

ORIGINAL ARTICLE

Supramodal Mechanisms of the Cognitive Control Network in Uncertainty Processing

Tingting Wu¹, Alfredo Spagna², Chao Chen³, Kurt P. Schulz⁴,
Patrick R. Hof^{5,6} and Jin Fan¹

¹Department of Psychology, Queens College, The City University of New York, Queens, NY 11367, USA,

²Department of Psychology, Columbia University in the City of New York, New York, NY 10025, USA,

³Departments of Biomedical Informatics, Stony Brook University, Stony Brook, NY 11794, USA, ⁴Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA, ⁵Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA and ⁶Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Address correspondence to Jin Fan, Department of Psychology, Queens College, The City University of New York, 65-30 Kissena Blvd., Queens, NY 11367, USA. Email: jin.fan@qc.cuny.edu

Abstract

Information processing under conditions of uncertainty requires the involvement of cognitive control. Despite behavioral evidence of the supramodal function (i.e., independent of sensory modality) of cognitive control, the underlying neural mechanism needs to be directly tested. This study used functional magnetic imaging together with visual and auditory perceptual decision-making tasks to examine brain activation as a function of uncertainty in the two stimulus modalities. The results revealed a monotonic increase in activation in the cortical regions of the cognitive control network (CCN) as a function of uncertainty in the visual and auditory modalities. The intrinsic connectivity between the CCN and sensory regions was similar for the visual and auditory modalities. Furthermore, multivariate patterns of activation in the CCN predicted the level of uncertainty within and across stimulus modalities. These findings suggest that the CCN implements cognitive control by processing uncertainty as abstract information independent of stimulus modality.

Key words: cognitive control, cognitive control network, information theory, supramodal brain mechanisms, uncertainty

Introduction

The human brain receives and processes a vast amount of information from multiple sensory modalities (e.g., visual and auditory), but only a fraction of the information reaches the conscious mind. Cognitive control is the psychological mechanism that coordinates mental operations under conditions of uncertainty for the selection and prioritization of information to be processed (Fan 2014; Wu et al. 2018; Wu, Wang, et al. 2019b). An information theory account of cognitive control proposes that the role of the mechanism is to encode and process abstract information regarding uncertainty that can be quantified in

units of information entropy (Fan 2014). Behavioral evidence of significant correlation between conflict effects in visual and auditory tasks suggests a supramodal mechanism of cognitive control (Spagna et al. 2015; Spagna, Kim, et al. 2018b). However, the neural framework that supports the supramodal mechanism of cognitive control is still not well understood.

Neuroimaging studies have shown the common involvement of a large-scale network in cognitive control across sensory modalities and cognitive domains (see Wu, Chen, et al. 2019a, for our recent meta-analysis). This cognitive control network

(CCN) (Niendam et al. 2012; Fan et al. 2014; Wu et al. 2018; Wu, Chen, et al. 2019a) is composed of three subnetworks: 1) the frontoparietal network that includes the frontal eye field (FEF), supplementary eye field, middle frontal gyrus (MFG), areas near and along the intraparietal sulcus (IPS), and superior parietal lobule (Corbetta 1998; Fan et al. 2014); 2) the cingulo-opercular network comprising anterior cingulate cortex (ACC) and anterior insular cortex (AIC) (Dosenbach et al. 2007, 2008); and 3) a subcortical network involving thalamus and basal ganglia (Rossi et al. 2009; Fan et al. 2014; Koziol 2014). The involvement of the CCN in uncertainty processing has been demonstrated for the visual (e.g., Farah et al. 1989; Corbetta 1998; Colby and Goldberg 1999; Macaluso et al. 2002; Pessoa et al. 2003; Reynolds and Chelazzi 2004; Knight 2007; Silver and Kastner 2009; Fan et al. 2014; Wu et al. 2018) and auditory (e.g., Salmi et al. 2009; Westerhausen et al. 2010; Green et al. 2011; Donohue et al. 2012; Kong et al. 2014; Lee et al. 2014; Costa-Faidella et al. 2017) modalities. Furthermore, within-subject conjunction analyses of specific cognitive control functions (e.g., conflict processing and response inhibition) revealed spatially overlapping regions in the CCN associated with uncertainty processing in both the visual and auditory modalities (Laurens et al. 2005; Roberts and Hall 2008; Walther et al. 2010; Walz et al. 2013; Mayer et al. 2017). However, the potential mechanisms for the CCN to encode and process information uncertainty from different modalities in overlapping regions need to be further examined.

From a systems biology perspective, uncertainty is encoded in protocols that constrain the neural architecture of the CCN to functionally connect the subnetworks and regions (Doyle and Csete 2011). It is unknown whether uncertainty in information received from different modalities is encoded and processed by the CCN as multiprotocol representations comprising distinct patterns of modality-specific neuronal coupling between regions or as cross-protocol representations supported by the same pattern of neuronal coupling between regions. Moreover, individual regions within the CCN may encode information as either multimodal representations supported by distinct populations of modality-specific neurons or as more abstract cross-modal representations supported by a single modality-independent group of neurons (Fairhall and Caramazza 2013; Simanova et al. 2014; Kaplan et al. 2015; Corradi-Dell'Acqua et al. 2016; Handjaras et al. 2016; Wurm et al. 2016; Alizadeh et al. 2017; Jung et al. 2018). Little is known about whether the CCN encodes information from multiple modalities into abstract neural representations of uncertainty to support the supramodal function of cognitive control.

This study used functional magnetic resonance imaging (fMRI) together with visual and auditory perceptual decision-making tasks in a within-subjects design to examine the neural mechanisms for the supramodal processing of uncertainty in the CCN. The level of uncertainty (and the corresponding cognitive control load) was manipulated in the tasks by varying the stimulus congruency and was quantified as bits of information entropy. The two tasks incorporated a delayed-response approach to control for potential motor confounds. The hypothesis that the CCN supports the supramodal processing of uncertainty would be supported by findings of 1) variations in activation in regions of the CCN as a function of the level of uncertainty that are independent of sensory modality; 2) cross-protocol representations of uncertainty manifested in similar architectures of intrinsic connectivity within the CCN and between the CCN and modality-specific sensory regions for the visual and auditory tasks; and most importantly 3)

cross-modal representations of uncertainty evidenced by patterns of activation in regions of the CCN that predict the level of uncertainty both within and across sensory modalities.

Materials and Methods

Participants

Participants included 30 adult volunteers (15 females and 15 males; mean \pm standard deviation [SD] age: 26.3 ± 4.5 years, range = 19–37 years) with no history of brain injury or psychiatric and neurological disorders. Data from nine additional participants were excluded from the analyses due to missing files ($n=1$) and poor behavioral performance (i.e., overall hit rate $<80\%$; $n=8$). A conservative exclusion criterion of $<80\%$ overall hit rate in both the visual and auditory tasks was chosen to assure data quality. A low hit rate in the tasks may indicate that participants were not paying attention to the task or were not be able to hear the tones clearly in the auditory task due to the noisy environment during the MRI scan. The study was approved by the institutional review boards of The City University of New York and the Icahn School of Medicine at Mount Sinai (ISMMS). Written informed consent was obtained from all study participants. The participants were compensated for their participation in the study.

Sequential Majority Function Tasks

The original majority function task was developed to parametrically manipulate the level of uncertainty during cognitive control (Fan et al. 2008, 2014). This task was modified to present the stimuli set sequentially rather than simultaneously to accommodate the auditory stimuli. The visual (SMFT-V) and auditory sequential majority function tasks (SMFT-A) each consisted of two 954-s runs that all began and ended with 30-s period of fixation. Each run contained 72 trials.

Each trial in the SMFT-V (Fig. 1a) began with a 1-s fixation period, followed by the sequential presentation of 5 black arrows at the center of the screen. Each arrow was displayed for 400 ms followed by a 100-ms inter-stimulus interval (ISI) of fixation. A majority of the five sequentially presented arrows were either left or right pointing in each trial (majority factor: left vs. right). The ratio of arrows pointing in the majority direction versus the minority direction varied among 5:0, 4:1, and 3:2. Left- and right-pointing arrows in the 4:1 and 3:2 trials were presented in random order. The 5-arrow sequence was followed by a period of fixation varying from 2.1 to 5.1 s (mean = 3.1 s), after which a red probe arrow was displayed. The direction of the probe arrow matched the majority direction of the 5-arrow sequence in half of the trials. Participants were instructed to press a button with their right index finger when the direction of the probe arrow matched the majority direction of the 5-arrow sequence and to make no response if the direction of the probe did not match the majority direction. The responses window was 2.5 s starting at probe onset. The matching trials served as catch trials to monitor whether participants maintained attention to the task and were able to hear the stimuli clearly. Once a response was made, the probe arrow was replaced by fixation for the remaining time of the response window and was then followed by a post-trial delay period (jittered from 1 to 8 s, mean = 3 s). Each trial lasted from 9 to 16 s in duration (mean = 12 s). Each run included 24 trials per ratio (i.e., 5:0, 4:1, 3:2), evenly split between trials with left and right majority direction. Trial order in each

