Behavioral/Cognitive

Cognitive Control and the Salience Network: An Investigation of Error Processing and Effective Connectivity

Timothy Ham,[1] **Alex Leff,**[2] **Xavier de Boissezon,**[3,4,5] **Anna Joffe,**[1] **and David J. Sharp**[1]

[1]Computational, Cognitive and Clinical Neuroimaging Laboratory, Centre for Neuroscience, Division of Experimental Medicine, Imperial College London, Hammersmith Hospital Campus, London W12 0NN, United Kingdom, [2]Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom, [3]Inserm, Imagerie cerebrale et handicaps neurologiques UMR 825, F-31059 Toulouse, France, [4]Universite de Toulouse, UPS, Imagerie cerebrale et handicaps neurologiques UMR 825, CHU Purpan, F-31059 Toulouse Cedex 9, France, and [5]Centre Hospitalier Universitaire de Toulouse, Pole Neurosciences, CHU Purpan, F-31059 Toulouse Cedex 9, France

The Salience Network (SN) consists of the dorsal anterior cingulate cortex (dACC) and bilateral insulae. The network responds to behaviorally salient events, and an important question is how its nodes interact. One theory is that the dACC provides the earliest cortical signal of behaviorally salient events, such as errors. Alternatively, the anterior right insula (aRI) has been proposed to provide an early cognitive control signal. As these regions frequently coactivate, it has been difficult to disentangle their roles using conventional methods. Here we use dynamic causal modeling and a Bayesian model evidence technique to investigate the causal relationships between nodes in the SN after errors. Thirty-five human subjects performed the Simon task. The task has two conditions (congruent and incongruent) producing two distinct error types. Neural activity associated with errors was investigated using fMRI. Subjects made a total of 1319 congruent and 1617 incongruent errors. Errors resulted in robust activation of the SN. Dynamic causal modeling analyses demonstrated that input into the SN was most likely via the aRI for both error types and that the aRI was the only region intrinsically connected to both other nodes. Only incongruent errors produced behavioral adaptation, and the strength of the connection between the dACC and the left insulae correlated with the extent of this behavioral change. We conclude that the aRI, not the dACC, drives the SN after errors on an attentionally demanding task, and that a change in the effective connectivity of the dACC is associated with behavioral adaptation after errors.

Introduction

The Salience Network (SN) responds to behaviorally salient events (Seeley et al., 2007) and is thought important for the initiation of cognitive control (Menon and Uddin, 2010), the maintenance and implementation of task sets (Dosenbach et al., 2006; Nelson et al., 2008), and the coordination of behavioral responses (Medford and Critchley, 2010). The SN consists of three main cortical areas: the dorsal anterior cingulate cortex (dACC), the left and anterior right insula (aRI), and the adjacent inferior frontal gyri (Seeley et al., 2007). Increased SN activity is observed in situations in which it may be important to change behavior (Dosenbach et al., 2007). For example, errors are associated with robust SN activation (Carter et al., 1998; Holroyd et al., 2004) and often signal the need for behavioral adaptation (Rabbitt, 1966).

An influential theory states that the dACC monitors performance and signals the need for behavioral adaptation (Holroyd et al., 2004; Ridderinkhof et al., 2004). Electrophysiological studies have identified an error-related negativity (ERN), a very early response to errors (80–110 ms) (Gehring et al., 1993) thought to arise from the dACC (Dehaene et al., 1994; Debener et al., 2005). It is proposed that activity in the dACC signals the need for increased cognitive control (Ridderinkhof et al., 2004), and interactions between the dACC and lateral prefrontal structures implement subsequent behavioral changes (Ridderinkhof et al., 2004; Egner, 2009; Kouneiher et al., 2009).

An alternative theory proposes that the aRI is a "cortical outflow hub," coordinating changes in activity across multiple brain networks (Sridharan et al., 2008; Menon and Uddin, 2010; Bonnelle et al., 2012). Our group recently demonstrated that the structural integrity of the white matter connection between the aRI and the dACC predicts behavioral and physiological abnormalities after traumatic brain injury (Bonnelle et al., 2012). In addition, Granger causality analysis has provided some support for the view that the aRI causally influences activity in other brain networks (Sridharan et al., 2008), although there are methodological problems with using lag-based methods for inferring causality from fMRI data (Smith et al., 2011).

Our previous work has also demonstrated that SN activation is linked to the neural response to internally signaled errors (Ham et al., 2013). However, the causal interactions of nodes within the
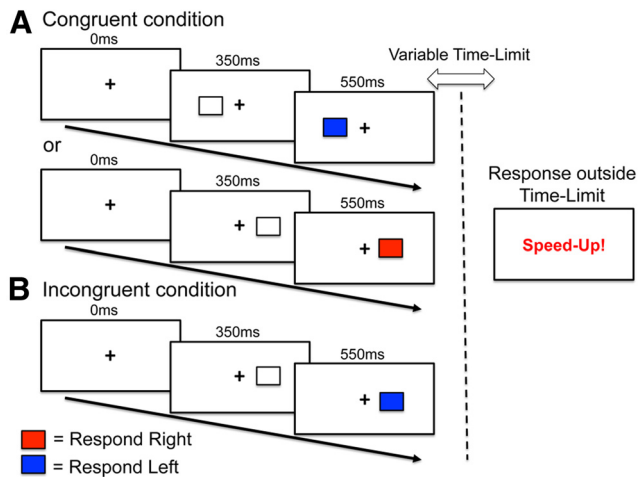
Figure 1. Schematic of the Simon task paradigm. Subjects responded with a right or left finger press for red and blue cues, respectively. The precue (☐) appeared to either the left or right side of the fixation cross for 200 ms before filling in with the cue color (either red or blue). **A**, The congruent condition, in which the spatial location of the cue corresponded to the side of the appropriate response press. **B**, The incongruent condition, in which the spatial location conflicted with the side of the response.

SN remain unclear, partly because activity in the dACC and anterior insulae tends to show tightly correlated neural activity. We investigated the causal relationships between these nodes in the SN using dynamic causal modeling (DCM) (Friston et al., 2003), a technique that infers effective connectivity from fMRI data. A group of healthy volunteers performed the Simon task, which has two different conditions (congruent and incongruent) leading to two distinct types of error. This allowed us to compare patterns of SN connectivity across distinct types of behaviorally salient events. Standard fMRI analysis was used to assess the pattern of relative activation within the SN, and DCM was used to test where input entered the SN and whether errors were associated with modulation of connections within the SN.

