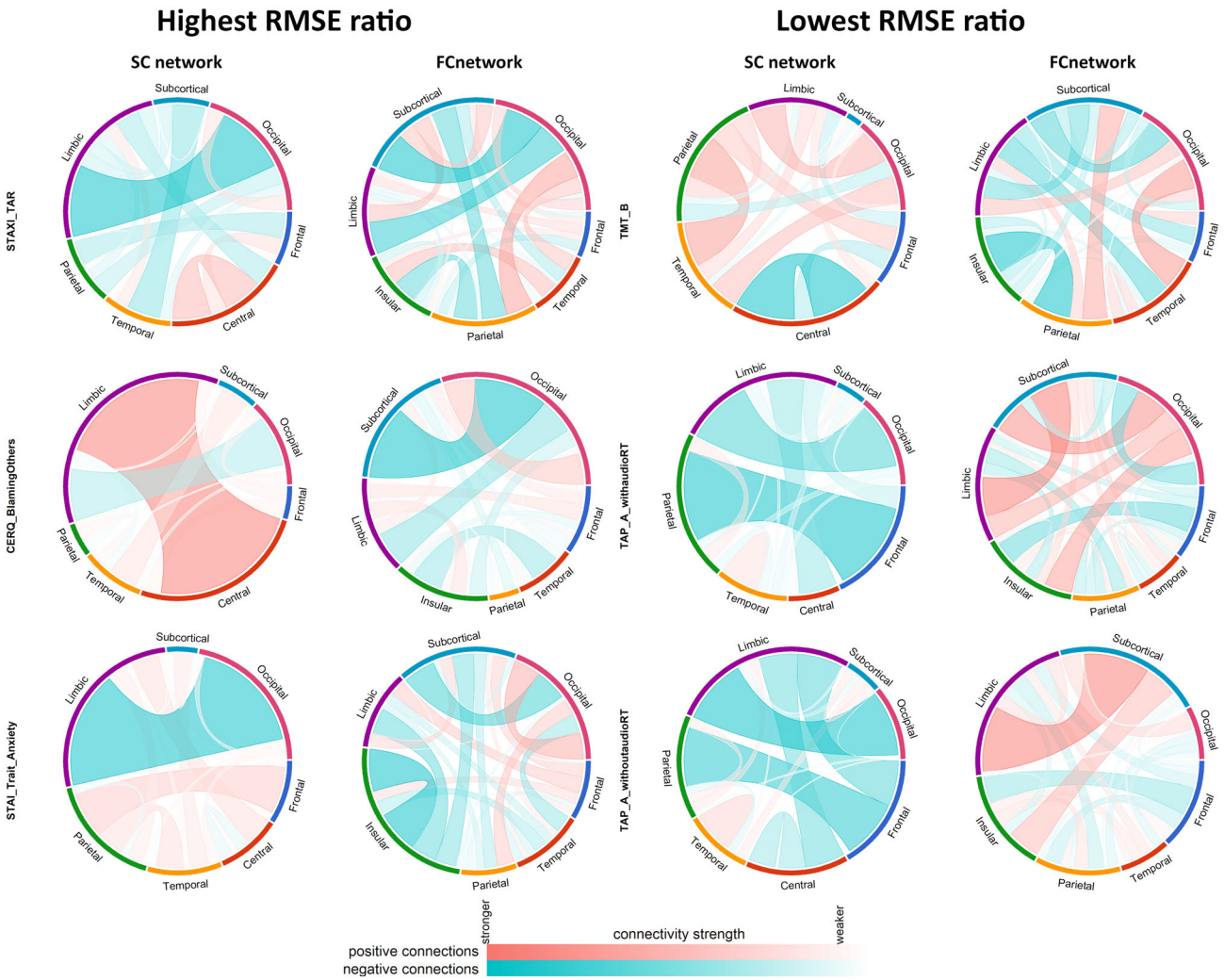

FIGURE 6 Chord diagrams showing the structural connectivity (SC) and functional connectivity (FC) averaged connectivity values between and within different regions, associated with the three behavioral measures with the highest and lowest $RMSE_{old}/RMSE_{young}$ ratios in the combined SC and resting-state functional connectivity (rsFC) model. The list of nodes corresponding to various regions in the SC and FC networks can be referred to in table 3 of Rolls et al. (2015) and at https://atlas.brainnetome.org/bnatlas.html, respectively

the other domains, produced disproportionately less accurate behavioral predictions in unseen old subjects than their younger counterparts. Hence, as we have hypothesized, BANs from the cognitive domain tend to fall toward the age-specific end of the continuum. Although those of personality traits fell toward the age-general end of the continuum as hypothesized, most other behavioral domains were similarly positioned at that end. Finally, we observed that the accuracy and young-to-old generalizability of behavioral predictions to be largely similar across imaging modalities.