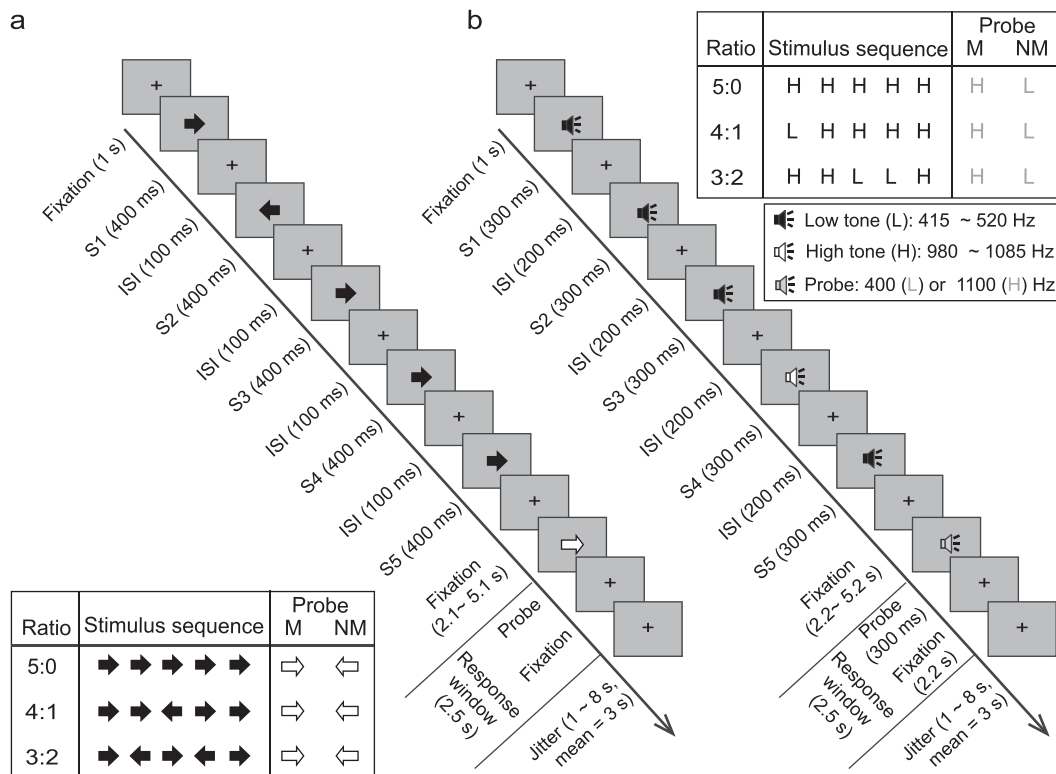


Figure 1. Schematic of the SMFT. (a) The visual task (SMFT-V). (b) The auditory task (SMFT-A). S, stimulus; ISI: inter-stimulus interval; M, matching condition; NM, nonmatching condition; H, high-pitched tone; L, low-pitched tone. Central panels: Timeline of the stimuli in one trial for each task. Participants were required to indicate whether the probe matched the majority direction or pitch of the stimulus sequence by pressing a response button only in the matching condition. Responses had to be made within a 2500-ms window. Corner Panels: Schematic of the 3 (Ratio: 5:0 vs. 4:1 vs. 3:2) \times 2 (Match: M vs. NM) design for each task. One possible stimulus sequence was illustrated for each ratio condition.

run was randomized across majority direction, ratio, and probe type.

The design and trial structure of the SMFT-A (Fig. 1b) were identical to the SMFT-V however, in the SMFT-A, the visual stimuli were replaced with auditory tones of 300 ms duration followed by a 200-ms ISI of fixation, and the post-probe fixation period was fixed to 2200 ms. The auditory stimuli consisted of eight high-pitched tones with frequencies ranging from 980 to 1085 Hz (i.e., 980, 995, 1010, 1025, 1040, 1055, 1070, and 1085 Hz) and eight low-pitched tones with frequencies in the 415 to 520 Hz range (i.e., 415, 430, 445, 460, 475, 490, 505, and 520 Hz). For the 5 tones in each trial, high-pitched and/or low-pitched tones were randomly sampled (without replacement) from the tone sets. The probe tone was either a high-pitched tone (1100 Hz) or a low-pitched tone (400 Hz). Participants were instructed to press a button with their right index finger when the pitch of the probe tone (i.e., high vs. low) matched the majority pitch of the 5-tone sequence and to make no response if the pitch of the probe tone did not match the majority pitch.

Estimation of Cognitive Control Load (in Information Entropy)

Information theory was applied to estimate the cognitive control load (i.e., amount of cognitive control needed to process uncertainty to reach a final decision) of the different conditions in the decision-making tasks. Cognitive control load was jointly determined by both the stimulus congruency and the cognitive strategy adopted to process the stimulus (Fan et al. 2008, 2014;

Fan 2014). The two most plausible strategies to search for the majority of items during the sequential presentation of stimuli make different demands on cognitive control: 1) an exhaustive search of each of the five arrows or tones sequentially prior to calculation of the majority, or 2) a self-terminating search that concluded once the majority could be determined (i.e., when 3 stimuli of the same direction/pitch in the sequence were presented). Each arrow or tone of the stimulus sequence entailed an independent decision between two alternatives (i.e., left vs. right pointing or high vs. low pitched) for both search strategies. The exhaustive search requires a constant 5 binary decisions per trial regardless of the congruency, while the self-terminating search requires an average of 3, 3.6, and 4.5 binary decisions to determine the majority in the 5:0, 4:1, and 3:2 ratio conditions, respectively. Uncertainty in a choice selection can be quantified as information entropy (H) in unit of bit as $H = \log_2(n)$, where n is the number of equiprobable alternative choices. Each binary decision corresponds to $\log_2(2) = 1$ bit of information entropy. Thus, cognitive control load is 5 bits of information entropy for all congruency conditions if the exhaustive search strategy is adopted, while it increases with the level of incongruence if the self-terminating search strategy is adopted, with respective means of 3, 3.6, and 4.5 bits of information entropy for the 3 ratio conditions, respectively.

The total cognitive control load for each trial gauges the demands imposed by binary decisions for each single stimulus but may also reflect demands imposed by evidence accumulation and working memory. The exhaustive search requires a constant updating of 5 items in working memory regardless of

the stimuli congruency. In contrast, the self-terminating search may require updating the judgment of the majority (i.e., the ratio) after processing each stimulus without maintaining the representation of each specific stimulus in working memory. Thus, the demands imposed by updating a subprocess of cognitive control (Miyake et al. 2000; Chen et al. 2019; Wu, Chen, et al. 2019a) may increase monotonically with the level of incongruence. Alternatively, the self-terminating search may engage working memory to maintain each stimulus in the series and involve a decision-making stage after processing each stimulus to determine whether the number of stimuli in the same direction/pitch in working memory has reached the majority. The decision-making stage engages the central executive component of working memory (Baddeley 2018), which is a cognitive construct that activates the CCN (Wu, Chen, et al. 2019a). In this case, the demand imposed by decision-making in working memory would also increase monotonically with the level of incongruency. Therefore, the total cognitive control load increases monotonically with the level of incongruency when taking both updating and working memory into account, with a range even greater than 1.5 bits that accounts for only the number of binary decisions.

This task has advantages over classical cognitive control tasks. The ~ 1.5 bits of cognitive control load increase in this task is greater than that found in such classical cognitive control tasks as the flanker conflict task (Eriksen and Yeh 1985) and the Stroop task (Stroop 1935), which engender less than or maximum 1 bit increase in uncertainty for the incongruent condition compared to the congruent condition (Fan 2014). This greater increase in cognitive control load together with a parametric design enhanced the detection power of brain regions and networks involved in cognitive control.

fMRI Data Acquisition

Participants were scanned on a 3T Siemens Magnetom Skyra system with a 16-channel phase-array coil (Siemens, Erlangen, Germany). All images were acquired in the axial plane parallel to the anterior commissure–posterior commissure (AC–PC) line. Head movement was minimized with foam padding. Four runs of 462 T_2^* -weighted gradient-echo echo-planar images (EPI) depicting the blood oxygenation level-dependent signal (time repetition [TR]=2000 ms, time echo [TE]=27 ms, flip angle = 77° , field of view [FOV]=240 mm, matrix size = 64×64 , voxel size = $3.75 \times 3.75 \times 4$ mm, 40 axial slices of 4 mm thickness with no skip) were acquired during the scan session, with the first two runs for the SMFT-V and the last two runs for the SMFT-A. Each run began with two dummy volumes before the onset of the task to allow for equilibration of T1 saturation effects. A high-resolution T1-weighted anatomical volume of the whole brain was acquired with a magnetization-prepared rapid gradient-echo sequence (TR=2200 ms, TE=2.51 ms, flip angle = 8° , FOV=240 mm, matrix size = 256×256 , voxel size = $0.9 \times 0.9 \times 0.9$ mm, 176 axial slices of 0.9 mm thickness with no skip).

Procedure

The task was compiled and run using E-Prime software (RRID: SCR_009567; Psychology Software Tools, Pittsburgh, PA). Participants practiced one run of each task, first on a desktop PC and then in an MRI simulator. During the actual scan, stimuli were projected onto a screen mounted at the back of the magnet

bore using a liquid crystal display projector. The length of the arrows in the SMFT-V was 2.5 cm on the screen with a viewing distance of 238 cm (visual angle = 0.6°). The auditory tones in the SMFT-A task were presented in both ears using MRI-compatible headphones (Avotec Audio System SS-3300 with a Persaio MRI noise cancellation system); the amplitude of all tones was constant. Participants responded with the index finger of the right hand using a fiber optic button system (BrainLogic, Psychology Software Tools).

