Differentiating the Functional Contributions of Resting Connectivity Networks to Memory Decision-making: fMRI Support for Multistage Control Processes

Ravi D. Mill[1], Ian Cavin[2], and Akira R. O'Connor[3]

Abstract

■ Neural substrates of memory control are engaged when participants encounter unexpected mnemonic stimuli (e.g., a *new* word when told to expect an *old* word). The present fMRI study (n = 18) employed the likelihood cueing recognition task to elucidate the role of functional connectivity (fcMRI) networks in supporting memory control processes engaged by these unexpected events. Conventional task-evoked BOLD analyses recovered a memory control network similar to that previously reported, comprising medial prefrontal, lateral prefrontal, and inferior parietal regions. These were split by their differential affiliation to distinct fcMRI networks ("conflict detection" and "confirmatory retrieval" networks). Subsequent ROI analyses clarified the functional significance of this connectivity differentiation, with "conflict" network-affiliated regions specifically sensitive to cue strength, but not to response confidence, and "retrieval" network-affiliated regions showing the opposite pattern. BOLD time course analyses corroborated the segregation of memory control regions into "early" conflict detection and "late" retrieval analysis, with both processes underlying the allocation of memory control. Response specificity and time course findings were generalized beyond task-recruited ROIs to clusters within the large-scale fcMRI networks, suggesting that this connectivity architecture could underlie efficient processing of distinct processes within cognitive tasks. The findings raise important parallels between prevailing theories of memory and cognitive control. ■

INTRODUCTION

Cognitive control enables flexible and adaptive memory use (Johnson, Hashtroudi, & Lindsay, 1993) and is preferentially heightened when expectations of encountering *old* or *new* stimuli are not met in the environment. In support of this, a recent fMRI study manipulated memory expectation and content by providing participants with Posner-like anticipatory cues (e.g., "likely old"; Posner, Snyder, & Davidson, 1980) that were mostly valid in predicting the mnemonic status of ensuing test probes (O'Connor, Han, & Dobbins, 2010). Using this "likelihood cueing" paradigm, the authors found that brain regions previously linked with the recovery of content for *old* probes ("retrieval success"; Vilberg & Rugg, 2008) showed greater activation when participants correctly rejected invalid cues relative to correctly endorsing valid ones. This "invalid cueing" effect prominently recruited prefrontal and parietal brain regions and occurred for both *old* and *new* probes. O'Connor et al.'s findings contribute to growing research scrutinizing the dedicated neural system underlying the engagement of memory control. However, the precise control subprocesses mediated by this neural system, as well as how these subprocesses interact with core retrieval processes in reaching a final memory decision, are in need of further elucidation.

The likelihood cueing paradigm is uniquely suited to address these aims, as it operationalizes memory control mechanisms that are left unconstrained in standard single item recognition. Control demands are reliably heightened in invalid cue trials, given the need to resolve the response conflict between the cued expectation and the probe-provoked memory analysis. Indeed, the prefrontal and parietal regions of the invalid cueing network identified in the O'Connor et al. paper have been repeatedly linked with the resolution of response conflict in fMRI studies in nonepisodic tasks, including medial pFC (mPFC; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), lateral pFC (LPFC; Koechlin, Ody, & Kouneiher, 2003), and inferior parietal lobule (IPL; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002). Activation in these regions is often elevated in parallel (Kerns et al., 2004; Botvinick, Braver, Barch, Carter, & Cohen, 2001), yet the potential for each to mediate different control subprocesses has been conjectured (Ridderinkhof et al., 2004; Miller & Cohen, 2001). One prominent account posits the need for two controlled processes to resolve response conflict: the initial detection of conflict followed by the allocation of control (Botvinick et al., 2001). However, direct fMRI evidence for regional differentiation among

[1]Rutgers University, [2]Ninewells Hospital and Medical School, Dundee, UK, [3]University of St Andrews

Journal of Cognitive Neuroscience 27:8, pp. 1617–1632
doi:10.1162/jocn_a_00808

these two functions has been mixed (Brass & von Cramon, 2002; cf. MacDonald, Cohen, Stenger, & Carter, 2000).

The equivocation in the fMRI control literature is contrary to the considerable electrophysiological evidence favoring a two-process control framework. A number of ERPs have been linked with aspects of control, including the P300 (detection of unexpected stimuli; Sutton, Braren, Zubin, & John, 1965) and the N200 (preresponse conflict; Näätänen & Gaillard, 1983), which are maximal at frontal and parietal electrode sites. Conjunct elicitation of these ERPs, such as the N2–P3 complex (Squires, Wickens, Squires, & Donchin, 1976), and their distinct spatiotemporal subcomponents, such as the early frontal N2b/P3a and later parietal N2c/P3b (Suwazono, Machado, & Knight, 2000), highlight the enmeshed nature of these ERP correlates of control. These "early" and "late" ERPs putatively map onto the conflict detection and control allocation processes implicated in the resolution of response conflict (Polich, 2007). Nevertheless, the potential involvement of these processes in the resolution of mnemonic conflict as induced by expectancy violations in the likelihood cueing paradigm has yet to be investigated.

Furthermore, the distributed topographies of ERPs underpinning the proposed two processes of conflict resolution highlight a limitation of fMRI studies of control in both memory and nonmemory domains. fMRI attempts at ascribing more specific control functions to frontoparietal regions have been hampered by a predominant reliance on task-evoked analytic methods that summarize activation changes in isolated brain regions. Such approaches obliquely support a "modular" view of cognition and, concordant with a wider transition toward a "network" view (Raichle, 2010), a greater emphasis on functional connectivity analyses might prove fruitful in refining neural models of control. Indeed, the distinct patterns of anatomical interconnectivity between subregions of mPFC and LPFC (Petrides & Pandya, 1999) and between LPFC and IPL (Cavada & Goldman-Rakic, 1989) suggest that control subprocesses might arise from the dynamics of large-scale networks in which these cortical regions are nested, rather than from isolated activation changes.

In this regard, analyses of functional connectivity in the resting-state (fcMRI) have already proved useful in formalizing brain networks underpinning memory and control processes. By computing the spontaneous BOLD correlations among brain voxels while participants were at rest (fcMRI), a number of studies have outlined a convergent network of prefrontal and parietal regions that elevate when on-task control demands increase (Yeo et al., 2011; Nelson et al., 2010; Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). This "conflict detection" network (also termed the "frontoparietal control" network) has been distinguished from another fcMRI network comprising distinct prefrontal, parietal, and medial-temporal lobe regions that is more directly involved in retrieval (also termed the "hippocampo-cortical" network and referred to hereafter as a "confirmatory retrieval" network; Yeo et al., 2011; Nelson et al., 2010; Vincent et al., 2006). Indeed, previous variants of the likelihood cueing paradigm have highlighted the selective connectivity of particular invalid cueing regions to similar "conflict" and "retrieval" networks (O'Connor et al., 2010). It is the aim of the present experiment to interrogate the precise functional significance of this differential network affiliation among invalid cueing regions and its correspondence with the previously described detection/allocation framework implicated in the resolution of response conflict. Such clarification would not only aid our understanding of task-evoked memory control activations but also yield reciprocal insight into the properties of the large-scale resting state networks from which these activations emerge.

We hence adopted a multianalytic approach, wherein analyses of task-evoked amplitudes and regional time courses were complemented by analyses of resting state functional connectivity. To enable a more precise elucidation of task-evoked activations, we incorporated two novel manipulations to the likelihood cueing paradigm: cue strength and response confidence. First, cue strength was varied by the presentation of "strong" old-worded ("likely old" and "unlikely old") and "weak" new-worded cues ("unlikely new" and "likely new"). Prior evidence suggests that old-worded cues instil stronger expectations than new-worded ones, even when both cue types are equally valid in predicting probe status (Dobbins, Jaeger, Studer, & Simons, 2012). Given that cues in the Dobbins et al. (2012) study were only of the type "likely new" or "likely old," we instantiated a more rigorous test of the relationship between cue wording and resultant expectation in an independent behavioral study involving both "likely" and "unlikely" variants of the old- and new-worded cues (see Results). The behavioral study validated our cue strength assumptions for the fMRI study, which held that invalid cueing regions also sensitive to cue strength are ideally disposed for the "early" detection of mnemonic expectation violations, given their unique access to cued expectations at the time to-be-judged probes appears. Second, we solicited response confidence to identify invalid cueing regions sensitive to confirmatory retrieval processes engaged after the expectancy violation has been detected. Correct invalid cueing decisions rendered with high confidence are likely to reflect the satisfactory result of controlled memory processing undertaken subsequent to an expectancy violation, as supported by prior links between high confidence and successful memory processing (Kim & Cabeza, 2007). These task manipulations aided the segregation of the invalid cueing network into regions affiliated with broader "conflict" and "retrieval" networks.

METHODS

Participants

Eighteen right-handed, native English-speaking participants (12 women; age 19–24 years) were included in the

main analysis. Two additional participants were excluded because of (i) technical difficulties during fMRI data acquisition and (ii) failure to complete all experimental runs. Two participants did not use the low confidence response option as instructed and were excluded from the confidence analyses (one gave 100% and the other 99% high confidence responses). Informed consent was obtained in a manner approved by the University Teaching and Research Ethics Committee at the University of St Andrews and the Tayside Committee on Medical Research Ethics, Ninewells Hospital and Medical School.