Materials and Methods

Participants
Thirty-five healthy subjects performed the Simon task (17 male, mean age 30.6 ± 8.6 years). Subjects gave written consent. The experiment was approved by the Hammersmith and Queen Charlotte's, and Chelsea Research ethics committee.

Simon task procedure
Error processing was investigated during performance of the Simon task. This is a choice reaction task that uses incongruent spatial and color cues to generate errors (Simon, 1969; Simon and Berbaum, 1990). Building on previous electrophysiological work (Christ et al., 2000), we used a version of the Simon task designed to produce large numbers of errors. Subjects were presented with a colored cue to the right or left of a fixation cross (Fig. 1). Cue color determined the direction of the required response: red signified a right-hand response, and blue signified a left-hand response. Spatial location and cue direction were either congruent or incongruent with respect to each other. In the incongruent condition, the prepotent response (to respond in the direction of the spatial location of the cue rather than the direction signaled by the color) must be inhibited. Errors occur when the subject's direction of response is not that signaled by the color. Errors during the incongruent condition were classified as incongruent errors, and errors during the congruent condition were classified as congruent errors. To produce more errors, subjects were encouraged to perform the task quickly. This encouragement came in the form of an audiovisual negative feedback cue that occurred when subjects responded outside of a variable time limit. This feedback was pre-

sented in the form of the words "Speed up" displayed on the screen accompanied by a 400 Hz auditory tone; both lasted 500 ms.

An adaptive staircase procedure was used to vary the response delay necessary to trigger feedback on congruent and incongruent conditions separately. The goal of this was to produce large numbers of both congruent and incongruent errors. At the start of the experiment, feedback was triggered if a subject's response was >500 ms after the presentation of the color stimulus. After the first 15 trials of a given condition, the percentage error rate was calculated after each trial, and the time limit was adjusted if the response fell outside the target range. For the congruent condition, a target error rate of 8–12% was used. The time limit was increased by 50 ms if the error rate was high, and decreased by the same amount if it was too low. Adaptation was performed separately for incongruent condition, aiming for an error rate of 17–25%. The time limit was adapted within a range of 400–1000 ms. For subsequent runs, the starting time limit was carried over from the previous run.

Two-thirds of trials were "congruent" and one-third "incongruent." All subjects performed six runs of 120 pseudo-randomly ordered trials with interstimulus intervals of 2.25 s. To increase task difficulty, we introduced a precue in the form of an empty rectangle that filled in after 200 ms with the color that indicated response direction. The precue has the effect of increasing the interference produced by a spatially incongruent color cue. The relative timing of the precue and color cues was designed to generate the maximum number of errors based on previous work (Christ et al., 2000). All subjects performed 120 trials (80 congruent and 40 incongruent) as training before scanning.

Behavioral analysis
Mean error rates and reaction times were calculated separately for congruent and incongruent conditions. In addition, behavior was investigated on trials surrounding errors. This allowed the effects of cognitive control to be observed as posterror slowing. The trials before an error ($N - 1$), the error trials (N), and up to three trials after an error ($N + 1$, $N + 2$, and $N + 3$) were investigated. Reaction times were compared with a baseline performance on trials for that condition (i.e., congruent or incongruent), calculated from the mean of the last 10 stable correct trials of a particular type. Stable correct trials were defined as correct and timely trials that had also been preceded by a correct timely trial. This was done to avoid contamination from the effects of on preceding trials, and also to control for slow fluctuations in subject's attention.

Image acquisition
MRI data were obtained using a Philips Intera 3.0 Tesla MRI scanner using Nova Dual gradients, a phased array head coil, and sensitivity encoding (SENSE) with an under sampling factor of 2. fMRI images were obtained using a T2*-weighted gradient-EPI sequence with whole-brain coverage (TR/TE = 2000/30; 31 ascending slices with thickness 3.25 mm, gap 0.75 mm, voxel size 2.19 × 2.19 × 4 mm, flip angle 90°, field of view 280 × 220 × 123 mm, matrix 112 × 87). Quadratic shim gradients were used to correct for magnetic field inhomogeneities within the brain. T1-weighted whole-brain structural images were also obtained in all subjects. Paradigms were programmed using Psychophysics toolbox (Matlab Psychtoolbox-3; www.psychtoolbox.org) and stimuli presented through an IFIS-SA system (In Vivo Corporation). Responses were recorded through a fiber optic response box (Nordicneurolab), interfaced with the stimulus presentation PC running Matlab. Sounds were presented using ear-defending headphones (MR Confon).

Whole-brain fMRI analysis
Statistical parametric mapping was performed using the SPM8 software (Wellcome Trust Centre for Neuroimaging; www.fil.ion.ucl.ac.uk/spm). All volumes from each subject and each run were realigned and unwarped, using the first volume from the first run as reference. For each subject, the T1 structural image was coregistered to the mean functional image. The volumes were then resliced with sinc interpolation. The volumes were then spatially normalized to the standard T2* template using SPM8 normalization software. The functional data were spatially smoothed, with a 8 mm full width at half-maximum isotropic Gaussian Kernel. A first-level fixed-effects statistical analysis was performed for every subject. To remove low-frequency drifts, the data were high-pass

filtered using a set of discrete cosine basis functions with a cutoff period of 128 s. Two types of event were modeled for congruent and incongruent trials separately: correct trials (appropriate response within the time limit) and errors (incorrect button press within the time limit). Parameter estimates were calculated for these events across all brain voxels using the general linear model, a synthetic hemodynamic response function, and its first temporal derivative. Four contrasts were examined: congruent errors > congruent correct; incongruent errors > incongruent correct; congruent errors > incongruent errors; and incongruent errors > congruent errors. Contrast images were performed at the individual level and then the combined analysis at the between-subject level using random effects (Friston et al., 1999). Final statistical images were thresholded using FWE threshold of $p < 0.05$. In addition, a conjunction analysis was performed to define regions common to both congruent error > congruent correct and incongruent error > incongruent correct contrasts (Nichols et al., 2005).

DCM analysis
We used the DCM10 package in SPM8 to perform the analyses described below. Congruent and incongruent trials had different frequencies (33.3% vs 66.6% of trials), were conceptually different (spatial incongruency vs congruency), and were behaviorally distinct (see below). Moreover, error and posterror trials in the two conditions were also behaviorally different. As a result, we analyzed each condition (congruent and incongruent) separately. This had the advantage of allowing us to test the reproducibility of results across different conditions. Moreover, to allow us to relate changes in the B-matrix to behavior, we compared the effects of trial type (correct vs error) within condition (congruent and incongruent) and thus against the appropriate baseline. The final stage of the analysis involved comparing across the two conditions.

