

Published in final edited form as:

Eur J Neurosci. 2015 January ; 41(2): 264–274. doi:10.1111/ejn.12764.

Role of the anterior insular cortex in integrative causal signaling during multisensory auditory–visual attention

Tianwen Chen^{1,*}, Lars Michels^{1,4,*}, Kaustubh Supekar¹, John Kochalka¹, Srikanth Ryali¹, and Vinod Menon^{1,2,3}

¹Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine Stanford, CA 94305, USA ²Program in Neuroscience, Stanford University School of Medicine Stanford, CA 94305, USA ³Department of Neurology & Neurological Sciences, Stanford University School of Medicine Stanford, CA 94305, USA ⁴Institute of Radiology University Hospital Zurich Zurich, Switzerland

Abstract

Coordinated attention to information from multiple senses is fundamental to our ability to respond to salient environmental events, yet little is known about brain network mechanisms that guide integration of information from multiple senses. Here we investigate dynamic causal mechanisms underlying multisensory auditory–visual attention, focusing on a network of right-hemisphere frontal–cingulate–parietal regions implicated in a wide range of tasks involving attention and cognitive control. Participants performed three ‘oddball’ attention tasks involving auditory, visual and multisensory auditory–visual stimuli during fMRI scanning. We found that the right anterior insula (rAI) demonstrated the most significant causal influences on all other frontal–cingulate–parietal regions, serving as a major causal control hub during multisensory attention. Crucially, we then tested two competing models of the role of the rAI in multisensory attention: an ‘integrated’ signaling model in which the rAI generates a common multisensory control signal associated with simultaneous attention to auditory and visual oddball stimuli versus a ‘segregated’ signaling model in which the rAI generates two segregated and independent signals in each sensory modality. We found strong support for the integrated, rather than the segregated, signaling model. Furthermore, the strength of the integrated control signal from the rAI was most pronounced on the dorsal anterior cingulate and posterior parietal cortices, two key nodes of saliency and central executive networks respectively. These results were preserved with the addition of a superior temporal sulcus region involved in multisensory processing. Our study provides new insights into the dynamic causal mechanisms by which the AI facilitates multisensory attention.

Keywords

oddball; multisensory attention; anterior insular cortex; multivariate dynamic system model; central executive network; salience network

Correspondence: Tianwen Chen and Vinod Menon, 401 Quarry Road, Stanford, CA 94305. tianwenc@stanford.edu (T.C.) and menon@stanford.edu (V.M.).

*Equal contribution

The authors declare no conflict of interest.

Introduction

Coordinated attention to salient events in our environment requires integration of information from multiple senses. Most studies to date have focused on localization of brain activation in response to multisensory stimuli and have implicated distributed brain areas spanning prefrontal, parietal and temporal cortices (Calvert et al., 2000; Bushara et al., 2001; Calvert et al., 2001; Macaluso et al., 2004; Teder-Salejarvi et al., 2005; Noesselt et al., 2007; Cappe et al., 2010). However, little is known about brain network mechanisms that guide integration of information from multiple senses (Molholm & Foxe, 2010).

A candidate brain region for facilitating coordinated attention from multiple modalities is the anterior insula (AI). The AI receives convergent input from multiple sensory modalities including auditory and visual systems (Mesulam & Mufson, 1982; Augustine, 1996; Bamiou et al., 2003; Butti & Hof, 2010; Nieuwenhuys, 2012). Of particular interest is the right-hemisphere (r)AI, which is consistently activated during a wide range of unisensory auditory and visual attention tasks (Crottaz-Herbette & Menon, 2006; Eckert et al., 2009; Sterzer & Kleinschmidt, 2010). Nevertheless, it is unknown whether the rAI plays an active role in integration of information during attention to stimuli from multiple senses.

Here we investigate causal network interactions underlying multisensory auditory–visual attention in the context of core right-hemisphere frontal–cingulate–parietal regions that have been implicated in a wide range of unisensory attention tasks (Corbetta & Shulman, 2002; Dosenbach et al., 2007; Dosenbach et al., 2008; Sridharan et al., 2008; Menon & Uddin, 2010; Supekar & Menon, 2012). These regions include the AI, ventrolateral prefrontal cortex (VLPFC), dorsal anterior cingulate cortex (dACC), dorsolateral prefrontal cortex (DLPFC) and posterior parietal cortex (PPC), key nodes of the salience network (SN) and the central executive network (CEN) (Seeley et al., 2007; Menon & Uddin, 2010). We used an oddball attention paradigm and three tasks, two of which involved detecting deviants in either the auditory or visual modalities (Crottaz-Herbette & Menon, 2006; Sridharan et al., 2008), and a third which utilized a multisensory task involving attention to simultaneously presented auditory–visual deviants (Figure 1). Causal interactions between frontal–cingulate–parietal nodes of the SN and CEN during multisensory attention were examined using multivariate dynamic systems analysis (Ryali et al., 2011; Supekar & Menon, 2012; Ryali et al., Under Review).

Based on previous findings that the rAI plays a causal role in attention to salient unisensory stimuli (Sridharan et al., 2008), and on converging evidence for its involvement in multisensory attention (Bushara et al., 2001; Bushara et al., 2003), we hypothesized that this region integrates control signals during multisensory attention. To test this we compared the strength of causal interactions to multisensory and unisensory ‘oddball’ stimuli under two competing hypotheses based on ‘integrated’ and ‘segregated’ signaling (Figure 2). We show that the rAI generates an integrated auditory–visual control signal, providing novel evidence that it is an integrative zone for multisensory attention. Furthermore, these results are preserved even after including an additional posterior superior temporal sulcus region implicated in multisensory processing (Baum et al., 2012; Noesselt et al., 2012).

Materials and methods

Participants

Eighteen right-handed individuals, all students at Stanford University, participated in the study. Data from three subjects were excluded: two subjects did not make responses in > 60% of trials with deviants in at least one oddball task, and one subject did not have complete data for all the three oddball tasks. The remaining 15 subjects were included for further analysis (nine men and six women, aged 18 – 33 years; mean \pm SD, 23.1 \pm 5.2 years; 15 right-handed). Handedness was assessed with the Edinburgh question naire (Oldfield, 1971). None of the participants showed any signs of a neurological disorder and they all had a normal hearing and normal or corrected-to-normal vision. Subjects refrained from caffeine, nicotine and alcohol for at least 12 hours prior to the recording. The study was approved by the Stanford University School of Medicine Human Subjects committee. All participants were recruited via advertisements on the campus of Stanford University. The experiments were undertaken with the understanding and written consent of each subject, and the study conformed to the 2013 World Medical Association Declaration of Helsinki.

Experimental design

The experimental design is shown in Figure 1. Each participant performed three different oddball tasks during fMRI scanning, two involving unimodal visual and auditory stimuli, respectively, and a third involving multisensory auditory–visual stimuli. The design and structure of all three tasks was similar with 160 ‘standards’ and 40 ‘deviants’ presented in randomized order. In the visual oddball task (\sim 7 min), blue and green circles of the same size were presented in the center of the screen. For half of the participants, blue circles were deviants, for the other half of participants green circles were deviants. Participants were instructed to respond as fast and accurately as possible after stimulus presentation. Half of the participants indexed their responses by a button press with their index finger in response to deviants and by a button press with middle finger in response to standards. The other half of the subjects responded in the opposite fashion. The auditory oddball task used an identical design except that standards and deviants consisted of low-frequency (1000 Hz) and high-frequency (2000 Hz) tones. Participants wore custom-built headphones designed to reduce the background scanner noise to \sim 80 dB. In the multisensory oddball task, stimuli consisted of simultaneous colored circles and binaural tones using the same stimuli as those used in the unimodal tasks. The assignment of deviants and standards was counterbalanced across participants, i.e., for one group a high tone was paired with a yellow circle as deviant, while for the other group a low tone was combined with a red circle. The choice of simultaneous presentations was similar to designs used in previous studies of multisensory integration (Teder-Salejarvi et al., 2005; Cappe et al., 2010). To ensure that participants attended to both stimulus modalities, they were given no prior information about the combination of auditory and visual stimuli used in the task. In all three tasks, stimuli were presented for 200 ms and the inter-trial period (blank screen) was 1800 ms.

To further investigate the behavioral aspects of simultaneous attention to auditory and visual stimuli, participants performed a fourth multisensory oddball task involving non-simultaneous stimulus presentation. The design of this task was similar to the other three

except that visual and auditory stimuli were presented non-simultaneously with a gap of 100 ms between them (Bushara et al., 2001). Specifically, in half of the trials the visual stimulus was presented 100 ms prior to auditory stimulus; in other half of the trials, the visual stimulus were presented 100 ms after the auditory stimulus (Bushara et al., 2001). Participants performed two runs of the same non-simultaneous multisensory task, and were given exactly the same instructions as in the simultaneous multisensory oddball tasks described above. All four tasks were randomized across participants but, due to a programming error, data from four of the 15 participants could not be used. Data from the remaining 11 participants were used here. Our hypothesis was that if participants were attending to only one sensory modality instead of both, their reaction time in the non-simultaneous multisensory task would be similar to those in the unisensory as well as the simultaneous multisensory tasks. On the other hand, if participants attended to both visual and auditory stimuli, their response would be delayed by 100 ms, the same as the gap between the non-simultaneously presented stimuli.