Behavioral Data Analysis

For the 30 participants, the percentage of correct responses was calculated separately for each task condition and analyzed with a 2 (Task: visual vs. auditory) \times 3 (Ratio: 5:0 vs. 4:1 vs. 3:2) \times 2 (Probe: matching vs. nonmatching) repeated measures analysis of variance (ANOVA). Reaction time (RT) was calculated for correct responses on matching trials only because no response was required on nonmatching trials. Individual RTs exceeding 3 SD from the mean of the corresponding condition were considered outliers and were excluded from the analyses of RT. Mean RT was analyzed using a 2 (Task: visual vs. auditory) \times 3 (Ratio: 5:0 vs. 4:1 vs. 3:2) repeated measures ANOVA. Follow-up comparisons were conducted with a Bonferroni correction. Confidence interval in within-subject design (Cousineau 2005) was computed for each condition.

Image Preprocessing and General Linear Modeling

Event-related fMRI analyses were conducted using statistical parametric mapping software (SPM 12, RRID: SCR_007037; Wellcome Trust Centre for Neuroimaging, London, UK). T1 images and EPI images were manually adjusted to align with the AC–PC plane, if necessary. For each participant, the EPI image volumes were realigned to the first volume in the series, corrected for the staggered acquisition of slices, coregistered to the T1 image using normalized mutual information coregistration, spatially normalized to a standardized template with a resampled voxel size of $2 \times 2 \times 2$ mm, and finally spatially smoothed with a Gaussian kernel of 8 mm full-width half-maximum.

Separate general linear models (GLMs) were conducted to identify brain regions that support cognitive control in the visual and auditory modalities. Participant-specific GLM-fitted beta weights to 6 regressors representing expected neural responses to the 2 (Match: matching vs. nonmatching) \times 3 (Ratio: 5:0 vs. 4:1 vs. 3:2) task conditions on trials with correct responses. The vectors for each condition were time-locked to the onsets of the first arrow or tone in the stimulus sequences and had a duration of 2.5 s for each event. Two additional vectors for the onset of probes of the matching and nonmatching responses were also generated. Nuisance vectors were entered when necessary for trials with incorrect responses separately in each task condition, as well as for the corresponding probes in each match condition, for a maximum of 8 nuisance regressors per run. These vectors were convolved with a standard hemodynamic response function (HRF) (Friston et al. 1998) to generate the corresponding regressors and were entered into the GLM. The six motion parameters generated during realignment and sessions (dummy coded) were also entered as covariates of no interest. Low-frequency signal drift was removed using a high-pass filter with a 128-s cutoff and serial correlations were estimated with an autoregressive AR(1) model. The GLM generated an image of parameter estimates (beta image) for each regressor.

Appropriate orthogonal polynomial contrasts were applied to the beta images to test for brain regions that 1) showed monotonic increases or decreases in activation as a function of cognitive control load (conjunction of 4:1 minus 5:0 and 3:2 minus 4:1 contrasts), 2) supported task performance in general (contrast of all stimuli-locked regressors minus baseline), and 3) were involved in motor-related responses (contrast of matching probes minus nonmatching probes). In addition, brain activation associated with the contrast of matching minus nonmatching stimulus sets was also examined to test whether motor programming and preparation impacts activation during the processing of the stimulus sequences. These analyses resulted in four images of contrast estimates for each participant.

The images of contrast estimates for all participants were entered into second-level group analyses conducted with separate random-effect GLMs to test the four brain effects of interest. Conjunction analyses were performed to test for common regions of activation in the visual and auditory tasks for each contrast of interest. Disjunction analyses were conducted to identify modality-specific activation by masking activation in one modality (e.g., auditory) when testing for activation in the other modality (e.g., visual). Significance for these one-tailed tests was set at an uncorrected voxel-wise level of $P < 0.001$ for the height threshold and an extent threshold of k contiguous voxels that was estimated by random field theory to correct for multiple voxel comparisons at a cluster-wise level of $P < 0.05$.

Region of Interest Analysis

Region of interest (ROI) analyses were conducted to test for variations in activation as a function of cognitive control load to determine the cognitive search strategy employed in the visual and auditory tasks. ROIs in the supplementary motor area extending to the ACC (SMA/ACC), left and right AIC, left and right FEF, and left and right IPS were defined from the results of the cross-modal conjunction analysis of monotonic changes in cognitive control-related activation, as the corresponding clusters showing a significant effect. The first eigenvariate of the beta values was extracted from all voxels within each ROI separately to model regional activation as a linear function of information entropy (H), as predicted by the self-terminating search strategy. The linear model predicted that $\text{activation} = b \cdot H + c$, where b represents the gain in activation for each 1 bit increase in information entropy and c represents baseline activation as a constant when entropy is 0 bit. This model was compared to the constant model (i.e., $\text{activation} = c$) predicted by the exhaustive search strategy. All regressors in the models were demeaned. The coefficients in each model were estimated using mixed effect regression, with the factors of “entropy” included as both fixed and random effects and participants entered as a random effect in each model. Model comparison was based on the Bayesian information criteria (BIC), with a smaller BIC indicating better model fitting.

Dynamic Causal Modeling

Dynamic causal modeling (DCM) was conducted using DCM12 implemented in SPM12 (Friston et al. 2014, 2016; Razi et al. 2017) to identify the network architecture of the CCN and associated sensory regions during cognitive control for each modality. Ten nodes identified in the GLM analyses were entered into each model, including eight regions in the CCN (i.e., left and right SMA/ACC, AIC, FEF, IPS) and two modality-specific sensory

regions for the SMFT-V (i.e., left and right visual areas) and the SMFT-A (i.e., left and right auditory areas). A fully connected base model was initially constructed with reciprocal connections between all possible pairs of regions and self-connection of each region (i.e., intrinsic feedback). The three ratio conditions (5:0 vs. 4:1 vs. 3:2) were entered in the model as both driving inputs to the sensory regions and modulatory inputs that influence effective connectivity between all pairs of regions. The prior probability for each input and intrinsic connection was set at 1.

The parameters of the intrinsic connections and driving inputs in the two fully connected models were estimated at the first level for each participant individually using spectral DCM (Razi et al. 2017). The base model was elaborated into a series of nested models that systematically varied the connections between pairs of regions; the posteriors and parameters for the nested models were inferred with Bayesian model reduction analysis (Friston et al. 2016). The first-level posteriors and parameters for all participants were entered into group-level analyses that used parametric empirical Bayes models to estimate the group-level posteriors of the nested models separately for SMFT-V and SMFT-A. Bayesian model averaging was then used to estimate the group-level models based on the group-level posteriors.

Four separate analyses were conducted to test the cross-modal similarity of the networks of inter-regional connectivity in the estimated group-level DCM for SMFT-V and SMFT-A. The community structure of the CCN for the two tasks was compared to test the cross-modal similarity in network architecture. The community structure of each network was constructed using the Brain Connectivity Toolbox (RRID: SCR_004841; brain-connectivity-toolbox.net) by subdividing the network into nonoverlapping groups of nodes when maximizing the numbers of within-group connections and minimizing the number of between-group connections (Rubinov and Sporns 2010). The cross-modal similarity of the inter-regional connectivity for the entire network was tested by calculating Pearson correlation coefficient of the posterior expectations for the 90 inter-regional connections in the group-level DCM for SMFT-V and SMFT-A. The cross-modal similarity of the inter-regional connectivity for the CCN was tested by calculating Pearson correlation coefficient of the posterior expectations for the 56 inter-regional connections in the group-level DCM for SMFT-V and SMFT-A. Pearson correlation coefficient were also calculated for the posterior expectations for the 32 connections between regions of the CCN and visual and auditory sensory nodes in the two group-level models.

Cross-modal Encoding Analysis

Single-Trial Brain Response Extraction

Whole-brain responses to each visual and auditory stimulus sequence (trial) were extracted individually using an “extract-one-trial-out” approach (Rissman et al. 2004; Choi et al. 2012; Kinnison et al. 2012; Wu et al. 2018; Wang et al. 2019). Specifically, the onset for a single stimulus sequence (trial) was removed from the vector for the corresponding task condition and was added as a new vector with a single onset for the trial of interest. The new vector was then convolved with the standard HRF as the single regressor for the trial of interest. A GLM with the original regressors (but without the trial of interest) and the new regressor for the trial of interest was estimated voxel-by-voxel based on normalized images that were not spatially smoothed to avoid the loss of fine-grained data. The resultant estimated

beta image for the new regressor represents the brain response for the single trial of interest. The process was repeated for every trial with a correct behavioral response. Trials with global mean beta values more than 3 SD from the mean were excluded from the prediction analyses below.

Within- and Cross-modal Prediction

Multivoxel pattern analyses (MVPAs) were conducted to test for supramodal representation of uncertainty in the CCN. The estimated beta images for the single-trial responses to auditory and visual stimuli were each divided into training and testing sets individually for each participant. The first set of images was used to train a decoder to predict the cognitive control load represented by the single-trial response; the performance of the decoder was then tested using the independent testing set of images. The training and testing sets of images were either both from the visual or auditory task (within-modal prediction) or from different tasks (cross-modal prediction) including both visual-to-auditory and auditory-to-visual predictions. The within- and cross-modal predictions were tested using four sets of features: 1) CCN regions that showed significant monotonic increase in activation in the conjunction analysis, including left and right SMA/ACC, left and right AIC, left and right FEF, and left and right IPS, which were identical to the nodes in the DCM; 2) sensory regions that included left and right visual and auditory areas defined using the all stimuli-locked regressors minus baseline contrasts; 3) CCN and sensory regions combined; and 4) the whole brain.