We identified the ROIs from each subject's first level univariate fMRI analysis. These were then used in the DCM analysis by fitting the data to a variety of different models to establish the connectivity of the SN. We did this in three stages: (1) we identified where the inputs to the SN entered the system using Bayesian model selection; (2) using Bayesian model averaging (BMA) for both congruent and incongruent conditions, we tested which connections within the SN were significantly modulated by both trial types (correct and error), and which were differentially modulated (correct vs error); and (3) we directly compared the resultant connection strengths across conditions to determine whether any connections were significantly modulated in the condition by trial type interaction (incongruent errors vs congruent errors), and then sought a behavioral correlation for any significantly modulated connection.

Identifying ROIs. Whole-brain cluster analysis was used on the group-level analyses described above to determine the peak voxels of interest for the given contrasts. The cluster analysis allowed us to establish the peak voxels within the dACC, aRI, and left anterior insula (aLI) at a group level for the contrasts of interest (i.e., congruent errors > congruent correct trials; and, incongruent errors > incongruent correct trials); 4 mm spherical ROIs were extracted for each contrast of interest adjusted for the equivalent F-contrast. The ROIs were initially centered on the peak voxels from the group-level contrasts. To accommodate interindividual variability and provide the optimal signal-to-noise ratio data in the time series for each region, the centers of the ROIs were allowed to move no more than 8 mm from the peak of the group level contrast depending on the individual pattern of activity. ROIs were extracted in an incremental fashion. The uncorrected p threshold started at 0.001 and increased in 0.001 increments until more than two suprathreshold voxels were included within a ROI. For each subject and contrast, if any of the three regions required thresholds above $p = 0.05$, the subject was excluded from further DCM analyses of that contrast. The exclusion criteria applied to seven subjects for incongruent errors and seven subjects for congruent errors; five of the subjects overlapped and were therefore excluded for both contrasts.

Model construction, family construction, and model estimation. For each condition (congruent and incongruent), DCM10 was used to create a series of models from the three node SN. Each model consists of three matrices, the "A," "B," and "C" matrices, each describing a different feature of model space: The "A" matrix represented the context-

independent intrinsic connections within the model (correct and incorrect trials treated equally). With a three node model, there were six possible connections described in the "A" matrices (i.e., dACC to aRI, aRI to dACC, dACC to aLI, aLI to dACC, aLI to aRI, and, aRI to aLI). In all of the models, the "A" matrices used were fully connected to each other to allow for average effects of trial type to be expressed (i.e., correct and error trials treated equally). The "A" matrices were not varied across models. The "B" and "C" matrices describe the two ways that the experimental conditions could affect regional neuronal activity.

In DCM, brain responses are evoked by known deterministic inputs (experimentally controlled stimuli) that embody designed changes in sensory stimulation or cognitive set. These experimental variables can change the predicted BOLD response in one of two ways. First, they can elicit responses through direct influences on specific network nodes (i.e., as inputs into the network and are expressed in the "C" matrix). The second class of inputs exert their effects vicariously, through a modulation of the coupling among nodes, in this case, the difference between correct versus incorrect trials, expressed in the "B" matrix.

The "C" matrix represented the potential inputs into the model after an error. These inputs arise elsewhere in the brain (e.g., from subcortical structures), although the current work focuses only on the cortical contributions to error processing. We wished to experimentally test where neural activity entered our three region network. Each node could have two possible states (i.e., an input or not). There were therefore $2^3 = 8$ mathematically possible combinations of "C" matrix required to thoroughly explore model space. The input arrangements possible were as follows: dACC alone, aRI alone, aLI alone, dACC and aRI in combination, dACC and aLI in combination, aLI and aRI in combination, all nodes, and no nodes. Because "no nodes" was biologically implausible, this option was excluded from further analysis, leaving seven remaining input arrangements

The "B" matrix represented the same connections as the "A" matrix, but here coding how effective connectivity between regions was affected by the two conditions (correct vs incorrect). These changes in the "hidden" neuronal states were considered the indirect influence of the errors on regional activity. In previous fMRI analyses, the "B" matrix has described how task demands have increased effective connectivity between regions (Friston et al., 2003). Just as with the "A" matrix, there were six possible connections described in the "B" matrices (i.e., dACC to aRI, aRI to dACC, dACC to aLI, aLI to dACC, aLI to aRI, and aRI to aLI). In the "B" matrix, each connection could exist in two states (i.e., modulated or unmodulated by trial type). There were therefore $2^6 = 64$ mathematically possible combinations of "B" matrix required to thoroughly explore model space.

We then combined the three matrices to into $7 \times 64 = 448$ models to allow comprehensive review of model space. We grouped the models together into families. Every family contained 64 models, each model within a family had the same input ("C" matrix), and all models had one of the 64 possible "B" matrices. The "B" matrices were equally distributed across the families, but the "C" matrices were distinct to each family. The families therefore represented seven groups with identical distributions of "B" matrices/modulated connections but distinct "C" matrices/input characters. Every model had a fully connected "A" matrix. In doing this, we created all arithmetically possible combinations of models and grouped them into nonoverlapping families that were distinguished only by their inputs. This process created 448 models in total (the family with no inputs was excluded from further analysis) that were then estimated in DCM10 (Friston et al., 2003; Penny et al., 2010).

Analysis 1: Where do inputs enter the SN? The 448 models were sorted into the seven input families, each containing 64 models. The families were compared using the random effects option of the family level Bayesian inference (Penny et al., 2010). This computes the frequency with which each family of models is used in the population from which the subjects were drawn. It also computes an exceedance probability (xp), which is the probability that a model family has the highest frequency. The sum of all seven families' xp equals 1. The threshold xp was set at >80%, as has been used in previous studies using this technique (Leff et al., 2008; Penny et al., 2010). The "winning" famil(ies) were then taken onto the next analysis, whereas the "losing" ones (those with little evidence) were rejected at this point.

Analysis 2: *What are the intrinsic SN connections and are they modulated by errors?* Two input families were taken forward into this analysis (128 models in total). We used a BMA analysis to identify the average connection strengths (weighted by subject, by model, on the model evidence) across all models (Penny et al., 2010). We did this to calculate both the average effect of trial type (correct and error trials treated equally within both congruent and incongruent conditions treated separately: the "A" matrix) and the differential effect of trial type (correct vs error trials, within both congruent and incongruent conditions treated separately: the "B" matrix). Here we are moving away from inferences based on the overall connectivity structure of the best model, toward asking which connections are significantly modulated by trial type (Stephan et al., 2010). The BMA provides the mean connection strength for each interregional connection within the "A" and "B" matrices for every subject. The starting point (prior) for the interregional connections is zero, a value that can be altered at the model estimation phase. The resulting connection strength values were then entered into a one-sample t test to determine whether their values had been significantly moved away from the starting prior with significance set at < 0.05 (Schofield et al., 2012).