Prior to the fMRI experiment, participants performed a brief trial scanning session to make certain that they could differentiate between the two tones and that they were able to hear the tones binaurally. Loudness levels were calibrated for each ear individually by having participants respond using a button press. Stimulus delivery, synchronization of stimulus presentation with fMRI scanning and response coding were controlled by E-Prime (Psychology Software Tools, www.pstnet.com).

MRI data acquisition

In order to minimize head movement, the head of each participant was secured using sponge pads. Functional images were recorded on a 3-Tesla GE Signa scanner using a T2*-weighted gradient spiral-in and spiral-out pulse sequence (Glover & Lai, 1998; Glover & Law, 2001). To reduce blurring and signal loss arising from field nonhomogeneities, an automated high-order shimming method based on spiral acquisitions was used before acquiring functional MRI scans. A total of 405 volumes were acquired from each subject. Each volume covered the whole brain and consisted of 28 axial slices (parallel to the anterior and posterior commissure) with a matrix size of 64×64 with an effective in-plane spatial resolution of $3.125 \times 3.125 \text{ mm}^2$ and a slice thickness of 4 mm with a 1 mm skip. The following acquisition parameters were used: time of repetition (TR), 2000 ms; time of echo (TE), 30 ms; flip angle (FA), 90° ; number of slices, 28; field of view, 200 mm; interleave, 1. During the same imaging session, a T1-weighted spoiled grass gradient recalled inversion recovery 3-D MRI sequence was acquired with the following parameters: TR, 35 ms; TE, 6 ms; FA, 45° slices in coronal plane, 124; matrix, 256×192 .

fMRI preprocessing

Spiral-in and spiral-out data were combined by using a weighted average of the two images, slice by slice. Weighting between the images for spiral-in and spiral-out acquisitions was determined by the intensities of the average image so that, in regions where the spiral-out average image had a lower intensity, the resultant image was weighted toward the spiral-in image and *vice versa*. In uniform regions, the combination reverts to a simple average of spiral-in and spiral-out images (Glover & Law, 2001). Images were reconstructed, by

inverse Fourier transforms, for each of the 405 time points into $64 \times 64 \times 28$ -image matrices.

fMRI data was preprocessed using SPM8 analysis software (<http://www.fil.ion.ucl.ac.uk/spm>). Images were realignment-corrected to correct for head motion, corrected for errors in slice-timing, spatially transformed to standard stereotaxic space [based on the Montreal Neurologic Institute (MNI) coordinate system], resampled every 2 mm using sinc interpolation, and smoothed with a 6-mm full-width half-maximum Gaussian kernel to decrease spatial noise prior to statistical analysis. Translational movement in millimeters (x , y and z) and rotational motion in degrees (pitch, roll and yaw) was calculated based on the SPM8 parameters for motion correction of the functional images in each subject. Across 15 subjects, mean ranges of x , y and z translation across all tasks were 0.29 ± 0.15 , 0.30 ± 0.15 and 0.78 ± 0.66 mm, and mean ranges of x , y and z rotation across all tasks were 0.01 ± 0.01 , 0.006 ± 0.003 and 0.005 ± 0.002 radians. The first five volumes (blank screen) were discarded (10 s) to minimize susceptibility artifacts and to allow for an equilibrium state.

fMRI analysis

Brain activations related to deviant stimuli were estimated for all three tasks. For each task and subject, a general linear model was used (Friston et al., 1995), which included regressors of interest for deviant stimuli and nuisance regressors for head motions. Both canonical hemodynamic response function (HRF) and its time-derivative were used to convolve the stimulus function to form the regressors for deviants. The significant activation patterns were determined using a voxel-wise height threshold of $P < 0.01$ and an extent threshold of $P < 0.01$ with family-wise error correction using a nonstationary suprathreshold cluster-size approach based on Monte-Carlo simulations (Nichols & Hayasaka, 2003). We also performed a multivoxel analysis on brain activation patterns in the three oddball tasks to show that multisensory and unisensory deviants were processed differently (See Multivariate Pattern Analysis in the Supplementary Materials).

Network nodes

Nodes were identified using an unbiased approach similar to our previous studies (Uddin et al., 2011; Supekar & Menon, 2012). The two main networks of interest, SN and CEN, were identified using independent component analysis (ICA) applied to resting-state fMRI data from a different group of participants (Supekar & Menon, 2012). ICA is a model-free, data-driven approach and has the flexibility to identify various independent spatial patterns and their associated temporal fluctuations (Beckmann et al., 2005). From the SN ICA maps we identified nodes in right AI, dACC and VLPFC. From the CEN ICA maps we identified nodes in right DLPFC and PPC. The anatomical location of these nodes is shown in Figure 3A as well as Table 1. All subsequent analyses were based on these canonical nodes of the SN and CEN.

Because of previous studies implicating the posterior superior temporal sulcus (STS) in multisensory processing (Menon, 2011; Gogolla et al., 2014), we conducted additional analyses including this region. A right STS node was determined based on activations during the simultaneous multisensory task. We examined the hypothesis that the rAI remains

a causal outflow hub, and the major region that generates integrated control signals, even when this cross-modal region is included in the model.

Finally, to demonstrate the specificity of effects in the rAI, we also conducted a parallel analysis using left hemisphere nodes (Supplementary Table S1) using the same selection procedures described above.

Multivariate dynamic systems (MDS) model

MDS is a state-space model for estimating causal interactions from fMRI data (Ryali et al., 2011). MDS estimates context-dependent causal interactions between multiple brain regions in latent quasi-neuronal state while accounting for variations in hemodynamic responses in these regions. MDS has been validated using extensive simulations (Ryali et al., 2011; Ryali et al., Under Review) and has been successfully applied to our previous studies (Cho et al., 2012; Supekar & Menon, 2012).

MDS models the multivariate fMRI time series by the following state-space equations:

$$\mathbf{s}(t) = \sum_{j=1}^J v_j(t) C_j \mathbf{s}(t-1) + \mathbf{w}(t) \quad (1)$$

$$\mathbf{x}_m(t) = [\mathbf{s}_m(t) \mathbf{s}_m(t-1) \dots \mathbf{s}_m(t-L+1)]' \quad (2)$$

$$y_m(t) = b_m \Phi \mathbf{x}_m(t) + \mathbf{e}_m(t) \quad (3)$$

In Equation (1), $\mathbf{s}(t)$ is a $M \times 1$ vector of latent quasi-neuronal signals at time t of M regions, A is an $M \times M$ connection matrix where in C_j is an $M \times M$ connection matrix ensued by modulatory input $v_j(t)$, and J is the number of modulatory inputs. The non-diagonal elements of C_j represent the coupling of brain regions in the presence of modulatory input $v_j(t)$. $C_j(m, n)$ denotes the strength of causal interactions from n -th region to m -th region for j -th type stimulus. A higher $C_j(m, n)$ indicates higher causal influences from region n to region m . Therefore, latent signals $\mathbf{s}(t)$ in M regions at time t is a bilinear function of modulatory inputs $v_j(t)$, corresponding to deviant or standard stimulus, and its previous state $\mathbf{s}(t-1)$. $\mathbf{w}(t)$ is an $M \times 1$ state noise vector whose distribution is assumed to be Gaussian distributed with covariance matrix $Q(\mathbf{w}(t) \sim N(0, Q))$. Additionally, state noise vector at time instances $1, 2, \dots, T$ ($\mathbf{w}(1), \mathbf{w}(2) \dots \mathbf{w}(T)$) are assumed to be identical and independently distributed (i.i.d.). Equation (1) represents the time evolution of latent signals in M brain regions. More specifically, the latent signals at time t , $\mathbf{s}(t)$, is expressed as a linear combination of latent signals at time $t-1$, external stimulus at time t ($u(t)$), bilinear combination of modulatory inputs $v_j(t)$, $j = 1, 2, \dots, J$ and its previous state, and state noise $\mathbf{w}(t)$. The latent dynamics modeled in Equation (1) gives rise to observed fMRI time series represented by Equations (2) and (3).

We model the fMRI time series in region m as a linear convolution of HRF and latent signal $\mathbf{s}_m(t)$ in that region. To represent this linear convolution model as an inner product of two

vectors, the past L values of $s_m(t)$ are stored as a vector. In equation (2), $\mathbf{x}_m(t)$ represents a vector with $L \times 1$ vector with L past values of latent signal at m -th region.

In Equation (3), $y_m(t)$ is the observed BOLD signal at t of m -th region. Φ is a $p \times L$ matrix whose rows contain bases for HRF. b_m is a $1 \times p$ coefficient vector representing the weights for each basis function in explaining the observed BOLD signal $y_m(t)$. Therefore, the HRF in m -th region is represented by the product $b_m \Phi$. The BOLD response in this region is obtained by convolving HRF ($b_m \Phi$) with the L past values of the region's latent signal ($\mathbf{x}_m(t)$) and is represented mathematically by the vector inner $b_m \Phi \mathbf{x}_m(t)$. product Uncorrelated observation noise $\mathbf{e}_m(t)$ with zero mean and variance σ_m^2 is then added to generate the observed signal $y_m(t)$. $\mathbf{e}_m(t)$ is also assumed to be uncorrelated with $\mathbf{w}(\tau)$, at all t and τ . Equation (3) represents the linear convolution between the embedded latent signal $\mathbf{x}_m(t)$ and the basis vectors for HRF. Here, we use the canonical HRF and its time derivative as bases, as is common in most fMRI studies.