Stimuli and Procedures

Participants first underwent structural scans and completed practice versions of the experimental tasks. This was followed by two resting scans either side of four on-task scans, comprising one run of a standard single-item recognition task and three runs of a likelihood cueing recognition task. For the unscanned study phase of the single-item recognition run, participants counted the syllables of 56 words (see Figure 1). The scanned test phase followed immediately, and participants rated 112 word probes (56 studied *old* probes and 56 *new* probes) as "old" or "new" and gave their confidence in this judgment ("high" or "low") with a single response. This run merely acted to familiarize participants with the recognition task before cues were introduced and is not analyzed further.

The cued runs used a variation of the likelihood cueing task in which participants make recognition judgments for single word probes presented alongside cues to their likely mnemonic status (see O'Connor et al., 2010). The study phase procedure was identical to that used in the single-item recognition run. In the ensuing scanned test phase, each probe was preceded by a cue that suggested that the probe was either "old" ("likely old" or "unlikely new") or "new" ("likely new" or "unlikely old") with an accuracy of 75%. Cue-probe jitter ranged from 0.5 to 4.5 sec in increments of 0.66 sec. Participants were fully informed of the cue accuracy probabilities but were not instructed as to which cues were likely strong or weak. Response mappings were made using a four-button response box. For each participant, a different set of words was randomly drawn from a pool of 2001 singular, common nouns from the English Lexicon Project (minimum log Hyperspace Analogue to Language frequency, 8.02; Balota et al., 2007).

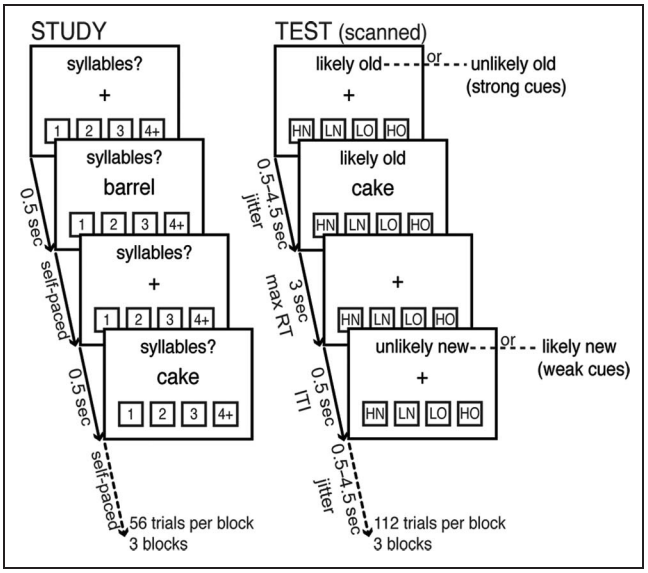

Figure 1. Design schematic for likelihood cueing recognition paradigm. Each study phase comprised a self-paced syllable counting task for 56 *old* words. The cued test phase followed immediately, with participants making "old" or "new" decisions for 112 word probes (56 studied *old* words and 56 unstudied *new* words) and registering their confidence in that decision with a single response. Response options included "HN" (high confidence new), "LN" (low confidence new), "LO" (low confidence old), and "HO" (high confidence old). Note that abbreviations of available responses options are provided here for illustrative purposes, and the response prompt shown to participants actually comprised an image of the four-button response box with "new" and "old" labels to the left and right, respectively, overlaid by directional arrows to denote variations in confidence. Cues to the likely mnemonic status of ensuing probes were presented at the start of each trial. These were randomly intermixed in each run and were of four types according to cue strength (strong old-worded cues and weak new-worded cues) and the response suggested (old-suggesting and new-suggesting cues). Participants were informed that cues were correct (i.e., valid) in predicting probe status on 75% of trials. This led to instances where the cue gave incorrect (i.e., invalid) predictions of probe status on 25% of trials. Probe onset was jittered from cue onset within a range of 0.5–4.5 sec in increments of 0.66 sec. Participants had 3 sec in which to respond ("max RT") and the intertrial interval (ITI) was 0.5 sec.

fMRI Acquisition and Preprocessing

Scanning was performed on a 3T Siemens Trio whole-body MRI scanner (Siemens Medical Solutions, Erlangen, Germany) using a standard 12-channel receive-only whole-head coil. On-task functional data were acquired using a descending echo-planar pulse sequence (repetition time = 2000 msec, echo time = 30 msec, 90° flip angle, 35 axial slices parallel to the AC–PC plane with 3.5 × 3.5 × 4 mm voxels, no interslice gap). Head motion was minimized using foam padding. The two resting state scans were carried out with participants fixating on a cross for the duration of each 6-min session. fcMRI images were acquired using a sequence with parameters identical to the on-task functional sequence. All BOLD data were processed with SPM8 (Wellcome Department of Imaging Neuroscience, London). Slice acquisition timing correction was carried out by temporally resampling relative to the middle slice collected, followed by rigid body motion correction. Functional volumes were then spatially normalized to a canonical echo-planar template using 12-parameter affine and cosine basis transformations and resampled to 3-mm isotropic voxels. Volumes were then spatially smoothed with a 6-mm Gaussian kernel.

fMRI Task-evoked Amplitude Summary Analysis

The amplitude summary analysis is the traditional method of rapid event-related fMRI analysis in which participants are treated as a random effect and volumes as a temporally correlated time series. Summary amplitudes were modeled by convolving a canonical hemodynamic response function with a series of delta functions marking the onset of each condition of interest. Cues were modeled as 0-sec duration events, and memory probes were modeled as 3-sec duration epochs from their respective onsets. Incorrect responses were grouped into a single variable of no interest and not further considered. The β parameter estimates of the best-fitting canonical hemodynamic response function for each condition were used in pairwise contrasts and stored as a separate image for each participant. These contrast images were tested against the null hypothesis of no difference between contrast conditions using one-tailed, repeated-measures t tests. The initial whole-brain invalid cueing contrast (invalid > validly cued probes) was thresholded at $p < .001$ (uncorrected) for five contiguous voxels; a typical threshold for recognition memory research. This contrast collapsed across cue strength and disregarded response confidence (high and low confidence correct responses were modeled together, and incorrect responses as a variable of no interest), meaning that the general effects of invalid cueing were unbiased with respect to cue strength and response confidence. The sensitivity of invalid cueing regions to these latter manipulations was then explored in ROIs inclusive to both the task-evoked invalid cueing effect and the functional connectivity analyses described below.

fcMRI Resting State Connectivity Analyses

Resting state functional connectivity (fcMRI) was examined by entering time courses extracted from 8-mm diameter seed ROIs (using the MARSBAR toolbox for SPM8; Brett, Anton, Valabregue, & Poline, 2002) as covariates of interest alongside 18 sources of nonspecific variance (six movement parameters; signal from spheres in the left lateral ventricle, in left hemisphere deep cerebral white matter and averaged across the whole brain; and the nine first derivatives of these covariates) in a general linear model. Seeds used to recover the "conflict detection" and "confirmatory retrieval" networks were centered on posterior and anterior maxima from the invalid cueing contrast, as in O'Connor et al. (2010).[1] The "conflict" network used the most posterior mPFC maximum ([0, 23, 52] in MNI space; BA 8), whereas the "retrieval" network used the most anterior mPFC maximum ([3, 53, 43] in MNI space; BA 9). fcMRI analyses were collapsed across pretask and posttask connectivity runs after model-free principal component analyses failed to find any differences across these runs. Resulting maps were thresholded identically to the on-task data and depict areas whose activation reliably covaries with the seed region, on a scan-by-scan basis, after nonspecific effects have been controlled. We also conducted analyses of amplitude and time course properties of aggregated clusters within each fcMRI network, which are detailed in the Functional Heterogeneity of the Invalid Cueing Response: Network-level Cluster Analysis section under Results.

Online Behavioral Validation Study

An online study was conducted in an independent sample of 202 participants (138 women, 62 men, 2 did not report sex; age range = 18–70 years, mean = 29.7 years, 4 did not report age) to validate the manipulation of cue strength by cue wording. This was conducted over the same period as the fMRI study, with informed consent procedures approved by the University Teaching and Research Ethics Committee at the University of St Andrews.

The online study comprised a single modified run of the likelihood cueing task used in the fMRI study. It was coded in JavaScript and presented to participants via their Internet browsers. In the study phase, participants counted the syllables of 60 words. The test phase followed immediately, in which a cue preceded each of 120 word probes (60 *old* and 60 *new*) by 1 sec. The cues were "likely old," "unlikely new," "likely new," "unlikely old," and a neutral cue ("?") and were presented with equal frequency. In all cases, cues were random and not predictive of the mnemonic status of the ensuing probe. Participants were not explicitly informed that the cues were uninformative and were instructed to "use this advice to help you make your mind up about whether or not you recognize each word." At test, they judged the mnemonic status of the probes and the confidence in this judgment ("sure old," "probably old," "guess old," "guess new," "probably new," "sure new"). Responses were self-paced and made using the mouse.

The aim of the online study was to verify whether old-worded cues ("likely old" and "unlikely old") led to stronger expectations than equivalent new-worded ones ("unlikely new" and "likely new"). Strength of expectation was assessed via response bias "c" estimated from the equal variance signal detection model (corrected for errorless responding; Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988). Greater deviations in c from 0 indicate greater bias, with sign denoting its direction, such that negative c values reflect a greater bias toward responding "old." To isolate the effects of cue wording on c, the validity manipulation was removed in the online study, that is, cue validity was set at chance 50% levels for all cue types. This modification was important in validating inferences of cue strength in the fMRI experiment, which were based solely on cue wording and the associated ease of constructing cues to aid recognition responding. Evidence that cue wording led to differing degrees of bias even when all cues were as uninformative as each other would provide clear evidence of the hypothesized relationship between

cue wording and strength. Additionally, a neutral cue ("???") was included to measure the cue-driven shift in expectation from an appropriate baseline (given that criterion placement in the neutral cue condition captured any uncontrolled sources of bias that varied between participants). The strength of cued expectation was hence calculated as the absolute deviation in criterion placement in each cue condition from that observed in the neutral "???" condition (i.e., adjusted bias for "likely old" and "unlikely new" cues was calculated as neutral *c* minus cued *c*, adjusted bias for "unlikely old" and "likely new" cues was cued *c* minus neutral *c*).