Analysis 3: *Are there distinct effects of errors in congruent and incongruent trials, and do these relate to behavior?* In the final Analysis 3, we investigated between-condition effects on any connections that were significantly modulated by trial type (correct vs error). Our previous work has shown that during the Simon task different trial types produce distinct behavioral profiles depending upon the condition (Ham et al., 2013). In Analysis 2, there is no common baseline across the conditions because the congruent correct and incongruent correct trials are different behaviorally and cognitively. Therefore, to test for a significant between-condition effect, a direct comparison of the within-condition changes in effective connectivity is needed. There was only one connection where this was the case (see Results). We performed a single, paired t test on connectivity values for this connection only. As above, the significance was set at < 0.05 and there were no multiple-comparison issues as only one parameter was tested. Last, we wished to determine whether connectivity values from this connection correlated with behavior. To do this, we used a Spearman rank coefficient test to assess any the relationship between a subjects' posterror slowing across all errors (both congruent and incongruent) and the alteration in effective connectivity associated with the errors.

Results

Behavior on the Simon task

All subjects were included in the behavioral analyses ($n = 35$). Behavioral performance was in keeping with previous studies (Christ et al., 2000). On average, subjects produced errors on $7.9 \pm 3.5\%$ of congruent trials and $19.3 \pm 7.0\%$ of incongruent trials. Although there were fewer incongruent trials, subjects made more errors during the incongruent than the congruent condition ($t = 2.45$, $df = 34$, $p = 0.019$).

As expected, response speed had a major effect on the accuracy of the incongruent condition but interestingly not for the congruent condition (Fig. 2). A 2×2 ANOVA showed a significant interaction between trial type and accuracy ($F = 127.2$, $p < 0.0005$). The interaction was the result of two factors: (1) subjects responded slower on correct incongruent trials than correct congruent trials ($t = 10.6$, $df = 34$, $p < 0.0001$), a demonstration of the Simon effect; and (2) subjects responded faster on incongruent errors than congruent errors ($t = 2.89$, $df = 34$, $p < 0.0001$). This balance between response speed and performance accuracy only occurred for the incongruent condition, where fast responses were more likely to be wrong because of a premature response to the spatially contradictory cue.

Incongruent errors, but not congruent errors, produce an adaptive change in behavior

Congruent and incongruent errors had different effects on subsequent behavior, which we studied by investigating the slowing
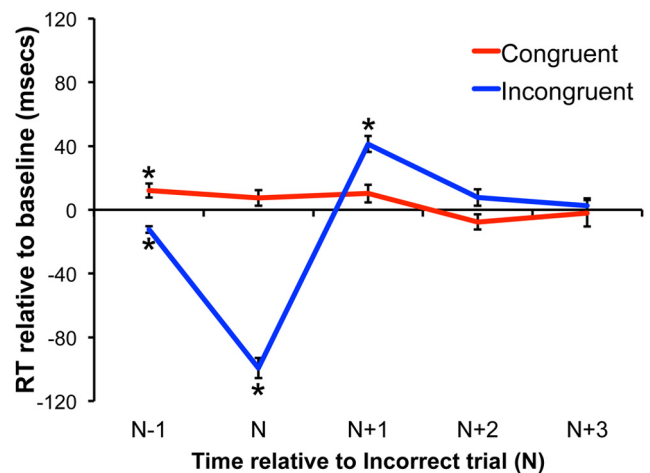

Figure 2. Adaptive behavioral changes after incongruent errors. This figure compares subjects' reaction times on trials around congruent and incongruent errors, relative to their baseline performance. *$p < 0.05$, trials that significantly differed from baseline performance for the preceding 10 correct trials of that type (i.e., congruent or incongruent).

of responses after errors. We tested whether the two conditions displayed different behavioral profiles by performing a 2×2 ANOVA examining the effect of trial type (error trial and posterror trial) and condition (congruent or incongruent) on reaction times. This ANOVA showed a significant interaction between trial type and condition ($F = 110$, $df = 1$, $p < 0.0005$). The interaction was because errors (N) during the incongruent condition were relatively fast compared with baseline ($t = 15.66$, $df = 34$, $p < 0.0005$), and posterror slowing was observed on the next trial ($N + 1$) ($t = 8.11$, $df = 34$, $p < 0.0005$). In contrast, during the congruent condition, errors were not abnormally fast ($p = 0.14$) and showed only a borderline relationship with posterror slowing ($p = 0.075$). Fast responses during the incongruent condition are very likely to result in errors, as acting prematurely on incomplete data is the essential cause of incongruent errors. In contrast, errors during the congruent condition are much less influenced by this as acting on both the premature spatial and later color information will result in the same response. The phenomenon of posterror slowing is a well-established measure of increased cognitive control (Falkenstein et al., 1991; Kerns et al., 2004; Ridderinkhof et al., 2004). Posterror slowing was only seen after incongruent errors (Fig. 2). It should be noted that the trials after incongruent errors were significantly slower than the mean previous trials of that type ($t = 8.11$, $df = 34$, $p < 0.0005$), as well as the correct trial immediately preceding the error ($t = 9.27$, $df = 34$, $p < 0.0005$) and not just the abnormally fast error trial ($t = 14.25$, $df = 34$, $p < 0.0005$). This demonstrates convincingly that the change in behavior does represent cognitive control and not simply a regression to the mean.