Equations (1)–(3) together represent a state-space model for estimating the causal interactions in latent signals based on observed multivariate fMRI time series. Furthermore, the MDS model also takes into account variations in HRF as well as the influences of modulatory and external stimuli in estimating causal interactions between the brain regions.

Estimating causal interactions between M regions specified in the model is equivalent to estimating the parameters C_j , $j = 1, 2, \dots, J$. In order to estimate values of C_j , the other unknown parameters Q , $\{b_m\}_{m=1}^M$ and $\{\sigma_m^2\}_{m=1}^M$ and the latent signal $\{s(t)\}_{t=1}^T$ based on the observations $\{y_m^s(t)\}_{m=1, s=1}^{M, S}$, $t = 1, 2, \dots, T$, where T is the total number of time samples and S is number of subjects, needs to be estimated. We use a variational Bayes approach for estimating the posterior probabilities of the unknown parameters of the MDS model given fMRI time series observations for each of the S subjects.

Causal interaction analysis

To prepare data for MDS analysis, the fMRI time-series from each node and subject was first linearly de-trended and then normalized by its standard deviation. For all nodes, time-series were extracted using the MarsBar toolbox in SPM8. Spherical regions of interest were defined as the sets of voxels contained in 6-mm (diameter) spheres centered on the MNI coordinates of each node. MDS was applied to estimate the causal interactions among five nodes for the deviant and the standard stimuli. The statistical significance of the causal interactions for deviant stimuli was assessed by using a nonparametric approach. Specifically, the empirical null distribution of the parameters in causal connection in MDS was constructed by generating surrogate datasets under the null hypothesis that there are no causal interactions between the regions. Those directed connections whose median (across subjects in the group) was significantly different from the median of the null distribution were identified using statistical tests based on their empirical null distributions. The causal connections were thresholded at $P < 0.01$ with Bonferroni correction for each experimental task separately (results are reported for deviants only). We also examined whether there were significant differences in causal interactions between experimental tasks (i.e. multisensory task compared to sum of unisensory tasks) with a threshold of $P < 0.01$ with

Bonferroni correction. Extensive computer simulations on previously published benchmark datasets, as well as more realistic neurophysiological models, have demonstrated that MDS can accurately estimate dynamic causal interactions in fMRI data (Ryali et al., 2011; Ryali et al., Under Review).

Network graph analysis

To further characterize the causal outflow pattern generated by MDS, we examined the 'out degree' in each node in the network. Out degree is defined as the number of causal outflow connections from a node in the network to any other node. A Wilcoxon signed-rank test was then applied on the key network metric, the out degree to identify those nodes whose network metrics were significantly different from the other nodes.

Results and statistical analyses

Behavior

Mean \pm SD task accuracy for oddball deviants was $89.3 \pm 11.3\%$ for the auditory, $90.5 \pm 7.6\%$ for the visual and $91.5 \pm 7.7\%$ for the multisensory task. A one-way repeated-measures ANOVA with the factor modality (auditory, visual and multisensory) revealed no significant main effect of stimulus modality ($F_{1,2} = 1.31, p = 0.26$) on accuracy. Median \pm SD reaction time (RT) for deviants was 474.3 ± 87.2 ms for the auditory, 438.6 ± 53.2 ms for the visual and 442.7 ± 76.9 ms for the multisensory task. There was no significant main effect of stimulus modality ($F_{1,2} = 0.045, p = 0.83$) on RT.

In the non-simultaneous multisensory oddball task, the median \pm SD RT for deviants across 11 subjects was 559.8 ± 90.1 ms, which was significantly higher than the median RTs in the auditory (485.5 ± 94.0 ms; $t_{10} = 3.66, p = 0.004$), visual (437.8 ± 47.9 ms; $t_{10} = 6.76, p = 5.0 \times 10^{-5}$) and simultaneous multisensory (456.7 ± 83.8 ms; $t_{10} = 4.51, p = 0.001$) tasks (Supplementary Figure S1). Importantly, RTs on the non-simultaneous and simultaneous tasks differed by 103.1 ms, a difference that was not statistically distinguishable from the 100-ms gap between the non-simultaneously presented visual and auditory stimuli ($t_{10} = 0.14, p = 0.89$). This analysis suggests that even with a gap of 100 ms participants were attending to both auditory and visual stimuli and that they were not using the strategy of attending to only one stimulus modality. This makes it highly unlikely that participants were attending to only one stimulus modality in the case of the simultaneous oddball task, where there was no gap between the visual and auditory stimuli.

Identification of SN and CEN nodes

To investigate the underlying causal networks during three attention tasks, we selected five unbiased nodes using regions of interest from the SN and CEN in the right hemisphere. The nodes were based on independent component analysis of resting-state fMRI data from a separate group of 22 adults (Supekar & Menon, 2012). Table 1 shows the MNI coordinates of the nodes used in the present study. These nodes were localized to the rAI, dACC, rVLPFC, rDLPFC and rPPC (Figure 3A), regions which have been frequently reported to be activated in a wide range of unisensory oddball attention tasks (Debener et al., 2005; Crottaz-Herbette & Menon, 2006). All five nodes showed significant activation during the

processing of deviant stimuli across the three attention tasks (Figure 3B, Supplementary Table S2). Next, we used multivoxel pattern analysis to investigate whether multisensory and unisensory deviants were processed differently (See Supplementary Materials). This analysis revealed that multisensory deviants were processed differently from both auditory and visual deviants, with prominent effects in the rAI and dACC (Supplementary Figure S2).

Multisensory attention-related dynamic causal interactions in the right SN and CEN

Our analyses focused on dynamic causal interactions elicited by deviant stimuli in each of the three attention tasks. MDS revealed significant causal interactions in the following links: rAI → dACC, rAI → rVLPFC and rAI → rPPC, with rAI → dACC having the highest causal influence ($P < 0.01$ with Bonferroni correction) during multisensory auditory–visual attention (Figure 4A). Similar causal connectivity profiles were observed during the unisensory auditory and visual attention (Figure 4B and C). Specifically, conjunction analysis revealed strong common causal interactions from the rAI to dACC, rVLPFC and rPPC in the multisensory and two unisensory tasks (Figure 4D).

To further characterize the critical role of rAI in multisensory attention, we performed graph-based network analysis on the causal outflow for each node and for each participant in the multisensory task. The causal outflow of each node was computed as the outflow degree (the number of significant causal outflow connections from a node to any other node in the network). The analysis revealed that during the multisensory attention task the rAI had the highest number of causal outflow connections (out degree) among all regions in the network (Figure 5). Particularly, rAI had significantly higher causal outflow than all other nodes during the multisensory attention task.

Because of previous studies implicating the STS in multisensory processing (Baum et al., 2012; Noesselt et al., 2012), we conducted additional analyses that included this region along with the five other nodes described above. We found a similar pattern of results as before (Figure 6A–D): the rAI remained the dominant region which had strong causal interactions with other nodes. Furthermore, the rAI remained a causal outflow hub, with significantly higher causal outflow than other nodes (Figure 7).

Integrated auditory–visual attention-related signaling in the right SN and CEN

We then investigated two competing models of attention-related signaling anchored in the AI: the integrated signaling model in which the AI generates a common multisensory control signal associated with simultaneous attention to auditory and visual oddball stimuli, and the segregated signaling model in which the AI generates two segregated and independent control signals corresponding to attention to each sensory modality (Figure 2B). We hypothesized that, if the AI functions as an integrated signaling system, the strength of causal interactions from the AI to dACC during simultaneous auditory–visual attention would be smaller than the sum of the strength of causal interactions associated independently with attention in each modality. We found that the strength of dynamic causal influences between rAI → dACC, as well as rAI → rPPC, was significantly lower during multisensory auditory–visual attention compared to sum of causal interactions during the

two unisensory attention tasks ($P < 0.01$ with Bonferroni correction; Figure 4E). Furthermore, neither $rAI \rightarrow dACC$ nor $rAI \rightarrow rPPC$ causal links showed significant differences in the strength of causal interaction when we compared the multisensory attention task to each of the unisensory auditory ($p = 0.23$) or the unisensory visual ($p = 0.27$) attention tasks. These results provide support for the integrated signaling model (Figure 2B). A similar pattern held when the $rSTS$ was included in the analysis (Figure 6E and F).

Multisensory attention-related dynamic causal interactions in the left SN and CEN

We next conducted a parallel analysis using left hemisphere nodes. A different pattern of causal interactions was observed for the five nodes in the left CEN and SN: although the left AI showed significant causal interactions with other regions across the three tasks, it was not the major locus of significant causal interactions with other regions. The $dACC$, left $VLPFC$ and left $DLPFC$ all showed significant common causal interactions with other regions (Supplementary Figure S3). The causal outflow analysis further revealed that left AI had significantly higher causal outflow than only the left PPC (Supplementary Figure S4). A similar pattern was observed when the left STS was included in the analysis (Supplementary Figures S5 and S6). Taken together, these results point to weaker and less stable effects in the left hemisphere and suggest that the right, rather than left, AI is the major causal hub for integrating multisensory information.

