

# Functional topography of the neocortex predicts covariation in complex cognitive and basic motor abilities

Ethan T. Whitman<sup>1,\*</sup>, Annchen R. Knodt<sup>1</sup>, Maxwell L. Elliott<sup>2</sup>, Wickliffe C. Abraham<sup>3</sup>, Kirsten Cheyne<sup>4</sup>, Sean Hogan<sup>4</sup>, David Ireland<sup>4</sup>, Ross Keenan<sup>5,6</sup>, Joan H. Leung<sup>7,8</sup>, Tracy R. Melzer<sup>5,9</sup>, Richie Poulton<sup>4</sup>, Suzanne C. Purdy<sup>5,7,8</sup>, Sandhya Ramrakha<sup>4</sup>, Peter R. Thorne<sup>5,8,10</sup>, Avshalom Caspi<sup>1,11,12,13,14</sup>, Terrie E. Moffitt<sup>1,11,12,13,14</sup>, Ahmad R. Hariri<sup>1</sup>

<sup>1</sup>Department of Psychology and Neuroscience, Duke University, Durham, NC 27710, USA,

<sup>2</sup>Department of Psychology, Center for Brain Science, Harvard University, Cambridge, MA 02138, USA,

<sup>3</sup>Department of Psychology, University of Otago, Dunedin 9016, New Zealand,

<sup>4</sup>Dunedin Multidisciplinary Health and Development Research Unit, Department of Psychology, University of Otago, Dunedin 9016, New Zealand,

<sup>5</sup>Brain Research New Zealand-Rangahau Roro Aotearoa, Centre of Research Excellence, Universities of Auckland and Otago, Auckland 1010, New Zealand,

<sup>6</sup>Christchurch Radiology Group, Christchurch 8014, New Zealand,

<sup>7</sup>School of Psychology, University of Auckland, Auckland 1142, New Zealand,

<sup>8</sup>Eisdell Moore Centre, University of Auckland, Auckland 1142, New Zealand,

<sup>9</sup>Department of Medicine, University of Otago, Christchurch 9016, New Zealand,

<sup>10</sup>School of Population Health, University of Auckland, Auckland 1142, New Zealand,

<sup>11</sup>Center for Genomic and Computational Biology, Duke University, Durham, NC 27710, USA,

<sup>12</sup>King's College London, Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, & Neuroscience, London SE5 8AF, UK,

<sup>13</sup>PROMENTA, Department of Psychology, University of Oslo, NO-0316 Oslo, Norway,

<sup>14</sup>Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC 27710, USA

\*Corresponding author: Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA. Tel: 978-810-0575;

E-mail: ethan.whitman@duke.edu

Although higher-order cognitive and lower-order sensorimotor abilities are generally regarded as distinct and studied separately, there is evidence that they not only covary but also that this covariation increases across the lifespan. This pattern has been leveraged in clinical settings where a simple assessment of sensory or motor ability (e.g. hearing, gait speed) can forecast age-related cognitive decline and risk for dementia. However, the brain mechanisms underlying cognitive, sensory, and motor covariation are largely unknown. Here, we examined whether such covariation in midlife reflects variability in common versus distinct neocortical networks using individualized maps of functional topography derived from BOLD fMRI data collected in 769 45-year-old members of a population-representative cohort. Analyses revealed that variability in basic motor but not hearing ability reflected individual differences in the functional topography of neocortical networks typically supporting cognitive ability. These patterns suggest that covariation in motor and cognitive abilities in midlife reflects convergence of function in higher-order neocortical networks and that gait speed may not be simply a measure of physical function but rather an integrative index of nervous system health.

**Key words:** cognition; fMRI; functional topography; gait speed; hearing.

## Introduction

Higher-order cognitive and lower-order sensorimotor abilities have been generally regarded as distinct and studied separately. However, there is growing evidence that they not only covary but also that this covariation increases across the lifespan. For example, there is evidence for covariation between sensorimotor and cognitive abilities during childhood (Welch and Dawes 2007; van der Fels et al. 2015), midlife (Rasmussen et al. 2019), and old age (van der Willik et al. 2021). In addition, this covariation tends to become more pronounced during aging (Baltes and Lindenberger 1997). Moreover, measures of sensory and motor abilities such as hearing tests and gait assessments are commonly used to predict cognitive decline and dementia risk in older adults (Studenski et al. 2011; Collyer et al. 2022; Kwok et al. 2022). Identifying brain mechanisms linking these seemingly disparate abilities can

further our understanding of not only the extent to which variability in fundamental aspects of human behavior reflect distinct and common features of brain function across the lifespan but also how they may be best leveraged in clinical applications.

Gait speed is a simple and widely used measure of biological aging measured by how quickly a person walks across a biosensor-equipped pad (Fritz and Lusardi 2009). As early as midlife, individual differences in gait speed covary with complex psychological processes including memory, executive functions, and cognitive ability (Rasmussen et al. 2019; Sunderaraman et al. 2019). Individuals suffering from mild cognitive impairment also show reduced gait stability compared with cognitively normal people (Montero-Odasso et al. 2012). However, the extent to which these processes are associated with common brain systems is unclear. There is evidence in older adults that gait speed is

associated with the strength of functional connectivity in higher-order brain networks supporting cognitive abilities (Yuan et al. 2015; Lo et al. 2017; Drobny et al. 2022). These same higher-order networks are thought to be especially sensitive to age (Fjell et al. 2015; Staffaroni et al. 2018; Varangis et al. 2019), bolstering the hypothesis that some effects of physical and cognitive aging reflect shared features of common brain systems.

Likewise, hearing ability as measured by pure tone audiometry is associated with cognitive ability during midlife and aging (Lindenberger and Baltes 1994; Baltes and Lindenberger 1997; Loughrey et al. 2018; Okely et al. 2021) and is predictive of future cognitive decline (Lin et al. 2011). However, there is some added complexity in assessing potential links between hearing and cognitive ability.

For example, pure tone audiometry is not a consistent indicator of functional hearing ability and some individuals who have no peripheral damage to the ear (i.e. peripheral hearing loss) may still report reduced hearing ability (Sardone et al. 2019), which could be more indicative of underlying brain health. Conversely, peripheral synaptic pathologies have been associated with normal pure tone thresholds (Kohrman et al. 2020). For this reason, hearing researchers often measure declining ability to differentiate speech in noisy environments, which is thought to be more affected by brain processing and is also associated with cognitive decline (Humes et al. 2013; Sardone et al. 2019; Humes 2020; Stevenson et al. 2022).

This ability to recognize speech in the presence of background or competing noise is measured as a speech reception threshold in noise (henceforth “SRT-hearing”). Crucially, measured SRT-hearing may show greater ecological validity than pure tone audiometry, as it captures the ability to hear and discern complex auditory information (Stevenson et al. 2022). As with gait speed, prior studies have associated SRT-hearing with functional connectivity in higher-order brain networks most associated with complex cognitive ability (Fitzhugh et al. 2021). These prior findings with gait speed and SRT-hearing suggest that physical and cognitive aging are reflected in shared features of common brain networks. However, they are limited by small sample sizes and only focus on a small number of networks. Thus, it remains unclear why basic sensorimotor functioning is so closely linked with cognition functioning especially in midlife and continuing into later life.

Intrinsic functional connectivity is a powerful tool to describe individual differences in network-level brain organization and their relationship to behavior (Shen et al. 2017; Chen et al. 2022). Traditional studies of functional connectivity use group-averaged atlases to assign anatomical brain regions to different functional networks. However, this assumes that the spatial layout of functional networks is identical from person to person. Recent technical advances have demonstrated that there is substantial variation between people in the spatial layout of functional neocortical networks (Laumann et al. 2015; Wang et al. 2015). This spatial variation is called *functional topography*. Traditional studies using functional connectivity ignore variation in functional topography, and this likely increases error in the calculation of functional connectivity. Importantly, emerging studies have demonstrated that functional topography reliably maps onto individual differences in behavior as well as onto developmental changes in early life (Kong et al. 2019; Cui et al. 2020, 2022; Keller et al. 2022). Thus, functional topography represents a novel measure that can reduce individual-level error in the estimation of neocortical network architecture and, subsequently, further capture covariation between brain and behavior.

Here, we leveraged measures of functional topography in a large population-representative birth cohort, the Dunedin Study, to examine the extent to which overlap in cognitive, motor, and sensory abilities reflect shared variability in neocortical network architecture. Specifically, we used Multi-Session Hierarchical Bayesian Modeling (Kong et al. 2019; MS-HBM) to generate reliable individualized estimates of functional topography at age 45. We then mapped functional topography onto variation in cognitive ability as measured by IQ, motor ability as measured by gait speed, and sensory ability as measured by SRT-hearing. Next, we tested our ability to predict variation in these three behaviors in Study members by training models on features of functional topography. Based on prior work, we hypothesized that IQ, gait speed, and SRT-hearing would be correlated within each Study member. We further hypothesized that higher IQ, faster gait speed, and better SRT-hearing would all be associated with relatively larger higher-order functional neocortical networks. Lastly, we hypothesized that variability in IQ, gait speed, and SRT-hearing would be predicted by overlapping patterns of individual differences in the functional topography of these networks.

## Materials and methods

### Study design and participants

Participants were members of the Dunedin Study, a population-representative birth cohort ( $n = 1,037$ ; 91% of eligible births; 52% male) born between April 1972 and March 1973 in Dunedin, New Zealand (NZ), and eligible based on residence in the province and participation in the first assessment at age 3 years (Poulton et al. 2015, 2022). The cohort represented the full range of socioeconomic status in the general population of NZ's South Island and, as adults, matches the NZ National Health and Nutrition Survey on key adult health indicators (e.g. body mass index, smoking, physical activity, physician visits) and the NZ Census of citizens the same age on educational attainment. The cohort is primarily White (93%). Assessments were carried out at birth and ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, 38, and most recently (completed April 2019) 45 years when 875 Study members completed neuroimaging. The NZ Health and Disability Ethics Committee approved the Study, and all Study members provided written informed consent. The concept and main analyses for this project were preregistered (link: [rb.gy/34xv8a](https://rb.gy/34xv8a)). All analyses were checked for reproducibility by an independent data analyst who used the manuscript, code, and an independent copy of the data to check all analyses.