Neuroimaging

Congruent and incongruent errors produce similar activation within the SN

The network of brain regions activated by errors was consistent with previous work and included significant activation of the SN (Fig. 3A,B) (Garavan et al., 2002; Ullsperger and von Cramon, 2003; Hester et al., 2004, 2005; Seeley et al., 2007; Ham et al., 2013). Importantly, similar activation was observed in the SN for errors during the congruent and incongruent conditions, despite the distinct effect of these errors on subsequent behavioral adaptation. In addition to activation within the SN, both types of error
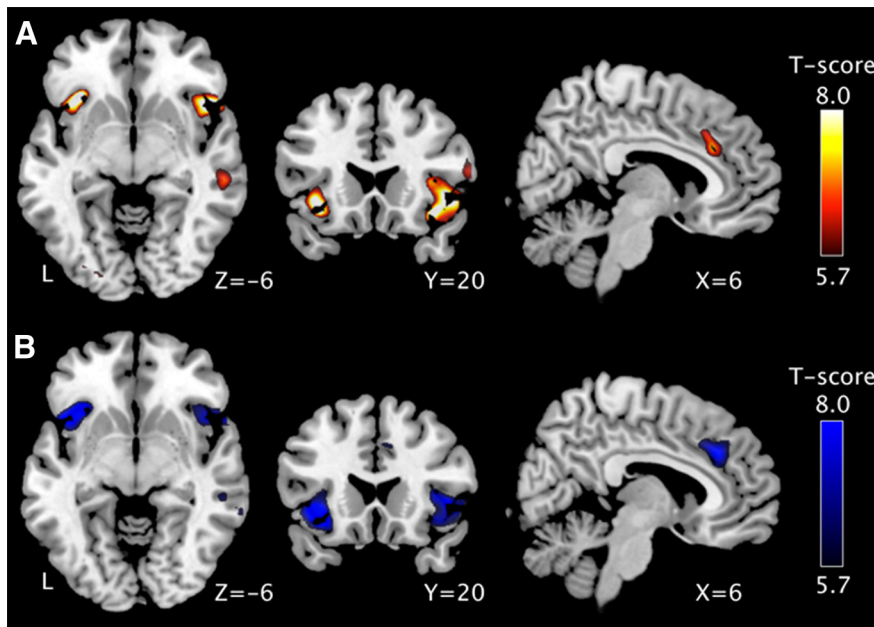
Figure 3. Congruent and incongruent errors activate the SN. **A**, Areas of significant brain activation associated with congruent errors compared with congruent correct trials (red-yellow). **B**, Areas of significant brain activation associated with incongruent errors compared with incongruent correct trials (light-dark blue). All images are thresholded ($p < 0.05$, FWE). Results are superimposed on the MNI 152 T1 1 mm brain template.

Table 1. FWE analysis of contrasts[a]

Anatomical region	Peak t-score	MNI coordinates		
		x	y	z
Congruent errors > congruent correct trials				
Right anterior insula/inferior frontal gyrus	10.42	46	16	−5
Left anterior insula/inferior frontal gyrus	9.78	−34	20	−11
Right supramarginal gyrus	8.63	60	40	19
Left supramarginal gyrus	7.99	54	−38	27
Anterior cingulate cortex	7.27	6	28	29
Left planum polare	6.90	−42	−4	15
Left occipital cortex	6.07	−30	−84	−9
Incongruent errors > incongruent correct trials				
Left anterior insula/inferior frontal gyrus	10.90	−34	22	−5
Right anterior insula/inferior frontal gyrus	9.17	44	44	1
Dorsal anterior cingulate cortex	8.51	6	30	33
Left supramarginal gyrus	8.42	−58	−46	31
Left frontal pole	8.05	−30	52	27
Right supramarginal gyrus	7.66	60	−40	29
Brainstem	6.58	2	−10	−25
Right frontal pole	6.19	30	50	25
Congruent errors > incongruent errors				
Left superior parietal lobule	6.74	−24	−44	51
Left occipital cortex	6.56	−38	−68	−17
Incongruent errors > congruent errors				
Nil				
Conjunction analysis: congruent errors > congruent correct and incongruent errors > incongruent correct				
Left anterior insula/inferior frontal gyrus	8.13	−37	22	−5
Right anterior insula/inferior frontal gyrus	7.71	46	22	−9
Dorsal anterior cingulate cortex	6.87	7	32	34
Right supramarginal gyrus	6.57	58	−37	33

[a]Whole-brain cluster analysis of the contrasts of interest. All clusters were thresholded ($p < 0.05$; FWE).

produced peaks of activation in the bilateral supramarginal and angular gyri, compared with correct trials (Table 1). Congruent errors were associated with additional activation in the left planum polare, right superior parietal lobule, and left occipital cor-

tex. Incongruent errors were associated with additional activation within the brainstem, right and left frontal poles. When congruent and incongruent errors were directly contrasted, small regions within the left superior parietal lobule and left occipital cortex regions showed greater activation for congruent errors that survived correction (Table 1). However, no regions survived correction on a contrast of activity in incongruent errors > congruent errors. Activation within the SN was similar for congruent and incongruent errors. The results of the conjunction analysis showed common activity for both types of errors within the dACC and both insulae as well as the right supramarginal gyrus (Table 1).

Dynamic causal modeling analyses
Analysis 1: Input into SN is through the right insula
For both conditions (congruent and incongruent), there was the most evidence for families with input into the aRI (78.1% xp for congruent and 62.5% for incongruent). The next most likely family of models had input into the left insulae (xp of 11.7% for congruent and 22.7% for incongruent). Combining these two families resulted in an xp >80%. Therefore, subsequent analysis focused on these models, and families with other types of input were rejected. (Fig. 4A,C). Within these families, there was no clear single winning model. For the congruent condition, the maximum xp of any model was 2.4%, which fell within the aRI input family. The mean xp across all models was 0.2 ± 0.3%, whereas the mean xp for the aRI input family was 1.2 ± 0.4%. For the incongruent condition, the maximum xp of any model was 2.8%, again within the aRI input family. The mean xp across all models was 0.2 ± 0.3%, whereas the aRI input family mean xp was 1 ± 0.4%. These seemingly small numbers are not surprising as 448 models were tested (64 models in seven input families) and the xp sums to one over all models. If all models had equal evidence, then their xp value would be ~0.2%. As no clear winning model was identified, we proceeded to the next stage of the analysis.

Analysis 2: What are the intrinsic SN connections and are they modulated by errors?
Analysis of the pattern of intrinsic connectivity within the SN demonstrated that the insulae were functionally connected to the dACC via the aRI. For the congruent condition, average connectivity across correct and error trials (represented in the "A" matrix) was significant for four connections: (1) from the aRI to the dACC ($t = 2.76$, $df = 26$, $p = 0.010$); (2) from the aLI to the aRI ($t = 2.18$, $df = 26$, $p = 0.0386$); (3) from the aRI to the aLI ($t = 2.81$, $df = 26$, $p = 0.0093$); and (4) from the dACC to the aRI ($t = 2.09$, $df = 36$, $p = 0.046$) (Fig. 4B, black arrows). For the incongruent condition, only two of these connections were significant: (1) from the aRI to the dACC ($t = 2.95$, $df = 26$, $p = 0.0066$); and (2) from the aLI to the aRI ($t = 2.52$, $df = 26$, $p = 0.0178$) (Fig. 4D, black arrows). For both conditions, there was no direct connection between the dACC and aLI.