Across most cognitive outcomes, the patterns of brain connectivity, especially in terms of SC, were observed to largely correspond to the cognitive control network (Niendam et al., 2012), which largely comprises of nodes from the frontoparietal and limbic regions (i.e., anterior cingulate cortex). As we have hypothesized, these BANs are highly age-specific alluding to significant age-related differences in the connectomes associated with cognition. These differences may be attributed to the neural compensatory and dedifferentiation processes that occur in the context of aging. In relation to the former, functional neuroimaging studies have typically observed an age-related increase in the activation of auxiliary brain networks during cognitive tasks. Such activation was explained to compensate for the age-associated degradation in brain structures and networks (Cabeza et al., 2018). In age-related neural dedifferentiation, task-based fMRI studies have generally observed an age-related decrease in the functional specificity of brain regions and networks, which have been attributed to the age-related decrease in segregation or increased interactions between large scale brain networks (Costanzo et al., 2015; Koen, Srokova, & Rugg, 2020). Both age-related processes are likely to contribute to the rewiring of connections associated with cognition, which ultimately resulted in differences in the connectivity correlates of cognition, across the lifespan. Generally, these age-related processes represent a shift toward less efficient recruitment of neural activity (Reuter-Lorenz & Park, 2014), and have been associated with worse cognitive abilities in the aging context (Meunier, Stamatakis, &

Tyler, 2014). This would also explain the significant relationship between age-group differences in behavioral scores and the young-to-old age-generalizability of their BANs.

Another possible explanation for the age-specific characteristics of these BANs may relate largely to the significant age-group differences in cognitive scores. Given these differences, some of the old-test subjects had cognitive scores which were beyond the range of scores observed in the young-train sample. As a result, the PLSR models had to extrapolate beyond the range of scores in the young-train sample and such extrapolations are more prone to prediction errors. Generally, this highlights the limitations of linear regression-based prediction approaches.

Although cognitive functions were generally underpinned by age-specific BANs, there were some interesting exceptions; we observed that BANs belonging to a subgroup of cognitive tests (i.e., RWT_catflu, RWT_letfluency, and WST_1) that assessed semantic knowledge demonstrated much higher young-to-old age-generalizability. Such knowledge has been shown not only to undergo less age-related decline relative to other aspects of cognition but also to continue to accumulate even in the later lifespan (Kavé & Halamish, 2015; Park et al., 2002). We speculate the age-general nature of BANs in this domain of cognition is intimately linked to the resilience of semantic knowledge in the context of age-related cognitive decline. We theorized that age-general BANs undergo relatively less rewiring of connections across the lifespan, possibly due to their resilient network structure. In particular, semantic knowledge is associated with a brain network that is widely distributed across the frontal, temporal and parietal lobes (Xu, Lin, Han, He, & Bi, 2016). This is fairly consistent with the patterns of rsFC observed to be associated with the three semantic knowledge tests in the current study. For instance, the temporal and parietal regions were observed to be relatively central in the network (see supplementary Figure 1). This semantic brain network has been shown to be resilient to some types of damage and perturbations (Rice, Caswell, Moore, Lambon Ralph, & Hoffman, 2018), in part due to its flexibility in varying regional activation and FC across domain-specific and -general subsystems (Jung, Rice, & Lambon Ralph, 2021).

In addition to the cognitive measures, we could observe that there were a few behavioral traits that were poorly predicted in the old-test sample, especially the ones relating to emotion regulation and coping strategies. This is likely related to the challenge in measuring these behavioral strategies in the elderly in a laboratory setting, leading to a larger degree of measurement error in the old compared to the young age groups. Older adults tend to exhibit different types of emotion regulation strategies than their younger counterparts, with the former being more prone to using suppression and avoidance of emotional situations and less inclined on engaging in reappraisal, rumination, and active coping when compared to young adults (Barrett Feldman, Lewis, & Haviland-Jones, 2016). Furthermore, the assessment of these behaviors requires participants to recall their habitual modes of dealing with certain scenarios rather than asking them to rate their procedural skills, which can be particularly challenging to the elderly and may lead to larger measurement errors in this age group.