### MRI acquisition

Study members were scanned using a MAGNETOM Skyra 3 T scanner (Siemens Healthcare GmbH) equipped with a 64-channel head/neck coil at the Pacific Radiology imaging center in Dunedin, New Zealand. High-resolution T1-weighted images were obtained using an MP-RAGE sequence with the following parameters: TR = 2,400 ms; TE = 1.98 ms; 208 sagittal slices; flip angle, 9°; FOV, 224 mm; matrix =  $256 \times 256$ ; slice thickness = 0.9 mm with no gap (voxel size  $0.9 \times 0.875 \times 0.875$  mm<sup>3</sup>); and total scan time = 6 min and 52 s. Functional MRI (fMRI) data were collected using 72 interleaved axial T2-weighted functional slices with 3-fold multi-band accelerated echo planar imaging (TR = 2,000 ms; TE = 27 ms; flip angle = 90 degrees; field of view = 200 mm; voxel size = 2 mm isotropic; slice thickness = 2 mm without gap). Resting-state and task fMRI data were collected as follows: (i) resting-state with eyes open and a gray screen displayed (8:16 min; 248 TRs), (ii) emotional face processing task (6:40 min, 200 TRs), (iii) color Stroop task (6:58 min, 209 TRs), (iv) monetary incentive delay task (7:44 min,

232 TRs), and (v) episodic memory task (5:44 min, 172 TRs). Details for each of the four tasks can be found in the [Supplemental Materials](#). Twenty Study members completed the entire scanning protocol a second time (mean days between scans = 79 days) allowing for calculation of test–retest reliability of all our MRI-derived measures.

## fMRI preprocessing

Resting-state and task fMRI data were concatenated into one time series to derive estimates of General Functional Connectivity (GFC), which we have shown enhances test–retest reliability and improves prediction of behavior in the Dunedin Study and other datasets ([Elliott et al. 2019](#)). The use of GFC was further motivated by strong convergent evidence that functional networks are largely invariant to task conditions and task-derived functional networks are similar to those derived from resting-state data ([Fair et al. 2007](#); [Gratton et al. 2018](#)). Further analyses validating GFC in this dataset and others have been previously described ([Elliott et al. 2019](#)). Briefly, data were analyzed using the Human Connectome Project minimal preprocessing pipeline ([Glasser et al. 2013](#)). T1-weighted anatomical images were skull-stripped, intensity-normalized, and nonlinearly warped into a study-specific average template in MNI space ([Avants et al. 2008](#); [Klein et al. 2009](#)). Functional time-series data were de-spiked, slice-time-corrected, and realigned to the first volume in the time-series using AFNI ([Cox 1996](#)). To limit distortion caused by in-scanner head motion, motion regressors were generated using 6 motion parameters and their first derivatives (to account for nonlinear effects) for a total of 12 motion regressors. Five components from white matter and cerebrospinal fluid were extracted ([Behzadi et al. 2007](#)) and used as nuisance regressors along with the mean global signal. Images were bandpass filtered to retain frequencies between 0.008 and 0.1 Hz. To reduce the influence of motion-related artifacts, we excluded all high-motion Study members and adhered to strict scrubbing of motion-infected timepoints. We investigated a range of framewise-displacement cutoffs using QC-RSFC plots to determine the optimal threshold for removing motion artifacts as recommended by [Power et al. 2014](#). We selected 0.35 mm framewise-displacement and 1.55 standardized DVARS as exclusion thresholds. Nuisance regression, bandpass filtering, censoring, and global-signal regression were performed using AFNI's 3dTproject. For task data, functional connectivity due to signal evoked by task structure was removed using a Finite Impulse Response model ([Fair et al. 2007](#)). Finally, time series data were then projected to a two-dimensional fs\_LR32k neocortical surface space made up of ~30,000 points or “vertices” that can be convoluted to improve anatomical correspondence between people ([Van Essen et al. 2012](#)).

Of the 875 Study Members who underwent MRI scanning, 769 were included in the current analyses after quality control procedures. In sum, 62 Study Members were excluded for excess head motion, 14 were excluded based on visual inspection of irregularities in a 360 ROI × 360 ROI functional connectivity matrix, 10 were missing one or more functional scans (i.e. rest and four tasks), 8 were excluded due to missing 3D-FLAIR sequences, 7 were excluded as MRI data were acquired with a 20-channel head coil instead of the 64-channel head coil to accommodate larger head circumferences, 4 were excluded due to incidental neurological findings, and 1 was excluded due to a missing fieldmap.

## Multi-session hierarchical Bayesian modeling

We implemented MS-HBM per the strategy of [Cui et al. \(2020\)](#), who applied this method in a dataset with a similar amount of

fMRI data to the Dunedin Study. Broadly, MS-HBM defines several parameters that are modeled in a variational Bayes expectation maximization algorithm to estimate network labels across the neocortical surface. Parameters required for this method are inter-region variability, inter-subject variability, and intra-subject variability as well as two tuning parameters that control the importance of the group average parcellation and smoothness.

To estimate inter-regional variability, MS-HBM first generates binarized connectivity profiles for 1,483 equally spaced vertices across the neocortical surface. Binarized profiles were defined here as the 10% of vertices across the neocortical surface with the strongest functional connectivity to each of these 1,483 vertices. A group average network parcellation was then derived using the average binarized connectivity profiles of all Study members. To estimate intra-subject variability with only one scan per Study member, time series data were split into two halves of 17:41 min each, as recommended by [Kong et al. \(2019\)](#). Next, to estimate inter-subject variability, we used a resampling method as we did not have a validation dataset and computing across all Study members was highly computationally expensive. We randomly resampled 50 sets of 200 Study members from the total datasets available. We calculated inter-subject variability across each of these 50 sets and averaged these estimates. Finally, we selected the following tuning parameters that were optimized using data from the Human Connectome Project ([Kong et al. 2019](#)) due to similarities in our preprocessing procedures: smoothness prior:  $c = 40$ ; group spatial prior:  $\alpha = 200$ . The smoothness prior ( $c$ ) controls the penalty assigned for assigning two adjacent vertices to different networks. The group spatial prior ( $\alpha$ ) controls the weight of the group spatial prior. In other words, a parcellation estimated with a high group spatial prior will be more similar to the group average.

Given the above parameter estimates, MS-HBM generates an individual-specific parcellation from each Study member's GFC time series data using a variational Bayes expectation maximization algorithm ([Kong et al. 2019](#)). We chose to examine brain function using the Yeo 17-network parcellation, which represents an optimal solution to capturing the structure of correlations between neocortical regions and allows for ready comparison with existing literature ([Yeo et al. 2011](#)). Specifically, these 17 networks were derived by applying a clustering algorithm to functional connectivity strength between evenly distributed vertices across the cortex ([Yeo et al. 2011](#)). To evaluate how “well” the resulting individualized parcellations captured variation in BOLD signal, we calculated functional homogeneities for all individualized parcellations and compared them to homogeneities from template network parcellation ([Yeo et al. 2011](#)). Specifically, functional homogeneity is the average BOLD timeseries correlation between all pairs of vertices assigned to the same network. Therefore, higher homogeneity means that, on average, vertices within the same network are more functionally connected.

To estimate test–retest reliability, we generated parcellations using data from each of two separate scanning sessions in a subset of 20 Study members who were scanned twice. We first estimated test–retest reliability of individual network surface areas using a two-way mixed-effects intraclass correlation coefficient (ICC) with session modeled as a fixed effect and subject as a random effect ([Shrout and Fleiss 1979](#)). We also calculated test re-test reliability of whole brain topographic organization by computing the Dice coefficient between timepoints as well as all pairs of unique Study members. The Dice coefficient is a metric to determine the overlap between two sets of categorical features

(Destrieux et al. 2010; Birm et al. 2013; Kong et al. 2019, 2021). The Dice coefficient can be calculated using the following formula:

$$\text{Dice (parcellation A, parcellation B)} = \frac{2 \times (\text{\#of vertices that overlap between parcellations A and B})}{(\text{total number of vertices in parcellation A} + \text{total number of vertices in parcellation B})}$$

The Dice coefficient is equal to 1 if there is complete overlap between two parcellations and equal to 0 if there is no overlap between parcellations.

## Cognitive function

Full-scale IQ was assessed at age 45 using the Wechsler Adult Intelligence Scale-IV (WAIS-IV; Wechsler 2008). Full-scale IQ was used in our primary analyses. Full-scale IQ summarizes different cognitive processes including attention, working memory, visuospatial reasoning, and processing speed that may be differentially associated with specific functional brain networks. Thus, we tested if individual indices of the WAIS-IV were associated with distinct profiles of network topography. We present separate analyses using each of these indices in the [Supplemental Materials](#).

## Gait speed

The gait speed of Study members was assessed using the GAITRite Electronic Walkway with 2 m acceleration and deceleration before and after the walkway. Gait speed was assessed under three conditions: usual gait speed (walking at a normal pace from a standstill), maximum gait speed (walk as fast as safely as possible), and dual-task gait speed (walk at a normal pace while reciting alternate letters of the alphabet out loud). Study members completed each gait condition twice and the mean speed was taken for each condition. Gait speed was highly correlated across all three conditions ([Supplemental Materials](#)). For our main analysis, we used the average of the usual gait speed and maximum gait speed conditions to remove any cognitive effects present in the dual-task condition. We also completed all gait-speed analyses using the mean of all three conditions to compare with previously published results from this cohort (Rasmussen et al. 2019). These results can be found in the [Supplemental Materials](#).