We next investigated whether the connection strengths between nodes of the SN were modulated by errors. For the incongruent condition, the connection from the dACC to aLI

significantly increased in strength after errors ($t = 2.81$, $df = 26$, $p = 0.008$; Fig. 4D, red arrow). For the congruent condition, no modulation of any connection strengths were observed.

Analysis 3: Are there distinct effects of errors in the congruent and incongruent conditions, and do these relate to behavior? Finally, we directly compared the effect of errors on the connection strengths of the SN in the congruent and incongruent conditions. This confirmed a specific modulation of the dACC to aLI connection during the incongruent condition ($t = 2.65$, $df = 24$, $p = 0.014$ for the comparison of incorrect > correct across the two conditions). As only incongruent trials were associated with posterror slowing, we next tested whether the change in this connection's effective connectivity related quantitatively to how much behavior changed after the error. Averaged across all errors for every subject, the change in effective connectivity strength of the dACC to aLI connection positively correlated with the extent of posterror slowing (Spearman correlation coefficient = 0.392, $p = 0.035$, two-tailed). This suggests that, in situations in which behavioral adaptation is required, the dACC shows increased interaction with the left insula, which correlates with the subsequent change in behavior.

Figure 4. Inputs and connections of the SN. The likelihood of families of models with input into specific nodes or combinations of nodes are illustrated for congruent (*A*) and incongruent (*C*) conditions. The *xp* for each family are shown after Bayesian Model Averaging. Schematic representations are shown of the winning models for congruent (*B*) and incongruent (*D*) conditions. Blue circles represent nodes within the SN; black arrows, significant intrinsic connectivity between nodes; and red arrow, increased effective connectivity between the dACC and aLI after incongruent errors in *D*. Large arrows beneath nodes contain the *xp* that that node provides the input to the network. The numbers next to the arrows represent the intrinsic connection strength (black) and the change in effective connectivity (red). Effective and intrinsic connectivity is expressed in 10^{-3} Hertz.

Discussion

The SN responds to behaviorally important events (Seeley et al., 2007; Menon and Uddin, 2010). Within the SN, the dACC and insulae often coactivate, making it difficult for conventional neuroimaging to clearly identify distinct functional roles (Ullsperger et al., 2010). Using DCM, we expanded upon our previous work (Ham et al., 2013) by investigating the causal interactions of the three main cortical nodes within the SN. Our results suggest that the aRI plays a central role in the SN response to errors. Input into the SN was most likely to come through the aRI, and this was the only node showing intrinsic connectivity to the other two parts of the SN. This is in keeping with the proposal that the aRI acts as a "cortical-outflow hub" regulating activity in other brain regions (Sridharan et al., 2008; Menon and Uddin, 2010).

A large body of previous work suggests the dACC is a key structure involved in cognitive control (Miltner et al., 1997; MacDonald et al., 2000; Botvinick et al., 2004; Holroyd et al., 2004). Electrophysiological work has demonstrated an ERN, which occurs 80–110 ms after an error (Falkenstein et al., 1991). Source localization suggests that the ERN originates in the dACC (Dehaene et al., 1994; Debener et al., 2005), which has led to the conclusion that the dACC provides the first cortical signal used for error detection (for a review, see Falkenstein et al., 2000). Our experimental design allowed us to probe the relationship between adjustments in cognitive control, measured by the amount of posterror slowing, and SN activity. Although input to the SN appears to come through the anterior insulae, our results do support an important role for the dACC in the implementation of
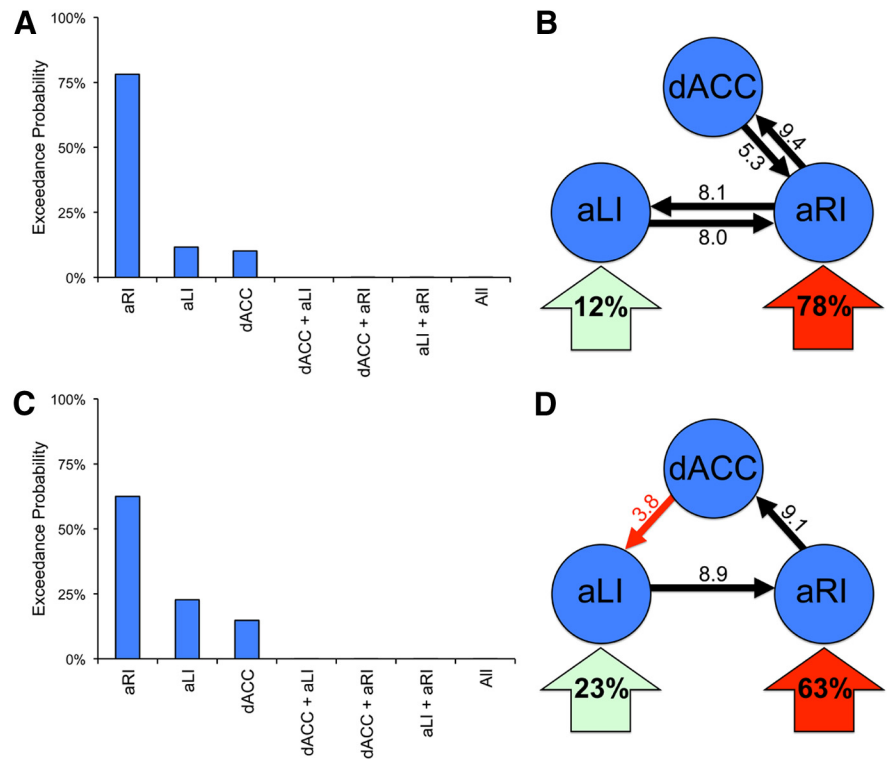
cognitive control. Changes in effective connectivity between the dACC and aLI correlated with posterror slowing, suggesting that the interactions of the dACC are important for moment-to-moment adjustments in behavioral control.

We adopted a hierarchical approach to analyzing the organization of the SN. This involved a stepwise process that builds toward identifying the most likely network configuration. In the initial step, we compared "families" of networks with inputs to different nodes. By far the most likely input route was the through the aRI. The importance of this region's influences on the rest of the network was reinforced by subsequent analysis of intrinsic connectivity. The aRI is known to be structurally and functionally connected to a range of cortical regions involved in various aspects of cognitive control. The aRI is functionally connected to networks responsible for adaptive behavior, including the SN (Seeley et al., 2007), as well as other parts of the frontoparietal control network (Vincent et al., 2008). Tractography evidence shows that the region has direct white matter connections to other key regions within these networks, including the dACC (van den Heuvel et al., 2009), the temporoparietal junction (Kucyi et al., 2012), and the inferior parietal lobe (Uddin et al., 2010). This connectivity makes the anterior insula well placed to perform its putative role of reorienting attention (Ullsperger et al., 2010), evaluating (Eckert et al., 2009) and switching between cognitive resources in response to salient events (Uddin and Menon, 2009).