Interestingly, we also observed a handful of BANs, such as those relating to anger, personality and socio-cognitive-related traits, which produced more accurate predictions in the old-test than in the young-test subjects. The counterintuitive predictive performance across age groups may seem somewhat puzzling. Nevertheless, this could happen when two conditions are satisfied. In the first, these behavioral predictions can be predicted with similar accuracy in the young and old, using the same set of connectomes. For behavioral phenotypes that showed minimal between-age-group differences (e.g., anger traits), this could point to the resilient network structure of their BANS, hence requiring little or no age-related organization or compensation. In particular, the nodes of a brain network commonly associated with negative affect have been shown to be relatively age-resilient both in terms of function and structure (Mather, 2016). As for behavioral traits that showed large between-age-group differences (e.g., NEOFFI_Conscientiousness), it is likely that their associated connectomes are highly specific to these traits, such that there is very little flexibility or capacity for it to undergo age-related reorganization and compensation to stabilize the age-related behavioral changes. The second condition relates to the relatively larger measurement error associated with these behavioral measurements in the young than in the old sample, thus leading to more accurate predictions in the latter. One plausible source of measurement error can be traced to the instability of certain dispositional traits. For instance, one study showed that young adults exhibited significantly greater variability in extraversion, agreeableness, and emotional stability across a 1 to 2-week period as compared to older adults (Noftle & Fleeson, 2010). In line with these previous findings, BANs associated with these three traits tend to produce worse predictions in the young than in the old.

Contrary to our hypothesis, the young-to-old age-generalizability of SC and rsFC predictions is largely comparable. Similarly, neither the SC nor rsFC modality is consistently superior to the other in behavioral predictions in both age groups and the differences in the overall accuracy of these SC and rsFC-based predictions are fairly minute. Furthermore, the combination of both modalities into a single regression-based prediction model did not necessarily produce better predictions than the SC-only or rsFC-only models. Taken together, this suggests that the SC and rsFC models shared a large overlap in predictive information. It should be noted that this overlap in predictive information does not necessarily mean that SC and rsFC patterns are tightly coupled. As shown in Figure 6, the profile of structural and functional connections shared very little resemblance. This is not surprising, since previous research has shown that FC between a pair of regions can thrive even with few or no direct structural connections, as long as indirect structural connections exist between the regions (Damoiseaux & Greicius, 2009). This minimal dependence of direct structural connections opens up many possible functional connections between pairs of regions, ultimately culminating in more densely connected BANs in the rsFC modality. This could possibly explain the significantly sparser PLSR solutions obtained (i.e., smaller k values) in the SC than in the rsFC modality.

Generally, we observed that the BANs in the affect and coping domains produced relatively less accurate predictions within the same young age group than in the other behavioral domains. This could be

due to the heterogeneous nature of these measured constructs. For instance, depression is one such heterogeneous construct. Two persons with the same score on the Hamilton depression scale, can present with very different symptom profiles. These different profiles would consequently be linked to different sets of connectomes (Drysdale et al., 2017). Therefore, a single connectome model alone, would not adequately capture the heterogeneous symptom profiles of depression. It is thus not surprising that the Hamilton depression score is among one of the least accurately predicted measures revealed in our study.

Presently, there has been a great deal of attention in using rsFC features in edge-based behavioral prediction models, and far less research has explored the use of SC-based features in these prediction models. Despite the highly comparable performance of SC and rsFC behavioral prediction models shown here, there are certain advantages in using SC- over rsFC-based features in behavioral prediction models. First, SC, which is typically derived from diffusion-weighted images, is less susceptible to head motion artifacts. Had this study focused solely on the SC modality, we would not have excluded the 15 subjects with excessive head motion during their rsfMRI scans. Second, edge-based prediction models utilizing SC features tend to be computationally "cheaper" due to the use of low-granularity atlases (e.g., AAL) to map out the structural connectome, and the relatively sparse connectivity matrices (i.e., containing mostly zeros)—arising from the fact that it is anatomically impossible for WM fibers to connect between all possible pairs of brain nodes.

Our findings on age-specific networks would caution extrapolating brain-cognitive functions associations from one age group to another, as we have shown these cognitive functions are usually underpinned by age-specific BANs. Furthermore, research on these neurocognitive correlates tends to be predominantly carried out in the older adult population, possibly due to its relevance to dementia, resulting in a relative paucity of similar studies on younger adults. To this end, it may be tempting, but ill-advised, to generalize findings from the old to young. Doing so will result in highly inaccurate conclusions. Beyond brain-based behavioral prediction research, these implications also weigh heavily on intervention studies utilizing brain stimulation techniques to target a particular brain region or network to enhance cognitive functions in a specific age group.