To test the specificity of gait speed, we also conducted post hoc, exploratory analyses using a measure of balance as reflected by the maximum time Study members could stand on one leg with their eyes closed up to 30 s. A Study member's balance score was the maximum time across three trials. Further details can be found in the [Supplemental Materials](#).

## Listening task

Study members completed the Listening in Spatialized Noise—Sentences test (LiSN-S) in a soundproof booth. Stimuli were presented using Sennheiser 215 headphones attached to a Mini PCM2704 external sound card. The LiSN-S generates a three-dimensional auditory environment in four different conditions. During the task, target sentences are superimposed with distractor sentences. The distractor sentences were presented at 55 decibels sound pressure level (dB SPL). Study members repeated target sentences out loud and were automatically scored according to the number of correct words in each sentence. The program began with target sentences presented at 62 dB SPL and intensity levels were adjusted according to performance: the intensity was adjusted down if > 50% of the words in a sentence were correct and adjusted up if < 50% of the words were correct. The first several sentence presentations were considered practice, with each presentation lowered in 4-dB increments until performance

dropped below 50% accuracy, after which increments decreased to 2 dB. Practice sessions were not included in the final scores.

The test condition continued until the average of the positive and negative-going reversals was  $\geq 3$  and the standard error of these midpoints was < 1 dB. If Study members did not reach this point, the test condition simply ended after 30 sentence presentations. Speech reception thresholds (SRT) were considered the lowest intensity at which a Study member could repeat 50% of words correctly. We used the low cue speech reception threshold, where the distractor speaker and target speaker had the same identity and were presented in the same location in the auditory environment, in our primary analyses. We selected low cue speech reception threshold as it is the most challenging condition and may have greater age-related variation. Crucially, lower scores on this measure indicate *better* hearing, as it reflects the intensity at which a person can successfully distinguish between distractor and target sentences. Further, we tested the other speech reception threshold from the LiSN-S task and measures of pure tone audiometry (see [Supplemental Materials](#)). We also tested results from an adaptation of the Digit Triplets Test (King 2011; Van den Borre et al. 2021), a speech-in-noise task that uses white noise as a distractor instead of other speech (See [Supplemental Materials](#)). Results for these additional hearing measures can be found in the Supplemental Results and [Supplemental Fig. S1](#). We also repeated all analyses while controlling for overall hearing ability as measured by pure tone audiometry ([Supplemental Results](#), [Supplemental Fig. S2](#)).

## Primary analyses

We associated IQ, gait speed, and SRT-hearing with two broad features of functional topography: (i) total network surface area and (ii) spatial similarity. These are described in more detail below.

## Total network surface area

Total network surface areas were calculated for all 17 functional networks by summing the number of vertices assigned to a given functional network in each Study member. As functional topography maps were generated after normalizing all Study members to fs\_LR32k space, measures of network surface area already control for total neocortical surface area. We conducted 17 univariate regressions between each functional network and each behavioral measure (IQ, gait speed, and SRT-hearing) while controlling for sex and in-scanner head motion (average framewise displacement). We used a Bonferroni-corrected  $P$ -value of  $0.05/17 = 0.0029$  to determine statistical significance.

To more stringently test associations observed with total network surface areas, we also trained linear ridge regression models using network surface areas and tested their ability to predict IQ, gait speed, and SRT-hearing in unseen data using a split-half scheme. Specifically, we used a 2-fold nested cross-validation scheme to train our model (Cui et al. 2020). The outer fold was used to estimate the generalizability of the model and the inner fold was used to optimize the penalty parameters. For each behavioral measure (IQ, gait speed, and SRT-hearing) we generated a rank ordering of Study members and placed odd-ranking Study members into the training set and even-ranking Study members into the test set (i.e. unseen data). This train/test set split was the outer fold. This approach to data splitting ensured that training



and test sets would be closely matched on the behavioral outcome of interest.

We included an L2 regularization term to prevent model overfitting. This term was optimized during inner-loop cross-validation. Specifically, we again split the training set in half (training subset 1 and training subset 2). We then trained our ridge regression model using subset 1 to predict behavior in subset 2 using a variety of L2 regularization terms. We then repeated this procedure using subset 2 to predict subset 1. We calculated the accuracy of each prediction (Pearson's  $r$ ) and the mean absolute error (MAE) for each individual prediction. For each L2 term, we averaged the  $r$  and reciprocal of the MAE terms and summed these values to get an "accuracy" measure for each L2 term. The L2 term with the highest overall accuracy was carried forward to the outer fold step. Effects of sex and in-scanner head motion were regressed from our behavioral measures prior to model training. To prevent information leak, we regressed average in-scanner head motion and sex from the training set and applied the resulting beta weights to the test set. This approach ensured that our model only used information from the training set, including covariate regression, when calculating predictions in the test set.

We employed a permutation method to test our predictions for statistical significance. Specifically, we shuffled the behavioral scores within the training set 1,000 times and recomputed the prediction strength each time. We considered observed predictions statistically significant if they had prediction accuracy above the 95th percentile of null predictions and prediction error below the 5th percentile of null predictions. Finally, to ensure prediction accuracies were not due to our data splitting procedure, we generated 100 random outer fold splits and generated predictions for each of these random splits to compare with our estimates derived from rank-order splitting.

We utilized the Haufe transform to assess for relative feature importance in our multivariate prediction model (Haufe et al. 2014). Because multivariate regression models involve transforming predictor variables from a high-dimensional feature space into a lower-dimensional feature space, the multivariate beta weights cannot be interpreted as relative feature importance. Instead, we calculated the covariance between each predictive feature and outcome variables. These covariances can be thought of as "relative feature importance" and can allow for interpretation of our predictive model. Prior to Haufe transformation, we standardized total surface area measures and regressed the effects of sex and in-scanner head motion from each of our behavioral variables. We then calculated covariance between each network's standardized surface area and each predicted behavioral score. To maintain a consistent scale across behavioral measures, we divided each covariance measure by the variance of respective behavioral measures. To ensure that our feature importance estimates were not driven by idiosyncrasies in the data splitting procedure, we computed feature importance estimates across the aforementioned 100 random outer fold data splits. We report the average feature importance across these 100 predictions. Notably, we performed this analysis only using Study members in the training set to better interpret our trained regression model.

## Spatial similarity

To test behavioral associations with the spatial layout of networks, we again used a split-half approach to train kernel ridge regression models using brain wide topographic similarity and tested its ability to predict behavior in held-out data. These models predict behavior of a held-out Study member as the weighted average of

associations observed in Study members from the training set. The weights were determined by that held-out Study member's brain-wide topographic similarity to each of the training Study members, as quantified by the Dice coefficient (Kong et al. 2019; Chen et al. 2022). We used the Dice coefficient between Study members to estimate which Study members should show the most similarity in behavior. In other words, if two Study members have similar brain-wide functional topography, our models would predict that they have similar behavior. Additional details of this model can be found in the [Supplemental Materials](#). We used the same 2-fold nested cross-validation approach as above to train and test the regression models using spatial similarity. We again used similar permutation methods to test our predictions for statistical significance and for robustness to data splitting procedures.

## Secondary analyses—regional prediction

To assess for specific regional contributions to prediction accuracies, we assessed the accuracy of our kernel ridge regression models while training them using only topographic similarity from within 360 neocortical parcels (Glasser et al. 2016). By using a parcellation that is not based on our functional connectivity maps, we are able to extract regional patterns of functional topography across the cortex. Due to the misalignment of these parcellations, each of the 360 parcels contain individual-level variation in functional network boundaries. Thus, regional predictions of behavior are based only on the topography within a specific region of the cortex. This procedure allowed us to test for qualitative similarity of the neocortical regions most important for the predictions of IQ, gait speed, and SRT-hearing.

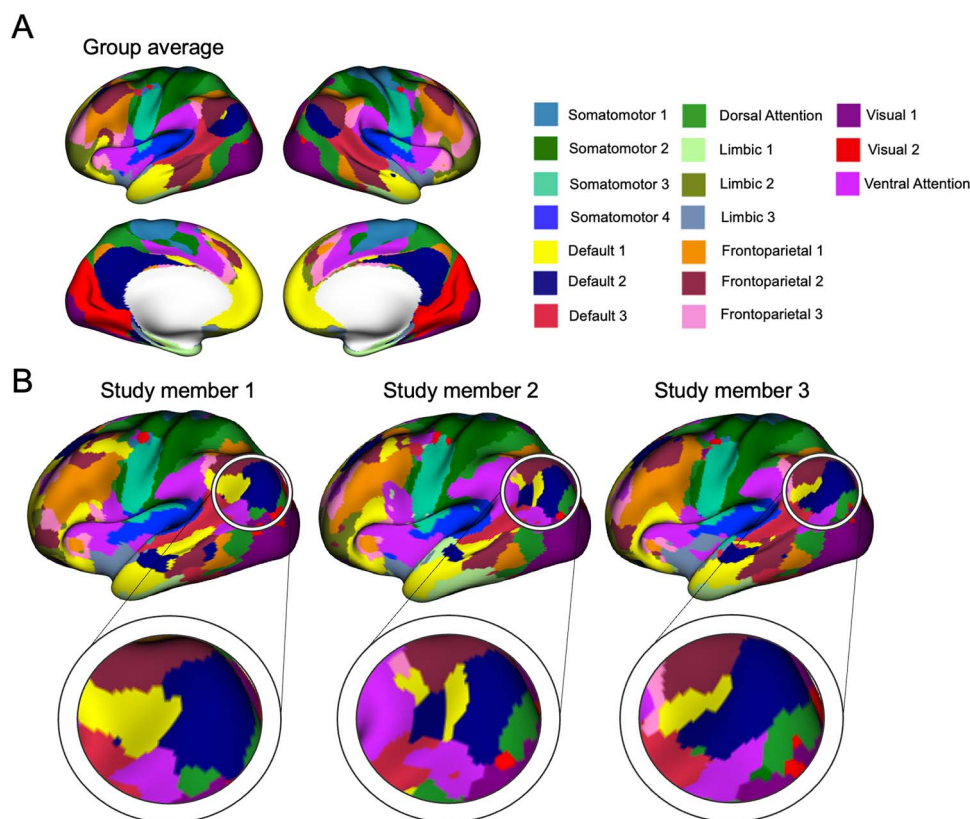
To formally test for correspondence between maps of parcel-wise prediction accuracies, we used a spatial permutation procedure known as a spin test (Váša et al. 2018). A spin test allows for tests of statistical significance by generating 10,000 random permutations of neocortical surface data while preserving the spatial covariance structure of the data. A truly random shuffling of neocortical data would create an unrealistically weak null distribution by generating biologically improbable distributions across the cortex (Alexander-Bloch et al. 2018). Thus, a spin test is a more conservative test of spatial correspondence. Correlations between two brain maps were considered statistically significant if they were higher than the 95th percentile of the set of null distributions generated through the spin test.