Tight control of the balance of activity in the SN and default mode network (DMN) appears important for efficient cognitive

function, as rapid deactivation of the DMN is required for focused attention (Weissman et al., 2006). Our group has previously demonstrated that damage to the connection between the aRI and the dACC after traumatic brain injury impaired DMN function (Bonnelle et al., 2012). Specifically, we showed that damage to the white matter tract connecting the aRI to dACC was associated with a failure to control activity in DMN and behavioral difficulty in rapidly switching actions. Consistent with this result, Granger causality analysis has provided evidence that the aRI plays a key role in switching between distinct brain states, which included deactivation of the DMN (Sridharan et al., 2008). The validity of this type of "lag-based" approach to the analysis of causal interaction in fMRI data have recently been called into question because of poor temporal resolution of the data (Smith et al., 2011). Because the DCM procedure our study used does not rely on a lag-based measure, our results provide more robust evidence for the causal interactions of nodes within the SN.

Previously, the amplitude of dACC and lateral prefrontal activation has been related to the extent behavioral adaptation (Gehring and Knight, 2000; Gehring and Fencsik, 2001; Klein et al., 2007; Li et al., 2008). Our work extends these findings by showing a relationship between the change in dACC effective connectivity and the magnitude of behavioral adaptation. During performance of the Simon task, congruent and incongruent errors occur for different reasons and have different behavioral consequences. Incongruent errors usually result from subjects acting upon incomplete information, or "jumping the gun." Subjects respond early to the spatial location of the cue rather than waiting for the cue color, which appears later. Congruent errors cannot be explained in this way, as responding to the location and color of the cue would both produce the same (correct) response. These errors are more likely the result of momentary lapses in attention (Robertson et al., 1997; van Driel et al., 2012). The difference in the cause of each error type is reflected in subsequent behavior. Posterror slowing is only seen after incongruent errors because it is only adaptive to slow behavior after incongruent errors (for a detailed discussion of the paradigm, see Ham et al., 2013).

On average, congruent errors produced no change in effective connectivity and resulted in no change in behavior. In contrast, increased effective connectivity after incongruent errors was associated with significant posterror slowing. Therefore, despite similar levels of SN activation in each condition, changes in effective connectivity within the network encode information about whether behavior needs to be adapted to improve subsequent performance. The observed increase in effective connectivity between the dACC and aLI for incongruent but not congruent errors suggests that the dACC exerts greater influence over the aLI when increased cognitive control is required.

There are different types of behaviorally salient errors, and not all activate the SN (Holroyd et al., 2009; Ham et al., 2013). In the current study, we investigated errors that can be viewed as internally signaled because explicit external feedback is not provided. In a previous study, we investigated the timing errors, which occur when slow responses result in unexpected external feedback (Ham et al., 2013). The SN was not activated by these timing errors, even though they produced clear behavioral change on subsequent trials. However, a similar increase in activity was observed for both error types in the pars operculari, a region adjacent to the anterior insulae. This result suggested that the SN is not required for all types of cognitive control adjustment, and that future work should examine causal interactions between the

pars opercularis and the rest of the SN in a variety of situations where increased cognitive control is required.

One potential limitation of this study is that we focused our analysis on the main cortical nodes within the SN (Dosenbach et al., 2007; Seeley et al., 2007). Other brain regions are activated by errors, but these were not included in the analyses. We think that this focus is justified because the SN forms such an important component of the error processing system, and increasing the number of nodes in DCM analyses has significant computational consequences that make comprehensively analyzing the total model space problematic. This approach is supported by the conjunction analysis of responses to congruent and incongruent errors, which shows that common activation to different types of errors is present in only one region outside the SN (the right supramarginal gyrus). It is nevertheless clear that brain regions outside the SN make important contributions to error processing, including brainstem dopaminergic and thalamic nuclei (Hollerman and Schultz, 1998). Future studies should investigate the interaction of the SN with other brain regions, including the contribution of subcortical inputs, although this work will need to carefully control potential artifacts that can confound analysis of brainstem fMRI signals (Limbrick-Oldfield et al., 2012). A further important limitation is that fMRI is insensitive to rapid interactions between network nodes. DCM on fMRI data assess the longer-lasting time-varying properties of the systems (i.e., the "hidden states" produced by changes in effective connectivity, to provide insight into network hierarchy) (Stephan et al., 2010). Future studies could apply DCM to modalities with higher temporal resolutions to study high-frequency interactions.

In conclusion, we provide evidence that the aRI plays a central role in the response of the SN to errors. Input was most likely to come through this node after errors on an attentionally demanding task, and the region was unique in showing strong interactions with all of the rest of the network. We also provide evidence that changes in the effective connectivity of the dACC are important for moment-to-moment adjustments in cognitive control of behavioral response.

References

Bonnelle V, Ham TE, Leech R, Kinnunen KM, Mehta MA, Greenwood RJ, Sharp DJ (2012) Salience network integrity predicts default mode network function after traumatic brain injury. Proc Natl Acad Sci U S A 109:4690–4695. CrossRef Medline

Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci 8:539–546. CrossRef Medline

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280:747–749. CrossRef Medline

Christ S, Falkenstein M, Heuer H, Hohnsbein J (2000) Different error types and error processing in spatial stimulus-response-compatibility tasks: behavioural and electrophysiological data. Biol Psychol 51:129–150. CrossRef Medline

Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK (2005) Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J Neurosci 25:11730–11737. CrossRef Medline

Dehaene S, Posner MI, Tucker DM (1994) Localization of a neural system for error-detection and compensation. Psychol Sci 5:303–305. CrossRef

Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE (2006) A core system for the implementation of task sets. Neuron 50:799–812. CrossRef Medline

Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104:11073–11078. CrossRef Medline

Eckert MA, Menon V, Walczak A, Ahlstrom J, Denslow S, Horwitz A, Dubno JR (2009) At the heart of the ventral attention system: the right anterior insula. Hum Brain Mapp 30:2530–2541. CrossRef Medline

Egner T (2009) Prefrontal cortex and cognitive control: motivating functional hierarchies. Nat Neurosci 12:821–822. CrossRef Medline

Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1991) Effects of cross-modal divided attention on late ERP components: II. Error processing in choice reaction tasks. Electroencephalogr Clin Neurophysiol 78:447–455. CrossRef Medline