Discussion

The oddball paradigm, which involves detection of deviants embedded in a stream of identical standard stimuli, has been widely used to probe unisensory attention (Debener et al., 2002; Kiehl & Liddle, 2003; Yago et al., 2004; Crottaz-Herbette & Menon, 2006). Based on this paradigm, we developed a multisensory attention task in which participants were required to attend to deviants presented simultaneously in the auditory and visual modalities. We used this multisensory auditory–visual task, along with two unisensory auditory and visual oddball tasks, to investigate dynamic brain mechanisms underlying auditory–visual integration and attention.

Our analysis of dynamic causal interactions focused on five right-hemisphere frontal–cingulate–parietal regions, encompassing key nodes of the SN and CEN, which are known to be involved in unisensory attention oddball tasks (Linden et al., 1999; Kiehl & Liddle, 2003; Crottaz-Herbette & Menon, 2006). We first examined the hypothesis that the rAI would show strong causal influences on right-hemisphere $dACC$, $VLPFC$, $DLPFC$ and PPC , during multisensory auditory–visual attention. Consistent with this hypothesis, we found that the rAI plays a dominant role in generating dynamic causal control signals that influence other cortical regions during multisensory attention. We next examined two competing models of rAI function: one model posits that the rAI generates an integrated multisensory control signal while the alternative model posits that the rAI generates two segregated and independent control signals corresponding to each of the modalities. Our findings provide strong support for the integrative signaling model and highlight a key role for the AI in attention to simultaneous auditory and visual events.

Dynamic causal influences of the rAI in unisensory and multisensory attention

Most previous research on attention to unisensory and multisensory stimuli has focused on activation profiles in multiple prefrontal and parietal areas (Bushara et al., 2001; Molholm et al., 2006; Driver & Noesselt, 2008; Talsma et al., 2010; Otto et al., 2013). Consistent with these studies we found significant activation of the five key frontal–cingulate–parietal regions encompassing the AI, dACC, PPC and VLPFC and DLPFC. No previous studies have, however, investigated dynamic causal mechanisms underlying multisensory attention. To address this gap we used MDS, a novel multivariate state space approach, to estimate dynamic causal interactions between five key frontal–cingulate–parietal regions of the SN and CEN. MDS has several key advantages over traditional methods for estimating causal interactions in fMRI data (Roebroeck et al., 2005; Seth, 2010). As with other methods used in noninvasive brain imaging, causal interactions here are based on the ability to predict current responses from the past observations (Roebroeck et al., 2005; Ryali et al., 2011). Notably, MDS estimates stimulus-specific causal interactions in latent neuronal signals, rather than in the recorded fMRI signals, after taking into account inter-regional variations in hemodynamic response, and it does require testing of an exponentially large number of models (Ryali et al., 2011; Ryali et al., Under Review). This makes it ideal for identifying attention-related dynamic causal interactions associated specifically with deviant oddball stimuli and networks that include a relatively large number of nodes.

MDS revealed that the rAI has significant dynamic causal influences on all other frontal–cingulate–parietal regions during the processing of simultaneously presented auditory–visual deviant oddball stimuli. The strongest influences were observed on the dACC, VLPFC and PPC (Figure 4A). The rAI showed the highest causal outflow among all nodes examined (Figure 5). While the pattern of casual interactions differed to some extent between the multisensory, auditory and visual tasks, strong causal influences of the rAI on the dACC, VLPFC and PPC were common across all three tasks (Figure 4D). These results suggest that the multisensory attention task shares several common causal fronto-cingulate-parietal pathways with unisensory tasks. Critically, each of these common pathways involves the AI.

The role of the rAI has been less well studied than the dACC and VLPFC, even within the context of unisensory attention tasks, as most previous research has primarily been focused on the role of the dACC in attention. For example, a combined EEG–fMRI analysis found modality-specific effects with increased dACC connectivity with Heschl's and superior temporal gyri during an auditory oddball task, and on the striate cortex during a visual oddball task (Debener et al., 2002; Debener et al., 2005; Crottaz-Herbette & Menon, 2006). Dipole modeling of event-related potentials based on source locations determined from fMRI activations showed that the dACC was a strong generator of the N2b-P3a attention-related components in both modalities. These results provided evidence for top-down attentional modulation of sensory processing by the dACC. However, as in most previous such studies, the role of the rAI was not modeled. Our findings here, based on a more direct analysis of causal interactions associated within specific brain networks, suggest a revised model of attentional control with a primary role for the rAI in signaling the dACC. Crucially, this process was common across unisensory and multisensory oddball stimuli.

Notably, it was the rAI rather than the VLPFC which exerted the strongest causal influences on the dACC. Our analysis thus provides new insights into the relative roles of the AI versus adjacent VLPFC in multisensory attention. Previous electrophysiological studies of multisensory processing in monkeys have generally focused on recordings from the VLPFC but not the AI (Romanski, 2007, 2012a, b). In almost all previous neuroimaging studies of multisensory attention, the AI and VLPFC are strongly co-activated (Anderson et al., 2010), making it difficult to disentangle their relative roles. Using nodes within the rAI and VLPFC that were selected independent of task activations, our study suggests that it is the rAI that has stronger causal effects on the VLPFC rather than the other way around. Thus, causal dynamics functionally differentiate the rAI and VLPFC and further validate the key role of the rAI as a dynamic causal hub during multisensory auditory–visual attention. This finding is important because it suggests that, as with unisensory stimuli, the rAI plays an important role in detecting and orienting attention to salient multisensory stimuli (Menon & Uddin, 2010).

Differential rAI responses during multisensory attention

We found significant differences in brain activation patterns elicited by deviant multisensory, compared to unisensory, stimuli. Crucially, these differences were particularly strong in the rAI and dACC (Supplementary Figure S2). Notably, differences in brain activation were observed even though participants did not differ in accuracy or reaction times between the unisensory and multisensory tasks. This raises the question of whether participants were attending to deviant stimuli only in one sensory modality. Several lines of evidence suggest otherwise. First, as noted above, there were significant differences in brain response patterns to multisensory and unisensory oddball stimuli. Second, participants had no prior knowledge about the combination of auditory and visual stimuli used in the multisensory task. Third, during synchronous presentation, auditory stimuli have a 12-ms advantage in transduction even though responses to visual stimuli tend to be faster (Spence, 2009). Crucially, in such situations, individuals can show both facilitation and inhibition to the same multisensory event (Sinnett et al., 2008; Spence, 2009). Finally, in the same participants, median RT during the (simultaneous) multisensory task was much longer than in a non-simultaneous multisensory task, where a gap of 100 ms was introduced between the auditory and visual stimuli. Furthermore, median RTs to simultaneous and non-simultaneous deviants differed by 100 ms, which was exactly the gap between the visual and auditory stimuli in the non-simultaneous multisensory task. If participants were attending to only one stimulus modality instead of both, their RTs in the non-simultaneous multisensory task would be similar to that in the unisensory as well as simultaneous multisensory tasks. On the other hand, if participants were attending to both visual and auditory stimuli, their response would be delayed by precisely 100 ms. Indeed, this is exactly what we found. This additional analysis suggests that even with a gap of 100 ms participants were attending to both auditory and visual stimuli, and that they were not using the strategy of attending to only one stimulus modality to make a response. Taken together, these results suggest that it is highly unlikely that participants were attending to only one stimulus modality in the simultaneous multisensory task. These observations are relevant for the interpretation of findings from dynamic causal analysis below because they suggest that auditory–visual

stimuli in the multisensory task are attended to, and processed, differently from unisensory tasks.

Integrative role of the right AI in multisensory attention

Building on the findings described above, we then tested our central hypothesis that the rAI is an integrative zone for processing simultaneous auditory–visual attentional signals. To accomplish this, we examined two competing models of integrated versus segregated signaling based on dynamic causal interactions between the rAI, dACC and other core regions of the fronto-cingulate-parietal networks. The integrated signaling model posits that the rAI generates a common multisensory control signal associated with simultaneous attention to auditory and visual oddball stimuli. In contrast, the alternative segregated signaling model posits that the AI generates two segregated and independent control signals corresponding to attention in each sensory modality. We hypothesized that if the AI functions as an integrated signaling system, the strength of causal interactions from the AI to dACC during simultaneous auditory–visual attention would be smaller than the sum of the strength of independent causal interactions associated with attention in each modality (integrated signaling model, Figure 2B).

Consistent with the integrated signaling model, we found that the strength of dynamic causal interactions between the rAI on the dACC as well as the PPC was significantly lower for the multisensory, compared to the sum of independent unisensory, oddball stimuli. Furthermore, despite having to process two different sensory modalities simultaneously, the strength of the causal control signal from the rAI during simultaneous auditory–visual attention was no different than control signals generated during auditory and visual unisensory attention alone. Together, these results indicate that the rAI generates an integrated multisensory control signal rather than two independent unisensory control signals and that these control signals influence both the SN and CEN.

Finally, to examine the specificity of rAI signaling effects, we conducted additional analyses including the STS, another brain area that has been widely implicated in multisensory processing (Baum et al., 2012; Noesselt et al., 2012). Interestingly, we found that while the rSTS has significant causal influences on the rAI in the unisensory and the multisensory tasks, the rAI remained the major causal outflow hub in all three tasks. Integrated signaling effects were sparse and highly specific to the rAI (rAI → dACC, rAI → rPPC, and rAI → rVLPFC) and not any other links. Crucially, the STS, another putative site of auditory–visual interaction, did not show such integrated effects. These findings clarify the differential roles of the rAI and STS in multisensory processing and further emphasize the specific and crucial role of the rAI in integrating multisensory information.