## Results

### Cohort characteristics

Attrition analyses revealed no significant differences in either childhood IQ or socioeconomic status between the full cohort, those still alive, those seen at age 45, or those scanned at age 45 (Supplemental Materials, [Supplemental Figs. S3–S4](#)). Of the 875 Study members completing the neuroimaging protocol at age 45, 769 had GFC data passing quality control that were included in the primary analyses. Study members with usable GFC data did not significantly differ from the full age 45 sample in IQ at age 45 ( $t = 1.44$ ,  $P = 0.15$ ) or sex distribution ( $X^2 = 0.21$ ,  $P = 0.88$ ). Thus, the current analyses continue to reflect effects in a population-representative cohort.

IQ, gait speed, and SRT-hearing were all normally distributed in the subsample of 769 Study members with high-quality GFC data ([Supplemental Fig. S5A](#)). IQ, gait speed, and SRT-hearing were all significantly correlated (IQ and gait:  $r = 0.38$ ,  $P < 0.001$ ; IQ and SRT-hearing:  $r = -0.31$ ,  $P < 0.001$ ; gait and SRT-hearing:  $r = -0.20$ ,



**Fig. 1.** Example of functional topography in the Dunedin Study cohort. (A) Group average functional parcellation. Legend shows colors corresponding to each of 17 neocortical networks. (B) Example of functional topography variation in three Study members. Left hemisphere is shown to portray overall group correspondence, whereas a region of the temporo-parietal junction is highlighted to show individual variation in topography.

$P < 0.001$ ; Supplemental Fig. S5B). This is consistent with previously reported associations between these variables in the full Dunedin Study cohort at age 45 (Rasmussen et al. 2019).

### Functional topography homogeneity and test-retest reliability

Consistent with previous work (Kong et al. 2019, 2021; Cui et al. 2020), we found that individualized parcellations showed reliable differences between Study members and boosted average functional homogeneity (i.e. the average correlation in BOLD signal between all pairs of vertices within each network) compared with the template parcellation. Each individual Study member showed higher homogeneity from their individualized parcellation compared with the template parcellation (Fig. 1, Supplementary Fig. S6).

High quality GFC data were available from 19 of the 20 Study members who repeated the scanning protocol. We found that total network surface areas showed fair test-retest reliability (mean network ICC = 0.47). We further observed strong within-person similarity between timepoints (Dice coefficient = 0.81), which was greater than between-person similarity (Dice coefficient = 0.72; Supplemental Fig. S7). These results are consistent with previous test-retest analyses of functional topography in other datasets (Kong et al. 2019, 2021; Cui et al. 2020). Thus, our maps capture reliable individual-level features of functional topography.

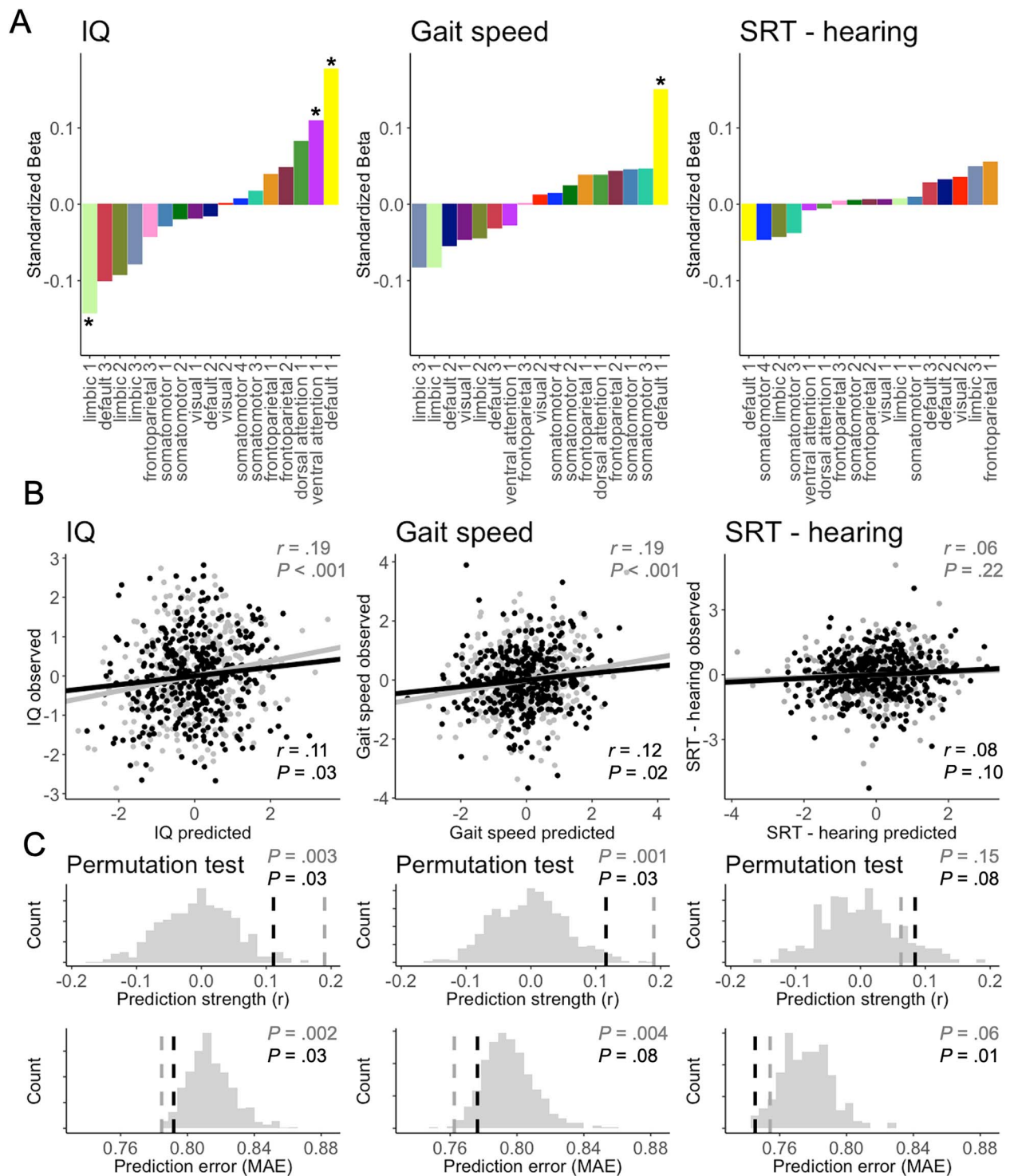
### Total network surface area

We first tested associations between total network surface area for each of the 17 functional networks and IQ, gait speed, and

SRT-hearing (Fig. 2A). IQ and gait speed but not SRT-hearing exhibited a broadly similar pattern of associations with network surface area. Notably, higher IQ and faster gait speed were associated with relatively larger default mode network 1, but relatively smaller default mode network 2 or 3 and limbic networks. No statistically significant associations with SRT-hearing were observed. A complete list of results is presented in Supplemental Table S1. Post hoc analyses revealed that balance showed a similar profile of univariate associations to gait speed. All index scores of the WAIS-IV showed similar profiles of univariate associations to full-scale IQ (see Supplemental Materials, Supplemental Figs. S8A and S9A). Thus, the associations observed between functional topography and IQ more closely reflect full-scale IQ and not any one domain of cognition.

To gauge the similarity between profiles of associations, we computed correlations between resulting beta weights with IQ, gait speed, and SRT-hearing. Although the pattern of associations appeared highly consistent between IQ and gait speed ( $r = 0.76$ ,  $P < 0.001$ ), there was no such consistency between IQ and SRT-hearing ( $r = -0.31$ ,  $P = 0.22$ ), or between gait speed and SRT-hearing ( $r = -0.40$ ,  $P = 0.11$ ).

