Dynamics of large-scale cortical networks reveal the cognitive control of episodic memory

James E. Kragel[1,2], Sean M. Polyn[1]

[1]*Department of Psychology*
[2]*Neuroscience Graduate Program*
Vanderbilt University, Nashville, Tennessee 37240, USA

Running title: Modeling of large-scale networks in free recall

Corresponding Author:
Sean M. Polyn
Vanderbilt University, Department of Psychology
PMB 407817
2301 Vanderbilt Place
Nashville, TN 37240
e-mail: sean.polyn@vanderbilt.edu
phone: (615) 322–2536
fax: (615) 343–8449

Number of pages: 48

Number of figures: 5

Number of tables: 0

Number of words Summary: 149

Number of characters: 55679

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

Author Contributions J.E.K. and S.M.P. designed the experiment, J.E.K. collected the data, J.E.K. carried out data analysis, and J.E.K. and S.M.P wrote the manuscript.

Acknowledgments: This research was supported by a Vanderbilt Discovery grant, NIH UL1 TR000445, and National Science Foundation grant 1157432 to S.M.P.

Summary

Distributed large-scale brain networks are thought to exert cognitive control to flexibly guide episodic memory, but the specific functions of these networks remain unclear. We estimated large-scale network activity using independent component analysis of fMRI data as participants performed a memory search task. A short delay following a study list was either unfilled, encouraging the maintenance of episodic information, or contained disruptive math distraction, encouraging reinstatement of the target episodic context. A computational model was used to formalize hypotheses linking large-scale network activity to cognitive control mechanisms supporting maintenance or contextual reinstatement. In the absence of distraction, dorsal attention network activity signaled contextual reinstatement, but after distraction, a frontoparietal control network was linked to this process, showing functional coupling with other large-scale networks, and with posterior medial temporal lobe. These novel behavioral correlates to large-scale brain activity provide insight into the mechanisms facilitating the cognitive control of human memory.

Introduction

Cognitive neuroscientific models of human memory propose that the medial temporal lobe (MTL) system, and in particular, the hippocampus, is critically important for the formation and retrieval of episodic memories of one's past experience (Moscovitch, 1992; Cohen and Eichenbaum, 1993; McClelland et al., 1995; Diana et al., 2007; Ranganath and Ritchey, 2012). In these theories, the MTL is often described as a repository of associative structures that support the reactivation of neuronal ensembles that represent the episodic past. Neural networks spanning prefrontal, parietal and temporal cortices interact with the hippocampus and other MTL structures to flexibly retrieve memories in a task-appropriate fashion (Miller, 1991; Moscovitch and Winocur, 1992; Buzsáki, 1996; Polyn and Kahana, 2008; Budson and Solomon, 2011). In the past few decades, we have learned a great deal about the nature of these interactions, but the computational mechanisms that enable large-scale cortical networks to control the human memory system remain unclear. However, computational models of human cognition suggest two core mechanisms that may underlie these interactions: First, an integrative mechanism allowing cortical regions to maintain neural representations in the face of distracting, task-irrelevant activity, and second, a cuing mechanism that allows neural representations in cortex to trigger the retrieval of episodic memories.

In the model of prefrontal function proposed by Miller and Cohen (2001), an integrative mechanism allows prefrontal circuitry to maintain task-relevant information in an active state, allowing prefrontal cortex (PFC) to guide processing in posterior cortical regions in a task-appropriate manner. This mechanism supports working memory, in that it allows information to be maintained even in the presence of distracting, task-irrelevant neural signals (Miller et al., 1996). This functional property of PFC has been most reliably associated with neural activity in dorsolateral PFC (DLPFC) in a large number of neuroimaging studies (Cohen et al., 1997; Postle, 2006; Sakai et al., 2002; Postle, 2005; Clapp et al., 2010). Recently, Feredoes et al. (2011)

demonstrated that stimulation of DLPFC during distraction elicits increased activity in posterior cortical regions that represent information in working memory, providing causal evidence for its role in top-down control. These findings support the idea that recurrent neural activity between the prefrontal cortex and posterior regions, such as parietal cortex, is important for guiding one's attention to relevant information, and then holding that information in mind (Corbetta and Shulman, 2002; Hasselmo and Stern, 2006).

A number of theories suggest that frontal and parietal cortical regions are critically involved in the cognitive control of episodic memory retrieval, and suggest particular mechanisms allowing these regions to guide the construction and utilization of retrieval cues in a task-appropriate manner (Moscovitch and Winocur, 2002; Becker and Lim, 2003; Simons and Spiers, 2003; Polyn and Kahana, 2008). In neural network models of episodic memory, such as the Complementary Learning Systems (CLS) model of McClelland et al. (1995), a retrieval cue is a distributed pattern of cortical activity that can be used to prompt the hippocampal memory system to reactivate associated episodic memory traces (Norman and O'Reilly, 2003). This reactivated episodic information specifies the spatiotemporal context of a particular past experience, and is thought to be related to the subjective experience of revisiting the past, sometimes referred to as *mental time travel* (Tulving, 1993). These ideas are supported by neuroscientific studies showing that patterns of cortical activity observed during episodic retrieval match the patterns seen during the original study event (Danker and Anderson, 2010). Neuroimaging studies demonstrating the reactivation of category-specific information (Polyn et al., 2005; Morton et al., 2013), task-specific information (Johnson and Rugg, 2007; Polyn et al., 2012), and perceptual information (Wheeler et al., 2000; Bosch et al., 2014), support the generality of this mechanism. Recently, researchers have shown that event-specific patterns are reactivated during episodic retrieval, in prefrontal, temporal, and parietal cortices (Staresina et al., 2013; Kuhl and Chun, 2014). Furthermore, the reactivation of event-specific patterns of cortical activity has associated with increased hippocampal activation (Ritchey et al., 2013; Bosch et al., 2014; Gordon et al., 2014), consistent

with theories positing that hippocampal associative structures are critically involved in episodic retrieval.

There is growing consensus in the neuroimaging literature that cognitive control operations are associated with activity in widespread networks of brain regions spanning frontal, parietal, and temporal cortices, with distinct networks associated with distinct aspects of cognitive control (Corbetta and Shulman, 2002; Fox et al., 2006; Vincent et al., 2006; Dosenbach et al., 2008). These networks reveal themselves in analyses of the covariance structure of neuroimaging data; sub-regions within a network may be anatomically distant, but show highly correlated timecourses of neural activation while a participant is at rest (Vincent et al., 2008), as well as during cognitive task performance (Smith et al., 2009). This has motivated the use of statistical techniques such as independent component analysis (ICA) to identify these networks and relate changes in the activation of a particular network to cognitive demands (e.g., Kragel and Polyn, 2015). The dorsal attention network (DAN), which includes regions in dorsal parietal cortex and the frontal eye fields, has been associated with the goal-directed guidance of attention to visuo-spatial information (Corbetta and Shulman, 2002). The ventral attention network (VAN), on the other hand, which includes regions in inferior prefrontal and parietal cortex, as well as the dorsal anterior cingulate, is more engaged during stimulus-driven attentional capture. The DAN is often contrasted with the default mode network (DMN), which has been associated with internally directed attention, mind-wandering, and memory retrieval (Buckner et al., 2008). The DMN spans the lateral and medial temporal lobes, the angular gyri, and midline regions in prefrontal and posterior parietal cortex, and shows a great deal of overlap with a putative recollection network implicated in the reactivation of episodic information (Okada et al., 2012).

The nature of the interactions between these large-scale brain networks is an area of active research. It has been proposed that the DAN and DMN act in opposition to one another (Fox et al., 2005), with suppression of the DMN requisite for performance on cognitively demanding tasks (Anticevic et al., 2012). This is

consistent with the finding that the DMN is suppressed when participants must maintain information in working memory in the face of distraction (Chadick and Gazzaley, 2011). However, recent studies examining interactions between these large-scale networks suggest that this opposition is not observed during retrieval of information from long-term memory (Fornito et al., 2012; Kragel and Polyn, 2015). During internally directed retrieval, a frontoparietal control network (FPCN) spanning lateral prefrontal cortex and the anterior inferior parietal lobe, has been proposed to serve as an interface between the attentional system reflected in DAN activity with the memory system reflected in DMN activity (Vincent et al., 2008).