RESULTS

Behavioral Validation of Manipulations

Invalid Cueing

Table 1 provides a summary of all ensuing results sections. We first confirmed that participants in the fMRI study were incorporating cues into their memory evaluations. A 2 (Mnemonic status: *old* or *new*) × 2 (Cue validity: valid or invalid) repeated-measures ANOVA was conducted on accuracy (see Figure 2A). There were significant main effects of Cue validity, $F(1, 17) = 44.16$, $p < .001$, $\eta_p^2 = .722$, and Mnemonic status, $F(1, 17) = 5.65$, $p = .029$,

Table 1. Summary of Presented Analyses and Results

Results Section	*Analysis Details*	*Summary*
1. Behavioral Validation of Manipulations	Decision accuracy under invalid cueing, response bias across cue strength, and decision accuracy under ratings of response confidence	Invalid cues reduced decision accuracy; strong cues (i.e., "old"-worded cues) led to greater shifts in response bias than weak cues (i.e., "new"-worded cues); and higher confidence increased accuracy
2. fMRI Task-Evoked Amplitude Analysis: Recovering the Invalid Cueing Network	Task-evoked whole-brain amplitude contrast: invalid > valid cue trials	Invalid cues increased activation in an established network prominently encompassing medial frontal, lateral frontal, and parietal regions
3. fcMRI Resting State Connectivity Analysis	Whole-brain correlation of mPFC invalid cueing seed regions with all other voxels in the brain, computed during the pretask resting condition	Invalid cueing regions were split by their affiliation to two fcMRI networks: "conflict detection" and "confirmatory retrieval" networks
4. Functional Heterogeneity of the Invalid Cueing Response: ROI Analysis	Invalid cueing amplitude differentials (invalid–valid, ICDs) and activation time courses were calculated for ROIs identified in both the task-evoked invalid cueing effect (Section **2**) and the fcMRI networks (Section **3**) and compared across cue strength and confidence conditions	Invalid cueing ROIs affiliated with the "conflict" fcMRI network were sensitive to manipulations of cue strength (strong > weak) and peaked "early," whereas those affiliated with the "retrieval" fcMRI network were sensitive to confidence (high > low) and peaked "late"
5. Functional Heterogeneity of the Invalid Cueing Response: Network-level Cluster Analysis	Extracted specificity amplitudes (based on ICDs as in Section **4**) and averaged time courses of clusters within "conflict" and "retrieval" fcMRI networks (defined in Section **3**) were compared across cue strength and confidence conditions	Clusters nested in the "conflict" network were specifically sensitive to cue strength (but not confidence) and peaked "early," whereas clusters nested in the "retrieval" network were specifically sensitive to confidence (but not cue strength) and peaked "late"
6. Supplementary Information (Mill, Cavin, & O'Connor, 2015)	Replication of fcMRI analyses (Sections **3** and **5**) with different "conflict" and "retrieval" network seed pairs taken from two previous studies (Yeo et al., 2011; Nelson et al., 2010)	Invalid cueing regions were similarly split by networks recovered from both these prior seed pairs as the main fcMRI analyses, and network clusters also showed identical amplitude specificity and time course properties (i.e., early peaking "conflict" clusters sensitive to cue strength and late peaking "retrieval" clusters sensitive to confidence)

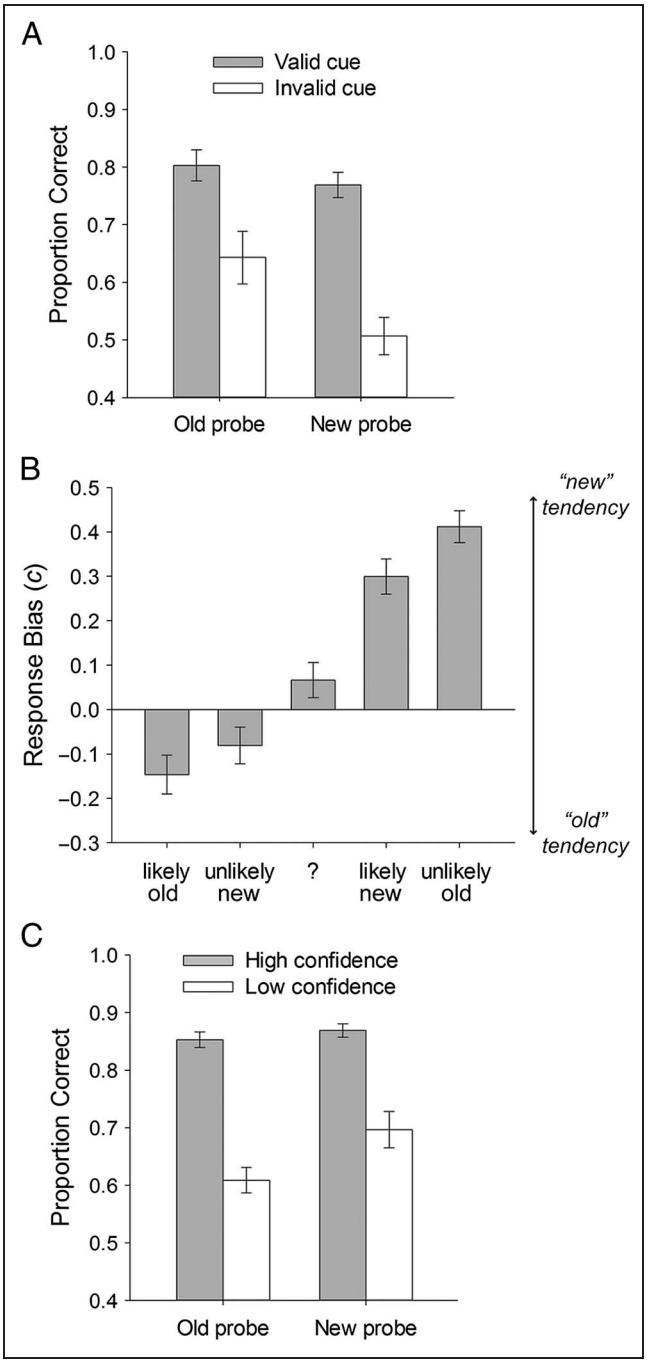

Figure 2. Behavioral validation results. (A) Invalid cueing. Plot of proportion correct by mnemonic status and cue validity. Gray bars represent accuracy for valid cues, and white bars represent accuracy for invalid cues. (B) Cue strength. Plot of response criterion (*c*) estimates from the equal variance signal detection model for cue types presented in the online study: likely old, unlikely new, neutral cue ("?"), likely new and unlikely old. Positive *c* values indicate an increased likelihood of responding "new," and negative *c* values indicate an increased likelihood of responding "old" (as indicated by the arrows to the right of the graph). (C) Response confidence. Plot of proportion correct by mnemonic status and response confidence. Gray bars represent accuracy for high confidence responses, and white bars represent accuracy for low confidence responses. In all plots, error bars represent *SEM*.

η_p^2 = .250, such that valid cues led to greater accuracy than invalidly cues (valid M = .79, SD = .06; invalid M = .57, SD = .12) and *old* probes led to greater accuracy than *new* probes (*old* M = .72, SD = .13; *new* M = .64, SD = .06). A significant interaction suggests that invalid cueing reduces accuracy for *new* probes more than *old* probes, $F(1, 17) = 11.50$, $p = .003$, $\eta_p^2 = .404$. Similar item-wise asymmetries in behavioral cueing effects have been reported previously (Jaeger, Cox, & Dobbins, 2012). Despite the interaction, post hoc t tests recovered invalid cueing effects on accuracy for both *old* probes, $t(17) = 4.71$, $p < .001$, $d = 1.11$, and *new* probes, $t(17) = 7.22$, $p < .001$, $d = 1.70$. As expected, valid cues lead to greater decision accuracy than invalid ones.