Single-trial beta values for each voxel were transformed into z-scores across all trials in the training and testing sets for each participant. The z-scores for the training sets were entered as predictors in the models. The ratio of arrows pointing in the majority versus minority direction was coded trial-by-trial as a rank scale (rank 1=3:2 ratio; rank 2=4:1 ratio; rank 3=5:0 ratio) and served as the target in the models. A ranking support vector machine (SVM-R; [Joachims 2002](#)) was trained to predict the rank of cognitive load based on brain activation. A separating hyperplane, defined by a weight vector w computed across all predictors, was learned to maximize the margins between different ranks constrained by the assumption of the rank order. A linear kernel SVM-R was adopted because there were no assumptions of nonlinear relationships among features in the study. Higher absolute values for the learned weight vector in the linear kernel of SVM indicated greater importance of the corresponding feature (voxel) to the prediction. The performance of the trained decoder was evaluated on the testing set by applying the learned weight vector to each trial of the testing set to compute ranking scores, with a greater score indicating a higher rank. Computed ranking scores for pairs of trials with different ranks were compared to the actual ranking scores to test whether the trials were classified correctly. The overall performance (i.e., prediction accuracy) of each model was calculated as the percentage of the total number of possible pairs of trials in different ranks that were correctly classified. The chance level of prediction accuracy was 50%. The SMV-R was implemented using the SVM^{rank} toolbox ([Joachims 2006](#); https://www.cs.cornell.edu/people/tj/svm_light/svm_rank.html).

Within-modal predictions for the SMFT-A and SMFT-V were examined with 8-fold cross-validation tests that randomly subdivided the trials in each task into eight nonoverlapping subsets, in which each of the subsets served as the testing set and the remaining seven subsets were concatenated as the training set.

Prediction accuracy and the learned weight vectors were averaged across the eight training and testing routines conducted for each model. Cross-modal predictions were tested using all trials in one task (e.g., SMFT-V) as the training set and all trials in the other task (e.g., SMFT-A) as the testing set. The learned weight vectors were extracted from the voxel-wise cross-modal predictions for each feature set and were entered into group-level GLM to identify common activation patterns across participants that predicted uncertainty. The weight vectors learned from the visual-to-auditory prediction and the auditory-to-visual prediction were analyzed separately. The absolute value of the weight of a voxel represents its contribution to the classification, while the sign of the weight indicates the condition that voxel supports, with positive and negative values corresponding to higher and lower cognitive control load conditions, respectively.

One-sample t-tests (one-tailed) were conducted to assess each prediction for accuracy higher than chance level (50%). The difference in accuracy between within-modal and cross-modal predictions using the CCN feature was assessed with two paired t-tests (two-tailed) that compared 1) visual-to-visual prediction versus visual-to-auditory prediction and 2) auditory-to-auditory prediction versus auditory-to-visual prediction. Planned comparisons used two paired t-tests (one-tailed) to contrast prediction accuracy for each within- and cross-modal prediction to test the hypothesis that 1) the combined CCN and sensory feature set is greater than the CCN feature set alone and 2) the whole brain feature set is greater than the combined feature set in terms of prediction accuracy. Bonferroni correction was applied to correct the significance level for multiple comparisons.

Results

Response Accuracy and RT as a Function of Entropy

The mean accuracy, including both the hit rate in matching trials and the correct rejection rate in nonmatching trials, was 89.9% (range 83.3–100%) in the most difficult condition (the 3:2 matching condition of the SMFT-A), indicating that participants maintained attention to the task and were able to see/hear the stimuli clearly during the scan. [Supplementary Table 1](#) presents the mean \pm SD accuracy and RT of responses for each ratio condition of the SMFT-V and the SMFT-A. The ANOVA testing accuracy revealed a significant main effect of Task, $F(1, 29) = 10.74$, $P = 0.003$, with higher accuracy on the visual task (mean \pm standard error: $96.6 \pm 0.8\%$) than the auditory task ($96.2 \pm 1.0\%$). The main effect of Ratio was also significant, $F(2, 58) = 27.12$, $P < 0.001$, with significantly lower accuracy in the 3:2 condition ($92.7 \pm 1.1\%$) than in the 4:1 ($96.6 \pm 0.7\%$, $P < 0.001$) and 5:0 ($96.8 \pm 0.9\%$, $P < 0.001$) conditions, but the difference between the 4:1 and 5:0 conditions was not significant ($p > 0.99$). The main effect of Match was not significant, $F(1, 29) = 3.20$, $P = 0.084$, but there was a significant Task-by-Match interaction, $F(1, 29) = 4.66$, $P = 0.039$, with a significant effect of Task on matching trials, $F(1, 29) = 11.41$, $P = 0.002$, but not on nonmatching trials, $F(1, 29) = 2.322$, $P = 0.138$. Task-by-Ratio, $F(2, 58) < 1$, Match-by-Ratio, $F(2, 58) < 1$, and Task-by-Match-by-Ratio interactions were not significant, $F(2, 58) = 3.14$, $P = 0.056$. The ANOVA testing RT revealed a significant main effect of Task, $F(1, 29) = 121.79$, $P < 0.001$, but not main effect of Ratio, $F(2, 58) = 1.05$, and Task-by-Ratio interaction, $F(2, 58) < 1$. Responses were significantly faster on the visual task (665 ± 22 ms) than on the auditory task (842 ± 28 ms).

Cross-modal Involvement of the CCN in Cognitive Control: GLM Results

Significant monotonic increases in activation as a function of cognitive load were found bilaterally in all cortical regions of the CCN for both the SMFT-V and SMFT-A (Fig. 2a,b and Supplementary Table 2). No region showed a significant monotonic decrease in activation. Conjunction analysis of the monotonic increase in activation as a function of cognitive load on the SMFT-V and SMFT-A identified significant modality-independent activation in cortical regions of the CCN, including the SMA/ACC, AIC, FEF, MFG, and IPS (Fig. 2c and Supplementary Table 3). The disjunction analysis found no significant modality-specific activation associated with cognitive control load for either task. All identified regions of the CCN were inside the clusters of the CCN identified in our meta-analytic study (Wu, Chen, et al. 2019a) highlighted as the green contours in the section views of Figure 2. The monotonic increase in activation as a function of cognitive control load for each ROI is illustrated in Figure 3. The regional activation fits the linear model significantly better than the constant model for all cortical regions of the CCN (Supplementary Table 4).

The all regressor minus baseline contrast identified significant modality-specific activation for both the SMFT-V and SMFT-A. General task-related activation was found in all cortical regions of the CCN in both the visual and auditory modalities (Fig. 4a,b, respectively), as well as in extrastriate visual areas in mid and inferior occipital gyri, fusiform gyrus, mid temporal gyrus, and superior and inferior parietal lobules in the visual modality (Supplementary Table 5) and primary and secondary auditory cortices and putamen in the auditory modality (Supplementary Table 6). Disjunction analyses identified significant visual-specific activation in extrastriate visual areas (Fig. 4c and Supplementary Table 7) and auditory-specific activation in primary and secondary auditory cortices and putamen (Fig. 4d and Supplementary Table 7). General task-related deactivation was revealed in regions of the default mode network for both tasks, with deactivation in visual-specific areas in the auditory task and deactivation in auditory-specific areas in the visual task.

The contrast of matching minus nonmatching probe identified significant motor-related activation in primary motor cortex extending ventrally to precentral gyrus, mid and posterior insular cortex, ventral ACC extending to mid cingulate cortex, thalamus, putamen, caudate nucleus, and the ventral part of cerebellum for both the SMFT-V and SMFT-A (Supplementary Fig. 1a,b). The activation was bilaterally distributed except in primary motor cortex, which was limited to the left hemisphere for the right-handed response. The conjunction analysis revealed a pattern of activation identical to those found separately for the individual modalities (Supplementary Fig. 1c and Supplementary Table 8). The contrast of matching minus nonmatching stimulus sequence identified no significant activation.

Cross-modal Similarity in Effective Connectivity of the CCN: DCM Results

The estimated group-level models of effective connectivity revealed highly interconnected cortical regions of the CCN that were also connected to sensory input regions for both the SMFT-V and SMFT-A (Fig. 5a,b). The community structure of intrinsic connectivity was composed of the same four modules for the auditory and visual tasks: 1) left and right SMA/ACC, 2) left

and right AIC, 3) left and right FEF and IPS, and 4) sensory regions in left and right visual areas for the SMFT-V or left and right auditory areas for the SMFT-A. This community structure reflects the protocol of network architecture for uncertainty processing. The inter-regional connectivity in the entire network was significantly correlated across the two tasks ($r=0.81$, $n=90$, $P < 0.001$). The inter-regional connectivity in the CCN (excluding the sensory regions) was significantly correlated across the two tasks ($r=0.81$, $n=56$, $P < 0.001$). The similarity in connectivity between specific regions of the CCN and sensory input regions for the two modalities was significant even after excluding the intra-CCN connections and connections between left and right sensory regions ($r=0.71$, $n=32$, $P < 0.001$).