The goal of the current work is to more precisely specify the cognitive operations associated with neural activity in these large-scale brain networks. In order to do this, we tracked the activity in these large-scale brain networks as participants performed a memory search task designed to pit the cognitive operations of active maintenance and memory retrieval against one another. Specifically, participants studied a list of words, which was followed by a short delay period. In one set of trials, the delay was unfilled, facilitating control operations supporting the active maintenance of information related to the last few studied items. In the rest of the trials, participants performed a math distraction task during the delay, challenging active maintenance, and encouraging the engagement of control operations triggering episodic memory retrieval. Dynamics of large-scale brain networks were characterized during this delay, both in terms of their response to distracting task performance, and in terms of their interactions with one another, and with an MTL region implicated in episodic retrieval. A neuro-cognitive model of memory search (Kragel et al., 2015) allowed us to test specific hypotheses linking activity in large-scale networks to particular cognitive control operations. This modeling framework specifies the linkage from neural signal, to cognitive operation, to behavioral observation, allowing one to assess the validity of a specific neuro-cognitive linking hypothesis in terms of whether the neural signal improves the ability of the model to predict the behavior of the participant, on a trial-by-trial basis.

Results

Using distraction to challenge the memory system

We used a variant of the free-recall paradigm to examine the effects of distraction on the ability of the human memory system to retrieve recently learned information (Figure 1a). Participants recalled a similar percentage of studied items after performing the distraction task (66.4 $\pm$0.04 SEM) as after an unfilled delay interval (65.5 $\pm$0.04 SEM); this difference was not statistically significant ($t_{19} = 0.73$, $p = 0.47$).

Consistent with prior work (Postman and Phillips, 1965; Glanzer and Cunitz, 1966), the presence of distraction during the delay interval led to a reduction in the memorability of end-of-list items. As shown in Figure 1b, the terminal list item was more likely to be recalled during unfilled, relative to filled, trials ($t_{19} = 4.96$, $p < 0.0001$). The effect of end-of-list distraction can be seen in terms of recall initiation (Fig.1c). On trials with an unfilled delay, recall was initiated with the terminal list item on 20.5 ($\pm$0.05 SEM) percent of trials, and with the start-of-list item on 22.1 ($\pm$0.05 SEM) percent of trials. When the distraction task was performed during the delay interval, the recency effect was attenuated and the primacy effect was enhanced: Recall was initiated with the terminal list item on 11.9 ($\pm$0.03 SEM) percent of trials, and with the start-of-list item on 32.4 ($\pm$0.05 SEM) percent of trials. A repeated measures ANOVA revealed a significant interaction between condition (filled vs. unfilled delay) and recall initiation ($F_{1,19} = 5.45$, $p = 0.03$).

Examining the effect of distraction on large-scale brain network activity

We used ICA to identify a set of large-scale brain networks, and to estimate the contributions of these networks to the observed blood oxygenation level-dependent (BOLD) response. In order to determine how the networks identified in this study relate to other large-scale networks identified in the literature, we carried out a spatial correspondence analysis to identify components whose spatial maps were consistent with the

spatial maps of large-scale networks identified by Yeo et al. (2011), in a study examining resting-state functional connectivity in 1,000 participants.

This analysis revealed six ICs that demonstrated a high degree of spatial correspondence with intrinsic networks of interest (Figure 2a). A single component, IC9, had a high degree of spatial overlap with the DAN ($r = 0.4448, p < 0.00001$). Another component, IC8, was consistent with the ventral attention network (VAN; $r = 0.2977, p < 0.00001$). Two ICs were spatially consistent with the frontoparietal control network (FPCN): IC21 was predominantly right-lateralized ($r = 0.3368, p < 0.00001$), and IC49 was predominantly left-lateralized ($r = 0.3435, p < 0.00001$). Two networks were spatially consistent with the default mode network (DMN): IC27 ($r = 0.3038, p < 0.00001$), which included posterior cingulate cortex and midline cortical regions, and IC53 ($r = 0.157, p < 0.00001$), which included left angular gyrus and middle frontal gyrus.

Consistent with prior work showing that challenging working memory with distraction engages networks spanning lateral prefrontal cortex (Sakai et al., 2002; Jha et al., 2004), we found a number of ICs with frontoparietal coverage that exhibited consistent changes in activation during the delay period. A linear regression was used to characterize the differential response profile for these ICs between filled and unfilled delay conditions.

The DAN (IC9) was reliably more engaged during the filled delay than the unfilled delay ($t_{19} = 3.31, p = 0.008$), with activation for the unfilled delay less than the implicit baseline of the regression ($t_{19} = -2.57, p = 0.023$). The VAN (IC8), on the other hand, increased in activity during both unfilled and filled delay periods, relative to baseline (unfilled delay: $t_{19} = 4.28, p = 0.0024$; filled delay: $t_{19} = 4.16, p = 0.0011$).

The FPCN (IC21, IC49) components were reliably engaged for both types of delay interval: IC21 (right

FPCN), unfilled delay: $t_{19} = 3.15, p = 0.010$; filled delay: $t_{19} = 5.85, p < 0.0001$. IC49 (left FPCN), unfilled delay: $t_{19} = 3.84, p = 0.0033$; filled delay: $t_{19} = 7.65, p < 0.0001$. Furthermore, both FPCN components showed reliably more engagement when the delay was filled, compared to unfilled (IC21: $t_{19} = 3.26, p = 0.0083$; IC49: $t_{19} = 5.14, p = 0.0004$).

Consistent with prior work examining the effect of task-based distraction during memory maintenance, the DMN components (IC27, IC53) showed suppression in the presence of distraction (Wager et al., 2013). IC53 decreased in activation during the filled delay ($t_{19} = -3.58, p = 0.003$), but remained near baseline on unfilled lists ($t_{19} = -0.86, p = 0.40$). IC27 showed decreased engagement during the filled, compared to the unfilled delay period ($t_{19} = -2.79, p = 0.017$). This effect was driven by an increase in activity during unfilled delay ($t_{19} = 2.22, p = 0.046$).

Functional connectivity between large-scale networks and the MTL shifts with distraction

These shifts in activation of the large-scale brain networks during the delay could reflect a variety of cognitive control operations related to the maintenance and retrieval of task-relevant mnemonic information. A recent study by Zanto et al. (in press) demonstrates that even with only a single stimulus to be maintained across a short delay, MTL memory systems are engaged if participants anticipate (and encounter) distraction during that delay. Thus, shifts in large-scale network activity may indicate not only the top-down filtering of distracting information, but also the engagement of MTL-based circuitry to support memory performance.

As described above, we observed increased activation of DAN under conditions of distraction, which may reflect the attentional system processing the visually presented distraction stimuli. DMN activity was suppressed during this distraction, but was higher during an unfilled delay, perhaps reflecting attention to internally maintained representations of the studied material. A cingulo-opercular network, which partially

overlaps with the VAN examined in the current study, has been proposed regulate the relative activation of the DMN and DAN (Sridharan et al., 2008). Furthermore, the FPCN has been proposed to serve as an interface between DMN and DAN, in tasks where long-term memory structures are important for task performance (Vincent et al., 2008). These ideas motivated us to examine the functional interactions of these networks during the delay interval of the current experiment.

To characterize the contributions of the MTL during the delay period, we constructed a functional ROI that spanned bilateral parahippocampal gyrus and hippocampus. In previous work, this ROI was shown to be involved in the reinstatement of contextual information that guides memory search (Kragel et al., 2015). Examination of the response of this region during the delay period revealed a sustained increase in activation during filled ($t_{19} = 2.35,\ p = 0.029$) but not unfilled ($t_{19} = 1.05,\ p = 0.31$) delay periods (Fig. 3a). To examine whether this MTL activation in the presence of distraction is related to its interactions with putative control networks, we conducted a correlational psychophysiological interaction (cPPI) analysis (Fornito et al., 2012) to examine task-based changes in functional connectivity during the delay period.

Multiple large-scale networks showed reliable delay-dependent interactions with one another, and with the MTL ROI (Fig. 3b). Specifically, the left lateralized FPCN and the VAN showed increased functional coupling to the MTL during the filled, relative to the unfilled delay periods (FPCN: $t_{19} = 5.31$, $p = 0.0004$; VAN $t_{19} = 6.42$, $p = 0.0004$). The left FPCN also exhibited relative increases in functional connectivity to both the VAN ($t_{19} = 3.35$, $p = 0.024$), as well as a subnetwork of the DMN anchored in the PCC ($t_{19} = 2.90$, $p = 0.048$) during the filled delay periods, relative to the unfilled delay periods. These interactions are consistent with prior work implicating a cingulo-opercular network (which overlaps considerably with the VAN) in regulating the relative engagement of the DMN and the DAN (Sridharan et al., 2008), as well as prior work proposing that the FPCN serves as an interface between the attentional systems of the DAN, and

the memory systems of the DMN (Vincent et al., 2008).

We carried out a secondary analysis separately contrasting functional connectivity during the filled and unfilled delay periods to baseline estimates of functional connectivity. The functional interactions observed during the unfilled delay were not reliably different from baseline. In the filled delay period, we observed decreased functional coupling between the two subnetworks of the DMN, relative to baseline ($t_{19} = -3.67$, $p = 0.034$). Given that the cPPI analysis accounts for differences in activity associated with performance of the distraction task, as well as any potential sustained responses during the delay period, these shifts in functional connectivity likely reflect task-context-dependent shifts in the interactions of these large-scale networks (Friston et al., 1997).