Cue Strength

The online study assessed whether "old"-worded cues led to stronger expectations, as indexed by response bias *c*. Figure 2B shows response bias for the five cue types presented. As anticipated, graded *c* parameters were observed such that, relative to the neutral cue ("?" M = 0.06, SD = 0.56), "likely old" and "unlikely new" cues increased the tendency to respond "old" ($M = -0.15$, SD = 0.62 and $M = -0.08$, SD = 0.58 respectively) and the "unlikely old" and "likely new" cues reduced this tendency (M = 0.41, SD = 0.51 and M = 0.30, SD = 0.56 respectively). Absolute bias values were adjusted according to *c* placement in the neutral "???" condition (see Methods for more details) and entered in a 2 (Cue wording: old ["likely old" and "unlikely old"] or new ["unlikely new" and "likely new"]) × 2 (Bias direction: old-suggesting ["likely old" and "unlikely new"] or new-suggesting ["likely new" and "unlikely old"]) repeated-measures ANOVA. A significant main effect of Cue wording confirmed that "strong" old-worded cues shift response bias reliably more than "weak" new-worded cues, $F(1, 201) = 11.27$, $p < .001$, $\eta_p^2 = .053$ (old-worded cue M = 0.28, SD = 0.39; new-worded cue M = 0.19, SD = 0.38). Neither a significant main effect of Bias direction nor an interaction was observed, $F(1, 201) = 2.49$, $p = .116$, $\eta_p^2 = .012$ and $F < 1$, respectively. These findings validate the subsequent classification of old-worded cues as strong cues eliciting greater expectation than the correspondingly weak new-worded cues.[2]

Response Confidence

We also verified that increased response confidence reflects increased accuracy in the fMRI sample, which would render it appropriate as a measure of controlled memory analysis undertaken after an expectation violation. A 2 (Mnemonic status: *old* or *new*) × 2 (Confidence: high or low) repeated-measures ANOVA was conducted on accuracy (see Figure 2C). There was neither a significant main effect of Mnemonic status, $F(1, 15) = 4.18$, $p = .059$, $\eta_p^2 = .218$, nor a significant interaction, $F(1, 15) = 4.50$, $p = .051$, $\eta_p^2 = .231$. Crucially, there was a significant main effect of

Confidence, $F(1, 15) = 129.66$, $p < .001$, $\eta_p^2 = .896$, with high confidence responses associated with greater accuracy (high confidence $M = .86$, $SD = .04$; low confidence $M = .65$, $SD = .07$), which validates its use as an index of fine-grained memory analysis.

fMRI Task-evoked Amplitude Analysis: Recovering the Invalid Cueing Network

The invalid cueing contrast (invalid > validly cued probes; see Figure 3A) yielded extensive activation in pFC, including bilateral dorsolateral pFC (DLPFC; ~BA 9/46), mPFC (~BA 8/9), and bilateral insula and inferior frontal gyrus (~BA 47; see Table 2). Additionally, activation in the IPL was observed bilaterally in supramarginal gyrus (~BA 40), and in the right hemisphere extending into angular gyrus (~BA 7). Relative to previous studies using this paradigm (e.g., O'Connor et al., 2010), the extent of the temporo-parietal invalid cueing response was reduced, although the strong pFC activation remained intact. In subsequent analyses, the task-evoked invalid cueing map presented in Figure 3A was segregated according to resting state network affiliation, and the network-based functional heterogeneity was systematically explored at the regional level.

fcMRI Resting State Connectivity Analysis

We first independently mapped the "conflict detection" and "confirmatory retrieval" fcMRI networks within our sample. The seeds used and the networks recovered are shown in Figure 3B. The two networks are highly similar in extent and threshold to those recovered in previous invalid cueing studies (O'Connor et al., 2010) and stand-alone fcMRI examinations (e.g., Yeo et al., 2011; Nelson et al., 2010; Vincent et al., 2008). The "conflict" network encompassed more lateral frontal regions and extended bilaterally from the posterior midline pFC regions surrounding its seed (~BA 8) through DLPFC and along middle frontal gyrus (~BA 9/46) to frontopolar regions (~BA 10). The posterior aspect of the "conflict" network

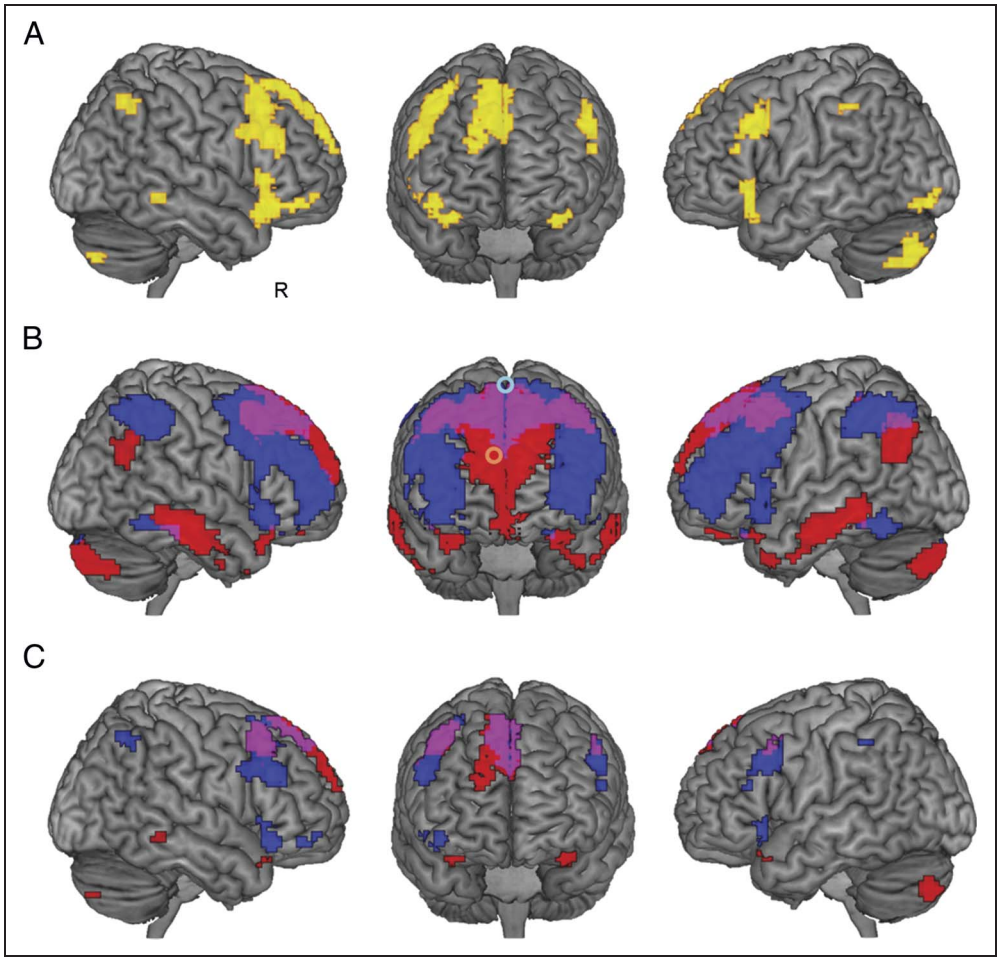

Figure 3. Task-evoked activations recovering the invalid cueing network and resting state network overlap. (A) Regions demonstrating significant activation in the invalid cueing contrast (invalid > valid cue trials; $p < .001$, 5 contiguous voxels). "R" = right sagittal view for all panels. (B) Recovered resting state networks with relevant seed regions marked ($p < .001$, 5 contiguous voxels). The "conflict detection" network (shown in blue) used the posterior local maximum of the mPFC invalid cueing activation (an 8-mm-diameter sphere centered on [0, 23, 52] in MNI space; denoted by the blue ring in the middle panel), whereas the "confirmatory retrieval" network (shown in red) used a seed located in the anterior mPFC maximum (an 8-mm-diameter sphere centered on [3, 53, 43] in MNI space; denoted by the red ring in the middle panel). Overlap between these networks is shown in purple. (C) Regions active in the invalid cueing contrast ($p < .001$, 5 contiguous voxels) masked by their affiliation to the "conflict" (shown in blue; $p < .001$, 5 contiguous voxels) or "retrieval" (shown in red; $p < .001$, 5 contiguous voxels; overlap in purple) resting state networks.

Table 2. Regions Demonstrating Significant Increases in Activation for Correct Responses to Invalidly Cued Probes versus Validly Cued Probes

Region	*Lat.*	*BA*	*x*	*y*	*z*	*Vox.*	*Z Score*
pFC							
IFG/insula	R	47	27	20	−17	250	5.27
IFG/insula	L	47	−33	20	−17	123	4.83
Superior medial/SFG	L/R	8/9/32	9	41	55	463	4.72
MFG/IFG	L	44	−54	20	34	109	4.26
MFG/IFG	R	9/44/46	36	23	55	257	3.93
IFG	L	45	−54	29	19	5	3.51
Cerebellum							
Posterior lobe	L	–	−27	−82	−44	172	4.03
Posterior lobe	R	–	30	−79	−47	19	3.82
Posterior lobe	R	–	12	−79	−29	12	3.51
Occipital							
IOG/fusiform	L	19	−39	−91	−11	24	3.96
Parietal							
AG/SMG	R	7/40	36	−61	49	41	3.63
SMG	L	40	−45	−37	46	8	3.49
Temporal							
MTG	R	21	66	−40	−8	12	3.45

Listed regions are SPM clusters containing at least five significant voxels. *x*, *y*, and *z* coordinates refer to cluster maxima. Lat. = Laterality; BA = approximate Brodmann's area; Vox. = number of significant voxels; IFG = inferior frontal gyrus; SFG = superior frontal gyrus; MFG = middle frontal gyrus; IOG = inferior occipital gyrus; AG = angular gyrus; SMG = supramarginal gyrus; MTG = middle temporal gyrus. Coordinates are in MNI space.

was largely restricted to bilateral supramarginal gyrus (~BA 40) with a small region recruited in a bilateral posterior region of inferior temporal gyrus (~BA 37). The "retrieval" network recruitment in pFC included a large swathe along the midline extending from superior frontal gyrus (~BA 8/9/10) ventrally into anterior cingulate (~BA 32) and laterally into DLPFC (~BA 9). In the posterior aspect, the "retrieval" network recruited precuneus (~BA 31) extending ventrally into posterior cingulate (~BA23). Lateral posterior recruitment to the "retrieval" network included angular gyrus within the IPL (~BA 39) and a region extending down the middle temporal gyrus toward the temporal pole (~BA 20/21).