In complementary analyses, ridge regression models trained using total network surface areas (henceforth: surface area prediction models) were able to predict variation in IQ (fold 1:  $r = 0.19$ ,  $P < 0.001$ ; fold 2:  $r = 0.11$ ,  $P = 0.03$ ) and gait speed (fold 1:  $r = 0.19$ ,  $P < 0.001$ , fold 2:  $r = 0.12$ ,  $P = 0.02$ ), but not SRT-hearing (fold 1:  $r = 0.06$ ,  $P = 0.22$ , fold 2:  $r = 0.08$ ,  $P = 0.10$ ; Fig. 2B). Our permutation procedure revealed that our predictions were significantly higher than would be expected by chance for IQ (fold 1:  $p_{\text{perm}} = 0.003$ , fold 2:  $p_{\text{perm}} = 0.03$ ) and gait speed (fold 1:  $p_{\text{perm}} = 0.001$ , fold 2:



**Fig. 2.** Network surface area analyses. (A) Results from univariate correlations between network surface area and IQ, gait, and SRT-hearing. Colors correspond to functional networks identified in Fig. 1(A). Correlations marked with a star are statistically significant after Bonferroni correction across the 17 comparisons. (B) Results from surface area prediction models for IQ, gait speed, and SRT-hearing. Scatterplots represent the correlation between observed scores and predicted scores according to ridge regression models. Fold 1 is shown in gray, whereas fold 2 is shown in black. (C) Permutation tests of observed predictions compared with 1,000 null predictions. (Top panel) Histograms show null distribution of prediction strengths and vertical lines show the observed prediction strengths. (Bottom panel) Histograms show the null distribution of prediction error and vertical lines show the observed prediction error. Fold 1 is shown in gray vertical lines, whereas fold 2 is shown in black.

$p_{\text{perm}} = 0.03$ ), but not for SRT-hearing (fold 1:  $p_{\text{perm}} = 0.15$ , fold 2:  $p_{\text{perm}} = 0.08$ ). The predictions for IQ also had less error than would be expected by chance (fold 1:  $p_{\text{perm}} = 0.002$ , fold 2:  $p_{\text{perm}} = 0.03$ ). Predictions for gait speed (fold 1:  $p_{\text{perm}} = 0.004$ , fold 2:  $p_{\text{perm}} = 0.08$ )

and SRT-hearing also had relatively low error (fold 1:  $p_{\text{perm}} = 0.06$ , fold 2:  $p_{\text{perm}} = 0.01$ ) (Fig. 2C). These predictions were robust to variation in data-splitting procedures (Supplemental Fig. S11). Secondary analyses found that surface area prediction models were



able to modestly predict balance as well as most WAIS-IV index scores (Supplemental Materials, Supplemental Figs. S8B and S9B).

Lastly, Hauke-transformed predictive feature scores derived for each functional network were broadly consistent between IQ and gait speed ( $r=0.74$ ,  $P=0.001$ ), but not between IQ and SRT-hearing ( $r=-0.28$ ,  $P=0.26$ ), or gait speed and SRT-hearing ( $r=-0.37$ ,  $P=0.14$ ). In addition, surface areas for default mode network 1 had strong positive feature importance for both IQ and gait, whereas limbic networks had strong negative feature importance for both IQ and gait (Supplemental Fig. S12).

## Spatial similarity

Next, we used a more fine-grained approach to testing the associations between functional topography and behavior. Using total network surface areas is a somewhat coarse way of testing associations with functional topography, as it does not consider the *shape* of functional networks. Furthermore, this coarse measure of total network surface area had only fair test re-test reliability in our dataset. Thus, the following analysis is based on the *shape* of individualized functional networks—a more detailed and reliable strategy of capturing functional topography. For this analysis, we again used a split-half approach to train kernel ridge regression models, this time using brain-wide topographic similarity as calculated using the Dice coefficient (henceforth: spatial similarity prediction models; Fig. 3A). We found that these spatial similarity prediction models were able to predict IQ (fold 1:  $r=0.34$ ,  $P<0.001$ , fold 2:  $r=0.23$ ,  $P<0.001$ ) and gait speed ( $r=0.19$ ,  $P<0.001$ , fold 2:  $r=0.10$ ,  $P=0.04$ ), but not SRT-hearing (fold 1:  $r=0.03$ ,  $P=0.58$ , fold 2:  $r=0.05$ ,  $P=0.32$ ) in held-out Study members (Fig. 3B). Prediction was stronger than would be expected by chance for IQ (fold 1:  $p_{\text{perm}}<0.001$ , fold 2:  $p_{\text{perm}}<0.001$ ) and gait speed (fold 1:  $p_{\text{perm}}<0.001$ , fold 2:  $p_{\text{perm}}<0.03$ ), but not SRT-hearing (fold 1:  $p_{\text{perm}}=0.28$ , fold 2:  $p_{\text{perm}}=0.14$ ). In addition, predictions for IQ (fold 1:  $p_{\text{perm}}<0.001$ , fold 2:  $p_{\text{perm}}<0.001$ ) had less error than would be expected by chance. Gait speed had relatively low levels of error (fold 1:  $p_{\text{perm}}=0.001$ , fold 2:  $p_{\text{perm}}=0.23$ ), but SRT-hearing prediction had error similar to the null distribution (fold 1:  $p_{\text{perm}}=0.64$ , fold 2:  $p_{\text{perm}}=0.09$ ) (Fig. 3C). Finally, these results were robust to variation in data splitting (Supplemental Fig. S13). Secondary analyses found that spatial similarity prediction models were also able to modestly predict balance as well as most WAIS-IV index scores (Supplemental Materials, Supplemental Figs. S8B and S9B). As has been observed previously (Kong et al. 2019; Cui et al. 2020), predictions using similarity tended to have stronger accuracy and less error compared with predictions using only total surface area (Supplemental Fig. S14).

## Regional prediction

To identify specific regional contributions to our observed predictions, we repeatedly assessed the accuracy of our spatial similarity prediction models while training them using only functional topography from within each of 360 anatomically derived neocortical parcels (Fig. 4A). Specifically, we calculated the Dice coefficient similarity between all pairs of participants for each neocortical parcel. We then used each of these 360 similarity matrices to train our spatial similarity prediction models. Each resulting prediction strength thus indicated how well functional topography from within that anatomical parcel alone could predict behavior. This procedure also allowed us to test for qualitative similarity of the neocortical regions most important for the predictions of IQ, gait speed, and SRT-hearing. Broadly, the strongest regional predictions for IQ and gait speed reflected variability in the functional topography of the temporoparietal

junction and superior temporal gyrus, regions assigned to various default mode subnetworks in our group average, as well as lateral frontal cortex. SRT-hearing was modestly predicted from variability in lateral frontal and temporoparietal parcels (Fig. 4B).

We also compared the prediction strength of each individual parcel with the prediction strength using global functional topography (i.e. across the entire cortex). We observed that the prediction of IQ based on global topography was stronger than prediction from the topography of any single parcel, whereas a small number of parcels predicted gait speed more strongly than did global topography. There was no difference between prediction strengths of regional or global topography for SRT-hearing (Fig. 4C).

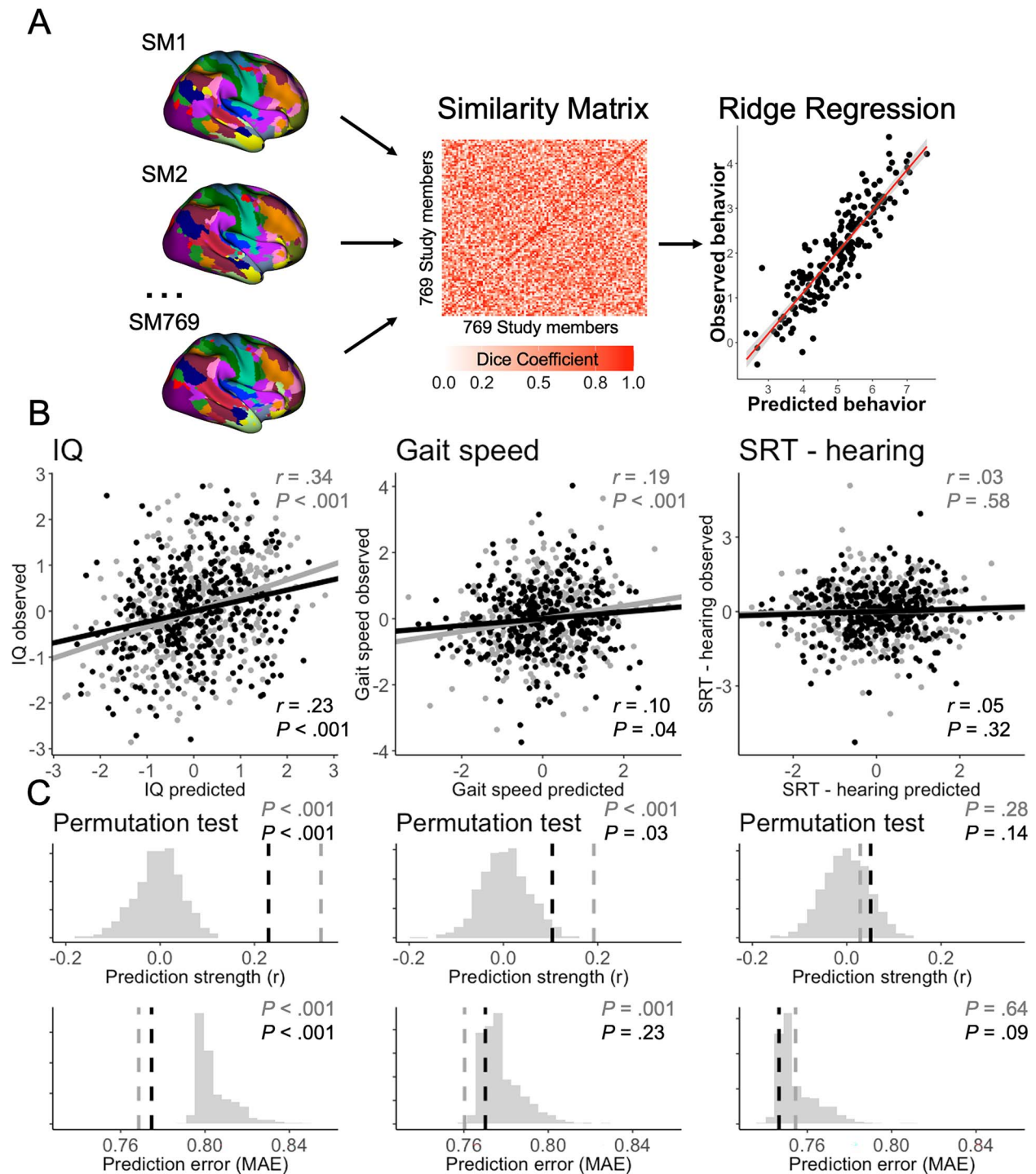
We used a spin test to formally evaluate correspondence between regional prediction maps for IQ, gait speed, and SRT-hearing. This test created 10,000 random permutations of values in the 360 neocortical parcels while preserving their overall spatial covariance structure, thus generating a more realistic and stricter null distribution to test the observed association (Váša et al. 2018). The spatial patterns of regional prediction for gait speed and SRT-hearing both modestly aligned with IQ (IQ-gait speed fold 1:  $r=0.12$ ,  $p_{\text{spin}}=0.01$ ; IQ-gait speed fold 2:  $r=0.21$ ,  $p_{\text{spin}}<0.001$ ; IQ-SRT-hearing fold 1:  $r=0.17$ ,  $p_{\text{spin}}<0.001$ ; IQ-SRT-hearing fold 2:  $r=0.14$ ,  $p_{\text{spin}}=0.007$ ). However, regional prediction patterns for gait speed and SRT-hearing did not align with each other (gait speed-SRT-hearing fold 1:  $r=0.03$ ,  $p_{\text{spin}}=0.30$ ; gait speed-SRT-hearing fold 2:  $r=0.06$ ,  $p_{\text{spin}}=0.09$ ; Fig. 4D).