Linking large-scale networks to cognitive operations

The preceding results demonstrate that a brief period of distraction causes recruitment of multiple large-scale brain networks that interact with each other, and with a posterior MTL region. In order to more clearly relate activation in these networks to behavioral performance of participants in this task, we developed a computational model describing a set of interacting cognitive operations sufficient to produce the recall behavior observed in this task (Kragel et al., 2015). The framework associated with the model allows one to formalize specific neuro-cognitive linking hypotheses, whereby a particular neural signal is used to control a corresponding cognitive operation in the model. If this linkage improves the ability of the model to predict the specific sequence of recalled items produced by participants, this is taken as support for the hypothesis that the neural signal is related to that cognitive operation.

We examined two cognitive mechanisms that could be engaged during the delay period, and whose operation might be reflected in the activation of large-scale brain networks: First, a contextual disruption (CD) process

that signals whether mnemonic information is disrupted by distraction, or is successfully maintained across the delay, and second, a contextual retrieval (CR) process that signals whether associative structures are being used to reactivate prior contextual states (a process often called *mental time travel*; Tulving 1993). Engagement of each of these mechanisms during the delay produces a distinct behavioral signature, in terms of how the participant initiates recall: Successful maintenance of mnemonic information through the delay is associated with initial recall of recent (end-of-list) items. Figure 4a shows how the predictions of the model change as distraction becomes more disruptive to the maintained retrieval cue. Successful contextual retrieval is associated with initial recall of the primary (start-of-list) items. Figure 4b shows how model predictions change in terms of the fidelity of this retrieved contextual information.

A baseline version of the model was fit to the behavioral data from the filled-delay and unfilled-delay conditions, and parameter estimates from this model were examined to determine whether there was evidence that the CD and CR mechanisms were more strongly engaged during the filled-delay trials than the unfilled-delay trials. We carried out a differential evolution Markov chain Monte Carlo approach to estimate the posterior distributions of the different model parameters (Table S2), and used an encompassing prior approach (Wetzels et al., 2010) to sample from these posterior distributions and determine whether the model provided evidence for enhanced engagement of these cognitive operations in the presence of distraction, as quantified by the Bayes factor (BF) statistic. Due to the uniform priors assumed in these analyses, the maximum evidence that can be attained is a Bayes factor of 2.

There was little support for the hypothesis that contextual disruption was increased in the filled-delay condition: Only 27.9% of sampled models were consistent with this prediction ($BF = 0.55$). This may be related to the general difficulty of maintaining an end-of-list contextual state for 12 seconds in the unfilled-delay condition. Despite the lack of evidence for a general increase in CD between the two conditions,

it is still possible that this mechanism is differentially engaged on different trials within each condition, and this differential engagement may correspond to trial-by-trial fluctuations in a particular neural signal. On the other hand, there was strong evidence that contextual retrieval was enhanced after distraction: 97.9% of sampled models supported this prediction ($BF = 1.93$).

We developed two families of neurally informed computational models of memory search in order to test whether the recruitment of large-scale cortical networks was related to cognitive processes that determine whether information encoded in either recent (end-of-list), or past (start-of-list), contextual states is accessible during memory retrieval. The CD model family allowed delay-period neural activity to influence whether an internal representation of context was maintained or disrupted during the delay (Fig. 4a). In this family of models, delay-period network activity recorded on individual trials influenced the contextual disruption parameter, allowing us to test whether activity in a particular network signaled the likely recall of recently learned information. The CR model family was used to determine whether network activity predicts the ability of the memory system to access individual items studied at the beginning of the study episode (Fig. 4b). In this family of models, delay-period network activity recorded on individual trials influenced the contextual retrieval parameter, allowing us to test whether activity in a particular network signaled the likely recall of more temporally distant information. The performance of a given neurally informed model is reported in terms of a deviance statistic (D) which quantifies the improvement in predictive power of the neurally informed model relative to a neurally naive baseline model.

Activation of a single network was linked to the context disruption mechanism: IC9, which is highly spatially consistent with the DAN, was related to context disruption in simulations of both filled-delay and unfilled-delay trials (Fig. 5a). Engagement of this network improved the ability of the CD model to predict recall initiation following both unfilled ($D = 4.03$, $p = 0.044$) and filled ($D = 3.95$, $p = 0.047$) delay

periods. This model described a positive relationship between the context disruption parameter, β_{ri}, and observed activation in this network during both delay periods ($\nu_u = 0.042$, $\nu_f = 0.016$). Figure 5c (green bars) shows how activity in this network relates to recall initiation: Increased DAN activity corresponds to a decreased likelihood of initiating recall with an end-of-list item, without a corresponding increase in the likelihood of initating with a start-of-lsit item. Activity in a subnetwork of the DMN (IC27) also reflected contextual disruption, but only in the unfilled-delay condition ($D = 4.78$, $p = 0.029$, $\nu_u = 0.977$). Neural signal estimated from the remaining networks of interest did not improve the ability of models within this family to predict recall sequences ($ps > 0.36$).

Activation in the FPCN (IC49) signaled reinstatement of start-of-list contextual information on filled-delay trials (Fig. 5b). Increased recruitment of this network predicted individuals would initiate recall from the beginning of the list ($\nu_f = 0.052$). Likelihood-ratio tests revealed improved fitness of this model, relative to the neurally naive baseline model, in terms of predicting recall sequences during the filled-delay trials ($D = 5.69, p = 0.017$). Figure 5c (orange bars) shows how FPCN activity relates to recall performance; when FPCN activity is high during the delay, there is a slight decrease in the likelihood of initiating recall with an end-of-list item, with a corresponding increase in the likelihood of initiating recall with a start-of-list item. Activity in the previously mentioned DAN also reflected contextual reinstatement, but only in the unfilled-delay condition ($D = 4.96$, $p = 0.026$).

In our analysis of network interactions (Fig. 3b), we found that the FPCN increased its functional coupling to a MTL region of interest that was associated with contextual retrieval in prior work Kragel et al. (2015). This MTL region also showed enhanced engagement during filled-delay periods (Fig. 3a). We tested whether activity in this region was related to the CD and CR model mechanisms in these data. Activity in this region was related to the CR mechanism in the filled-delay trials ($D = 6.20$, $p = 0.013$, $\nu_f = 0.051$) but not in the

unfilled-delay trials ($D = 2.97$, $p = 0.085$, $\nu_u = 0.035$). Neural signal from this region was informative for the CD model as well, improving prediction on filled-delay ($D = 6.51$, $p = 0.011$, $\nu_f = 0.1188$) but not unfilled-delay ($D = 3.15$, $p = 0.076$, $\nu_u = 0.031$) trials.

Figure 5b (bottom panel) demonstrates how neural signal in this region relates to recall initiation on filled-delay trials. When MTL activity was low during the delay period (gray bars), participants sometimes initiated with an item from the beginning of the list, and sometimes with an item from the end of the list. However, when MTL activity was high during the delay (blue bars), the likelihood of initiating recall with start-of-list item increased (consistent with the CR model; Fig. 4b), and the likelihood of initiating with an end-of-list item dropped (consistent with the both CD and CR models; Fig. 4a & b).

While MTL regions are strongly represented in the DMN, our model-based approach dissociated MTL activity from DMN activity. Whereas a number of neuroimaging studies suggest that the DMN supports episodic recollection (Vincent et al., 2006; Okada et al., 2012) and autobiographical memory retrieval (Spreng and Grady, 2009; Svoboda et al., 2006), activity in this network was not informative for the contextual retrieval mechanism in the model (Fig. 5b). Increased DMN activity during the delay reliably indicated a substantially diminished likelihood of initiating recall with an end-of-list item (Fig. 5c, red bars), which was more informative for the CD model than for the CR model. Taken together, these findings are consistent with the idea that the MTL proper mediates contextual retrieval, while activation of the broader cortico-hippocampal DMN reflects whether attention is directed to these internal representations generated by the MTL subsystem (Buckner et al., 2008).

Discussion

In the present work we characterized the response of large-scale brain networks, and regional activation within the MTL, during a delay period following a study episode, which was either unfilled, to encourage maintenance of study items from the end of the study episode, or contained a brief period of distraction task performance, to encourage retrieval of the temporal contextual from the beginning of the study episode. Consistent with other neural investigations of memory in the presence of distraction (Clapp et al., 2010; Wager et al., 2013), we found that distraction elicited enhanced recruitment of the FPCN and a number of other brain networks thought to be involved in cognitive control. Furthermore, we found that activity in posterior MTL (including the hippocampus) showed functional coupling to a subset of these networks after this disruption. Using a model-based approach to cognitive neuroscience (Palmeri, 2014), we related these neural signals to particular cognitive operations, formalized within a simplified neural network model of memory search.