Cross-modal Representation of Uncertainty in the CCN: Cross-decoding Analysis Results

Accuracy rates of the within- and cross-modal predictions of cognitive control load are shown in Figure 6. The accuracy was significantly higher than the chance level threshold for all predictions (all P s < 0.001), indicating that both CCN and sensory regions supported within-modal and cross-modal representations of uncertainty. The difference between within-modal and cross-modal predictions in accuracy using the CCN feature was marginally significant for the within-visual versus visual-to-auditory comparison ($68.3 \pm 1.6\%$ vs. $59.9 \pm 1.0\%$, $t_{28}=1.99$, $P=0.056$) and was significant for the within-auditory versus auditory-to-visual comparison ($62.6 \pm 1.6\%$ vs. $59.7 \pm 1.0\%$, $t_{28}=6.01$, $P < 0.001$). The accuracy based on the combined CCN and sensory feature set was significantly higher than the accuracy based on the CCN feature set only for the within-auditory prediction ($66.8 \pm 1.7\%$ vs. $62.6 \pm 1.6\%$, $t_{28}=4.53$, $P < 0.001$), but not for the within-visual prediction ($68.9 \pm 1.6\%$ vs. $68.3 \pm 1.6\%$, $t_{28}=0.631$, $P=0.267$), visual-to-auditory prediction ($59.6 \pm 1.2\%$ vs. $59.9 \pm 1.0\%$, $t_{28}=0.376$, $P=0.627$), and auditory-to-visual prediction ($59.5 \pm 1.2\%$ vs. $59.7 \pm 1.0\%$, $t_{28}=0.306$, $P=0.619$). Prediction accuracy based on the whole-brain feature set was significantly higher than accuracy based on the combined CCN and sensory feature set for the within-auditory prediction ($69.6 \pm 1.6\%$ vs. $66.8 \pm 1.7\%$, $t_{28}=2.65$, $P=0.026$), visual-to-auditory prediction ($63.3 \pm 1.3\%$ vs. $59.6 \pm 1.2\%$, $t_{28}=3.65$, $P=0.002$), and auditory-to-visual prediction ($64.1 \pm 1.4\%$ vs. $59.5 \pm 1.2\%$, $t_{28}=4.00$, $P < 0.001$), but not for the within-visual prediction ($70.8 \pm 1.7\%$ vs. $68.9 \pm 1.6\%$, $t_{28}=1.69$, $P=0.203$).

The distributions of the weight vector for voxel-wise cross-modal predictions based on the CCN set and the combined CCN and sensory set are presented in Supplementary Figure 1a and b, respectively. The sparse distribution of voxels with weights significantly different from zero in Supplementary Figure 1a suggests that the fine-grained multivoxel representation of uncertainty within regions of the CCN varies substantially between participants. As shown in Supplementary Figure 2c, the distribution of the weight vector for whole-brain predictions indicates that regions of the CCN and modality-specific sensory areas both contributed significantly to predictions of cognitive load.

Discussion

Supramodal Involvement of the CCN in Cognitive Control

The common involvement of regions in the CCN in the processing of information uncertainty in both auditory and visual modalities confirms the crucial role of the network

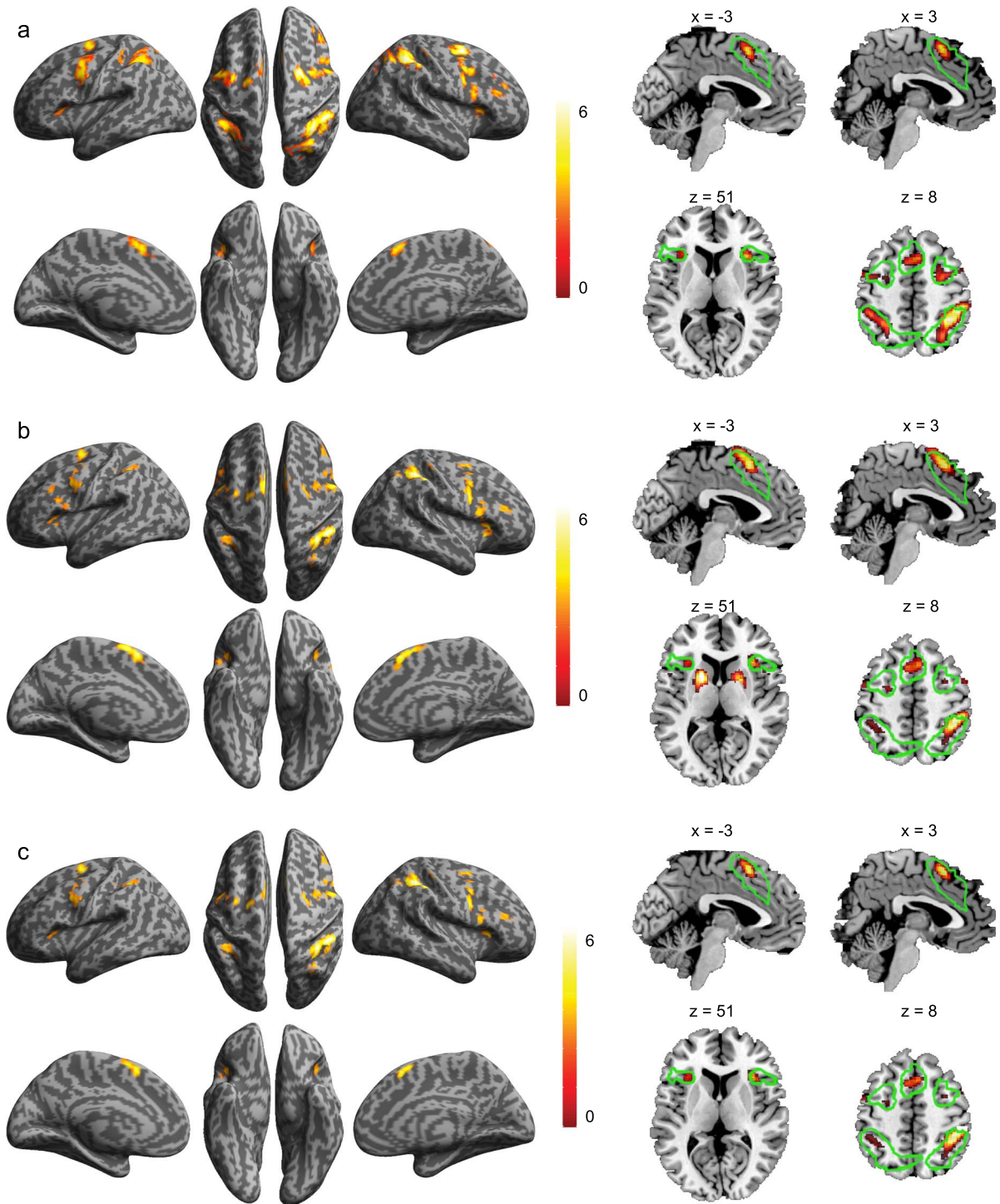


Figure 2. Brain regions involved in uncertainty processing. Brain regions that showed a significant monotonic increase in activation as a function of cognitive load for the SMFT-V (a) and for the SMFT-A (b) and the conjunction across the two tasks (c). The color bar indicates T value. Left column: Surface view. Right column: Sagittal view (top) and axial view (bottom), with the green contours showing the cluster of regions of the CCN identified in our previous meta-analytic study.

in the supramodal function of cognitive control. Task-related activation in cortical regions of the CCN increased monotonically as escalating information uncertainty made increasing demands on cognitive control. Moreover, the conjunction of the monotonic activation gains for visual and auditory information

processing showed in the same regions of the CCN, suggesting that this network encodes modality-independent representations of uncertainty that are integral to cognitive control. The extensive connectivity among regions of the CCN and between these regions and sensory input regions forms the neural

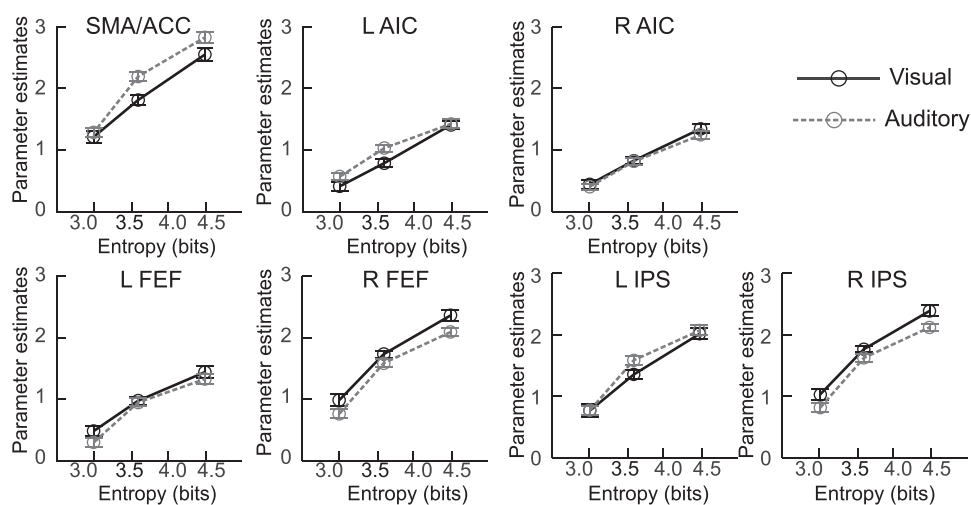


Figure 3. Activation of the regions of interest as a function of entropy for SMFT-V and SMFT-A. L, left; R, right; SMA/ACC, supplementary motor areas extending to anterior cingulate cortex; AIC, anterior insular cortex; FEF, frontal eye field; MFG, mid frontal gyrus; IPS, areas near and along the intraparietal sulcus. Error bar indicates 95% confidence interval in the within-subject design.