Falkenstein M, Hoormann J, Christ S, Hohnsbein J (2000) ERP components on reaction errors and their functional significance: a tutorial. Biol Psychol 51:87–107. CrossRef Medline

Friston KJ, Holmes AP, Price CJ, Büchel C, Worsley KJ (1999) Multisubject fMRI studies and conjunction analyses. Neuroimage 10:385–396. CrossRef Medline

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19:1273–1302. CrossRef Medline

Garavan H, Ross TJ, Murphy K, Roche RA, Stein EA (2002) Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. Neuroimage 17:1820–1829. CrossRef Medline

Gehring WJ, Fencsik DE (2001) Functions of the medial frontal cortex in the processing of conflict and errors. J Neurosci 21:9430–9437. Medline

Gehring WJ, Knight RT (2000) Prefrontal-cingulate interactions in action monitoring. Nat Neurosci 3:516–520. CrossRef Medline

Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (1993) A neural system for error-detection and compensation. Psychol Sci 4:385–390. CrossRef

Ham TE, de Boissezon X, Leff A, Beckmann C, Hughes E, Kinnunen KM, Leech R, Sharp DJ (2013) Distinct frontal networks are involved in adapting to internally and externally signaled errors. Cereb Cortex 13:703–713. CrossRef Medline

Hester R, Fassbender C, Garavan H (2004) Individual differences in error processing: a review and reanalysis of three event-related fMRI studies using the GO/NOGO task. Cereb Cortex 14:986–994. CrossRef Medline

Hester R, Foxe JJ, Molholm S, Shpaner M, Garavan H (2005) Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. Neuroimage 27:602–608. CrossRef Medline

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309. CrossRef Medline

Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom L, Mars RB, Coles MG, Cohen JD (2004) Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. Nat Neurosci 7:497–498. CrossRef Medline

Holroyd CB, Krigolson OE, Baker R, Lee S, Gibson J (2009) When is an error not a prediction error? An electrophysiological investigation. Cogn Affect Behav Neurosci 9:59–70. CrossRef Medline

Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. Science 303:1023–1026. CrossRef Medline

Klein TA, Endrass T, Kathmann N, Neumann J, von Cramon DY, Ullsperger M (2007) Neural correlates of error awareness. Neuroimage 34:1774–1781. CrossRef Medline

Kouneiher F, Charron S, Koechlin E (2009) Motivation and cognitive control in the human prefrontal cortex. Nat Neurosci 12:939–945. CrossRef Medline

Kucyi A, Moayedi M, Weissman-Fogel I, Hodaie M, Davis KD (2012) Hemispheric asymmetry in white matter connectivity of the temporoparietal junction with the insula and prefrontal cortex. PLoS One 7:e35589. CrossRef Medline

Leff AP, Schofield TM, Stephan KE, Crinion JT, Friston KJ, Price CJ (2008) The cortical dynamics of intelligible speech. J Neurosci 28:13209–13215. CrossRef Medline

Li CS, Huang C, Yan P, Paliwal P, Constable RT, Sinha R (2008) Neural correlates of post-error slowing during a stop signal task: a functional magnetic resonance imaging study. J Cognitive Neurosci 20:1021–1029. CrossRef Medline

Limbrick-Oldfield EH, Brooks JC, Wise RJ, Padormo F, Hajnal JV, Beckmann CF, Ungless MA (2012) Identification and characterisation of midbrain nuclei using optimised functional magnetic resonance imaging. Neuroimage 59:1230–1238. CrossRef Medline

MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS (2000) Dissociating

the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835–1838. CrossRef Medline

Medford N, Critchley HD (2010) Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. Brain Struct Funct 214:535–549. CrossRef Medline

Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214:655–667. CrossRef Medline

Miltner WHR, Braun CH, Coles MGH (1997) Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. J Cogn Neurosci 9:788–798. CrossRef

Nelson LD, Bernat EM, Holroyd CB, Gehring WJ, Patrick CJ (2008) Loss and error information impact feedback-locked brain potentials in a gambling task. Int J Psychophysiol 69:208. CrossRef

Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005) Valid conjunction inference with the minimum statistic. Neuroimage 25:653–660. CrossRef Medline

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. PLoS Comput Biol 6:e1000709. CrossRef Medline

Rabbitt PM (1966) Errors and error correction in choice-response tasks. J Exp Psychol 71:264–272. CrossRef Medline

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:443–447. CrossRef Medline

Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J (1997) "Oops!": performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 35:747–758. CrossRef Medline

Schofield TM, Penny WD, Stephan KE, Crinion JT, Thompson AJ, Price CJ, Leff AP (2012) Changes in auditory feedback connections determine the severity of speech processing deficits after stroke. J Neurosci 32:4260–4270. CrossRef Medline

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349–2356. CrossRef Medline

Simon JR (1969) Reactions toward the source of stimulation. J Exp Psychol 81:174–176. CrossRef Medline

Simon JR, Berbaum K (1990) Effect of conflicting cues on information processing: the "Stroop effect" vs the "Simon effect." Acta Psychol 73:159–170. CrossRef Medline

Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW (2011) Network modelling methods for FMRI. Neuroimage 54:875–891. CrossRef Medline

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. Proc Natl Acad Sci U S A 105:12569–12574. CrossRef Medline

Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ (2010) Ten simple rules for dynamic causal modeling. Neuroimage 49:3099–3109. CrossRef Medline

Uddin LQ, Menon V (2009) The anterior insula in autism: under-connected and under-examined. Neurosci Biobehav Rev 33:1198–1203. CrossRef Medline

Uddin LQ, Supekar K, Amin H, Rykhlevskaia E, Nguyen DA, Greicius MD, Menon V (2010) Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. Cereb Cortex 20:2636–2646. CrossRef Medline

Ullsperger M, von Cramon DY (2003) Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci 23:4308–4314. Medline

Ullsperger M, Harsay HA, Wessel JR, Ridderinkhof KR (2010) Conscious perception of errors and its relation to the anterior insula. Brain Struct Funct 214:629–643. CrossRef Medline

van den Heuvel MP, Mandl RC, Kahn RS, Hulshoff Pol HE (2009) Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. Hum Brain Mapp 30:3127–3141. CrossRef Medline

van Driel J, Ridderinkhof KR, Cohen MX (2012) Not all errors are alike: θ

and α EEG dynamics relate to differences in error-processing dynamics. J Neurosci 32:16795–16806. CrossRef Medline

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. J Neurophysiol 100:3328–3342. CrossRef Medline

Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. Nat Neurosci 9:971–978. CrossRef Medline