## Discussion

We used a recently developed approach to reliably capture individual differences in the organization of functional networks across the neocortex (i.e. functional topography) in a large population-representative birth cohort now in midlife. We then leveraged this information to help better understand commonly observed correlations between seemingly disparate abilities, namely cognitive functioning as captured by IQ, and sensorimotor functioning as captured by gait speed and SRT-hearing. First, we found evidence for considerable variation in the topographic organization of common functional neocortical networks across this population-representative cohort, bolstering prior work with samples of convenience (Kong et al. 2019; Cui et al. 2020; Keller et al. 2022). Next, we found evidence for shared variation in the functional topography associated with IQ and gait speed but not SRT-hearing. Although IQ and gait speed would appear to be quite different behaviors, they mapped onto highly similar aspects of functional topography. This suggests that covariation between cognitive and motor abilities during midlife is partially driven by shared variation in the topography of functional networks distributed across the neocortex.

More specifically, we found that higher IQ and faster gait speed both mapped onto increased surface area of higher-order functional neocortical networks typically associated with cognitive functions and not lower-order somatomotor neocortical networks typically associated with motor functions. This is in line with prior findings that gait speed in older adults most strongly correlates with the functional connectivity strength of higher-order functional networks (Yuan et al. 2015; Lo et al. 2017). However, these prior studies utilized functional connectivity rather than functional topography so we cannot directly compare our results to these prior studies. In our analyses, variation in IQ

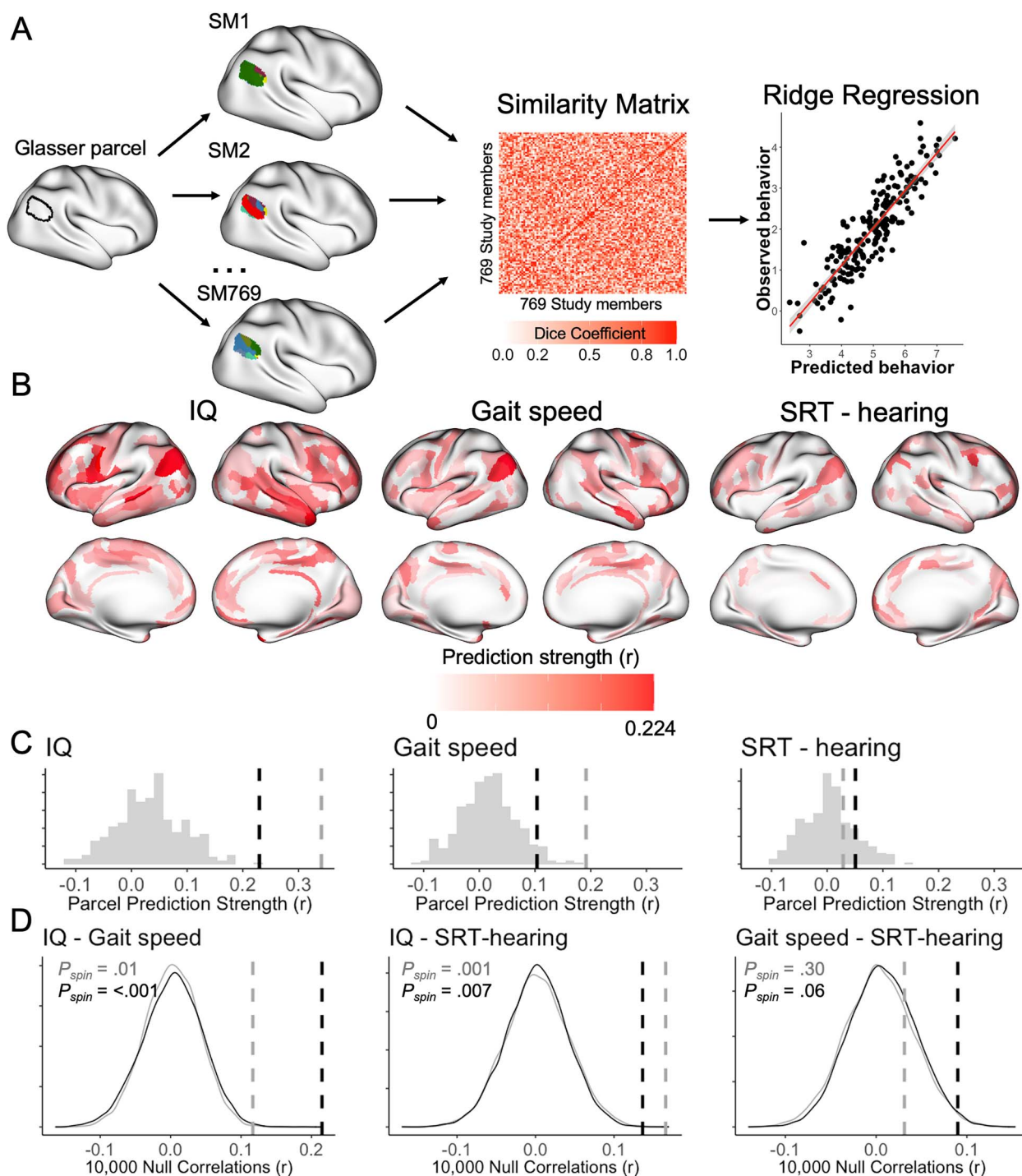




**Fig. 3.** Whole brain spatial similarity analyses. (A) Schematic depicting the organization of our ridge regression model. Brain-wide similarity (Dice coefficient) was calculated between all unique pairs of Study members and stored in a similarity matrix. This matrix was then used to train our ridge regression model to predict IQ, gait speed, and SRT-hearing in unseen Study members. (B) Results of spatial similarity prediction models for IQ, gait speed, and SRT-hearing. Scatterplots show the correlations between observed scores for IQ, gait speed, and SRT-hearing and the predicted scores according to our ridge regression models. Fold 1 is shown in gray, whereas fold 2 is shown in black. (C) Permutation testing of spatial similarity prediction models. (top panel) histograms show the null distribution of prediction strengths from 1,000 null predictions. Vertical lines represent observed prediction strengths. (Bottom panel) Histograms show the null distribution of prediction error from 1,000 null predictions. Vertical lines represent observed prediction error. Fold 1 is shown in gray, whereas fold 2 is shown in black.

and gait speed most strongly reflected the functional topography of the default mode network. Thus, slower gait speed in midlife appears to reflect changes in default mode network organization. Notably, the default mode network is thought to be critical during

aging, cognitive decline, and dementia onset (Buckner et al. 2005). Specifically weaker functional connectivity of the default mode network has been associated with aging, cognitive decline, and poorer health in adults (Sambataro et al. 2010; Smith et al. 2015;



**Fig. 4.** Regional prediction analyses. (A) Schematic depicting our method for regional prediction. For each of the 360 Glasser parcels, we generated a new similarity matrix based on topographic similarity between all unique pairs of Study members within only that Glasser parcel. We then tested the ability for topography in that parcel alone to predict IQ, gait speed, and SRT-hearing. (B) Prediction strengths for 360 Glasser parcels for IQ, gait speed, and SRT-hearing. Darker red indicates stronger prediction based on topography within that parcel. Parcels that had negative prediction values (i.e. no predictive strength) have no color. (C) Comparison of regional predictions with whole brain predictions. Histograms show distribution of predictions from each parcel and the vertical lines show prediction strength based on the whole brain. Fold 1 is shown in gray, whereas fold 2 is shown in black. (D) Cross-parcel correlations in predictions of IQ, gait speed, and SRT-hearing (vertical lines), overlaid on null density plots generated with spin tests. Fold 1 is shown in gray, whereas fold 2 is shown in black. The leftmost plot shows the correlation between parcel-wise prediction of IQ and gait speed, the middle plot shows the correlation between parcel-wise prediction of IQ and SRT-hearing, and the rightmost plot shows the correlation between parcel-wise prediction of gait speed and SRT-hearing.

Staffaroni et al. 2018). The observed topographic patterns of associations with IQ and gait speed may index poorer overall health at midlife that may predispose people to accelerated biological aging (Elliott et al. 2021), and could reflect the early signs of aging itself. This would align with the “last in, first out” hypothesis of brain aging (Douaud et al. 2014), which posits that the last brain networks to develop during early life (i.e. default mode network) are the first to degrade during later life. Given the relative youth of our cohort compared with most aging studies (i.e. age 45 vs. age 65+), the cross-sectional associations with the default mode network could reflect early stages of accelerated aging that first manifest in higher-order, later-developing networks. Additionally, those with the most extreme age acceleration at age 45 may be on a pathway to later frailty (Elliott et al. 2021). It is possible that early signs of this geriatric syndrome could manifest in higher-order functional network organization.