Within this modeling framework, we specified two core computational mechanisms central to cognitive neuroscientific theories of the cognitive control of memory. The first was an integrative mechanism that allows neural circuitry to maintain task-relevant information; in the current model, this maintained information was a context representation targeting the most recent studied materials. A model parameter was introduced to allow a given neural signal to indicate whether this information was successfully maintained through the delay, or whether it was disrupted during performance of a mathematical distraction task. The data supported a model in which activity in the DAN indicates the success or failure of this maintenance operation, regardless of whether participants performed a distraction task during the delay. It has been proposed that the DAN and DMN work in opposition to one another, controlling the relative balance of attention to external or internal information, respectively (Fox et al., 2005; Anticevic et al., 2012). However, recent work brings

this stark contrast into question, establishing situations in which DMN can become positively coupled with frontoparietal networks (Fornito et al., 2012; Kragel and Polyn, 2015; Spreng et al., 2010; 2014). We found that a subnetwork of the DMN was also informative regarding contextual disruption, but only during unfilled delay intervals. Prior work implicating the DMN in episodic search tasks (Vincent et al., 2006; Sestieri et al., 2011; Fornito et al., 2012; Kragel and Polyn, 2013) and in the maintenance of internal representations (Buckner et al., 2009; Vilberg and Rugg, 2012) may help us to interpret this result. Maintaining a contextual representation for 12 seconds may be quite demanding even in the absence of external distraction; if increased DMN activity reflected mind wandering during the delay interval, this would presumably be disruptive to a maintained contextual representation (Sahakyan and Kelley, 2002).

The second core cognitive control operation involved the engagement of episodic retrieval processes to reinstate the temporal context of the study list. Prior empirical and theoretical studies suggest that such a contextual reinstatement operation will tend to preferentially support items from the beginning of the targeted study episode (Laming, 1999; Tan and Ward, 2000; Tulving, 2008; Davelaar, 2013). The data supported a model in which the FPCN supports contextual reinstatement, after the memory system was challenged by distraction. From this study alone, we cannot determine whether activity in the FPCN directly reflects the successful retrieval of contextual information, or whether it reflects the engagement of control operations that facilitate the reinstatement of contextual information in posterior regions (e.g., in hippocampus or parahippocampal cortex; Kragel et al. 2015). The latter view is consistent with recent neuroimaging work suggesting that frontoparietal activity supports the maintenance of task-relevant information in posterior regions (Riggall and Postle, 2012), whereas the former view is consistent with a number of studies suggesting that task-relevant information may be maintained in the frontoparietal circuits themselves (Cole et al., 2013; Lee et al., 2013; Waskom et al., 2014; Etzel et al., in press).

The data also supported a model in which DAN activity indicated the engagement of contextual reinstatement in the unfilled-delay trials. In another free-recall experiment, Kragel and Polyn (2015) observed increased functional coupling between the DAN and subnetworks within the DMN during memory search, around the time of successful recall events (and, like here, in the absence of task-relevant visual stimuli). These results suggest that the DAN, usually associated with top-down attention to visuospatial material, may have a direct role in the self-guided retrieval of mnemonic information. One possibility is that in the absence of disruption, the DAN is able to facilitate contextual retrieval, but after distraction the contextual representation is more damaged, and the FPCN must be engaged to mediate contextual retrieval.

Theories of FPCN function suggest that it mediates cognitive control by facilitating communication between cortical modules requisite for the performance of a particular task (Cole et al., 2013; Sneve et al., 2013; Cole et al., 2014), consistent with observations that FPCN exhibits dynamic changes in functional connectivity with various cortical systems over short time scales Monti et al. (2014); Zalesky et al. (2014); Allen et al. (2014); Yu et al. (2015). Our analysis of network interactions during the delay revealed that after distraction, the FPCN shows increased functional coupling with the DMN and a posterior MTL region, which in turn show increased coupling with the VAN. The VAN is highly overlapping with cingulo-opercular networks that have been shown to couple with the DMN during episodic memory tasks (Sestieri et al., 2014), and contains prefrontal regions implicated in the controlled retrieval of episodic information (Wagner et al., 2001; Badre and Wagner, 2005). As mentioned above, it is possible that these widespread interactions represent the cognitive system's response to the contextual disruption wreaked by the math distraction task, a strategic attempt to reconstruct a contextual cue that will target the desired mnemonic information.

The utilization of cognitive models to interpret neural phenomena represents a growing trend in neuroscientific studies of behavior (e.g., Mitchell et al., 2008; Purcell et al., 2010; Mack et al., 2013; Turner et al., 2013a;

Khaligh-Razavi and Kriegeskorte, 2014; Turner et al., in press). These neuro-cognitive models allow one to formulate specific hypotheses linking neural signals to particular cognitive mechanisms, which precisely specify the behavioral consequences of observing that neural signal. This approach can be contrasted with the method of cognitive subtraction (Friston et al., 1996), which, while quite powerful, and of great utility to cognitive neuroscientists, is limited in terms of the nature of the hypotheses that can be tested. In the current work, we examined the relative validity of models linking neural signals to contextual disruption and contextual retrieval operations in a computational model of free recall. The model is a member of a class usually referred to as retrieved-context models, in which a distributed representation of temporal context is constructed during an episode, and is reinstated when that episode is remembered (Howard and Kahana, 2002; Sederberg et al., 2008; Polyn et al., 2009). A growing number of studies support the theoretical importance of the temporal reinstatement operation at the heart of the model, including studies of memory misattributions during list learning (Gershman et al., 2013), the semantic organization of memory search (Polyn et al., 2005; Morton et al., 2013), and the temporal organization of remembered items (Manning et al., 2011; Kragel et al., 2015).

Our results reveal the potential cognitive underpinnings of neural activity in several large-scale brain networks, and in the MTL, during a delay interval leading up to a period of memory search. In this neuro-cognitive model, neural signals are associated with cognitive operations, and the validity of these linkages are assessed in terms of the increased ability of the model to predict human behavior. We find that activity in the DAN indicates contextual disruption in the presence of perceptual distraction, but indicates contextual reinstatement in the absence of distraction. After distraction, contextual reinstatement is indicated by activation in the FPCN and MTL, which show functional coupling with the VAN and DMN. These results provide a glimpse into the shifting dynamics and complex interactions of large-scale brain networks in the service of the cognitive control of episodic memory retrieval.

Experimental Procedures

Experimental Design

Twenty individuals (14 female) with an average age of 21.64 ± 0.88 years (mean $\pm$ SD) provided written, informed consent to participate in the study for monetary compensation, in accordance with the procedures of the Vanderbilt University Institutional Review Board. Participants performed up to 18 trials of a free-recall task, while fMRI data was collected. Each trial consisted of three distinct phases: encoding, delay, and free recall. During the encoding period, participants studied a list of 16 items. Items were presented for 2.5 sec, during which subjects rated the size of the displayed stimulus (i.e., determining whether the stimulus will fit into a shoe box) using an MRI compatible response box. Following the presentation of each stimulus, a fixation cross was displayed during an inter-stimulus interval ranging between 0.5-4.5 s. On each trial, one of two delay period conditions followed the encoding period. For unfilled trials, a fixation cross remained on the screen for 12 s; this delay condition occurred on 50% of trials, pseudo-randomly ordered such that three consecutive lists of the same type did not occur. On filled trials, participants performed six seconds of a self-paced arithmetic task. During this task, mathematical statements (e.g., Figure 1a) were presented in the center of the screen. Participants made a true/false judgment for each mathematical statement, using two buttons on the manual response box. Following the delay, a row of asterisks was displayed on the screen for a duration of 0.5 s, indicating the start of the free-recall period. Subjects were instructed to vocally recall items studied on the most recent list, in any order. Audio recordings of speech were collected using a scanner-safe microphone (Resonance Technologies, Inc.), and were annotated with Penn TotalRecall (`http://memory.psych.upenn.edu/TotalRecall`).

Image Acquisition

Imaging was performed at the Vanderbilt University Institute of Imaging Science, collected with a 3T Philips Intera Achieva magnet. Functional images were collected using an interleaved gradient echo $T2^*$-weighted pulse sequence with BOLD contrast (TR=2000 msec, TE=30 msec, flip angle=75^o, voxel size=3.0x3.0x3.6 mm, FOV=192 mm). 30 oblique slices, oriented parallel to the AC-PC plane were collected during functional imaging. Whole brain structural scans were acquired using an MP-RAGE sequence (TR=2500 msec, TE=4.38 msec, flip angle=8^o, voxel size=1.0x1.0x1.0mm, FOV=256 mm).