The current study is subjected to some limitations. First, we had a relatively small sample of older participants. This meant it would not be appropriate for us to derive BANs from the older sample and assess the generalizability of these BANs in the young sample—which would have been a very meaningful addition to the study. Second, we excluded disproportionately more subjects in the old age group due to excessive head motion. While it is not unusual that older adults generally produce greater head motion during MRI scans (Saccà et al., 2021), the larger number of exclusions would contribute to a selection bias in the old age group. Third, we explored only two extreme age groups with relatively narrow age ranges; there is a large intermediate age range (i.e., 36–58) not accounted for in this study. Although it is convenient and tempting to assume that the age generalizability or specificity of the BANs, would vary across this intermediate age range in a graded manner, it remains to be verified by future research. Fourth, we had an unbalanced male to female ratio across the datasets (i.e., 29% female subjects in the young-age group against 50% female participants in the old-age group); nevertheless, a recent study using a classification machine-learning approach based on rsfMRI data could not find enough evidence that would support a clear sexual differentiation of the human brain (Weis et al., 2020). Furthermore, the relatively small sample size, especially for the older participants sample, meant that splitting our sample by gender would not be feasible. Nevertheless, we encourage future studies to compare gender effects in the BANs model performance across different age ranges. Finally, this study is essentially a cross-sectional comparison between the young and old. Such cross-sectional comparisons may not accurately reflect the longitudinal aging processes and would be subjected to cohort-related confounds such as socioeconomic conditions, educational attainment, and nutrition.

CONFLICT OF INTEREST

Both the authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The raw behavioral and neuroimaging data are publicly available (Babayan et al., 2019). The preprocessed connectivity matrices and selection of behavioral data analyzed in this study are available at https://osf.io/u6zck/.

ETHICS STATEMENT

The study protocol conformed to the Declaration of Helsinki and was approved by the ethics committee at the medical faculty of the University of Leipzig (reference number 154/13-ff).

ORCID

Junhong Yu https://orcid.org/0000-0002-2563-9658

REFERENCES

Allen, E. A., Erhardt, E. B., Damaraju, E., Gruner, W., Segall, J. M., Silva, R. F., … Calhoun, V. D. (2011). A baseline for the multivariate comparison of resting-state networks. *Frontiers in Systems Neuroscience*, *5*, 1–23. https://doi.org/10.3389/fnsys.2011.00002

Babayan, A., Erbey, M., Kumral, D., Reinelt, J. D., Reiter, A. M. F., Röbbig, J., … Villringer, A. (2019). Data descriptor: A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Scientific Data*, *6*. https://doi.org/10.1038/sdata.2018.308

Barrett Feldman, L., Lewis, M., & Haviland-Jones, M. J. (2016). *Handbook of emotions* (4th ed.). New York, NY: The Guilford Press.

Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X. N., & Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage*, *102*(P2), 345–357. https://doi.org/10.1016/j.neuroimage.2014.07.067

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., … Rajah, M. N. (2018). Maintenance, reserve and compensation: The cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, *19*(11), 701–710. https://doi.org/10.1038/s41583-018-0068-2

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238–251. https://doi.org/10.1016/j.neuron.2014.05.014

Costanzo, E. Y., Villarreal, M., Drucaroff, L. J., Ortiz-Villafañe, M., Castro, M. N., Goldschmidt, M., … Guinjoan, S. M. (2015). Hemispheric specialization in affective responses, cerebral dominance for language, and handedness. *Behavioural Brain Research*, *288*, 11–19. https://doi.org/10.1016/j.bbr.2015.04.006

Damoiseaux, J. S., & Greicius, M. D. (2009). Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, *213*(6), 525–533. https://doi.org/10.1007/s00429-009-0208-6

Damoiseaux, J. S. (2017). Effects of aging on functional and structural brain connectivity. *NeuroImage*, *160*, 32–40. https://doi.org/10.1016/j.neuroimage.2017.01.077

Dhollander, T., & Connelly, A. (2016). A novel iterative approach to reap the benefits of multi-tissue CSD from just single-shell (+b=0) diffusion MRI data. *Proceedings of the 24th Aanual meeting of the International Society of Magnetic Resonance in Medicine*, Singapore. Vol. 24, pp. 3010.

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28–38. https://doi.org/10.1038/nm.4246

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., … Laird, A. R. (2016). The human Brainnetome atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex*, *26*(8), 3508–3526.