Notably, whereas we observed the strongest associations with the functional topography of the default mode network, our prediction models performed best when using information from all 17 neocortical networks. Thus, although the default mode network may have greater relative importance, the shared variance between IQ and gait speed reflects the functional organization of the entire neocortex.

In contrast, we did not observe similar associations between SRT-hearing and functional topography. This is somewhat surprising given that SRT-hearing was correlated with cognitive ability at a similar magnitude as gait speed in our cohort, and all three measures were similarly normally distributed. This null finding was true for SRT-hearing as well as other hearing variables and was robust to variation in overall hearing ability (see [Supplemental Materials](#) for additional analyses). It is important to note that very few Study members demonstrated clinically significant hearing loss at midlife. Age and peripheral hearing loss significantly contribute to SRT-hearing (Besser et al. 2015) and normal hearing has been shown to nullify the association between SRT-hearing and cognitive ability (Glyde et al. 2013). Although many studies have found associations between sensory functioning and cognitive decline, these studies have typically focused on participants older than 65 years, who would typically have mild or greater hearing loss (Lindenberger and Baltes 1994; Lin et al. 2011; Loughrey et al. 2018). A possible explanation for the absence of associations between SRT-hearing and functional topography is that sensory functioning is less closely tied to the organization of neocortical networks in midlife compared with cognitive and motor functioning. Indeed, when younger (25–69 years) and older (70–103 years) adults are directly compared, older adults show stronger associations between sensory and cognitive ability compared with younger adults (Baltes and Lindenberger 1997). Finally, prior research in the Dunedin Study has found that biological aging in midlife is more weakly associated with sensory functioning compared with cognitive and motor functioning (Elliott et al. 2021). Thus, if the profile of associations between functional topography, IQ, and gait speed does reflect midlife stages of biological aging, this profile may not yet reflect variation in hearing ability.

In addition to these novel findings, our analyses extend prior studies of functional topography in several ways. First, our individualized maps of functional topography had good test–retest reliability comparable with prior work (Kong et al. 2019, 2021), including with parcellation methods other than MS-HBM (Cui et al. 2020). Second, we found associations between functional topography and cognitive ability that have effect sizes and prediction performances that are similar to prior work in other datasets

(Kong et al. 2019; Cui et al. 2020; Keller et al. 2022). Third, we replicated findings that measures of network spatial organization tend to outperform summary metrics of total network surface areas in predicting behavior (Kong et al. 2019). Taken together, these findings provide further evidence for the utility of functional topography as a novel technique for reliably capturing individual differences in brain function and their mappings out to behavior.

Our study is not without limitations. First, neuroimaging, gait speed assessment, and hearing measurements were only conducted at one time point, precluding longitudinal analyses. Although we have previously found that gait speed is associated with longitudinal biological aging (Rasmussen et al. 2019), we are not yet able to describe longitudinal changes in gait speed itself, hearing ability, or functional topography. However, repeat testing of neuroimaging, gait speed, IQ, and hearing in this cohort will begin at age 52 in 2023. Second, the Dunedin Study cohort was established 5 decades ago, bringing an inherent limit on sample size. Thus, we faced power constraints typical of cross-sectional analyses of brain-behavior associations (Gratton et al. 2022; Marek et al. 2022). It is partially for this reason that we selected a strategy (MS-HBM) designed to reduce individual error inherent in traditional analyses of brain function. Although our replications of prior brain-behavior findings from other large datasets are reassuring (Kong et al. 2019), we must await additional independent replications of our novel findings. Moreover, our population-representative dataset may generalize more readily than convenience samples that suffer from healthy volunteer bias. In addition, our preregistration and reproducibility-check strategies prevent p-hacking, which is known to contribute to replication failures. Finally, SRT-hearing was associated with working memory and attention. Although this makes the measure a more ecologically valid assessment of hearing ability, it also may draw more heavily on cognition. However, we still did not observe any significant associations with topography and measures of pure hearing ability ([Supplemental Materials, Supplemental Figs. S1–S2](#)).

Taken together, our results present a profile of brain-wide functional organization that correlates with both cognitive and motor functioning during midlife. These convergent patterns across the functional topography of IQ and gait speed provide a plausible biological basis for why gait speed captures individual differences in midlife cognitive function (Rasmussen et al. 2019), overall health (Smith et al. 2015), and, potentially, early signs of accelerated aging and a pathway to later frailty. This suggests that gait speed may not be simply a measure of physical function but rather an integrative index of nervous system health.

## Acknowledgments

We thank the Dunedin Study members, Unit research staff, and Study founder Phil Silva. The Dunedin Unit is within the Ngāi Tahu tribal area who we acknowledge as first peoples, tangata whenua (translation: people of this land).

## Author contributions

Ethan Whitman (Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing—original draft, Writing—review and editing), Annchen Knodt (Conceptualization, Formal analysis, Methodology, Writing—original draft, Writing—review and editing), Maxwell Elliott (Conceptualization, Formal analysis, Methodology, Validation, Writing—review and editing), Cliff



Abraham (Data curation, Project administration, Writing—review and editing), Kirsten Cheyne (Data curation, Project administration, Writing—review and editing), Sean Hogan (Data curation, Project administration, Writing—review and editing), David Ireland (Data curation, Project administration, Writing—review and editing), Ross Keenan (Data curation, Project administration, Writing—review and editing), Joan Leung (Data curation, Project administration, Writing—review and editing), Tracy Melzer (Conceptualization, Project administration, Writing—review and editing), Richie Poulton (Data curation, Funding acquisition, Project administration, Writing—review and editing), Suzanne Purdy (Data curation, Project administration, Writing—review and editing), Sandhya Ramrakha (Data curation, Project administration, Writing—review and editing), Peter Thorne (Data curation, Project administration, Writing—review and editing), Avshalom Caspi (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing), Terrie Moffitt (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing), Ahmad Hariri (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing).

## Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

## Funding

This research received support from US-National Institute on Aging grants R01AG069939, R01AG032282, and R01AG049789 and UK Medical Research Council grant MR/P005918/1. The Dunedin Multidisciplinary Health and Development Research Unit is supported by the New Zealand Health Research Council (Programme Grant 16-604) and New Zealand Ministry of Business, Innovation and Employment (MBIE).

**Conflict of interest statement:** The authors report no biomedical financial interests or potential conflicts of interest.

## Data and code availability

Dunedin Study data are available via managed access (<https://moffittcaspi.trinity.duke.edu/research>). All code used in these analyses are available upon request.

## References

- Alexander-Bloch AF, Shou H, Liu S, Satterthwaite TD, Glahn DC, Shinohara RT, Vandekar SN, Raznahan A. On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage*. 2018;178(September):540–551.
- Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*. 2008;12(1):26–41.
- Baltes PB, Lindenberger U. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol Aging*. 1997;12(1):12–21.
- Behzadi Y, Restom K, Liao J, Liu TT. A Component based noise Correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*. 2007;37(1):90–101.
- Besser J, Festen JM, Goverts ST, Kramer SE, Pichora-Fuller MK. Speech-in-speech listening on the LiSN-S test by older adults with good audiograms depends on cognition and hearing acuity at high frequencies. *Ear Hear*. 2015;36(1):24–41.
- Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, Kirk GR, Nair VA, Meyerand ME, Prabhakaran V. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *NeuroImage*. 2013;83(December):550–558.
- Buckner RL, Snyder AZ, Shannon BJ, LaRossa G, Sachs R, Fotenos AF, Sheline YI, Klunk WE, Mathis CA, Morris JC, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J Neurosci*. 2005;25(34):7709–7717.
- Chen J, Tam A, Kebets V, Orban C, Ooi LQR, Asplund CL, Marek S, Dosenbach NUF, Eickhoff SB, Bzdok D, et al. Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nat Commun*. 2022;13(1):2217.
- Collyer TA, Murray AM, Woods RL, Storey E, Chong TT-J, Ryan J, Orchard SG, Brodtmann A, Srikanth VK, Shah RC, et al. Association of Dual Decline in cognition and gait speed with risk of dementia in older adults. *JAMA Netw Open*. 2022;5(5):e2214647.
- Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29(3):162–173.
- Cui Z, Li H, Xia CH, Larsen B, Adebimpe A, Baum GL, Cieslak M, Gur RE, Gur RC, Moore TM, et al. Individual variation in functional topography of association networks in youth. *Neuron*. 2020;106(2):340–53.e8.
- Cui Z, Pines AR, Larsen B, Sydnor VJ, Li H, Adebimpe A, Alexander-Bloch AF, Bassett DS, Bertolero M, Calkins ME, et al. Linking individual differences in personalized functional network topography to psychopathology in youth. *Biol Psychiatry*. 2022;92(12):973–983.
- Destrieux C, Fischl B, Dale A, Hagren E. Automatic Parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*. 2010;53(1):1–15.
- Douaud G, Groves AR, Tamnes CK, Westlye LT, Duff EP, Engvig A, Walhovd KB, James A, Gass A, Monsch AU, et al. A common brain network links development, aging, and vulnerability to disease. *Proc Natl Acad Sci USA*. 2014;111(49):17648–17653.
- Droby A, Varangis E, Habeck C, Hausdorff JM, Stern Y, Mirelman A, Maidan I. Effects of aging on cognitive and brain inter-network integration patterns underlying usual and dual-task gait performance. *Front Aging Neurosci*. 2022;14:956744.
- Elliott ML, Knodt AR, Cooke M, Kim J, Melzer TR, Keenan R, Ireland D, Ramrakha S, Poulton R, Caspi A, et al. General functional connectivity: shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks. *NeuroImage*. 2019;189(April):516–532.
- Elliott ML, Caspi A, Houts RM, Ambler A, Broadbent JM, Hancox RJ, Harrington H, Hogan S, Keenan R, Knodt A, et al. Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. *Nature Aging*. 2021;1(3):295–308.
- Fair DA, Schlaggar BL, Cohen AL, Miezin FM, Dosenbach NUF, Wenger KK, Fox MD, Snyder AZ, Raichle ME, Petersen SE. A method for using blocked and event-related fMRI data to study 'resting state' functional connectivity. *NeuroImage*. 2007;35(1):396–405.