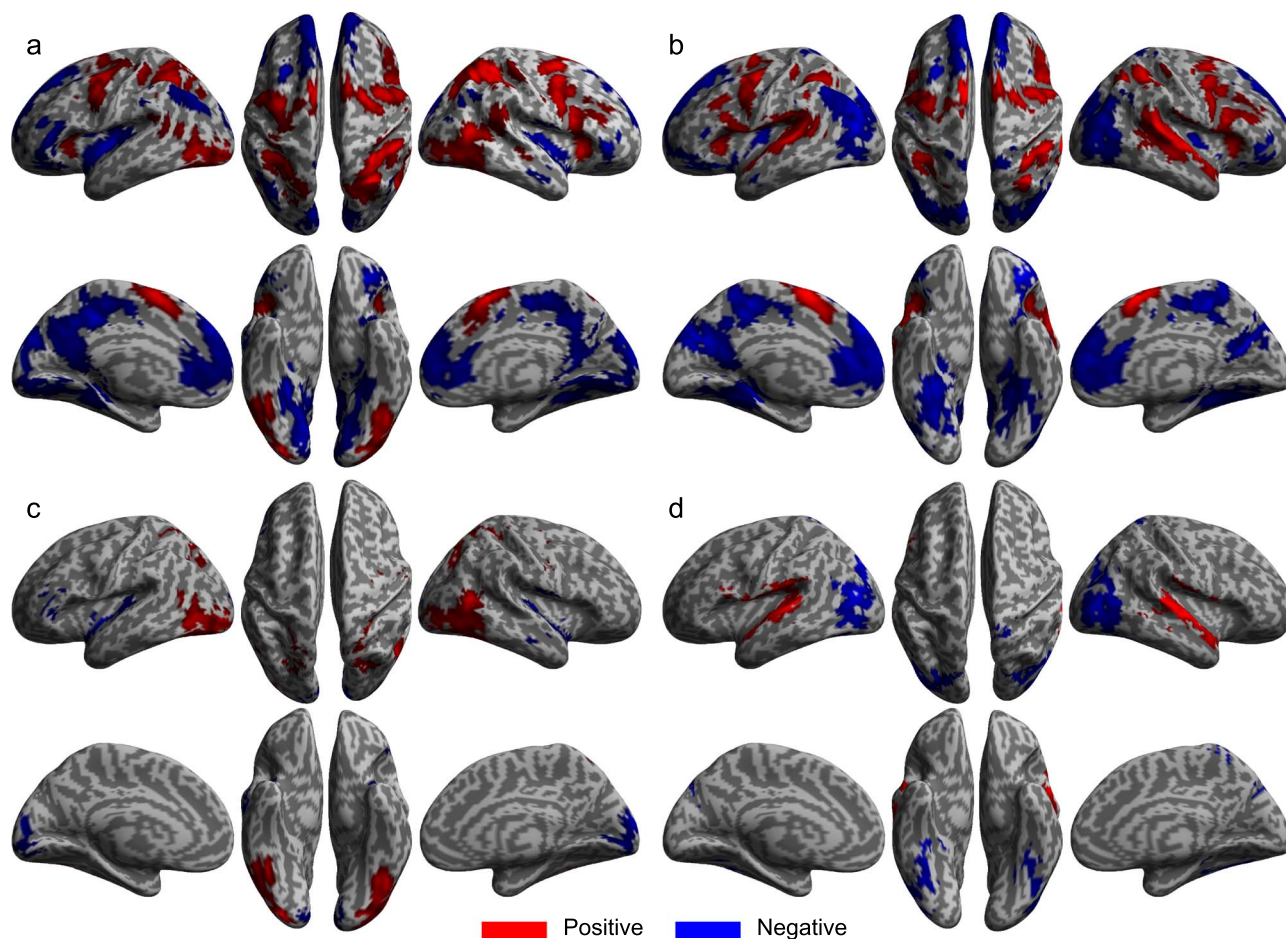


Figure 4. General and specific brain activation for SMFT-V and SMFT-A. Brain regions that showed significant activation change for all stimulus-locked regressors minus baseline for the SMFT-V (a) and for the SMFT-A (b). Brain regions that showed significant visual-specific (c) and auditory-specific (d) activation. Red: Region that showed an increase in activation. Blue: Region that showed a decrease in activation.

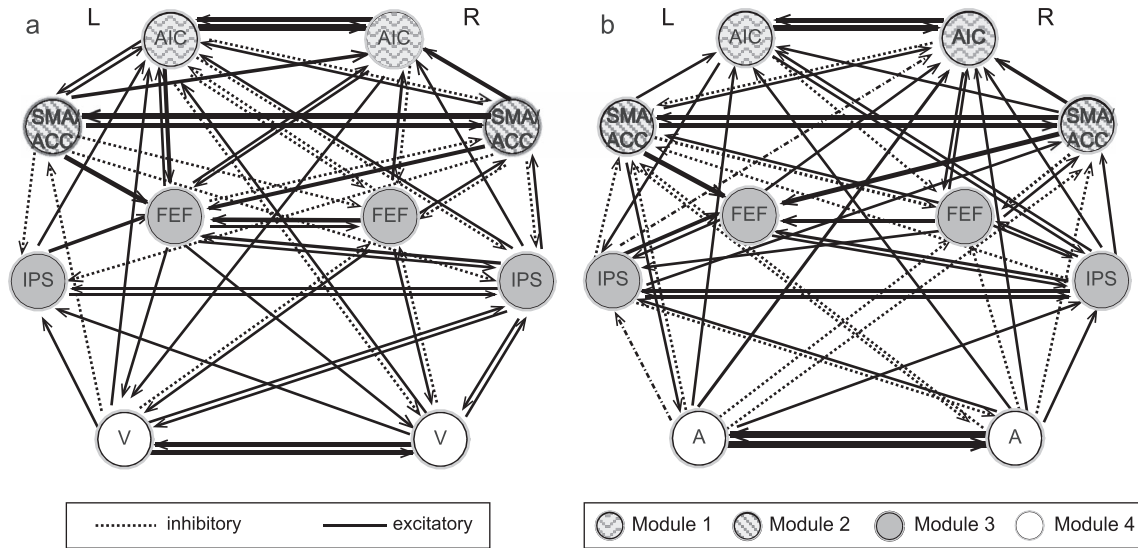


Figure 5. Intrinsic connectivity across regions of the CCN and modality specific sensory regions. (a) Connectivity for the SMFT-V. (b) Connectivity for the SMFT-A. V, visual areas; A, auditory areas; L, left; R, right. The strength of each connection is indicated by the thickness of the arrow.

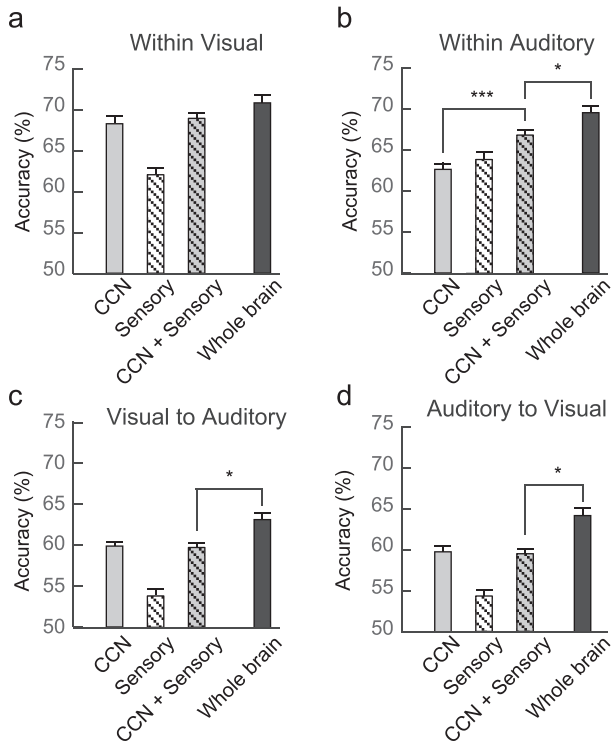


Figure 6. Within- and cross-modality prediction accuracy. (a) Within-visual prediction. (b) Within-auditory prediction. (c) Visual-to-auditory prediction. (d) Auditory-to-visual prediction. Error bar indicates 95% confidence interval in the within-subject design.

framework to integrate multiple information streams into modality-independent and domain-general representations of information uncertainty.

The results of both the GLM and the ROI analyses of the cognitive search strategy employed in the visual and auditory tasks provide clues regarding the nature of the abstract information

represented in regions of the CCN. The monotonic increases in task-related activation in cortical regions of the CCN were best explained by a linear model of increasing information entropy consistent with a self-terminating search strategy in both visual and auditory modalities. Each bit increase in information entropy or average content processed by cognitive control required an increase of one constant unit of neural resources that was quantifiable in the hemodynamic responses (Borst and Theunissen 1999; Averbeck et al. 2006; Fan 2014; Fan et al. 2014; Wu et al. 2018). Thus, the monotonic increase in regional activation in the CCN represents higher-order abstract coding of the linear increase in information uncertainty across the different task conditions rather than the lower-level coding of the sensory and perceptual properties of the stimuli that were immutable across trials.

The findings provide no evidence that the contributions of the CCN to cognitive control in the visual and auditory tasks were related to potential motor effects. The possibility that an increase in CCN activation may reflect the enhanced motor programming demands that usually accompany an increase in cognitive control is an unresolved issue in the field. The tasks in this study were designed with a delayed response with no response for half the trials, instead of choice reaction time mapping, to temporally segregate activation related to cognitive control and motor response generation. Activation in regions of the CCN for uncertainty processing during the stimulus sequence also did not significantly differ between matching trials that required a motor response and nonmatching trials that required no response. Motor programming and preparation did not vary across trials and therefore could not account for the monotonic increases in CCN activation.

Connectivity that Underlies Cross-protocol Representation of Uncertainty in the CCN

The cross-protocol representation of uncertainty in the dynamic coupling between regions of the CCN across cognitive control load levels is supported by a stable intrinsic architecture that is independent of sensory modality. Direct causal modeling

revealed that uncertainty is represented at the network level in protocols constraining a core neural architecture in the CCN that has comparable functional couplings with key visual and auditory sensory nodes. The current results demonstrate that regions of the CCN form a stable network structure that receives modality-specific inputs and further processes the information through the intrinsic connectivity of the network as proposed in previous studies (Sridharan et al. 2008; Menon and Uddin 2010; Cocchi et al. 2013; Power et al. 2013; Nee and D'Esposito 2016; Wu et al. 2018).