Fountain-Zaragoza, S., Samimy, S., Rosenberg, M. D., & Prakash, R. S. (2019). Connectome-based models predict attentional control in aging adults. *NeuroImage*, *186*, 1–13. https://doi.org/10.1016/j.neuroimage.2018.10.074

Gao, M., Wong, C. H. Y., Huang, H., Shao, R., Huang, R., Chan, C. C. H., & Lee, T. M. C. (2020). Connectome-based models can predict processing speed in older adults. *NeuroImage*, *223*, 117290. https://doi.org/10.1016/j.neuroimage.2020.117290

Gao, S., Greene, A. S., Constable, R. T., & Scheinost, D. (2019). Combining multiple connectomes improves predictive modeling of phenotypic measures. *NeuroImage*, *201*, 116038. https://doi.org/10.1016/j.neuroimage.2019.116038

Hyon, R., Youm, Y., Kim, J., Chey, J., Kwak, S., & Parkinson, C. (2020). Similarity in functional brain connectivity at rest predicts interpersonal closeness in the social network of an entire village. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(52), 33149–33160. https://doi.org/10.1073/pnas.2013606117

Itahashi, T., Kosibaty, N., Hashimoto, R.-I., & Aoki, Y. Y. (2021). Prediction of life satisfaction from resting-state functional connectome. *Brain and Behavior*, *11*, e2331. https://doi.org/10.1002/brb3.2331

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jiang, R., Calhoun, V. D., Zuo, N., Lin, D., Li, J., Fan, L., … Sui, J. (2018). Connectome-based individualized prediction of temperament trait scores. *NeuroImage*, *183*, 366–374. https://doi.org/10.1016/j.neuroimage.2018.08.038

Jung, J. Y., Rice, G. E., & Lambon Ralph, M. A. (2021). The neural bases of resilient semantic system: Evidence of variable neuro-displacement in cognitive systems. *Brain Structure and Function*, *226*(5), 1585–1599. https://doi.org/10.1007/s00429-021-02272-1

Kavé, G., & Halamish, V. (2015). Doubly blessed: Older adults know more vocabulary and know better what they know. *Psychology and Aging*, *30*(1), 68.

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, *76*(5), 1574–1581. https://doi.org/10.1002/mrm.26054

Kelly, C., & Castellanos, F. X. (2014). Strengthening connections: Functional connectivity and brain plasticity. *Neuropsychology Review*, *24*(1), 63–76. https://doi.org/10.1007/s11065-014-9252-y

Knight, M., & Knight, M. (2006). The Affective Neuroscience of Aging and Its Implications for Cognition. In T. Canli (Ed.), *Biology of personality and individual differences* (pp. 159–183). New York, NY: The Guilford Press.

Koen, J. D., Srokova, S., & Rugg, M. D. (2020). Age-related neural dedifferentiation and cognition. *Current Opinion in Behavioral Sciences*, *32*, 7–14. https://doi.org/10.1016/j.cobeha.2020.01.006

Li, J., Kong, R., Liégeois, R., Orban, C., Tan, Y., Sun, N., … Yeo, B. T. T. (2019). Global signal regression strengthens association between resting-state functional connectivity and behavior. *NeuroImage*, *196*, 126–141. https://doi.org/10.1016/j.neuroimage.2019.04.016

Mather, M. (2016). The affective neuroscience of aging. *Annual Review of Psychology*. https://doi.org/10.1146/annurev-psych-122414-033540

Meunier, D., Stamatakis, E. A., & Tyler, L. K. (2014). Age-related functional reorganization, structural changes, and preserved cognition. *Neurobiology of Aging*, *35*(1), 42–54. https://doi.org/10.1016/j.neurobiolaging.2013.07.003

Ng, K. K., Lo, J. C., Lim, J. K. W., Chee, M. W. L., & Zhou, J. (2016). Reduced functional segregation between the default mode network and the executive control network in healthy older adults: A longitudinal study. *NeuroImage*, *133*, 321–330. https://doi.org/10.1016/j.neuroimage.2016.03.029

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Noftle, E. E., & Fleeson, W. (2010). Age differences in big five behavior averages and variabilities across the adult life span: Moving beyond retrospective, global summary accounts of personality. *Psychology and Aging*, *25*(1), 95.