Image Processing

Preprocessing was carried out using the SPM8 (Wellcome Trust Centre for Neuroimaging, University College Longon, England) software package, as follows: The first four functional images of each session were discarded to allow scanner signal to equilibrate. The remaining functional volumes from each run were realigned to the first functional scan, in order to correct for head motion. The mean of this realigned series was computed and coregistered to the T1 structural image. Anatomical segmentation and registration were performed using the VBM8 toolbox, (C. Gaser, Department of Psychiatry, University of Jena, Germany; http://dbm.neuro.uni-jena.de/vbm8). Gray matter, white matter, and cerebrospinal fluid images were spatially aligned using DARTEL (Ashburner, 2007) to create a group-averaged anatomical template. We applied nonlinear transformations to register the group data to Montreal Neurological Institute (MNI) space. The deformations from this procedure were applied to the coregistered functional data, and the resulting output images were resampled to 3-mm isotropic and smoothed with an 8-mm FWHM Gaussian kernel.

Independent Component Analysis

The Group ICA of fMRI toolbox (GIFT, `http://icatb.sourceforge.net/`, version 2.0d) was used to identify spatially independent, temporally synchronous networks of voxels (Calhoun et al., 2001; 2009). First, temporal concatenation of data from all participants was performed to construct a single dataset that underwent dimensionality reduction using a two-stage principal component analysis procedure (Calhoun et al., 2001). The minimum description length (Li et al., 2007) criteria was used to optimize the number of independent sources within the dataset, which was found to be 61. The Infomax algorithm (Bell and Sejnowski, 1995) was used to generate a timecourse of BOLD signal change and a spatial map for each IC for the concatenated group data, which were back-reconstructed for each individual participant (Calhoun et al., 2001; Meda et al., 2009).

In order to determine significance of spatial distributions with each IC, we standardized back-reconstructed spatial maps of each IC (Beckmann et al., 2005), demonstrating the contribution of each voxel to the overall IC timeseries. An average (i.e., across runs) standardized spatial map was computed for each subject and entered into a one-sample *t*-test. Voxels were considered significant if they exceeded an uncorrected threshold of $p < 0.001$, in conjunction with a cluster extent exceeding 50 voxels.

We implemented a spatial correspondence analysis in order to identify ICs that reflect estimates of functional networks of interest, as opposed to additional sources (e.g., head motion, ventricular pulsation, or additional neural sources). For each IC, we computed Pearson product-moment correlation between the group-averaged spatial map (thresholded at $t < 3.75$) and templates corresponding to 4 cortical networks of interests. These templates were constructed by Yeo et al. (2011) and correspond to binary maps of the DAN, VAN, FPC, and DMN. ICs demonstrating a strong one-to-one correspondence with networks of interest were considered for additional analysis.

Estimation of delay period network activity

Task-relevance of large-scale network activity was tested using a general linear model (GLM) framework. Events were modeled using the canonical hemodynamic response function and its temporal derivative. In the first GLM, separate regressors were used to model the study, delay, recall, and speech periods. Where relevant, filled and unfilled trials were modeled separately. For study trials, subsequently recalled and forgotten items were modeled using unique regressors. Delay-period regressors were constructed with stimulus onsets at the beginning of the delay period, with a duration of 6 s. We constructed two regressors to model recall-period activity for each trial type of interest. A sustained recall regressor of 75 s in duration accounted for block-level effects during memory search. Transient recall-related activity was modeled by constructing a regressor from delta functions occurring at the onset of vocalization. The speech task was modeled with a single regressor, time-locked to the onset of stimulus presentation. Additional nuisance regressors were included to account for motion parameters and differences in mean BOLD signal across sessions. Linear contrasts were used to obtain subject-specific estimates for each effect of interest, which were then entered into a second-level, random-effects analysis. The second GLM was used to estimate delay-period activity at the single trial level, in order to predict recall behavior using cognitive models of memory search. The parameters of this GLM were identical to the first, except separate regressors were used to estimate activity for each delay period of interest.

Network Interaction Analysis

Task-related functional interactions between large-scale networks, as well as signal from independently defined regions of interest within the MTL, were estimated using a correlational psychophysiological interaction (cPPI) analysis (Fornito et al., 2012). Our task-related functional connectivity analysis aimed to identify interactions between distinct cortical networks that change as a function of task-demands (e.g., in response

to the disruption of memory processes). While the goal of this analysis is to identify synchronized network dynamics, coordinated activation in separate networks of the brain may arise due to many circumstances, including independent task-based coactivation. To account for such effects, we computed an interaction term using a deconvolution technique (Gitelman et al., 2003). Three separate cPPI analyses were performed, constructing interaction terms from regressors estimating delay-period dependent changes in activity. The first analyses identified task-dependent changes in connectivity due to the unfilled delay, the filled delay, and differences across the two delay periods. In each analysis, we computed the partial correlation between each these interactions terms for each pairwise combination of networks of interest, while accounting for covariation due to task-based activation, previously described motion parameters, and activation in any of the networks of interest. Correlation measures were Fisher z-transformed prior to testing for significance using a one-sample t test, with a threshold of $p < 0.05$, FDR corrected (Benjamini and Hochberg, 1995).

Computational Modeling

In order to test the validity of hypotheses linking activity in large-scale brain networks to particular cognitive operations, we utilized a neurally informed variant of the Context Maintenance and Retrieval model of human memory (Kragel et al., 2015). This approach involved constructing a baseline model, and two neurally informed families of models. The baseline model was fit to behavioral data but was not informed by neural signal. The two neurally informed model variants included a Contextual Disruption (CD) model, in which a given neural signal determined the value of a model parameter controlling how much the model's representation of temporal context was disrupted during the end-of-list distraction period, and a Contextual Reinstatement (CR) model, in which a neural signal determined the fidelity with which the system retrieved the representation of temporal context associated with the beginning of the most recent study list.

An optimization procedure was used to find the set of parameters that allowed each model variant (baseline,

CD, CR) to best predict behavioral performance. A likelihood statistic quantified the goodness-of-fit for each model. For each recall event, the model computed a set of scores corresponding to the probability of reporting each studied item, and the probability of recall termination. The log-transformed probability of each recall event was collected, and summed across all recall events in the experiment. This log-likelihood score corresponds to the likelihood that a model gave rise to the observed data, and was used as the basis for the model comparison statistics reported below. Despite being fit to individual recall events, all model variants provided a good fit to the overall behavioral performance as measured by a set of benchmark summary statistics describing recall initiation, transitions between recalled items, recall termination, and overall memorability as a function of serial position on the study list. For more detail regarding the basic model mechanisms, the reader is referred to Kragel et al. (2015). We describe a handful of modifications to the model as implemented in the current simulations.

We allowed a number of model parameters to vary between the two conditions (unfilled delay and filled delay), under the assumption that this experimental manipulation influenced the cognitive processes controlled by these parameters. Distraction task performance was assumed to influence both contextual disruption (controlled by parameter β_{ri}) and the fidelity of contextual reinstatement (controlled by parameter β_s), so these parameters were allowed to take distinct values for each condition. Additionally, in order to capture any condition-level differences in overall recall performance, we allowed the parameter controlling recall termination probability (ξ_d) to vary by condition.

The rule determining the probability of recall termination took a different form than the model examined by Kragel et al. (2015). Here, the probability of stopping recall, p_{stop}, was calculated as an function of output position ($\mathbf{r}$):

$$p_{stop} = \xi_s + e^{\xi_d r}, \tag{1}$$

where ξ_d is a free parameter controlling the rate at which the termination probability increases with output position. ξ_s is a parameter that reflects the probability that no items are recalled, and was fixed to a value of 5^{-3} for all simulations.

A probabilistic decision competition was simulated for each recall event, in which the relative support for each item determines the probability of that item being recalled. The relative support for each item $\mathbf{s}$ is subjected to exponential scaling prior to the recall decision:

$$\mathbf{s} = e^{\frac{2\mathbf{s}}{\tau}}, \tag{2}$$

where the free parameter τ describes the ferocity of the competition between items during recall. This decision rule provides some minimum degree of support for all studied items, allowing us to remove a noise parameter (α) from the model.

In order to account for the potential influence of semantic associations on recall behavior, we included an additional free parameter, s, that controlled the strength of pre-experimental associations between items with similar semantic meanings:

$$\mathbf{M}^{CF}_{i,j} = s\mathbf{M}^{sem}_{i,j}, \tag{3}$$

where $\mathbf{M}^{sem}_{i,j}$ is a cosine similarity measure of the semantic relatedness of any pair of items i and j. These

semantic relatedness scores were computed from the semantic model Word Association Spaces (WAS; Steyvers et al. 2004), in which a vector representation of each word is constructed from a set of free-association norms (Nelson et al., 2004).