Overlap between the invalid cueing contrast map and the two connectivity maps is shown in Figure 3C. Notable convergence between the two maps was observed in mPFC, bridging the locations of the two seeds, although neither of the local maxima around which seed ROIs were constructed were in regions of overlap. Task-evoked activation in middle and inferior frontal gyrus was largely restricted to the "conflict" network, although there were pockets of network overlap in bilateral DLPFC (~BA 9). ROIs identified from these overlap maps were used to assess how resting network affiliation of invalid cueing regions relates to their on-task functional heterogeneity.

Functional Heterogeneity of the Invalid Cueing Response: ROI Analysis

To reiterate, the present experiment allowed control demands to be varied and assessed at early and late stages of each trial. Manipulation of cue wording generated differing degrees of bias, reflecting differing strengths of expectation. Cue strength hence served as an index of expectancy-induced response conflict, with strong invalid cues generating greater conflict than weak invalid ones, as a result of the greater bias toward the incorrect response elicited by the former. Furthermore, the confidence measure allowed more targeted identification of brain regions underlying the controlled retrieval analysis than that afforded by a reliance on response accuracy alone.

These manipulations laid the foundation for the ROI analysis, which directly probed how the differential connectivity of invalid cueing regions to the "conflict" and "retrieval" fcMRI networks related to on-task sensitivity to cue strength and response confidence. First, regions

from the invalid cueing map that also overlapped with the "conflict" and "retrieval" networks were identified (see Figure 3C and previous Results section), and their differential responses to cue strength (strong > weak) and confidence level (high > low) were extracted. The first ROI pair consisted of the mPFC ROIs used to seed the fcMRI analyses: the "conflict"-affiliated mPFC region in the posterior aspect [0, 23, 52] (~BA 8) and the "retrieval"-affiliated mPFC region in its anterior aspect [3, 53, 43] (~BA 9; see Figure 4A). The second ROI pair used the left hemispheric LPFC cluster: The "conflict" LPFC region was this time in the most ventral aspect [−54, 20, 34] (~BA 44) and the "retrieval" LPFC region was in the dorsal aspect [−51, 17, 43] (~BA 9/44; see Figure 4B). Note that, although the mPFC ROIs and the ventrolateral pFC ROI were each solely affiliated with the "conflict" network,

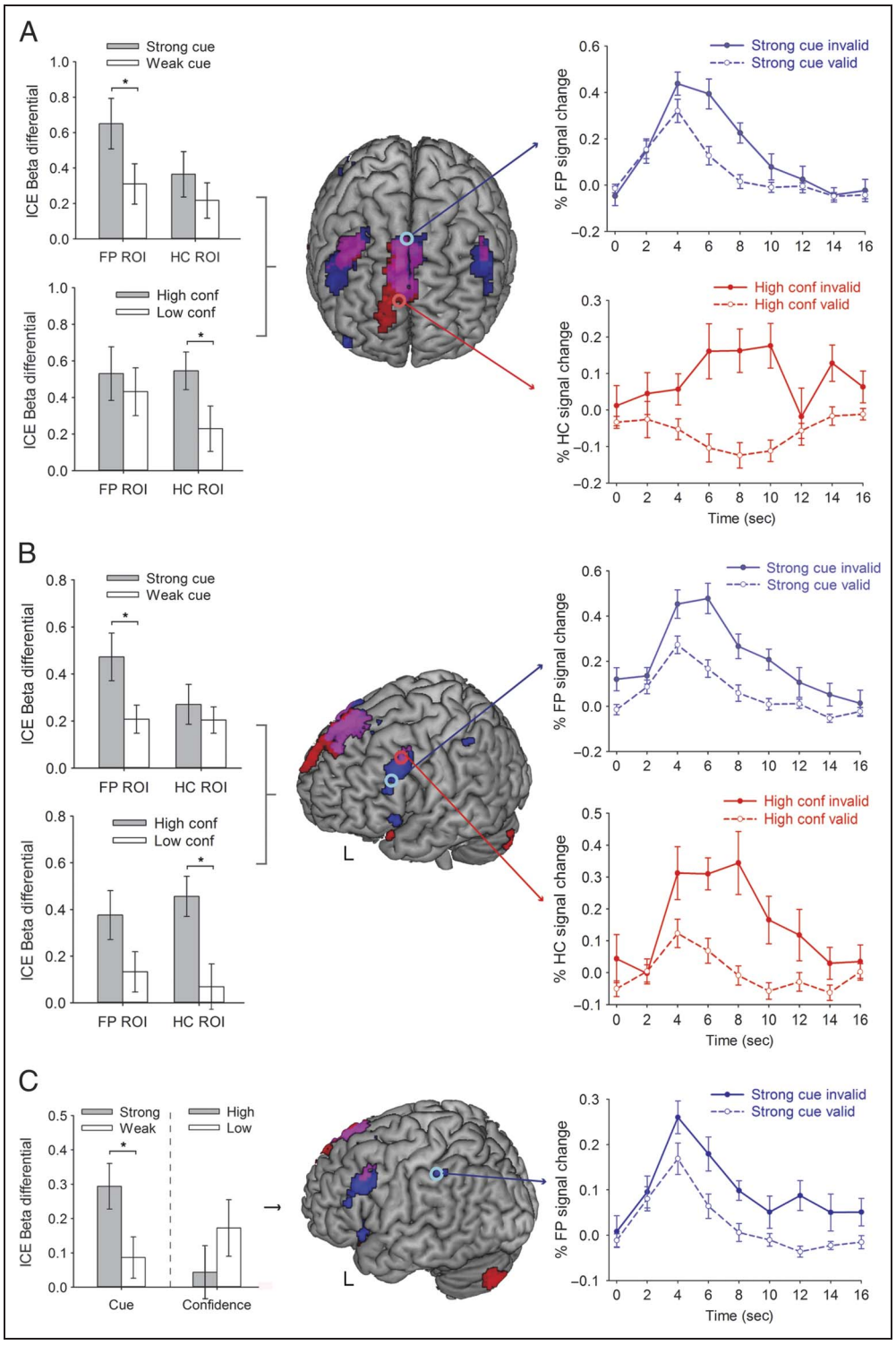

Figure 4. ROI analyses demonstrating functional differentiation of the invalid cueing clusters between cue strength and response confidence. Functional differentiation is shown in (A) mPFC, (B) LPFC, and (C) IPL invalid cueing clusters by their affiliation to the "conflict" or "retrieval" connectivity networks. The left panel shows response conflict sensitivity plots for each region's invalid cueing β differential: cue strength (strong invalid cue − weak invalid cue; for which "conflict" ROIs are selective) above response confidence (high confidence invalid cue − low confidence invalid cue; for which "retrieval" ROIs are selective). Error bars represent *SEM*, and asterisks represent significant conflict sensitivity effects ($p < .05$). The center panel shows the "conflict" ROI (blue ring) and the "retrieval" ROI (red ring) overlaid on the relevant brain sections (note that the IPL cluster was solely affiliated to the "conflict" network). The right panel shows the extracted BOLD time courses for each invalid cueing cluster averaged across participants and plotted for a 0–16 sec portion of the total 24-sec postprobe duration that was extracted. For each invalid cueing region, the "conflict" ROIs' response to strong cues is presented above the "retrieval" ROIs' response to high confidence, with separate lines within each sensitivity plot for valid and invalid cue trials (as conveyed by the plot legends). Time course error bars represent *SEM*. "L" = left sagittal view.

the DLPFC ROI lay within both fcMRI networks. However, this was the only region proximal to the ventrolateral ROI that was both active in the invalid cueing contrast and recruited by the "retrieval" network. This DLPFC ROI was hence included in subsequent analyses as a retrieval fcMRI region. A final invalid cueing ROI was identified in the IPL [−45, −37, 46; ~BA 40], which was wholly affiliated to the "conflict" network (see Figure 4C).

For these invalid cueing ROI pairs, response amplitudes from each "conflict"-affiliated ROI and each "retrieval"-affiliated ROI were extracted by binning the data according to (i) cue strength and (ii) response confidence. Invalid cueing effects were summarized by subtracting validly from invalidly cued probe amplitudes, yielding an invalid cueing differential (ICD). These were compared to establish regions whose amplitudes were more sensitive to the violation of strong than weak cues and regions whose amplitudes were more sensitive to the countermanding of expectation with high than low confidence. Time courses were then extracted from ROIs for contrasts that yielded significant differences in their ICDs to establish whether differences in invalid cueing activations were linked to temporal profiles consistent with "early" and "late" memory control processes.

mPFC Response Amplitudes and Time Courses

Figure 4A depicts ICDs for each mPFC ROI according to cue strength and response confidence. Amplitude differentials for the "conflict"-affiliated mPFC region were significantly greater for strongly cued items ($M = 0.65$, $SD = 0.61$) than weakly cued ones ($M = 0.31$, $SD = 0.48$), $t(17) = 2.33$, $p = .032$, $d = 0.55$. However, no significant difference was observed in the "retrieval"-affiliated mPFC region's response to the different cues (strong cues: $M = 0.36$, $SD = 0.54$; weak cues: $M = 0.22$, $SD = 0.42$), $t < 1$. For the same mPFC ROIs, the opposite sensitivity pattern was observed when responses were binned by response confidence. The "retrieval"-affiliated mPFC region yielded significantly larger ICDs for high ($M = 0.55$, $SD = 0.41$) than low confidence responses ($M = 0.23$, $SD = 0.50$), $t(15) = 2.28$, $p = .038$, $d = 0.57$. Conversely, the "conflict"-affiliated mPFC region showed no difference in its sensitivity to high and low confidence responses (high confidence: $M = 0.53$, $SD = 0.59$; low confidence: $M = 0.43$, $SD = 0.52$), $t < 1$. The same broad region of mPFC that increases when probes are invalidly cued hence shows a posterior-anterior dissociation: the posterior, "conflict"-affiliated region is exclusively sensitive to cue strength and the anterior, "retrieval"-affiliated region is exclusively sensitive to confidence.