- Fitzhugh MC, Schaefer SY, Baxter LC, Rogalsky C. Cognitive and neural predictors of speech comprehension in noisy backgrounds in older adults. *Lang Cogn Neurosci*. 2021;36(3):269–287.
- Fjell AM, Sneve MH, Grydeland H, Storsve AB, de Lange A-MG, Amlien IK, Røgeberg OJ, Walhovd KB. Functional connectivity change across multiple cortical networks relates to episodic memory changes in aging. *Neurobiol Aging*. 2015;36(12):3255–3268.
- Fritz S, Lusardi M. White paper: "walking speed: the sixth vital sign". *J Geriatr Phys Ther*. 2009;32(2):46–49.
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, et al. The minimal preprocessing pipelines for the human connectome project. *NeuroImage*. 2013;80(October):105–124.
- Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, et al. A multi-modal Parcellation of human cerebral cortex. *Nature*. 2016;536(7615):171–178.
- Glyde H, Cameron S, Dillon H, Hickson L, Seeto M. The effects of hearing impairment and aging on spatial processing. *Ear Hear*. 2013;34(1):15–28.
- Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, Gilmore AW, Nelson SM, Coalson RS, Snyder AZ, Schlaggar BL, et al. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*. 2018;98(2):439–52.e5.
- Gratton C, Nelson SM, Gordon EM. Brain-behavior correlations: two paths toward reliability. *Neuron*. 2022;110(9):1446–1449.
- Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J-D, Blankertz B, Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*. 2014;87(February):96–110.
- Humes LE. Associations between measures of auditory function and brief assessments of cognition. *Am J Audiol*. 2020;29(4):825–837.
- Humes LE, Busey TA, Craig J, Kewley-Port D. Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten Percept Psychophys*. 2013;75(3):508–524.
- Keller AS, Pines AR, Sydnor VJ, Cui Z, Bertolero MA, Barzilay R, Alexander-Bloch AF, Byington N, Chen A, Conan GM, et al. Personalized functional brain network topography predicts individual differences in youth cognition. *bioRxiv*. 2022. <https://doi.org/10.1101/2022.10.11.511823>.
- King SM. *Development and evaluation of a New Zealand digit triplet test for auditory screening* (Master's Thesis). Christchurch: University of Canterbury; 2011.
- Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang M-C, Christensen GE, Collins DL, Gee J, Hellier P, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*. 2009;46(3):786–802.
- Kohrman D, Wan G, Cassinotti L, Corfas G. Hidden hearing loss: a disorder with multiple etiologies and mechanisms. *Cold Spring Harb Perspect Med*. 2020;10(1). <https://doi.org/10.1101/cshperspect.a035493>.
- Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N, Zuo X-N, Holmes AJ, Eickhoff SB, et al. Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cereb Cortex*. 2019;29(6):2533–2551.
- Kong R, Yang Q, Gordon E, Xue A, Yan X, Orban C, Zuo X-N, Spreng N, Ge T, Holmes A, et al. Individual-specific areal-level Parcellations improve functional connectivity prediction of behavior. *Cereb Cortex*. 2021;31(10):4477–4500.
- Kwok SS, Nguyen X-MT, Wu DD, Mudar RA, Llano DA. Pure tone audiometry and hearing loss in Alzheimer's disease: a meta-analysis. *Front Psychol*. 2022;12:788045.
- Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen M-Y, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, et al. Functional system and areal organization of a highly sampled individual human brain. *Neuron*. 2015;87(3):657–670.
- Lin FR, Jeffrey Metter E, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L. Hearing loss and incident dementia. *Arch Neurol*. 2011;68(2):214–220.
- Lindenberger U, Baltes PB. Sensory functioning and intelligence in old age: a strong connection. *Psychol Aging*. 1994;9(3):339–355.
- Lo O-Y, Halko MA, Zhou J, Harrison R, Lipsitz LA, Manor B. Gait speed and gait variability are associated with different functional brain networks. *Front Aging Neurosci*. 2017;9(November):390.
- Loughrey DG, Kelly ME, Kelley GA, Brennan S, Lawlor BA. Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg*. 2018;144(2):115–126.
- Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, Donohue MR, Foran W, Miller RL, Hendrickson TJ, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*. 2022;603(7902):654–660.
- Montero-Odasso M, Muir SW, Speechley M. Dual-task complexity affects gait in people with mild cognitive impairment: the interplay between gait variability, dual tasking, and risk of falls. *Archives of Physical Medicine and Rehabilitation*. 2012;93(2):293–299.
- Okely JA, Akeroyd MA, Deary IJ. Associations between hearing and cognitive abilities from childhood to middle age: the National Child Development Study 1958. *Trends Hear*. 2021;25:233121652110537.
- Poulton R, Moffitt TE, Silva PA. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(5):679–693.
- Poulton R, Guiney H, Ramrakha S, Moffitt TE. The Dunedin study after half a century: reflections on the past, and course for the future. *J R Soc N Z*. 2022;1–20. <https://doi.org/10.1080/03036758.2022.2114508>.
- Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*. 2014;84(January):320–341.
- Rasmussen LJH, Caspi A, Ambler A, Broadbent JM, Cohen HJ, d'Arbeloff T, Elliott M, Hancox RJ, Harrington HL, Hogan S, et al. Association of neurocognitive and physical function with gait speed in midlife. *JAMA Netw Open*. 2019;2(10):e1913123.
- Sambataro F, Murty VP, Callicott JH, Tan H-Y, Das S, Weinberger DR, Mattay VS. Age-related alterations in default mode network: impact on working memory performance. *Neurobiol Aging*. 2010;31(5):839–852.
- Sardone R, Battista P, Panza F, Lozupone M, Griseta C, Castellana F, Capozzo R, Ruccia M, Resta E, Seripa D, et al. The age-related central auditory processing disorder: silent impairment of the cognitive ear. *Front Neurosci*. 2019;13:619.
- 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.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428.
- Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, Ugurbil K, Barch DM, Van Essen DC, Miller KL. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci*. 2015;18(11):1565–1567.
- Staffaroni AM, Brown JA, Casaletto KB, Elahi FM, Deng J, Neuhaus J, Cobigo Y, Mumford PS, Walters S, Saloner R, et al. The longitudinal trajectory of default mode network connectivity in healthy

- older adults varies as a function of age and is associated with changes in episodic memory and processing speed. *J Neurosci*. 2018;38(11):2809–2817.
- Stevenson JS, Clifton L, Kuźma E, Littlejohns TJ. Speech-in-noise hearing impairment is associated with an increased risk of incident dementia in 82,039 UK biobank participants. *Alzheimers Dement*. 2022;18(3):445–456.
- Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, Brach J, Chandler J, Cawthon P, Connor EB, et al. Gait speed and survival in older adults. *JAMA*. 2011;305(1):50–58.
- Sunderaraman P, Maidan I, Kozlovski T, Apa Z, Mirelman A, Hausdorff JM, Stern Y. Differential associations between distinct components of cognitive function and mobility: implications for understanding aging, turning and dual-task walking. *Front Aging Neurosci*. 2019;11:166.
- Van den Borre E, Denys S, van Wieringen A, Wouters J. The digit triplet test: a scoping review. *Int J Audiol*. 2021;60(12):946–963.
- van der Fels IMJ, Te Wierike SCM, Hartman E, Elferink-Gemser MT, Smith J, Visscher C. The relationship between motor skills and cognitive skills in 4-16 year old typically developing children: a systematic review. *J Sci Med Sport*. 2015;18(6):697–703.
- van der Willik KD, Licher S, Vinke EJ, Knol MJ, Darweesh SKL, van der Geest JN, Schagen SB, Ikram MK, Luik AI, Ikram MA. Trajectories of cognitive and motor function between ages 45 and 90 years: a population-based study. *J Gerontol A Biol Sci Med Sci*. 2021;76(2):297–306.
- van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T, Coalson T. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex*. 2012;22(10):2241–2262.
- Varangis E, Habeck CG, Razlighi QR, Stern Y. The effect of aging on resting state connectivity of predefined networks in the brain. *Front Aging Neurosci*. 2019;11:234.
- Váša F, Seidlitz J, Romero-Garcia R, Whitaker KJ, Rosenthal G, Vértes PE, Shinn M, Alexander-Bloch A, Fonagy P, Dolan RJ, et al. Adolescent tuning of association cortex in human structural brain networks. *Cereb Cortex*. 2018;28(1):281–294.
- Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ, Stoecklein S, Langs G, Pan R, Qian T, Li K, et al. Parcellating cortical functional networks in individuals. *Nat Neurosci*. 2015;18(12):1853–1860.
- Wechsler D. *Wechsler Adult Intelligence Scale*. 4th ed. Pearson Assessment; 2008.
- Welch D, Dawes PJD. Variation in the normal hearing threshold predicts childhood IQ, linguistic, and behavioral outcomes. *Pediatr Res*. 2007;61(6):737–744.
- Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 2011;106(3):1125–1165.
- Yuan J, Blumen HM, Verghese J, Holtzer R. Functional connectivity associated with gait velocity during walking and walking-while-talking in aging: a resting-state fMRI study. *Hum Brain Mapp*. 2015;36(4):1484–1493.