Given that in the unfilled delay condition participants would have extra time to encode the final list item, we introduced a parameter λ to enhance the associative strength between the contextual retrieval cue and the final item representation, relative to the other studied items (λ was set to zero in simulations of the filled condition).

$$\mathbf{M}^{CF}(t) = \mathbf{M}^{CF}(t-1) + (\phi(t) + \lambda)\mathbf{fc}^{\top}, \tag{4}$$

where t indexes list position. ϕ is a scaling factor that increases context-to-item associations for early list positions to capture the primacy effect (the enhanced recall performance associated with the first few list positions).

Neurally informed models of free recall

In this work we examined two neurally informed variants of the CMR model. Each neurally informed model was identical to the baseline model, with the sole addition of a neural scaling parameter which allowed neural signal recorded on a particular trial to influence a specific operation in the model.

In the contextual disruption (CD) model, neural signal recorded during the delay period determined how much the contextual representation in the model was disrupted during the delay period. In the baseline model, this was controlled by model parameter β_{ri}, which determines how much distraction-task-related information is integrated into the temporal context representation. Higher values of β_{ri} indicate more

disruption to the temporal context representation, as more irrelevant information is integrated into the contextual representation (this operation is described in more detail in Equation 2 of Kragel et al. 2015). In the CD model, neural signal causes β_{ri} to vary from trial to trial:

$$\beta_{ri}(t) = \beta_{ri} + \nu_{ri} N(t), \tag{5}$$

where ν_{ri} is a free parameter that scales the estimated neural activity during the delay-period, N, at time t. In order for the models to capture potential differences in the mapping of neural signals onto cognitive processes between the two conditions, we allowed these two free parameters to take on separate values during filled (i.e., $\beta_{ri,f}$ and $\nu_{ri,f}$) and unfilled (i.e., $\beta_{ri,u}$ and $\nu_{ri,u}$) trials.

In the context reinstatement (CR) model, neural signal recorded during the delay period determined the fidelity with which the system retrieved a representation of temporal context from the beginning of the study period. In the baseline model, this was controlled by model parameter β_s, which determines how much of this start-list information is integrated into the temporal context representation (using the same basic machinery described in Equation 2 of Kragel et al. 2015). In the CR model, neural signal causes β_s to vary from trial to trial:

$$\beta_s(t) = \beta_s + \nu_s N(t), \tag{6}$$

where ν_s is a free parameter that maps the estimated large-scale network activity, N, onto the context reinstatement parameter, β_s. As in the CD model, these parameters were separately estimated for both list types.

Parameter estimation

For a given version of the model, we estimated the set of parameters that best allowed the model to predict recall behavior, using a particle swarm optimization (PSO) algorithm (Eberhart and Kennedy, 1995). This technique was used to estimate the optimal parameters for the baseline model and for each neurally informed mode family (i.e., CR and CD models). During each application of the PSO algorithm, we constructed each swarm from 50 uniformly distributed random parameter sets, which explored the parameter space for up to 1000 generations or until a stopping criteria (the average change in model fitness over 50 generations) was less than a specified threshold (10^{-7}).

Model comparison and statistical inference

Once the optimal set of parameters was estimated for the baseline model and a neurally informed model, a likelihood ratio test (Wilks, 1938) was used to determine whether any improvement in model predictions for the neurally informed model was significant. This produced a test statistic, D (also known as deviance, twice the difference of the log-likelihood of the neural and baseline models), for each voxel within a priori regions of interest. Significance was determined by testing D on a χ^2_1 distribution.

In addition to performing inference on the best-fitting parameter set for each model of interest, we used a differential evolution Markov chain Monte Carlo approach to perform Bayesian estimation of model parameters (Turner et al., 2013b). For each model, we ran 20 chains for 500 iterations with a burn-in period of 50 samples per chain, resulting in a total of 9000 samples from the target distribution. The γ parameter was set to a value of $2.38/\sqrt{2d}$, where d was the dimensionality of the parameter space for each model. After ensuring convergence of the chains, 95% highest density intervals (Kruschke, 2010) were constructed for each parameter. Both Savage-Dickey tests (Wagenmakers et al., 2010), and encompassing

prior approaches (Wetzels et al., 2010) were used to evaluate model fitness and test hypotheses regarding differences in parameter values between the two delay conditions.

References

Allen E.A., Damaraju E., Plis S.M., Erhardt E.B., Eichele T., and Calhoun V.D. (2014). Tracking whole-brain connectivity dynamics in the resting state. Cerebral Cortex *24*, 663–676.

Anticevic A., Cole M.W., Murray J.D., Corlett P.R., Wang X.J., and Krystal J.H. (2012). The role of default network deactivation in cognition and disease. Trends in Cognitive Sciences *16*, 584–592.

Ashburner J. (2007). A fast diffeomorphic image registration algorithm. NeuroImage *38*, 95–113.

Badre D., and Wagner A.D. (2005). Frontal lobe mechanisms that resolve proactive interference. Cerebral Cortex *15*, 2003–2012.

Becker S., and Lim J. (2003). A computational model of prefrontal control in free recall: Strategic memory use in the california verbal learning task. Journal of Cognitive Neuroscience *15*, 821–832.

Beckmann C.F., DeLuca M., Devlin J.T., and Smith S.M. (2005). Investigations into resting-state connectivity using independent component analysis. Philosophical Transactions of the Royal Society of London B: Biological Sciences *360*, 1001–1013.

Bell A.J., and Sejnowski T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Computation *7*, 1129–1159.

Benjamini Y., and Hochberg Y. (1995). Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. Journal of Royal Statistical Society, Series B *57*, 289–300.

Bosch S.E., Jehee J.F.M., Fernández G., and Doeller C.F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. The Journal of Neuroscience *34*, 7493–7500.

Buckner R.L., Andrews-Hanna J.R., and Schacter D.L. (2008). The brain's default network. Annals of the New York Academy of Sciences *1124*, 1–38.

Buckner R.L., Sepulcre J., Talukdar T., Krienen F.M., Liu H., Hedden T., Andrews-Hanna J.R., Sperling R.A., and Johnson K.A. (2009). Cortical hubs revealed by intrinsic functional connectivity: Mapping, assessment of stability, and relation to alzheimer's disease. The Journal of Neuroscience *29*, 1860–1873.

Budson A.E., and Solomon P.R. (2011). Memory Loss: A Practical Guide for Clinicians (Philadelphia: Elsevier).

Buzsáki G. (1996). The hippocampo-neocortical dialogue. Cerebral Cortex *6*, 81–92.

Calhoun V.D., Adali T., Pearlson G.D., and Pekar J.J. (2001). A method for making group inferences from functional MRI data using independent component analysis. Human Brain Mapping *14*, 140–151.

Calhoun V.D., Liu J., and Adalı T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. NeuroImage *45*, S163–S172.

Chadick J.Z., and Gazzaley A. (2011). Differential coupling of visual cortex with default network or frontal-parietal network based on goals. Nature neuroscience *14*, 830–832.

Clapp W.C., Rubens M.T., and Gazzaley A. (2010). Mechanisms of Working Memory Disruption by External Interference. Cerebral Cortex *20*, 859–872.

Cohen J.D., Perlstein W.M., Braver T.S., Nystrom L.E., Noll D.C., Jonides J., and Smith E.E. (1997). Temporal dynamics of brain activation during a working memory task. Nature *386*, 604–608.

Cohen N.J., and Eichenbaum H. (1993). Memory, amnesia, and the hippocampal system (Cambridge, MA: MIT).

Cole M.W., Bassett D.S., Power J.D., Braver T.S., and Petersen S.E. (2014). Intrinsic and task-evoked network architectures of the human brain. Neuron *83*, 238–251.

Cole M.W., Reynolds J.R., Power J.D., Repovs G., Anticevic A., and Braver T.S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. Nature Neuroscience *16*, 1348–1355.

Corbetta M., and Shulman G. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature reviews. Neuroscience *3*, 201–215.

Danker J.F., and Anderson J.R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. Psychological Bulletin *136*, 87–102.

Davelaar E.J. (2013). A novelty-induced change in episodic (NICE) context account of primacy effects in free recall. Psychology *4*, 695–703.

Diana R.A., Yonelinas A.P., and Ranganath C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends in Cognitive Sciences *11*, 379–386.

Dosenbach N.U., Fair D.A., Cohen A.L., Schlaggar B.L., and Petersen S.E. (2008). A dual-networks architecture of top-down control. Trends in cognitive sciences *12*, 99–105.

Eberhart R., and Kennedy J. (1995). A new optimizer using particle swarm theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995. MHS '95. pp. 39–43.

Etzel J.A., Cole M.W., Zacks J.M., Kay K.N., and Braver T.S. (in press). Reward motivation enhances task coding in frontoparietal cortex. Cerebral Cortex .