Onoda, K., Ishihara, M., & Yamaguchi, S. (2012). Decreased functional connectivity by aging Is associated with cognitive decline. *Journal of Cognitive Neuroscience*, *24*(11), 2186–2198. https://doi.org/10.1162/jocn_a_00269

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*(1), 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Reuter-Lorenz, P. A., & Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review*, *24*(3), 355–370. https://doi.org/10.1007/s11065-014-9270-9

Rice, G. E., Caswell, H., Moore, P., Lambon Ralph, M. A., & Hoffman, P. (2018). Revealing the dynamic modulations that underpin a resilient neural network for semantic cognition: An fMRI investigation in patients with anterior temporal lobe resection. *Cerebral Cortex*, *28*(8), 3004–3016. https://doi.org/10.1093/cercor/bhy116

Rolls, E. T., Joliot, M., & Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *Neuroimage*, *122*, 1–5.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Saccà, V., Sarica, A., Quattrone, A., Rocca, F., Quattrone, A., & Novellino, F. (2021). Aging effect on head motion: A machine learning study on resting state fMRI data. *Journal of Neuroscience Methods*, *352*, 109084. https://doi.org/10.1016/j.jneumeth.2021.109084

Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols*, *12*(3), 506–518. https://doi.org/10.1038/nprot.2016.178

Smith, R. E., Tournier, J.-D., Calamante, F., & Connelly, A. (2012). Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*, *62*(3), 1924–1938. https://doi.org/10.1016/j.neuroimage.2012.06.005

Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., … Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*(January), 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

van den Heuvel, M. P., & Hulshoff, Pol H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, *20*(8), 519–534. https://doi.org/10.1016/j.euroneuro.2010.03.008

Veraart, J., Novikov, D. S., Christiaens, D., Ades-aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *NeuroImage*, *142*, 394–406. https://doi.org/10.1016/j.neuroimage.2016.08.016

Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., … He, Y. (2009). Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. *Human Brain Mapping*, *30*(5), 1511–1523. https://doi.org/10.1002/hbm.20623

Wang, Z., Goerlich, K. S., Ai, H., Aleman, A., Luo, Y., & Xu, P. (2021). Connectome-based predictive modeling of individual anxiety. *Cerebral Cortex*, *31*(6), 3006–3020. https://doi.org/10.1093/cercor/bhaa407

Weis, S., Patil, K. R., Hoffstaedter, F., Nostro, A., Yeo, B. T. T., & Eickhoff, S. B. (2020). Sex classification by resting state brain connectivity. *Cerebral Cortex*, *30*(2), 824–835. https://doi.org/10.1093/cercor/bhz129

Xu, Y., Lin, Q., Han, Z., He, Y., & Bi, Y. (2016). Intrinsic functional network architecture of human semantic processing: Modules and hubs. *NeuroImage*, *132*, 542–555. https://doi.org/10.1016/j.neuroimage.2016.03.004

Yan, C.-G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R. C., Di Martino, A., … Milham, M. P. (2013). A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage*, *76*, 183–201. https://doi.org/10.1016/j.neuroimage.2013.03.004

Yan, C.-G., Wang, X.-D., Zuo, X.-N., & Zang, Y.-F. (2016). DPABI: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, *14*(3), 339–351. https://doi.org/10.1007/s12021-016-9299-4

Yoo, K., Rosenberg, M. D., Hsu, W.-T., Zhang, S., Li, C.-S. R., Scheinost, D., … Chun, M. M. (2018). Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets. *NeuroImage*, *167*, 11–22. https://doi.org/10.1016/j.neuroimage.2017.11.010

Yu, J., Rawtaer, I., Fam, J., Feng, L., Kua, E.-H., & Mahendran, R. (2020). The individualized prediction of cognitive test scores in mild cognitive impairment using structural and functional connectivity features. *NeuroImage*, *223*, 117310. https://doi.org/10.1016/j.neuroimage.2020.117310

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., & Bullmore, E. T. (2010). Whole-brain anatomical networks: Does the choice of nodes matter? *NeuroImage*, *50*(3), 970–983. https://doi.org/10.1016/j.neuroimage.2009.12.027

Zonneveld, H. I., Pruim, R. H. R., Bos, D., Vrooman, H. A., Muetzel, R. L., Hofman, A., & Vernooij, M. W. (2019). Patterns of functional connectivity in an aging population: The Rotterdam Study. *NeuroImage*, *189*, 432–444. https://doi.org/10.1016/j.neuroimage.2019.01.041

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Yu, J., & Fischer, N. L. (2022). Age-specificity and generalization of behavior-associated structural and functional networks and their relevance to behavioral domains. *Human Brain Mapping*, *43*(8), 2405–2418. https://doi.org/10.1002/hbm.25759