In conclusion, we have shown that the rAI exerts strong causal influences on other nodes of the frontal–cingulate–parietal attention network during simultaneous processing of multisensory events and is a dynamic causal hub that facilitates multisensory auditory–visual attention. The AI serves as an important source of integrated attention-related signaling to the dACC, a brain region that can adaptively regulate behavior via its dense connectivity with the mid-cingulate motor and the supplementary motor areas (Menon & Uddin, 2010). We also found support for our central hypothesis that the rAI generates an

integrated multisensory control signal rather than two segregated and independent unisensory control signals. Crucially, these results were preserved even with the inclusion of a posterior superior temporal sulcus region involved in multisensory processing, further highlighting the specific and integrative role of the rAI in multisensory attention.

Our study expands significantly on neurophysiological studies in non-human primates which have reported that the AI is a convergence zone for inputs from multiple senses (Mufson et al., 1981; Mesulam & Mufson, 1982; Augustine, 1996; Nimchinsky et al., 1999; Butti & Hof, 2010; Nieuwenhuys, 2012). More recently, using optogenetic techniques in a mouse model, Gogolla and colleagues have provided strong evidence for multisensory integration in the insular cortex and shown that the integrative properties in this region rely on the maturation and strengthening of inhibitory circuits (Gogolla et al., 2014). The insula is thus well placed to integrate multisensory inputs and generate attention-related signals, and our findings provide novel insights into the dynamic causal mechanisms by which this region, together with its interconnected brain networks, facilitate multisensory attention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health (Grants NS071221, NS0860851 and K25HD074652). We thank Drs Weidong Cai and Daniel Abrams for helpful comments on this manuscript.

References

- Anderson JS, Ferguson MA, Lopez-Larson M, Yurgelun-Todd D. Topographic maps of multisensory attention. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:20110–20114. [PubMed: 21041658]
- Augustine JR. Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews*. 1996; 22:229–244. [PubMed: 8957561]
- Bamiou DE, Musiek FE, Luxon LM. The insula (Island of Reil) and its role in auditory processing. Literature review. *Brain Res Brain Res Rev*. 2003; 42:143–154. [PubMed: 12738055]
- Baum SH, Martin RC, Hamilton AC, Beauchamp MS. Multisensory speech perception without the left superior temporal sulcus. *Neuroimage*. 2012; 62:1825–1832. [PubMed: 22634292]
- Beckmann CF, DeLuca M, Devlin JT, Smith SM. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci*. 2005; 360:1001–1013. [PubMed: 16087444]
- Bushara KO, Grafman J, Hallett M. Neural correlates of auditory-visual stimulus onset asynchrony detection. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2001; 21:300–304. [PubMed: 11150347]
- Bushara KO, Hanakawa T, Immisch I, Toma K, Kansaku K, Hallett M. Neural correlates of cross-modal binding. *Nature neuroscience*. 2003; 6:190–195.
- Butti C, Hof PR. The insular cortex: a comparative perspective. *Brain structure & function*. 2010; 214:477–493. [PubMed: 20512368]
- Calvert GA, Campbell R, Brammer MJ. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol*. 2000; 10:649–657. [PubMed: 10837246]

- Calvert GA, Hansen PC, Iversen SD, Brammer MJ. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*. 2001; 14:427–438. [PubMed: 11467916]
- Cappe C, Thut G, Romei V, Murray MM. Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2010; 30:12572–12580. [PubMed: 20861363]
- Cho S, Metcalfe AW, Young CB, Ryali S, Geary DC, Menon V. Hippocampal-prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *Journal of cognitive neuroscience*. 2012; 24:1849–1866. [PubMed: 22621262]
- Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 2002; 3:201–215. [PubMed: 11994752]
- Crottaz-Herbette S, Menon V. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *Journal of cognitive neuroscience*. 2006; 18:766–780. [PubMed: 16768376]
- Debener S, Kranczioch C, Herrmann CS, Engel AK. Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*. 2002; 46:77–84. [PubMed: 12374648]
- Debener S, Makeig S, Delorme A, Engel AK. What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Brain Res Cogn Brain Res*. 2005; 22:309–321. [PubMed: 15722203]
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. *Trends in cognitive sciences*. 2008; 12:99–105. [PubMed: 18262825]
- Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE. Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:11073–11078. [PubMed: 17576922]
- Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*. 2008; 57:11–23. [PubMed: 18184561]
- Eckert MA, Menon V, Walczak A, Ahlstrom J, Denslow S, Horwitz A, Dubno JR. At the heart of the ventral attention system: the right anterior insula. *Hum Brain Mapp*. 2009; 30:2530–2541. [PubMed: 19072895]
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, Turner R. Analysis of fMRI time-series revisited. *Neuroimage*. 1995; 2:45–53. [PubMed: 9343589]
- Glover GH, Lai S. Self-navigated spiral fMRI: interleaved versus single-shot. *Magn Reson Med*. 1998; 39:361–368. [PubMed: 9498591]
- Glover GH, Law CS. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn Reson Med*. 2001; 46:515–522. [PubMed: 11550244]
- Gogolla N, Takesian AE, Feng G, Fagiolini M, Hensch TK. Sensory Integration in Mouse Insular Cortex Reflects GABA Circuit Maturation. *Neuron*. 2014
- Kiehl KA, Liddle PF. Reproducibility of the hemodynamic response to auditory oddball stimuli: a six-week test-retest study. *Hum Brain Mapp*. 2003; 18:42–52. [PubMed: 12454911]
- Linden DE, Prvulovic D, Formisano E, Vollinger M, Zanella FE, Goebel R, Dierks T. The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. *Cereb Cortex*. 1999; 9:815–823. [PubMed: 10601000]
- Macaluso E, George N, Dolan R, Spence C, Driver J. Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage*. 2004; 21:725–732. [PubMed: 14980575]
- Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in cognitive sciences*. 2011; 15:483–506. [PubMed: 21908230]
- Menon V, Uddin LQ. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*. 2010; 214:655–667. [PubMed: 20512370]
- Mesulam MM, Mufson EJ. Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *J Comp Neurol*. 1982; 212:1–22. [PubMed: 7174905]

- Molholm S, Foxe JJ. Making sense of multisensory integration. *The European journal of neuroscience*. 2010; 31:1709–1712. [PubMed: 20584173]
- Molholm S, Sehatpour P, Mehta AD, Shpaner M, Gomez-Ramirez M, Ortigue S, Dyke JP, Schwartz TH, Foxe JJ. Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *J Neurophysiol*. 2006; 96:721–729. [PubMed: 16687619]
- Mufson EJ, Mesulam MM, Pandya DN. Insular interconnections with the amygdala in the rhesus monkey. *Neuroscience*. 1981; 6:1231–1248. [PubMed: 6167896]
- Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat Methods Med Res*. 2003; 12:419–446. [PubMed: 14599004]
- Nieuwenhuys R. The insular cortex. A review. *Progress in brain research*. 2012; 195:123–163. [PubMed: 22230626]
- Nimchinsky EA, Gilissen E, Allman JM, Perl DP, Erwin JM, Hof PR. A neuronal morphologic type unique to humans and great apes. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:5268–5273. [PubMed: 10220455]
- Noesselt T, Bergmann D, Heinze HJ, Munte T, Spence C. Coding of multisensory temporal patterns in human superior temporal sulcus. *Frontiers in integrative neuroscience*. 2012; 6:64. [PubMed: 22973202]
- Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2007; 27:11431–11441. [PubMed: 17942738]
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Otto TU, Dassy B, Mamassian P. Principles of multisensory behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2013; 33:7463–7474. [PubMed: 23616552]
- Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage*. 2005; 25:230–242. [PubMed: 15734358]
- Romanski LM. Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. *Cereb Cortex*. 2007; 17(Suppl 1):i61–69. [PubMed: 17634387]
- Romanski, LM. *Frontiers in Neuroscience: Convergence of Auditory, Visual, and Somatosensory Information in Ventral Prefrontal Cortex*. In: Murray, MM.; Wallace, MT., editors. *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press LLC; 2012a.
- Romanski LM. Integration of faces and vocalizations in ventral prefrontal cortex: implications for the evolution of audiovisual speech. *Proceedings of the National Academy of Sciences of the United States of America*. 2012b; 109(Suppl 1):10717–10724. [PubMed: 22723356]
- Ryali S, Supekar K, Chen T, Menon V. Multivariate dynamical systems models for estimating causal interactions in fMRI. *Neuroimage*. 2011; 54:807–823. [PubMed: 20884354]
- Ryali S, Tu T, Chen T, Menon V. Multivariate dynamical systems-based estimation of causal brain interactions in fMRI: Group-level validation using benchmark data and neurophysiological models. Under Review.
- Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2007; 27:2349–2356. [PubMed: 17329432]
- Seth AK. A MATLAB toolbox for Granger causal connectivity analysis. *Journal of neuroscience methods*. 2010; 186:262–273. [PubMed: 19961876]
- Sinnett S, Soto-Faraco S, Spence C. The co-occurrence of multisensory competition and facilitation. *Acta psychologica*. 2008; 128:153–161. [PubMed: 18207117]
- Spence C. Explaining the Colavita visual dominance effect. *Progress in brain research*. 2009; 176:245–258. [PubMed: 19733761]
- Sridharan D, Levitin DJ, Menon V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:12569–12574. [PubMed: 18723676]