According to the control subprocesses hypothesized within the invalid cueing response, the detection of response conflict by the cue-sensitive ROI should be rapid, whereas the controlled analysis of memory evidence underpinned by the confidence-sensitive ROI should be slower. Time courses extracted from the "conflict"-affiliated mPFC ROI sensitive to cue strength and the "retrieval"-affiliated mPFC ROI sensitive to response confidence are presented in Figure 4A. The difference in peak latencies was tested by comparing the time-to-peak for the conflict ROI responsive to invalidly and strongly cued probes and the retrieval ROI responsive to invalidly cued probes to which a high confidence response was given (with the peak defined as the single highest point within the extracted 24-sec time course). The conflict response peaked significantly earlier (4.50 sec, $SD = 1.55$ sec) than the retrieval response (11.88 sec, $SD = 6.75$ sec), $t(15) = 4.22$, $p < .001$, $d = 4.22$. These analyses show that the coupling of subregions within the same task-evoked mPFC cluster to different fcMRI networks reflects dissociable control subfunctions within the invalid cueing response. The time course analyses suggest that this differentiation is reflected in each region's temporal properties, consistent with the hypothesized "early" and "late" control framework.

LPFC Response Amplitudes and Time Courses

Equivalent analyses were conducted on the LPFC ROI pair, with identical outcomes (see Figure 4B). The "conflict"-affiliated ROI's ICD was sensitive to cue strength (strong: $M = 0.47$, $SD = 0.43$; weak: $M = 0.21$, $SD = 0.25$), $t(17) = 2.35$, $p = .031$, $d = 0.55$, but not response confidence (high: $M = 0.37$, $SD = 0.42$; low: $M = 0.13$, $SD = 0.34$), $t < 1$. Conversely, the "retrieval"-affiliated ROI had an ICD insensitive to cue strength (strong: $M = 0.27$, $SD = 0.36$; weak: $M = 0.20$, $SD = 0.24$), $t(17) = 1.68$, $p = .114$, $d = .42$, but sensitive to response confidence (high: $M = 0.46$, $SD = 0.34$; low: $M = 0.07$, $SD = 0.39$), $t(15) = 2.59$, $p = .020$, $d = 0.65$.

The time courses also displayed the same pattern of peak latencies (see Figure 4B), such that the "conflict"-affiliated LPFC region's ICD to strong cues peaked significantly earlier (5.63 sec, $SD = 3.95$ sec) than the "retrieval" region's ICD to high confidence decisions (8.88 sec, $SD = 5.66$ sec), $t(15) = 2.23$, $p = .042$, $d = 0.56$. Identical patterns of connectivity-mediated response sensitivity and activation latency to those shown in the mPFC were therefore also present in the LPFC.

IPL Response Amplitudes and Time Courses

A final ROI analysis examined cue strength and response confidence activations in the IPL region sensitive to invalid cueing. As this region lay solely in the "conflict" network, it was expected that the ICDs would be sensitive to cue strength and not to response confidence. This expectation was confirmed by data presented in Figure 4C, with significant differences according to cue strength (strong: $M = 0.29$, $SD = 0.28$; weak: $M = 0.09$, $SD = 0.26$), $t(17) = 2.69$, $p = .015$, $d = 0.64$, but not response confidence (high: $M = 0.04$, $SD = 0.31$; low: $M = 0.17$, $SD = 0.33$), $t(15) = -1.07$, $p = .301$, $d = -0.27$. The time course was not

subjected to formal analyses as there was no corresponding "retrieval"-affiliated ROI within parietal cortex with which to contrast peak latencies. Nevertheless, examination of Figure 4C reveals an early peak consistent with the other "conflict" network regions previously described (M = 6.66 sec, SD = 4.55 sec).

Influence of Behavioral RTs

An alternative explanation for the time course differences seen in the ROIs is that these simply reflect the different RT profiles for the conditions in question. That is, words cued by strong cues might be responded to faster than words to which high confidence responses were given—a discrepancy which could manifest as a difference in the time taken for the respective time courses to peak. To investigate this, we calculated mean RTs for strongly cued items and items responded to with high confidence and compared them within participants. In fact, we found the opposite difference in RTs: Strongly cued items were responded to significantly more slowly (M = 1.74, SD = .22) than items responded to with high confidence (M = 1.55, SD = .20), $t(15) = 6.35$, $p < .001$, $d = 1.60$. This suggests that a difference in condition RT profiles is unlikely to account for the difference in time course peak latencies between "conflict" and "retrieval" network-affiliated ROIs.

Functional Heterogeneity of the Invalid Cueing Response: Network-level Cluster Analysis

Network Specificity Analysis

We have thus far argued for a functional dissociation between invalid cueing regions according to separate memory control processes mediated by separate fcMRI networks: the initial detection of expectation-induced conflict and later controlled retrieval analysis. Nonetheless the use of a selection of ROIs, however representative they are of their broader networks, necessarily ignores a great deal of data. To overcome this, we complemented the ROI analyses with analyses involving all clusters recovered in the two resting state networks. Once again, we expected "conflict" network clusters to be specifically sensitive to cue strength and not response confidence, with "retrieval" network clusters showing the opposite pattern.

We first defined independent "conflict" and "retrieval" network clusters by exclusively masking each resting state network by the other (each map thresholded at $p < .001$, 5 contiguous voxels). The masking procedure recovered 10 "conflict" clusters and 16 "retrieval" clusters. Response amplitudes were extracted on a participant-by-participant basis for each network cluster and binned according to cue strength and response confidence. Subtraction of validly from invalidly cued amplitudes yielded ICDs for each cluster. We then calculated two parameters to illustrate the specificity of "conflict" and "retrieval" cluster activations (averaged across participants) to conflict detection and controlled respectively: a cue specificity parameter and a confidence specificity parameter. These parameters were calculated from two t statistics based on the participant-wide differences between the network cluster ICDs t_{cue} (the paired-samples t statistic for strong cue ICD [I_{strong}] > weak cue ICD [I_{weak}]; shown in Equation 1 below, where σ represents the standard deviation of the differences between the two ICDs) and t_{conf} (the t statistic for high confidence ICD [I_{high}] > low confidence ICD [I_{low}]; Equation 2). The cue specificity parameter was then calculated by subtracting the absolute value of t_{conf} from t_{cue}, and the confidence specificity parameter was calculated by subtracting the absolute value of t_{cue} from t_{conf}.

$$t_{cue} = \frac{\sum \left(I_{strong} - I_{weak}\right)/n}{\sigma_{\left(I_{strong} - I_{weak}\right)}/\sqrt{n}} \quad (1)$$

$$t_{conf} = \frac{\sum \left(I_{high} - I_{low}\right)/n}{\sigma_{\left(I_{high} - I_{low}\right)}/\sqrt{n}} \quad (2)$$

To illustrate, consider hypothetical Cluster Y, which shows a greater ICD for high than low confidence trials, yielding a t_{conf} of 2.0. This region also displays a weaker ICD for strong than weak cues, yielding a t_{cue} of −1.0, meaning that cluster Y's confidence specificity parameter would be 1.0 and its cue specificity parameter would be −3.0. Cluster Y would hence be deemed confidence specific, although to a lesser extent than would be suggested by examination of the t_{cue} and t_{conf} parameters alone. This demonstrates that a positive specificity parameter reflects a hypothesized amplitude differential in one response factor greater than any amplitude differential (hypothesized or otherwise) in the second factor. These parameters afford a robust test of the proposed functional dissociation of "conflict" and "retrieval" networks.

Figures 5A and 5B render the specificity parameters at the cluster level (based on the averaged cluster ICDs across 16 participants) for the "conflict" and "retrieval" networks, respectively. These figures illustrate that the functional heterogeneity observed in the previously presented ROIs carries over to the network level, such that cue-sensitive clusters are most prominent in the "conflict" resting state network whereas confidence-sensitive clusters are most prominent in the "retrieval" network. A 2 (Specificity parameter: cue or confidence) × 2 (Resting state network: "conflict" or "retrieval") mixed factorial ANOVA for each cluster yielded a significant interaction term only, $F(1, 24) = 12.65$, $p = .002$, $\eta_p^2 = .345$ (the main effects of Sensitivity parameter and Resting state network were $F < 1$ and $F(1, 24) = 1.15$, $p = .294$, $\eta_p^2 = .046$ respectively). Consistent with this functional heterogeneity account, planned comparisons found that, for clusters within the "conflict" network, the cue specificity parameter (M = .95, SD = .90) was significantly greater than