Feredoes E., Heinen K., Weiskopf N., Ruff C., and Driver J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. Proceedings of the National Academy of Sciences *108*, 17510–17515.

Fornito A., Harrison B.J., Zalesky A., and Simons J.S. (2012). Competitive and cooperative dynamics of

large-scale brain functional networks supporting recollection. Proceedings of the National Academy of Sciences *109*, 12788–12793.

Fox M.D., Corbetta M., Snyder A.Z., Vincent J.L., and Raichle M.E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proceedings of the National Academy of Sciences of the United States of America *103*, 10046–10051.

Fox M.D., Snyder A.Z., Vincent J.L., Corbetta M., Essen D.C.V., and Raichle M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences of the United States of America *102*, 9673–9678.

Friston K.J., Buechel C., Fink G.R., Morris J., Rolls E., and Dolan R.J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. NeuroImage *6*, 218–229.

Friston K.J., Price C.J., Fletcher P., Moore C., Frackowiak R.S., and Dolan R.J. (1996). The trouble with cognitive subtraction. NeuroImage *4*, 97–104.

Gershman S.J., Schapiro A.C., Hupbach A., and Norman K.A. (2013). Neural context reinstatement predicts memory misattribution. The Journal of Neuroscience *33*, 8590–8595.

Gitelman D.R., Penny W.D., Ashburner J., and Friston K.J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. NeuroImage *19*, 200–207.

Glanzer M., and Cunitz A.R. (1966). Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behavior *5*, 351–360.

Gordon A.M., Rissman J., Kiani R., and Wagner A.D. (2014). Cortical reinstatement mediates the

relationship between content-specific encoding activity and subsequent recollection decisions. Cerebral Cortex (New York, N.Y.: 1991) *24*, 3350–3364.

Hasselmo M., and Stern C. (2006). Mechanisms underlying working memory for novel information. Trends in cognitive sciences *10*, 487–493.

Howard M.W., and Kahana M.J. (2002). A distributed representation of temporal context. Journal of Mathematical Psychology *46*, 269–299.

Jha A.P., Fabian S.A., and Aguirre G.K. (2004). The role of prefrontal cortex in resolving distractor interference. Cognitive, Affective, & Behavioral Neuroscience *4*, 517–527.

Johnson J.D., and Rugg M.D. (2007). Recollection and the reinstatement of encoding-related cortical activity. Cerebral Cortex *17*, 2507–2515.

Khaligh-Razavi S.M., and Kriegeskorte N. (2014). Deep supervised, but not unsupervised, models may explain it cortical representation. PLOS Computational Biology *10*, e1003915.

Kragel J.E., Morton N.W., and Polyn S.M. (2015). Neural Activity in the Medial Temporal Lobe Reveals the Fidelity of Mental Time Travel. The Journal of Neuroscience *35*, 2914–2926.

Kragel J.E., and Polyn S.M. (2013). Functional interactions between large-scale networks during memory search. Cerebral Cortex .

Kragel J.E., and Polyn S.M. (2015). Functional Interactions Between Large-Scale Networks During Memory Search. Cerebral Cortex *25*, 667–679.

Kruschke J.K. (2010). What to believe: Bayesian methods for data analysis. Trends in Cognitive Sciences *14*, 293–300.

Kuhl B.A., and Chun M.M. (2014). Successful remembering elicits event-specific activity patterns in lateral parietal cortex. The Journal of Neuroscience *34*, 8051–8060.

Laming D. (1999). Testing the idea of distinct storage mechanisms in memory. International Journal of Psychology *34*, 419–426.

Lee S.H., Kravitz D.J., and Baker C.I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nature Neuroscience *16*, 997–999.

Li Y., Adali T., and Calhoun V.D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. Human Brain Mapping *28*, 1251–1266.

Mack M.L., Preston A.R., and Love B.C. (2013). Decoding the brain's algorithm for categorization from its neural implementation. Current Biology *23*, 2023–2027.

Manning J.R., Polyn S.M., Baltuch G., Litt B., and Kahana M.J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. Proceedings of the National Academy of Sciences of the United States of America *108*, 12893–12897.

McClelland J.L., McNaughton B.L., and O'Reilly R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review *102*, 419–57.

Meda S.A., Stevens M.C., Folley B.S., Calhoun V.D., and Pearlson G.D. (2009). Evidence for anomalous network connectivity during working memory encoding in schizophrenia: An ICA based analysis. PLoS ONE *4*, e7911.

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

Miller E.K., Erickson C.A., and Desimone R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. Journal of Neuroscience *16*, 5154.

Miller R. (1991). Cortico-hippocampal interplay and the representation of contexts in the brain (Springer-Verlag).

Mitchell T.M., Shinkareva S.V., Carlson A., Chang K.M., Malave V.L., Mason R.A., and Just M.A. (2008). Predicting human brain activity associated with the meanings of nouns. Science *320*, 1191.

Monti R.P., Hellyer P., Sharp D., Leech R., Anagnostopoulos C., and Montana G. (2014). Estimating time-varying brain connectivity networks from functional MRI time series. NeuroImage *103*, 427–443.

Morton N.W., Kahana M.J., Rosenberg E.A., Baltuch G.H., Litt B., Sharan A.D., Sperling M.R., and Polyn S.M. (2013). Category-specific neural oscillations predict recall organization during memory search. Cerebral Cortex *23*, 2407–2422.

Moscovitch M. (1992). Memory and Working-with-Memory: A Component Process Model Based on Modules and Central Systems. Journal of Cognitive Neuroscience *4*, 257–267.

Moscovitch M., and Winocur G. (1992). The neuropsychology of memory and aging. In Handbook of Aging and Cognition, F.I.M. Craik, and T.A. Salthouse, eds. (Hillsdale, NJ: Earlbaum), pp. 315–372.

Moscovitch M., and Winocur G. (2002). The frontal cortex and working with memory. In Principles of frontal lobe function, D.T. Stuss, and R.T. Knight, eds. (New York: Oxford University Press), pp. 188–209.

Nelson D.L., McEvoy C.L., and Schreiber T.A. (2004). The University of South Florida free association, rhyme, and word fragment norms. Behavior Research Methods, Instruments and Computers *36(3)*, 402–407.

Norman K.A., and O'Reilly R.C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. Psychological Review *110*, 611–646.

Okada K., Vilberg K.L., and Rugg M.D. (2012). Comparison of the neural correlates of retrieval success in tests of cued recall and recognition memory. Human Brain Mapping *33*, 523–533.

Palmeri T.J. (2014). An exemplar of model-based cognitive neuroscience. Trends in Cognitive Sciences *18*, 67–69.

Polyn S.M., and Kahana M.J. (2008). Memory search and the neural representation of context. Trends in Cognitive Sciences *12*, 24–30.

Polyn S.M., Kragel J.E., Morton N.W., McCluey J.D., and Cohen Z.D. (2012). The neural dynamics of task context in free recall. Neuropsychologia *50*, 447–457.

Polyn S.M., Natu V.S., Cohen J.D., and Norman K.A. (2005). Category-specific cortical activity precedes retrieval during memory search. Science *310*, 1963–1966.

Polyn S.M., Norman K.A., and Kahana M.J. (2009). A context maintenance and retrieval model of organizational processes in free recall. Psychological Review *116*, 129–156.

Postle B.R. (2005). Delay-period activity in prefrontal cortex: one function is sensory gating. Journal of cognitive neuroscience *17*, 1679–1690.

Postle B.R. (2006). Working memory as an emergent property of the mind and brain. Neuroscience *139*, 23–38.

Postman L., and Phillips L.W. (1965). Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology *17*, 132–138.

Purcell B.A., Heitz R.P., Cohen J.Y., Schall J.D., Logan G.D., and Palmeri T.J. (2010). Neurally constrained modeling of perceptual decision making. Psychological Review *117*, 1113–1143.

Ranganath C., and Ritchey M. (2012). Two cortical systems for memory-guided behaviour. Nature Reviews Neuroscience *13*, 713–726.

Riggall A.C., and Postle B.R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. The Journal of Neuroscience *32*, 12990–12998.

Ritchey M., Wing E.A., LaBar K.S., and Cabeza R. (2013). Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. Cerebral Cortex *23*, 2818–2828.

Sahakyan L., and Kelley C.M. (2002). A contextual change account of the directed forgetting effect. Journal of Experimental Psychology Learning, Memory, and Cognition *28*, 1064–1072.

Sakai K., Rowe J.B., and Passingham R.E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nature Neuroscience *5*, 479–484.

Sederberg P.B., Howard M.W., and Kahana M.J. (2008). A context-based theory of recency and contiguity in free recall. Psychological Review *115*, 893–912.