- Sterzer P, Kleinschmidt A. Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Struct Funct.* 2010; 214:611–622. [PubMed: 20512379]
- Supekar K, Menon V. Developmental maturation of dynamic causal control signals in higher-order cognition: a neurocognitive network model. *PLoS Comput Biol.* 2012; 8:e1002374. [PubMed: 22319436]
- Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences.* 2010; 14:400–410. [PubMed: 20675182]
- Teder-Salejarvi WA, Di Russo F, McDonald JJ, Hillyard SA. Effects of spatial congruity on audio-visual multimodal integration. *Journal of cognitive neuroscience.* 2005; 17:1396–1409. [PubMed: 16197693]
- Uddin LQ, Supekar KS, Ryali S, Menon V. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *The Journal of neuroscience : the official journal of the Society for Neuroscience.* 2011; 31:18578–18589. [PubMed: 22171056]
- Yago E, Duarte A, Wong T, Barcelo F, Knight RT. Temporal kinetics of prefrontal modulation of the extrastriate cortex during visual attention. *Cogn Affect Behav Neurosci.* 2004; 4:609–617. [PubMed: 15849901]

Abbreviations

AI	anterior insula
CEN	central executive network
dACC	dorsal anterior cingulate cortex
DLPFC	dorsolateral prefrontal cortex
HRF	hemodynamic response function
MDS	multivariate dynamic systems
MNI	Montreal Neurologic Institute
PPC	posterior parietal cortex
r (prefix)	right-hemisphere
RT	reaction time
SN	salience network
STS	superior temporal sulcus
VLPFC	ventrolateral prefrontal cortex

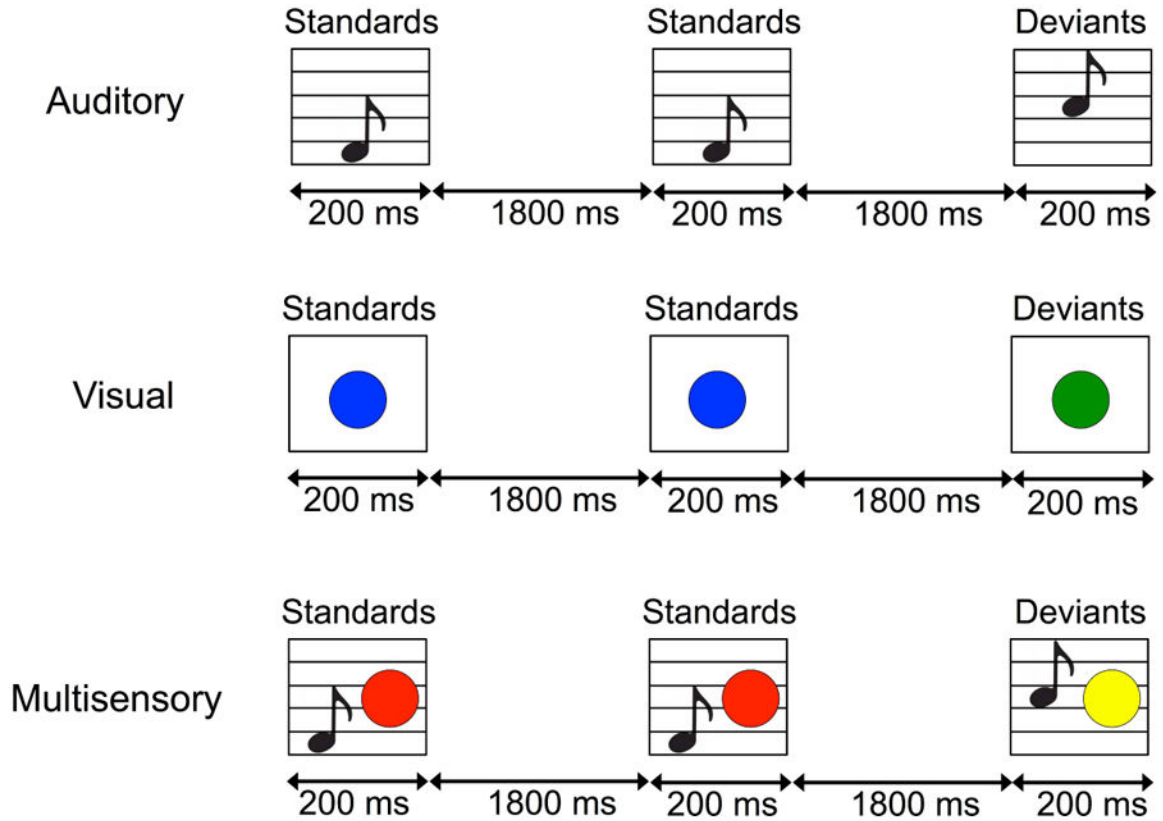


Figure 1.

Experimental Design. For all tasks, a total of 200 trials were presented (40 oddball deviants and 160 standards) in randomized order, after an initial blank screen of 10s. For the auditory oddball task, participants listened to 200-ms-long low-frequency (1000 Hz) and high-frequency (2000 Hz) tones. The inter-trial period (blank screen) was 1800 ms. For half of the participants, the low tone was the deviant stimulus (40 trials) while the high tone served as standard stimuli (160 trials). For the other half participants, the deviant and standard stimulus assignment was reversed. During the visual oddball task circles with different colors were presented for 200 ms at the center of the screen. The inter-trial period (blank screen) was 1800 ms. For half of the participants, the e.g. blue circle was the deviant stimulus, for the other half of participants the e.g. green circle served as the deviant stimulus. During the multisensory oddball task, participants saw a centrally presented circle and heard a binaural tone at the same time using the same stimuli as the unisensory tasks. The assignment of deviant and standard stimuli was counterbalanced across participants, i.e., for one group a low tone was paired with a red circle as deviant while for the other group a high tone was combined with a yellow circle.

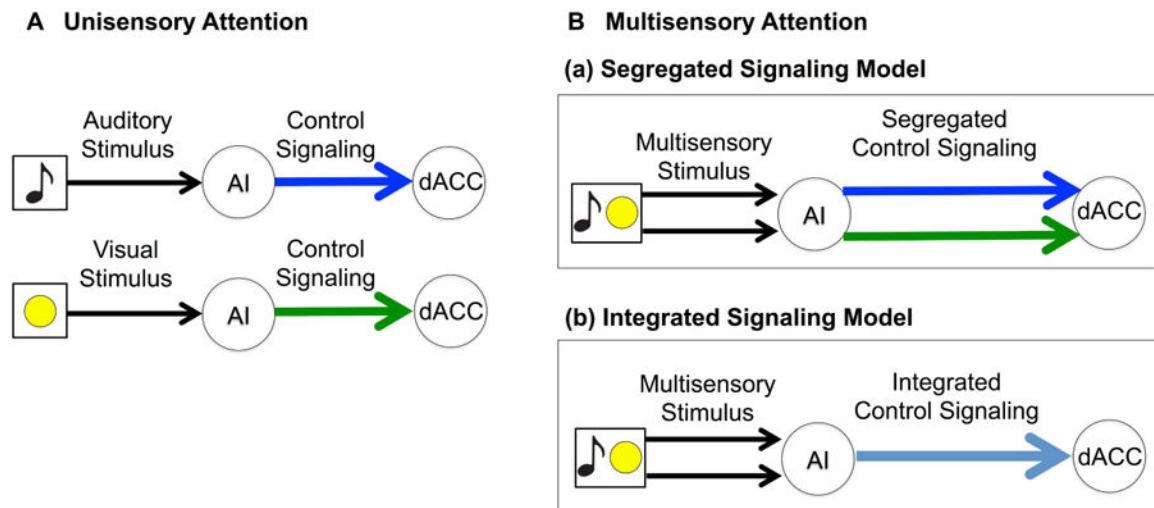


Figure 2.

Illustration of integrated versus segregated model of multisensory attention-related signals. (A) In each unisensory (auditory or visual) task, the deviant stimulus elicits a control signal from AI to dACC. (B) Models of segregated and integrated signaling in the multisensory task. (a) Segregated model. The model posits that if processing of the two deviant stimuli occur in a segregated manner, the overall control signal strength from AI to dACC would be equal to the sum of control signals elicited by each stimulus as in the unisensory tasks. (b) Integrated model. This model posits that if the two deviant stimuli are integrated by the AI, this region would generate an integrated control signal which is smaller than the sum of control signals elicited by each stimulus as in the unisensory tasks.