At the group level, the shared community structure of the CCN for the visual and auditory tasks confirms the stable intrinsic structure of the network. This common network structure can explain approximately 35–40% of the total network similarity effect and is consistent with previous studies demonstrating core network configurations that are independent of cognitive tasks (Krienen et al. 2014; Ito et al. 2017; Dixon et al. 2018). The similar intrinsic connectivity of the CCN across the two tasks suggests that the CCN functions independently of input modality. In addition, the core intrinsic structure of the CCN was similarly coupled with the auditory and visual sensory regions. This network architecture may enable the CCN to efficiently integrate and coordinate information from multiple modalities and domains (Spagna et al. 2015; Spagna, He, et al. 2018a; Wu, Chen, et al. 2019a).

Cross-modal Representation of Uncertainty in the CCN

The results of the cross-decoding analyses demonstrate that the supramodal function of cognitive control is also supported by cross-modal representations of uncertainty in regions of the CCN. The results of the MVPA corroborate the finding that information uncertainty in visual and auditory inputs is encoded in patterns of activation in regions of the CCN that predicted cognitive control load (i.e., level of information uncertainty) both within and across sensory modalities at a level that was significantly higher than chance. The finding of significant cross-modal prediction of cognitive load suggests that cortical regions of the CCN encode a unified neural representation of information uncertainty that is independent of the input modality and is generalizable across sensory modalities. The greater accuracy of the within-modal predictions compared to cross-modal predictions of cognitive control load indicates that the CCN also encodes a proportion of the modality-specific information.

The higher-order cross-modal representations of uncertainty encoded in the CCN incorporate the lower-level modality-specific information represented in sensory regions. Although the within-modality predictions based on the sensory feature set were significantly higher than chance level, combining the sensory and CCN features did not improve the prediction compared to the CCN feature alone. It suggests that patterns of activation in the CCN may integrate multiple lower-level modality-specific information streams into a unified cross-modal representation of uncertainty and that patterns of activation within sensory areas may partially mirror activation of the CCN via feedback modulation of top-down control.

The spatial distribution of weight vectors that represent the importance of each voxel in the predictions provides further evidence of how uncertainty is encoded by the CCN and other networks. Group-level analysis of the weight vectors for the whole-brain predictions based on the combined CCN and sensory features found spatially similar activation patterns across

participants confirming the contribution of these common regions to the representation of uncertainty. However, the large intersubject variability in the weight vectors for predictions based on the CCN feature set suggests that the fine-grained representation of uncertainty in regions of the CCN is unique to individual participants, which is consistent with previous findings of highly variable microstructure in regions of the CCN (e.g., ACC and AIC) across individuals (Menon et al. 2020). Finally, the slight but significant increase (3%) in the accuracy of whole-brain predictions compared to the combined CCN and sensory features may be attributable to regions in the default mode network (Raichle and Snyder 2007), which are anticorrelated with the CCN and may mirror supramodal representations of uncertainty in the CCN (Wu et al. 2018) with the effect of enhancing reliability of the brain in predicting behavior.

The CCN as a Unified Information Processing Entity

The role of the CCN in the processing and representation of information uncertainty may extend beyond the visual and auditory modalities tested in this study. This network has been implicated in the processing of information uncertainty in other sensory modalities (e.g., somatosensory modality) (Macaluso et al. 2002; Wang et al. 2019) and for other cognitive domains including working memory and response selection and inhibition (Lewis et al. 2000; Zhang and Li 2012; Wu et al. 2018). In this capacity, the CCN integrates a vast amount of information from multiple input modalities into higher-order abstract representations of uncertainty that are subjectively experienced as “thoughts,” “feelings,” and “awareness” (Craig 2009; Singer et al. 2009; Brass and Haggard 2010; Nelson et al. 2010; Craig 2011; Gu et al. 2013). It is suggested that the CCN is a unified information processing entity in the brain that is responsible for higher-order uncertainty processing independent of sensory modalities and cognitive domains (Fan 2014; Spagna, Kim, et al. 2018b).

The architecture of a single unified supramodal network for cognitive control rather than multiple modality-specific centers reflects an economic trade-off between the cost of limited capacity and the benefit of adaptively and efficiently coordination of information received from multiple modalities. Most regions of the CCN belong to a “rich club” of highly interconnected connector hubs that participate in multiple modality- and domain-specific networks in the brain (van den Heuvel and Sporns 2011; van den Heuvel et al. 2012). This organization may support efficient intranetwork information transfer and the adaptive reconfiguration of connectivity between the CCN and other networks to implement cognitive control. However, the heavy computational load imposed by the participation in multiple cognitive processes may lead to a bottleneck in these CCN regions, such as the AIC, which limits the capacity of cognitive control (Wu, Wang, et al. 2019b).

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Author Contributions

J.F. and T.W. designed the experiments. T.W. analyzed the data. All authors discussed the results and contributed to writing up the report.

Funding

National Institute of Mental Health of the National Institutes of Health grant number R01 MH094305. National Science Foundation (NSF) (grant NSF IIS 1855759 to J.F., C.C., and T.W.).

Notes

We thank Dr Michael I. Posner for his consultation on this project. We also thank Dr Thomas Lee for participating in discussions regarding the design and pilot testing of the task; Mr Alexander J. Dufford for his help in the collection of fMRI data; and Dr Laura Egan for her help with the description of the tasks. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of Interest: The authors declare no competing financial interests.

References

- Alizadeh S, Jamalabadi H, Schonauer M, Leibold C, Gais S. 2017. Decoding cognitive concepts from neuroimaging data using multivariate pattern analysis. *Neuroimage*. 159:449–458.
- Averbeck BB, Latham PE, Pouget A. 2006. Neural correlations, population coding and computation. *Nat Rev Neurosci*. 7:358–366.
- Baddeley AD. 2018. The concept of working memory: a view of its current state and probable future development. In: Baddeley AD, editor. *Exploring working memory: Selected works from Alan Baddeley*. London: Routledge/Taylor & Francis Group, pp. 99–106.
- Borst A, Theunissen FE. 1999. Information theory and neural coding. *Nat Neurosci*. 2:947–957.
- Brass M, Haggard P. 2010. The hidden side of intentional action: the role of the anterior insular cortex. *Brain Struct Funct*. 214:603–610.
- Chen Y, Spagna A, Wu T, Kim TH, Wu Q, Chen C, Wu Y, Fan J. 2019. Testing a cognitive control model of human intelligence. *Sci Rep*. 9:2898.
- Choi JM, Padmala S, Pessoa L. 2012. Impact of state anxiety on the interaction between threat monitoring and cognition. *Neuroimage*. 59:1912–1923.
- Cocchi L, Zalesky A, Fornito A, Mattingley JB. 2013. Dynamic cooperation and competition between brain systems during cognitive control. *Trends Cogn Sci*. 17:493–501.
- Colby CL, Goldberg ME. 1999. Space and attention in parietal cortex. *Annu Rev Neurosci*. 22:319–349.
- Corbetta M. 1998. Frontoparietal cortical networks for directing attention and the eye to visual locations: identical, independent, or overlapping neural systems? *Proc Natl Acad Sci U S A*. 95:831–838.
- Corradi-Dell'Acqua C, Tusche A, Vuilleumier P, Singer T. 2016. Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nat Commun*. 7:10904.
- Costa-Faidella J, Sussman ES, Escera C. 2017. Selective entrainment of brain oscillations drives auditory perceptual organization. *Neuroimage*. 159:195–206.
- Cousineau D. 2005. Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutor Quant Methods Psychol*. 1:42–45.
- Craig AD. 2009. How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci*. 10:59–70.
- Craig AD. 2011. Significance of the insula for the evolution of human awareness of feelings from the body. *Ann N Y Acad Sci*. 1225:72–82.
- Dixon ML, De La Vega A, Mills C, Andrews-Hanna J, Spreng RN, Cole MW, Christoff K. 2018. Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc Natl Acad Sci U S A*. 115:E1598–E1607.
- Donohue SE, Liotti M, Perez R, Woldorff MG. 2012. Is conflict monitoring supramodal? Spatiotemporal dynamics of cognitive control processes in an auditory Stroop task. *Cogn Affect Behav Neurosci*. 12:1–15.
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. *Trends Cogn Sci*. 12:99–105.
- Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME. 2007. Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci U S A*. 104:11073–11078.
- Doyle JC, Cséte M. 2011. Architecture, constraints, and behavior. *Proc Natl Acad Sci U S A*. 108:15624–15630.
- Eriksen CW, Yeh YY. 1985. Allocation of attention in the visual field. *J Exp Psychol Hum Percept Perform*. 11:583.
- Fairhall SL, Caramazza A. 2013. Brain regions that represent amodal conceptual knowledge. *J Neurosci*. 33:10552–10558.
- Fan J. 2014. An information theory account of cognitive control. *Front Hum Neurosci*. 8:680.
- Fan J, Guise KG, Liu X, Wang H. 2008. Searching for the majority: algorithms of voluntary control. *PLoS One*. 3:e3522.
- Fan J, Van Dam NT, Gu X, Liu X, Wang H, Tang CY, Hof PR. 2014. Quantitative characterization of functional anatomical contributions to cognitive control under uncertainty. *J Cogn Neurosci*. 26:1490–1506.
- Farah MJ, Wong AB, Monheit MA, Morrow LA. 1989. Parietal lobe mechanisms of spatial attention - modality-specific or supramodal. *Neuropsychologia*. 27:461–470.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg M, Turner R. 1998. Event-related fMRI: characterizing differential responses. *Neuroimage*. 7:30–40.
- Friston KJ, Kahan J, Biswal B, Razi A. 2014. A DCM for resting state fMRI. *Neuroimage*. 94:396–407.
- Friston KJ, Litvak V, Oswal A, Razi A, Stephan KE, van Wijk BCM, Ziegler G, Zeidman P. 2016. Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage*. 128:413–431.
- Green JJ, Doesburg SM, Ward LM, McDonald JJ. 2011. Electrical neuroimaging of voluntary audiospatial attention: evidence for a supramodal attention control network. *J Neurosci*. 31:3560–3564.
- Gu X, Hof PR, Friston KJ, Fan J. 2013. Anterior insular cortex and emotional awareness. *J Comp Neurol*. 521:3371–3388.
- Handjaras G, Ricciardi E, Leo A, Lenci A, Cecchetti L, Cosottini M, Marotta G, Pietrini P. 2016. How concepts are encoded in the human brain: a modality independent, category-based cortical organization of semantic knowledge. *Neuroimage*. 135:232–242.
- Ito T, Kulkarni KR, Schultz DH, Mill RD, Chen RH, Solomyak LI, MiWC. 2017. Cognitive task information is transferred between brain regions via resting-state network topology. *Nat Commun*. 8:1027.