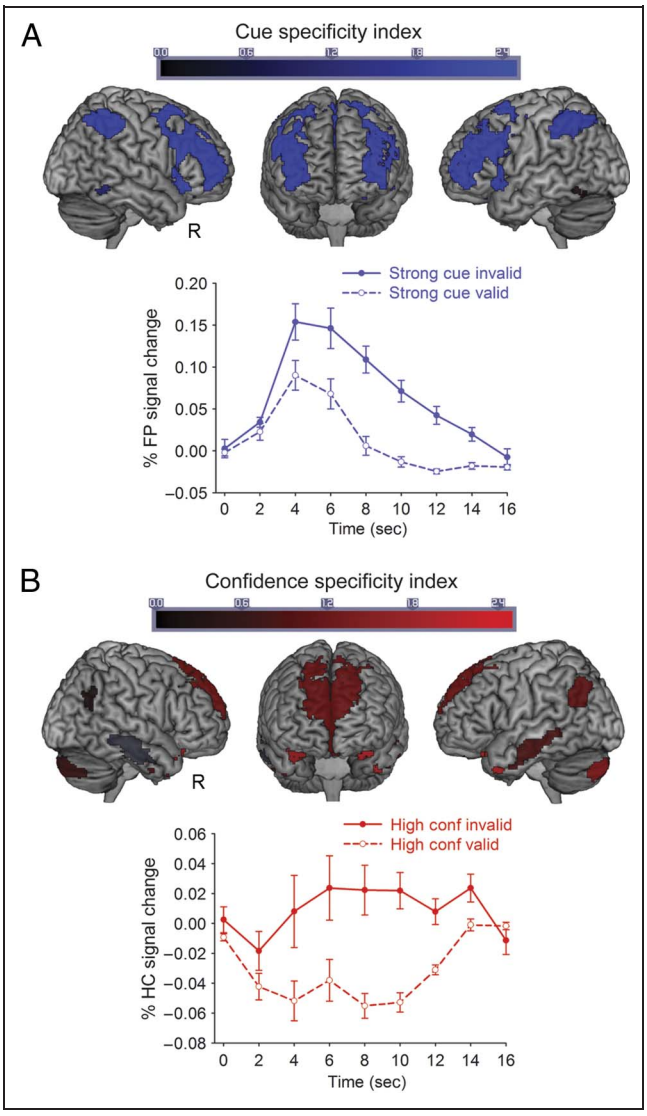

Figure 5. Network-level cluster analyses showing cue strength and response confidence specificities. Network-level activation maps showing (A) cue specificity of the "conflict" network-affiliated invalid cueing response and (B) confidence specificity of the "retrieval" network-affiliated invalid cueing response, as calculated from the ICDs for each network cluster averaged across participants. The color bar in each panel illustrates the color-coding of the overlaid specificity of regions within each network in respectively signaling cue strength and response confidence during invalidly cued trials. The calculation of these specificity parameters is outlined in Equations 1 and 2 in the Results section. The lower portion of each panel plots the relevant averaged network response, across constituent clusters, underlying the relevant invalid cueing specificity, with solid lines denoting the invalid cueing response and dashed lines denoting the valid cue response. Time course error bars represent *SEM*. "R" = right sagittal view.

the confidence specificity parameter ($M = -1.35$, $SD = .95$), $t\ (9) = 3.69$, $p = .005$, $d = 1.38$, while a numerical trend in the opposite direction was observed for clusters in the "retrieval" network (cue specificity: $M = -1.26$, $SD = 1.48$; confidence specificity: $M = .19$, $SD = 1.91$), $t(15) = 1.98$, $p = .067$, $d = 0.50$.

Network Time Course Analysis

As a final parallel to the ROI analyses, we examined time course peak latencies at the network level. Trial time courses were extracted for each cluster within the respective resting state networks, as is shown separately for the "conflict" and "retrieval" networks in the bottom panels of Figure 5A and B. Peak latencies were averaged across participants for each resting state network cluster and subjected to statistical analysis. We could then assess whether the peak latency differences in task-related activations observed in the ROI analysis could be expected over all clusters within an fcMRI network. Once again, we found the anticipated pattern of network heterogeneity, with "conflict" clusters (responding to strong invalid cues trials) showing significantly faster response latencies ($M = 8.40$, $SD = 1.82$) than "retrieval" clusters (responding to high confidence invalid trials; $M = 11.50$, $SD = 2.53$), $t(24) = 3.36$, $p = .003$, $d = 1.35$.

The network cluster analyses confirm that the functional heterogeneity prominent in ROIs selected from "conflict" and "retrieval" fcMRI networks persists at the network level. Furthermore, a supplementary set of fcMRI analyses highlighted the generalizability of our findings, such that similar patterns of network affiliation across invalid cue regions and associated network cluster specificity effects were observed even with fcMRI networks recovered from seeds taken from previous studies (see Supplementary Information [Mill et al., 2015] for further details).

As a final consideration, it is important to note that the cluster-level analyses underestimate the amplitude and time course findings described above within each fcMRI network. Given that we collapsed across clusters of unequal size, we underweighted voxels within both large mPFC clusters relative to the voxels comprising other smaller clusters. These mPFC clusters showed extremes of response (both amplitude and time course) in the hypothesized directions, meaning that voxels chosen at random from within each fcMRI network are likely to be more heterogeneous in their functional responses than would be predicted from these cluster-level data alone. Regardless, the fact that the activation patterns from the ROI analyses persist at the network level supports our assertions of distinct mnemonic control subprocesses mediated by independent resting state networks.[3]

DISCUSSION

Theories of cognitive control typically formalize two core components—detection of the need for control followed by engagement of controlled processing (Botvinick et al., 2001). Nonetheless, direct evidence of separable neural substrates underlying these two processes is rare (cf. MacDonald et al., 2000). In the present experiment, we empirically substantiated just such a segregation as applied to memory control. We employed a modified version of the likelihood cueing paradigm to specify more precise

control subprocesses for regions within the previously reported invalid cueing network (O'Connor et al., 2010). Our manipulations explicitly sought mnemonic analogues of the two key control processes likely involved when participants are invalidly cued: the initial detection of mnemonic expectancy violation (and associated response conflict) and the subsequent engagement of controlled retrieval processing. Conventional task-evoked amplitude analyses were combined with resting state connectivity methods to elucidate the neural substrates of these processes. The findings from these convergent analytic approaches are now discussed with reference to prior research in memory and cognitive control.

Task-evoked amplitude analyses recovered a network of brain regions that elevate for invalidly cued trials over validly cued ones, comprising similar regions of prefrontal and parietal cortices to those previously observed (O'Connor et al., 2010). Constituents of this network included mPFC, LPFC, and IPL—all regions that have been recurrently linked with aspects of controlled processing in memory (Nieuwenhuis & Takashima, 2011; Buckner, 2003) and response conflict (Ridderinkhof et al., 2004; Miller & Cohen, 2001). Much debate has centered on whether these isolated regions underpin dissociable control subprocesses; however, evidence in support of this modular view has been mixed (Brass & von Cramon, 2002; MacDonald et al., 2000). Rather, considering the functional networks to which these regions are connected offers an alternative avenue of clarification. In support of this, we found that the mPFC and LPFC regions of the invalid cueing network were split by their affiliation to either the "conflict" or "retrieval" resting state connectivity networks. This connectivity dissociation among memory control regions is consistent with that observed in a prior likelihood cueing study (O'Connor et al., 2010) and a recent study probing the neural substrates of source monitoring (Barredo, Oztekin, & Badre, 2015).

Our task manipulations then allowed us to directly link the observed connectivity differentiation with an on-task dissociation of the invalid cueing regions by their involvement in the two formalized control processes. We firstly varied cue strength by the presentation of old-worded and new-worded cues, which have been suggested to respectively instil strong and weak mnemonic expectations in prior research (Dobbins et al., 2012), as validated by the independent online study. Regions heightened by invalid cueing at the time when probes appear (to initiate the evaluation of memory) that are also sensitive to the strength of preceding cues are ideally disposed to detect violations of mnemonic expectancy, given their access to both the cued expectation and the probe-provoked memory analysis. We also elicited response confidence to probe neural sensitivity to secondary, confirmatory retrieval processes. Importantly, these manipulations were orthogonal, enabling us to collapse across the alternate sensitivity category for each of the cue strength and response confidence contrasts, and thereby maximized the utility of the unavoidably lower number of invalid cue trials. Consistent with our predictions, we observed an fcMRI-gated dissociation within both medial and lateral prefrontal invalid cueing regions: "conflict"-affiliated regions were sensitive to cue strength but not response confidence, whereas "retrieval"-affiliated regions were sensitive to confidence but not cue strength.

Our manipulations extended prior fMRI investigations in which control demands were assumed to be heightened for "low" compared to "high" confidence decisions in otherwise standard recognition task formats (Fleck, Daselaar, Dobbins, & Cabeza, 2006; Henson, Rugg, Shallice, & Dolan, 2000). In focusing on the low > high confidence contrast, these previous studies relied on a general functional inference as the basis of identifying memory control regions, in the absence of more systematic manipulation of control demands. As noted by Henson and colleagues (2000) themselves, this basic design recovers low confidence-sensitive brain regions that are broadly linked with some form of memory control, without enabling specification of the precise control subprocesses underpinned by these regions. The cueing manipulation employed in this study permitted a more thorough interrogation, such that regions linked with the general heightening of memory control were recovered by the invalid > valid cue contrast and were more precisely characterized through nested analyses of cue strength and response confidence. Indeed, the cueing manipulation led to a reversal of prior confidence contrasts, such that invalid cueing regions that additionally elevate for high compared to low confidence regions were linked with controlled memory processing undertaken after violations of memory expectation. To clarify, satisfactory resolution of the mnemonic conflict instilled by invalid cues requires controlled analysis of memory evidence, the results of which could conceivably have led to increased decision confidence. The combination of cueing and confidence manipulations therefore enabled more precise functional specification of brain regions previously linked in broader terms with aspects of memory control.