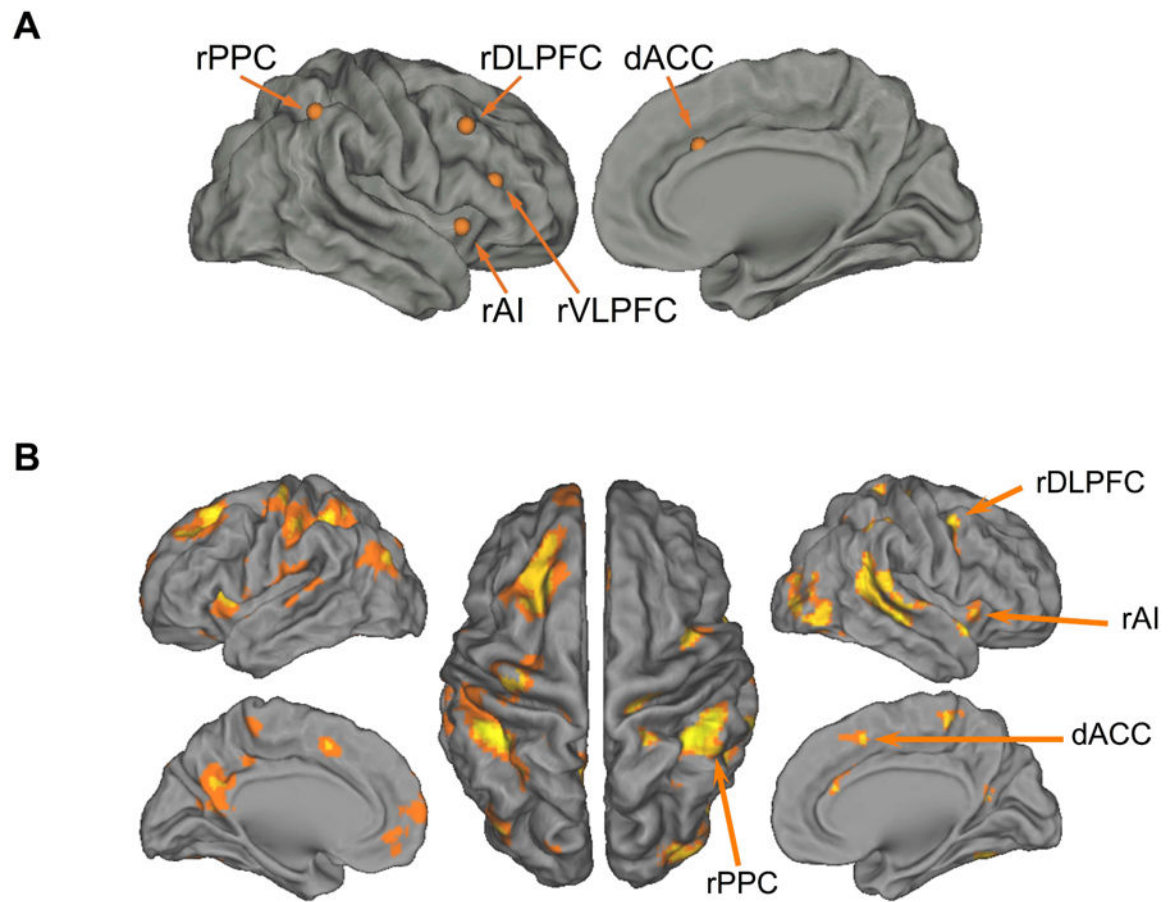


Figure 3.

Nodes within SN and CEN. (A) Five nodes are selected within the right hemisphere SN and CEN identified using intrinsic connectivity analysis in a separate group of participants. (B) Combined attention-related activation in multisensory, auditory and visual tasks. Both the SN and the CEN are activated by the deviant stimuli, consistent with the regions of interest selected using intrinsic connectivity analysis.

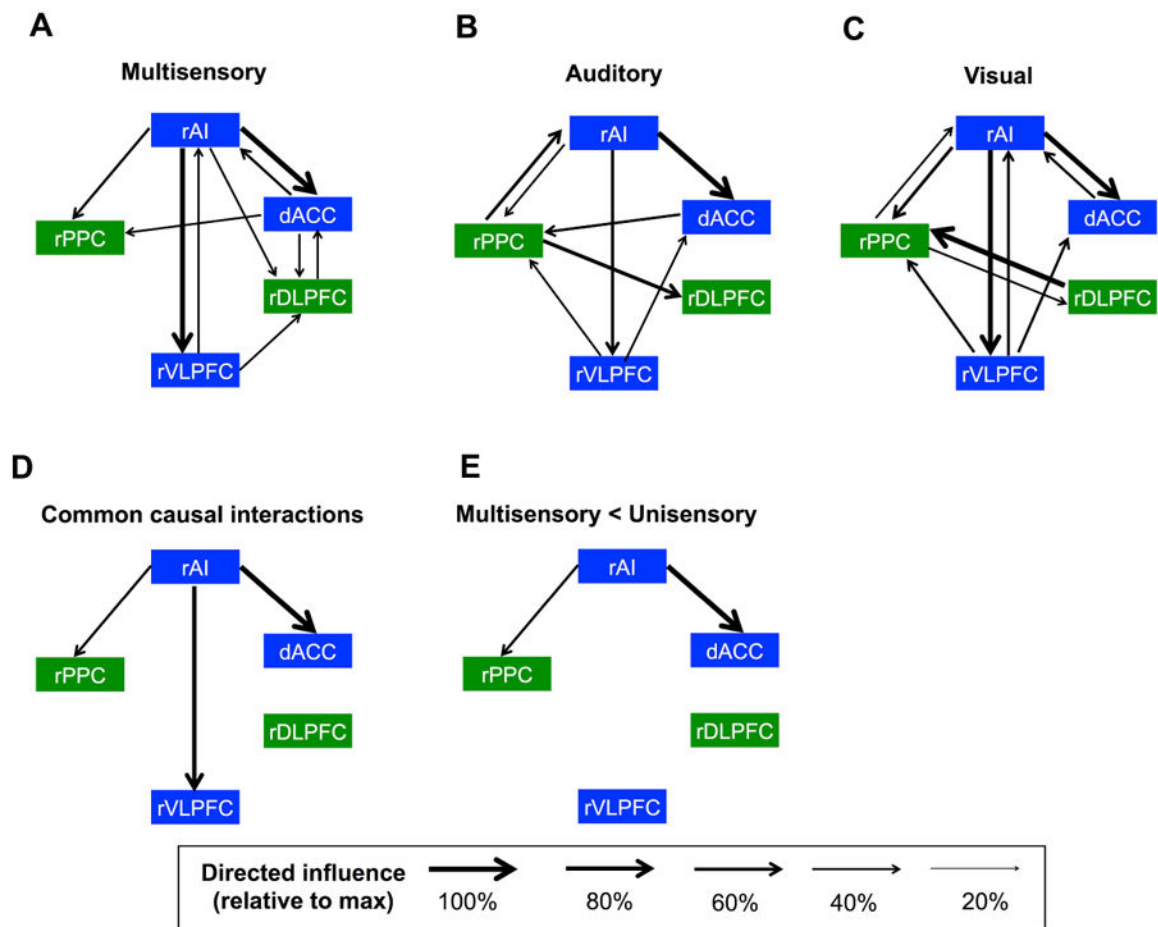


Figure 4.

Attention-related dynamic causal interactions between the five nodes of the SN and CEN in the right hemisphere. Significant causal interactions were observed between five key nodes in SN (blue) and CEN (green) in (A) multisensory, (B) auditory and (C) visual tasks. Across the three tasks, the AI was the dominant source of casual influence. Results are shown with $p < 0.01$ (Bonferroni-corrected). (D) Common causal interactions across the three attention tasks. (E) Sum of unisensory causal interactions between the AI and dACC were significantly stronger than multisensory causal interactions. Results are shown with $p < 0.01$ (Bonferroni-corrected).

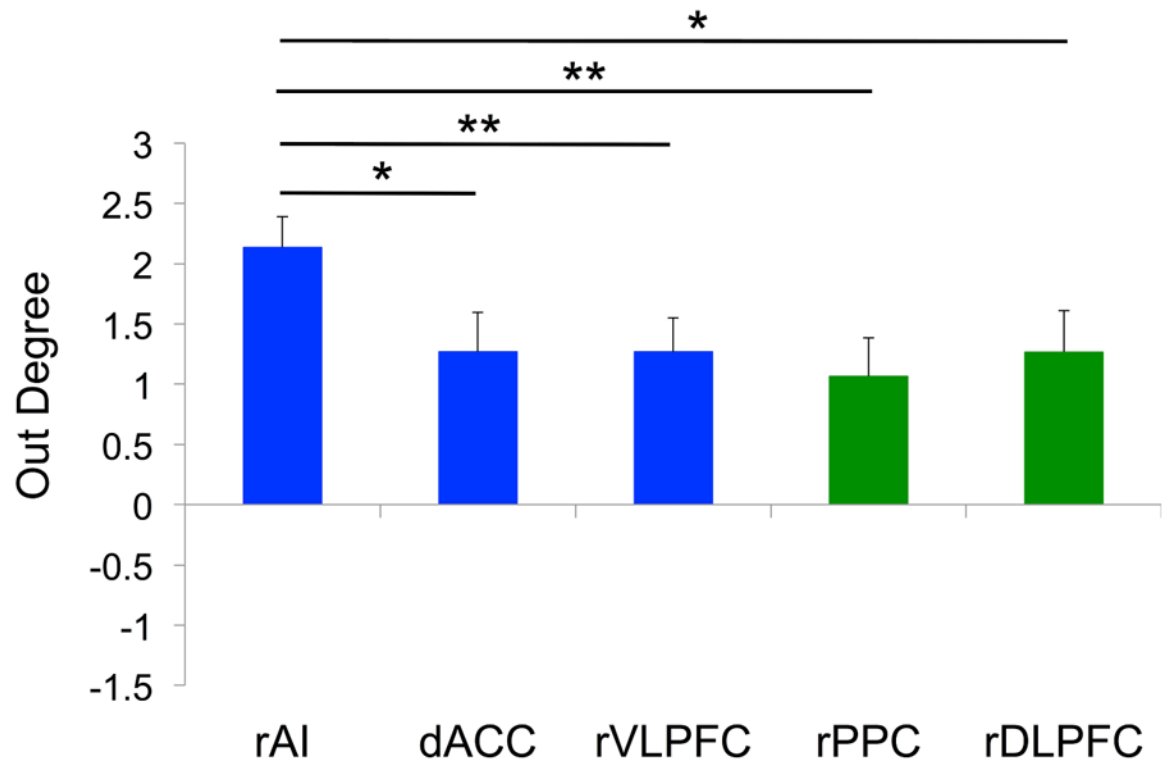


Figure 5. Causal outflow from SN and CEN nodes during multisensory attention. The rAI consistently showed the highest number of causal outflow connections (out degree) among all five frontal–cingulate–parietal regions. * $P < 0.05$, ** $P < 0.01$.