- Joachims T. 2002. Optimizing search engines using click-through data. Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; July 2002. Edmonton Alberta Canada: ACM. 133–142.
- Joachims T. 2006. Training linear SVMs in linear time. In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; August 2006. Philadelphia PA USA: ACM, pp. 217–226.
- Jung YL, Larsen B, Walther DB. 2018. Modality-independent coding of scene categories in prefrontal cortex. *J Neurosci*. 38:5969–5981.
- Kaplan JT, Man K, Greening SG. 2015. Multivariate cross-classification: applying machine learning techniques to characterize abstraction in neural representations. *Front Hum Neurosci*. 9:151.
- Kinnison J, Padmala S, Choi J-M, Pessoa L. 2012. Network analysis reveals increased integration during emotional and motivational processing. *J Neurosci*. 32:8361–8372.
- Knight RT. 2007. Neural networks debunk phrenology. *Science*. 316:1578.
- Kong L, Michalka SW, Rosen ML, Sheremata SL, Swisher JD, Shinn-Cunningham BG, Somers DC. 2014. Auditory spatial attention representations in the human cerebral cortex. *Cereb Cortex*. 24:773–784.
- Koziol LF. 2014. Cognitive control, reward, and the basal ganglia. In: *The myth of executive functioning*. 1st ed. Cham (Switzerland): Springer International Publishing, pp. 61–64.
- Krienen FM, Yeo BTT, Buckner RL. 2014. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos Trans R Soc Lond B Biol Sci*. 369:20130526.
- Laurens KR, Kiehl KA, Liddle PF. 2005. A supramodal limbic-paralimbic-neocortical network supports goal-directed stimulus processing. *Hum Brain Mapp*. 24:35–49.
- Lee AKC, Larson E, Maddox RK, Shinn-Cunningham BG. 2014. Using neuroimaging to understand the cortical mechanisms of auditory selective attention. *Hear Res*. 307:111–120.
- Lewis JW, Beauchamp MS, DeYoe EA. 2000. A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex*. 10:873–888.
- Macaluso E, Frith CD, Driver J. 2002. Supramodal effects of covert spatial orienting triggered by visual or tactile events. *J Cogn Neurosci*. 14:389–401.
- Mayer AR, Ryman S, Hanlon FM, Dodd AB, Ling JM. 2017. Look hear! The prefrontal cortex is stratified by modality of sensory input during multisensory cognitive control. *Cereb Cortex*. 27:2831–2840.
- Menon V, Guillermo G, Pinsk MA, Nguyen V-D, Li J-R, Cai W, Wassermann D. 2020. Quantitative modeling links in vivo microstructural and macrofunctional organization of human and macaque insular cortex, and predicts cognitive control abilities. *eLife*. 9:e53470.
- Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*. 214:655–667.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn Psychol*. 41:49–100.
- Nee DE, D’Esposito M. 2016. The hierarchical organization of the lateral prefrontal cortex. *Elife*. 5:e12112.
- Nelson SM, Dosenbach NU, Cohen AL, Wheeler ME, Schlaggar BL, Petersen SE. 2010. Role of the anterior insula in task-level control and focal attention. *Brain Struct Funct*. 214:669–680.
- Niendam TA, Laird AR, Ray KL, Dean YM, Glahn DC, Carter CS. 2012. Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn Affect Behav Neurosci*. 12:241–268.
- Pessoa L, Kastner S, Ungerleider LG. 2003. Neuroimaging studies of attention: from modulation of sensory processing to top-down control. *J Neurosci*. 23:3990–3998.
- Power JD, Schlaggar BL, Lessov-Schlaggar CN, Petersen SE. 2013. Evidence for hubs in human functional brain networks. *Neuron*. 79:798–813.
- Raichle ME, Snyder AZ. 2007. A default mode of brain function: a brief history of an evolving idea. *Neuroimage*. 37:1083–1090.
- Razi A, Seghier ML, Zhou Y, McColgan P, Zeidman P, Park H-J, Sporns O, Rees G, Friston KJ. 2017. Large-scale DCMs for resting-state fMRI. *Network Neuroscience*. 1:222–241.
- Reynolds JH, Chelazzi L. 2004. Attentional modulation of visual processing. *Annu Rev Neurosci*. 27:611–647.
- Rissman J, Gazzaley A, D’Esposito M. 2004. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage*. 23:752–763.
- Roberts KL, Hall DA. 2008. Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory stroop tasks. *J Cogn Neurosci*. 20:1063–1078.
- Rossi AF, Pessoa L, Desimone R, Ungerleider LG. 2009. The prefrontal cortex and the executive control of attention. *Exp Brain Res*. 192:489–497.
- Rubinov M, Sporns O. 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. 52:1059–1069.
- Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. 2009. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Res*. 1286:155–164.
- Silver MA, Kastner S. 2009. Topographic maps in human frontal and parietal cortex. *Trends Cogn Sci*. 13:488–495.
- Simanova I, Hagoort P, Oostenveld R, van Gerven MA. 2014. Modality-independent decoding of semantic information from the human brain. *Cereb Cortex*. 24:426–434.
- Singer T, Critchley HD, Preuschoff K. 2009. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci*. 13:334–340.
- Spagna A, He G, Jin S, Gao L, Mackie M-A, Tian Y, Wang K, Fan J. 2018a. Deficit of supramodal executive control of attention in schizophrenia. *J Psychiatr Res*. 97:22–29.
- Spagna A, Kim TH, Wu T, Fan J. 2018b. Right hemisphere superiority for executive control of attention. *Cortex*. 122:263–276.
- Spagna A, Mackie M-A, Fan J. 2015. Supramodal executive control of attention. *Front Psychol*. 6:65.
- Sridharan D, Levitin DJ, Menon V. 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *P Natl Acad Sci USA*. 105:12569–12574.
- Stroop JR. 1935. Studies of interference in serial verbal reactions. *J Exp Psychol*. 18:643.
- van den Heuvel MP, Kahn RS, Goni J, Sporns O. 2012. High-cost, high-capacity backbone for global brain communication. *Proc Natl Acad Sci U S A*. 109:11372–11377.
- van den Heuvel MP, Sporns O. 2011. Rich-club organization of the human connectome. *J Neurosci*. 31:15775–15786.

- Walther S, Goya-Maldonado R, Stippich C, Weisbrod M, Kaiser S. 2010. A supramodal network for response inhibition. *Neuroreport*. 21:191–195.
- Walz JM, Goldman RI, Carapezza M, Muraskin J, Brown TR, Sajda P. 2013. Simultaneous EEG-fMRI reveals temporal evolution of coupling between supramodal cortical attention networks and the brainstem. *J Neurosci*. 33:19212–19222.
- Wang X, Wu Q, Egan LJ, Gu X, Liu P, Gu H, Yang Y, Luo J, Wu Y, Gao Z, et al. 2019. Anterior insular cortex plays a critical role in interoceptive attention. *Elife*. 8:e42265.
- Westerhausen R, Moosmann M, Alho K, Belsby SO, Hamalainen H, Medvedev S, Specht K, Hugdahl K. 2010. Identification of attention and cognitive control networks in a parametric auditory fMRI study. *Neuropsychologia*. 48:2075–2081.
- Wu T, Chen C, Spagna A, Mackie M-A, Russell-Giller S, Xu P, Y-j L, Liu X, Hof PR, Fan J. 2019a. The functional anatomy of cognitive control: a domain-general brain network for uncertainty processing. *J Comp Neurol*. 528:1265–1292.
- Wu T, Dufford AJ, Egan LJ, Mackie M-A, Chen C, Yuan C, Chen C, Li X, Liu X, Hof PR, et al. 2018. Hick–Hyman law is mediated by the cognitive control network in the brain. *Cereb Cortex*. 28:2267–2282.
- Wu T, Wang X, Wu Q, Spagna A, Yang J, Yuan C, Wu Y, Gao Z, Hof PR, Fan J. 2019b. Anterior insular cortex is a bottleneck of cognitive control. *Neuroimage*. 195:490–504.
- Wurm MF, Ariani G, Greenlee MW, Lingnau A. 2016. Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb Cortex*. 26:3390–3401.
- Zhang S, Li CR. 2012. Functional networks for cognitive control in a stop signal task: independent component analysis. *Hum Brain Mapp*. 33:89–104.