Implicit in the proposed distinction between detection and allocation of control is an "early" and "late" temporal ordering of these processes. Our analyses confirmed that, within the prefrontal and parietal ROIs, the response of "conflict"-affiliated subregions to cue strength peaked significantly earlier than the "retrieval"-affiliated subregions' response to confidence. The observed temporal order of control parallels that postulated for contiguously elicited frontoparietal ERPs in the cognitive control literature, with the N2–P3 complex (Squires et al., 1976) and the P3a–P3b complex (Polich, 2007) all highlighted as potential substrates for the early–late control process dichotomy. Indeed, the late P3b has been dissociated from the early P3a by its selective correlation with RT (Conroy & Polich, 2007)—an association between a neural marker and an overt behavioral measure that mirrors our reported sensitivity of the later "retrieval"-affiliated activation with

response confidence. A similar processing dichotomy has been proposed to underlie the controlled processing of language (i.e., the comprehension of syntactically incongruous sentences; Hahne & Friederici, 1999) and even recognition memory (Jacoby, Kelley, & McElree, 1999). Indeed, it is worth highlighting the overlap between the early–late control processes delineated here and the dual processes of familiarity and recollection in recognition memory (Yonelinas, 2002; Henson et al., 2000). However, this characterization fails to capture the potentially broad applications of the reported processes and their neural substrates in signalling control in cognitive domains beyond memory.

Furthermore, the persistence of regional amplitude and time course differences at the network cluster level also yield insight into the general functional properties of the "conflict" and "retrieval" resting state networks themselves. The findings directly link the "conflict" network with the detection of mnemonic response conflict, as would be expected of a network previously observed to be sensitive to heightened control demands in both nonmemory (Vincent et al., 2008) and memory tasks (Spreng, Stevens, Chamberlain, Gilmore, & Schacter, 2010). The "retrieval" network is involved in the engagement of controlled memory analysis subsequent to conflict being detected, consistent with intrinsic connections of its prefrontal/parietal nodes with medial-temporal lobe regions linked with memory encoding and retrieval processes (Kahn, Andrews-Hanna, Vincent, Snyder, & Buckner, 2008; Vincent et al., 2006). Indeed, our findings contribute to an emerging network approach in the study of control, combining traditional task-evoked amplitude analyses with task-evoked methods of network localization, such as structural equation models of effective connectivity (Koechlin et al., 2003) and on-task functional connectivity (Barredo et al., 2015). We report similar prefrontal and parietal network interactions when localizing on the basis of task-independent resting state connectivity methods.

The findings also address the general lack of understanding as to how the dynamics of intrinsic connectivity networks—recovered by task-free correlation methods in the resting state—relate to the actual performance of a cognitive task (Buckner, Krienen, & Yeo, 2013). Indeed, conventional task-evoked fMRI analyses are restricted to the examination of suprathreshold task activations and hence prevent scrutiny of nuanced response patterns of larger networks. As outlined by Jernigan, Gamst, Fennema-Notestine, and Ostergaard (2003), this concealment arises as many regions displaying functionally consistent responding might nevertheless display "subthreshold" task-evoked activations, if the estimated effect sizes fail to reach the adopted criterion for significance. We therefore combined conventionally thresholded analyses with more unorthodox fcMRI-gated analyses (at the ROI and network cluster level) to examine the global distribution of invalid cueing effects across functional networks. The findings support a more general interpretation of the separate control functions subserved by the "conflict" and "retrieval" brain networks, beyond the specifics of the employed task.

Future research involving different imaging methods will be necessary to validate the described network-gated control processes and their underlying temporal dynamics. To this end, the improved temporal resolution afforded by simultaneous EEG-fMRI has already proven beneficial in the study of control (Debener et al., 2005) and would enable a direct test of the speculated correspondence between present fMRI activations and established ERPs. Nevertheless, the present findings highlight the interaction between retrieval-specific and more general "higher-order" processes in constraining evaluations of the past. Further investigation of the overlapping neural correlates of seemingly diverse psychological processes should continue to provide insight into the adaptive and flexible nature of cognition.

Acknowledgments

This work was supported by the Scottish Imaging Network: A Platform for Scientific Excellence (SINAPSE), who provided a PhD studentship to Ravi D. Mill.

Reprint requests should be sent to Akira R. O'Connor, School of Psychology and Neuroscience, University of St Andrews, St. Mary's College, South Street, St. Andrews, Fife, KY16 9JP, Scotland, UK, or via e-mail: aoconnor@st-andrews.ac.uk.

Notes

1. Prior research supports this anterior/posterior seed choice within mPFC, given similar differentiations between these mPFC subregions observed in independent studies of comparable retrieval and control networks (e.g., Yeo et al., 2011; Vincent et al., 2008); the reported relationship between task-evoked activation in posterior and anterior mPFC regions with conflict detection (e.g., Ridderinkhof et al., 2004) and retrieval processing (e.g., Cabeza & St Jacques, 2007), respectively; and the differential anatomical connectivity of anterior mPFC with retrieval-linked medial-temporal lobe regions (Petrides & Pandya, 2007) and posterior mPFC with attention- and conflict-linked frontal, parietal, and occipital regions (Petrides & Pandya, 1994).
2. Analyses of the online study also revealed that overall performance was reliably above chance and that reductions in old and new item accuracy under invalid cueing were equivalent to the fMRI task effects. Furthermore, analyses of a restricted online sample that was age- and gender-matched with the fMRI sample revealed identical effects of "strong" old-worded cues in eliciting greater bias shifts. Both findings highlight the generalizability of the online validation results to the main fMRI sample.
3. We also conducted network cluster analyses for those regions at the overlap of the "conflict" and "retrieval" fcMRI networks, which revealed a lack of clear amplitude specificity for either cue strength or confidence manipulations and an averaged network time course that was temporally interposed between the early conflict and late retrieval networks. These findings suggest that the overlap regions might serve to integrate the operation of conflict and retrieval networks in the service of memory control—a possibility in need of future exploration.

REFERENCES

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Barredo, J., Oztekin, I., & Badre, D. (2015). Ventral fronto-temporal pathway supporting cognitive control of episodic memory retrieval. *Cerebral Cortex, 25,* 1004–1019.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Brass, M., & von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cerebral Cortex, 12,* 908–914.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J. B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June.

Buckner, R. L. (2003). Functional-anatomic correlates of control processes in memory. *Journal of Neuroscience, 23,* 3999–4004.

Buckner, R. L., Krienen, F. M., & Yeo, B. T. (2013). Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience, 16,* 832–837.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Development of frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33,* 301–311.

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences, 11,* 219–227.

Cavada, C., & Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *Journal of Comparative Neurology, 287,* 422–445.

Conroy, M. A., & Polich, J. (2007). Normative variation of P3a and P3b from a large sample (N = 120): Gender, topography, and response time. *Journal of Psychophysiology, 21,* 22–32.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience, 25,* 11730–11737.

Dobbins, I. G., Jaeger, A., Studer, B., & Simons, J. S. (2012). Use of explicit memory cues following parietal lobe lesions. *Neuropsychologia, 50,* 2992–3003.

Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex, 16,* 1623–1630.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 194–205.

Henson, R. N., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience, 12,* 913–923.

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (Vol. xii, pp. 383–400). New York: Guilford Press.

Jaeger, A., Cox, J. C., & Dobbins, I. G. (2012). Recognition confidence under violated and confirmed memory expectations. *Journal of Experimental Psychology: General, 141,* 282–301.

Jernigan, T. L., Gamst, A. C., Fennema-Notestine, C., & Ostergaard, A. L. (2003). More "mapping" in brain mapping: Statistical comparison of effects. *Human Brain Mapping, 19,* 90–95.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., & Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *Journal of Neurophysiology, 100,* 129–139.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.

Kim, H., & Cabeza, R. (2007). Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. *Journal of Neuroscience, 27,* 12190–12197.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302,* 1181–1185.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). New York: Lawrence Erlbaum.

Mill, R. D., Cavin, I., & O'Connor, A. R. (2015). Supplementary information for: " Differentiating the functional contributions of resting connectivity networks to memory decision-making fMRI support for multi-stage control processes". Retrieved 10:27, Mar 24, 2015 (6) from http://dx.doi.org/10.6084/m9.figshare.1348808.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202.

Näätänen, R., & Gaillard, A. W. K. (1983). The N2 deflection of the ERP and the orienting reflex. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event-related potential research: Endogenous components*. Amsterdam: North Holland.

Nelson, S. M., Cohen, A. L., Power, J. D., Wig, G. S., Miezin, F. M., Wheeler, M. E., et al. (2010). A parcellation scheme for human left lateral parietal cortex. *Neuron, 67,* 156–170.

Nieuwenhuis, I. L., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research, 218,* 325–334.

O'Connor, A. R., Han, S., & Dobbins, I. G. (2010). The inferior parietal lobule and recognition memory: Expectancy violation or successful retrieval? *Journal of Neuroscience, 30,* 2924–2934.

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 17–58). Amsterdam: Elsevier.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: Comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience, 11,* 1011–1036.

Petrides, M., & Pandya, D. N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. *Journal of Neuroscience, 27,* 11573–11586.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118,* 2128–2148.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174.
Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences, 14,* 180–190.
Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443–447.
Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.
Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage, 53,* 303–317.
Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science, 193,* 1142–1146.
Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150,* 1187–1188.
Suwazono, S., Machado, L., & Knight, R. T. (2000). Predictive value of novel stimuli modifies visual event-related potentials and behavior. *Clinical Neurophysiology, 111,* 29–39.
Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia, 46,* 1787–1799.
Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology, 100,* 3328–3342.
Vincent, J. L., Snyder, A. Z., Fox, M. D., Shannon, B. J., Andrews, J. R., Raichle, M. E., et al. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. *Journal of Neurophysiology, 96,* 3517–3531.
Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106,* 1125–1165.
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517.