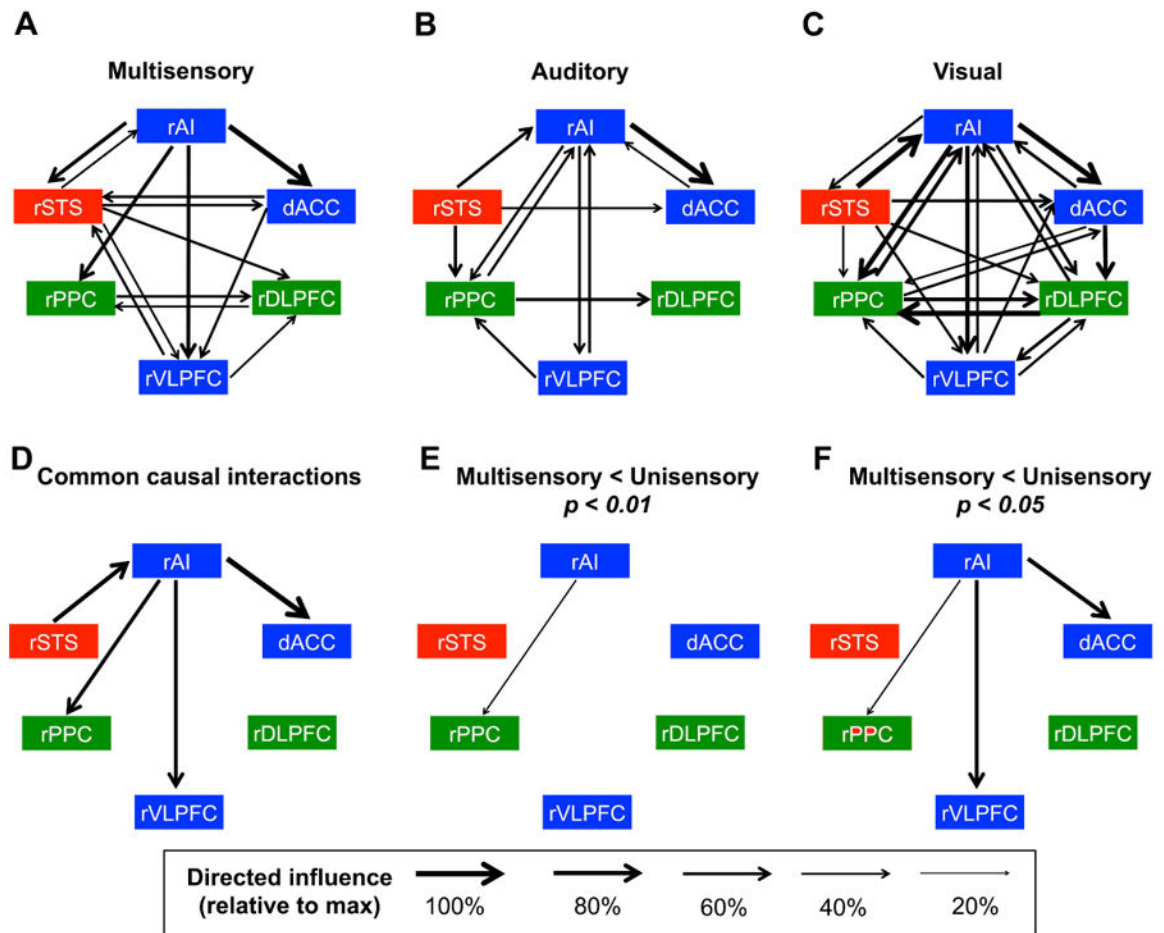


Figure 6.

Attention-related dynamic causal interactions between the five nodes of the SN and CEN, and an additional STS node in the right hemisphere. Significant causal interactions were observed between six nodes in SN (blue), CEN (green) and STS (red) in (A) multisensory, (B) auditory and (C) visual tasks. Across the three tasks, the AI was the dominant source of casual influence. Results are shown with $P < 0.01$ (Bonferroni-corrected). (D) Common causal interactions across the three attention tasks. (E) Sum of unisensory causal interactions between the AI and PPC were significantly stronger than multisensory causal interactions with $p < 0.01$ (Bonferroni-corrected). (F) Sum of unisensory causal interactions between the AI and dACC, AI and PPC, and AI and VLPFC were all significantly stronger than multisensory causal interactions with $p < 0.05$ (Bonferroni corrected).

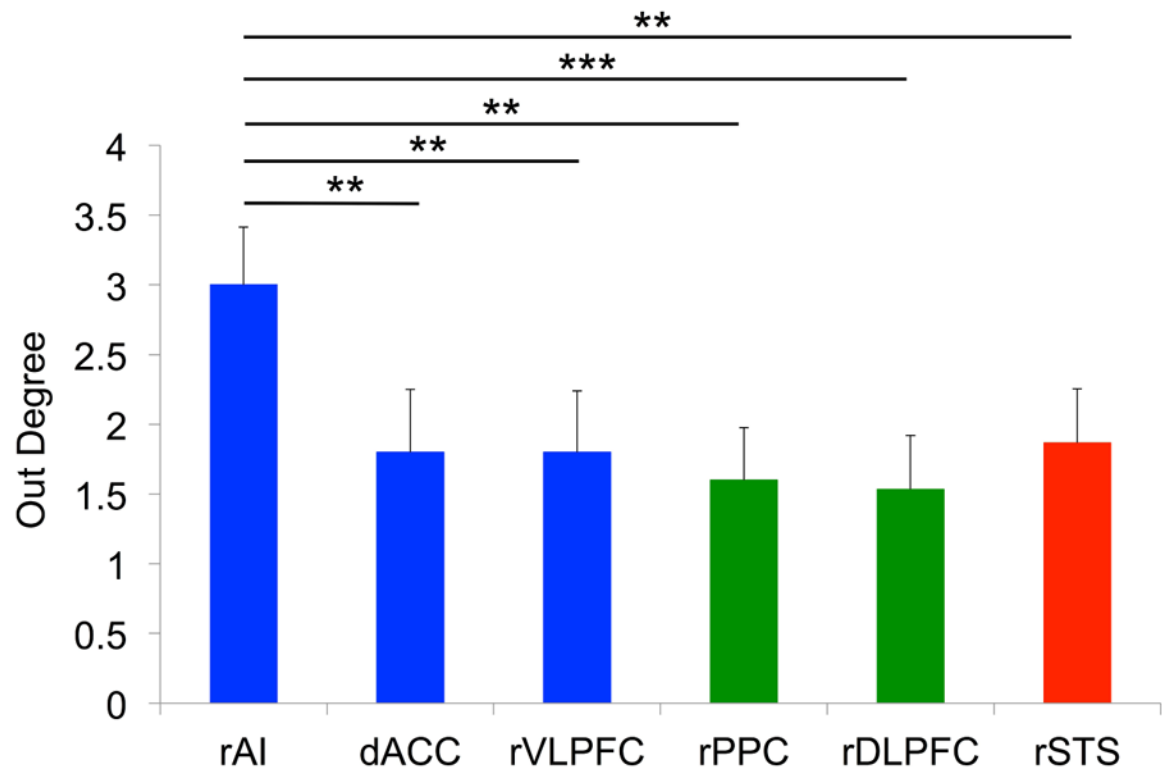


Figure 7. Causal outflow during multisensory attention in the five nodes of the SN and CEN, and an additional node in the STS in the right hemisphere, during multisensory attention. rAI consistently showed the highest number of causal outflow connections (out degree) among all five SN and CEN nodes and the rSTS node. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Table 1

Description of the six selected regions of interest

Region	Hemisphere	Brodmann area	MINI coordinates		
			x	y	z
SN					
AI	Right	13	37	16	-2
VLPFC	Right	46	42	26	14
dACC		31	7	18	33
CEN					
DLPFC	Right	8	50	18	44
PPC	Right	7	48	-52	50
STS	Right	22	51	-55	11