Sestieri C., Corbetta M., Romani G.L., and Shulman G.L. (2011). Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. The Journal of Neuroscience *31*, 4407–4420.

Sestieri C., Corbetta M., Spadone S., Romani G.L., and Shulman G.L. (2014). Domain-general signals in the cingulo-opercular network for visuospatial attention and episodic memory. Journal of Cognitive Neuroscience *26*, 551–568.

Simons J.S., and Spiers H.J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. Nature Reviews Neuroscience *4*, 637–648.

Smith S.M., Fox P.T., Miller K.L., Glahn D.C., Fox P.M., Mackay C.E., Filippini N., Watkins K.E., Toro R., Laird A.R., and Beckmann C.F. (2009). Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Science USA *106*, 13040–13045.

Sneve M.H., Magnussen S., Alnæs D., Endestad T., and D'Esposito M. (2013). Top–Down Modulation from Inferior Frontal Junction to FEFs and Intraparietal Sulcus during Short-term Memory for Visual Features. Journal of Cognitive Neuroscience *25*, 1944–1956.

Spreng R.N., DuPre E., Selarka D., Garcia J., Gojkovic S., Mildner J., Luh W.M., and Turner G.R. (2014). Goal-congruent default network activity facilitates cognitive control. The Journal of Neuroscience *34*, 14108–14114.

Spreng R.N., and Grady C.L. (2009). Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. Journal of Cognitive Neuroscience *22*, 1112–1123.

Spreng R.N., Stevens W.D., Chamberlain J.P., Gilmore A.W., and Schacter D.L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. NeuroImage *53*, 303–317.

Sridharan D., Levitin D.J., and Menon V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proceedings of the National Academy of Sciences *105*, 12569 –12574.

Staresina B.P., Cooper E., and Henson R.N. (2013). Reversible information flow across the medial temporal

lobe: The hippocampus links cortical modules during memory retrieval. The Journal of Neuroscience *33*, 14184–14192.

Steyvers M., Shiffrin R.M., and Nelson D.L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In Cognitive Psychology and its Applications: Festschrift in Honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer., A.F. Healy, ed. (Washington, DC: American Psychological Association), pp. 237–249.

Svoboda E., McKinnon M.C., and Levine B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. Neuropsychologia *44*, 2189–2208.

Tan L., and Ward G. (2000). A recency-based account of the primacy effect in free recall. Journal of Experimental Psychology Learning, Memory, and Cognition *26*, 1589–1626.

Tulving E. (1993). What is episodic memory? Current Directions in Psychological Science *2*, 67–70.

Tulving E. (2008). On the law of primacy. In Memory and Mind: A Festschrift for Gordon H. Bower, M.A. Gluck, J.R. Anderson, and S.M. Kosslyn, eds. (New York, NY: Taylor & Francis), chap. 3, pp. 31–48.

Turner B.M., Forstmann B.U., Wagenmakers E.J., Brown S.D., Sederberg P.B., and Steyvers M. (2013a). A Bayesian framework for simultaneously modeling neural and behavioral data. NeuroImage *72*, 193–206.

Turner B.M., Sederberg P.B., Brown S.D., and Steyvers M. (2013b). A Method for Efficiently Sampling From Distributions With Correlated Dimensions. Psychological methods *18*, 368–384.

Turner B.M., van Maanen L., and Forstmann B.U. (in press). Informing cognitive abstractions through neuroimaging: The neural drift diffusion model. Psychological Review .

Vilberg K.L., and Rugg M.D. (2012). The neural correlates of recollection: Transient versus sustained fMRI effects. The Journal of Neuroscience *32*, 15679–15687.

Vincent J.L., Kahn I., Snyder A.Z., Raichle M.E., and 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., and Buckner R.L. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. Journal of Neurophysiology *96*, 3517–3531.

Wagenmakers E.J., Lodewyckx T., Kuriyal H., and Grasman R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. Cognitive Psychology *60*, 158–189.

Wager T.D., Spicer J., Insler R., and Smith E.E. (2013). The neural bases of distracter-resistant working memory. Cognitive, Affective, & Behavioral Neuroscience *14*, 90–105.

Wagner A.D., Par-Blagoev E.J., Clark J., and Poldrack R.A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron *31*, 329–338.

Waskom M.L., Kumaran D., Gordon A.M., Rissman J., and Wagner A.D. (2014). Frontoparietal representations of task context support the flexible control of goal-directed cognition. The Journal of Neuroscience *34*, 10743–10755.

Wetzels R., Grasman R.P.P.P., and Wagenmakers E.J. (2010). An encompassing prior generalization of the Savage–Dickey density ratio. Computational Statistics & Data Analysis *54*, 2094–2102.

Wheeler M.E., Petersen S.E., and Buckner R.L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. Proceedings of the National Academy of Sciences of the United States of America *97*, 11125–11129.

Wilks S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. The Annals of Mathematical Statistics *9*, 60–62.

Yeo B.T.T., Krienen F.M., Sepulcre J., Sabuncu M.R., Lashkari D., Hollinshead M., Roffman J.L., Smoller J.W., Zöllei L., Polimeni J.R., Fischl B., Liu H., and Buckner R.L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology *106*, 1125–1165.

Yu Q., Erhardt E.B., Sui J., Du Y., He H., Hjelm D., Cetin M.S., Rachakonda S., Miller R.L., Pearlson G., and Calhoun V.D. (2015). Assessing dynamic brain graphs of time-varying connectivity in fMRI data: Application to healthy controls and patients with schizophrenia. NeuroImage *107*, 345–355.

Zalesky A., Fornito A., Cocchi L., Gollo L.L., and Breakspear M. (2014). Time-resolved resting-state brain networks. Proceedings of the National Academy of Sciences *111*, 10341–10346.

Zanto T.P., Clapp W.C., Rubens M.T., Karlsson J., and Gazzaley A. (in press). Expectations of Task Demands Dissociate Working Memory and Long-Term Memory Systems. Cerebral Cortex .

Figure Legends

Figure 1 Using free recall to investigate context-based mechanisms. **(a)** Free recall paradigm. **(b)** Mean proportion of items recalled and **(c)** probability of first recall for each delay condition. Shaded regions represent SEM.

Figure 2 Estimation of large-scale networks using ICA. **(a)** Spatial ICA was applied to temporally concatenated group fMRI data, resulting in identification of 61 ICs. A spatial correlation analysis was used to identify ICs corresponding to parcellation maps of the cortical surface (Yeo et al., 2011). Six ICs exhibited positive spatial correlations with networks of interest. DAN, dorsal attention network; VAN ventral attention network; FPCN, frontoparietal control network; DMN default mode network. **(b)** Large-scale networks and their response to distraction. Left, representative t-maps of group-level spatial IC loadings ($p < 0.001$, uncorrected; see Figure S1 for complete visualization). Right, average β parameters estimating IC activity during the filled and unfilled delay periods. $*$, $p < 0.05$ FDR corrected. a.u., arbitrary units.

Figure 3 Distraction causes engagement of MTL and shifts in functional connectivity. **(a)** A functional ROI within bilateral MTL exhibits increased activity in the presence of distraction. Mean β estimates of sustained delay-period activity within the MTL are depicted, with error bars depicting standard error. $*$, $p < 0.05$. **(b)** Reorganized functional connectivity in the presence of distraction. Significant differences ($p < 0.05$, FDR corrected) in functional connectivity are denoted by weighted lines. Black lines indicate increased functional connectivity during the filled, relative to the unfilled, delay intervals.

Figure 4 Relating large-scale network activity to cognitive mechanisms. **(a)** Behavioral correlates of contextual disruption (CD). A low value of the β_{ri} parameter indicates that a contextual retrieval cue is successfully maintained in the face of distraction, leading to an enhanced likelihood of initiating recall with an item from the end of the list. In contrast, a high value of β_{ri} indicates that the contextual retrieval cue is disrupted, leading to a more even spread of recall initiation across the list items. **(b)** Behavioral correlates of contextual retrieval (CR). A high value of the β_s parameter indicates the successful retrieval of the context representation associated with the beginning of the study list, leading to an enhanced likelihood of initiating recall with an item from the start of the list. In contrast, a low β_s value indicates failure to retrieve start-list context, leading to a more even spread of recall initiation across the list items.

Figure 5 Large-scale network and MTL activity reflect distinct computational mechanisms during the filled-delay condition. **(a)** Evidence for neurally informed context disruption models. **(b)** Evidence for context reinstatement models. Deviance statistics above the critical value (denoted in the dashed line) are significant at $p < 0.05$. **(c)** Changes in recall initiation associated with delay-period activation. Left, average activity during the filled delay. The dashed and dotted vertical lines denote the onset of the delay and recall periods, respectively. Right, average probability of recall initiation. Data are partitioned at the 80% quantile. Error bars depict SEM.

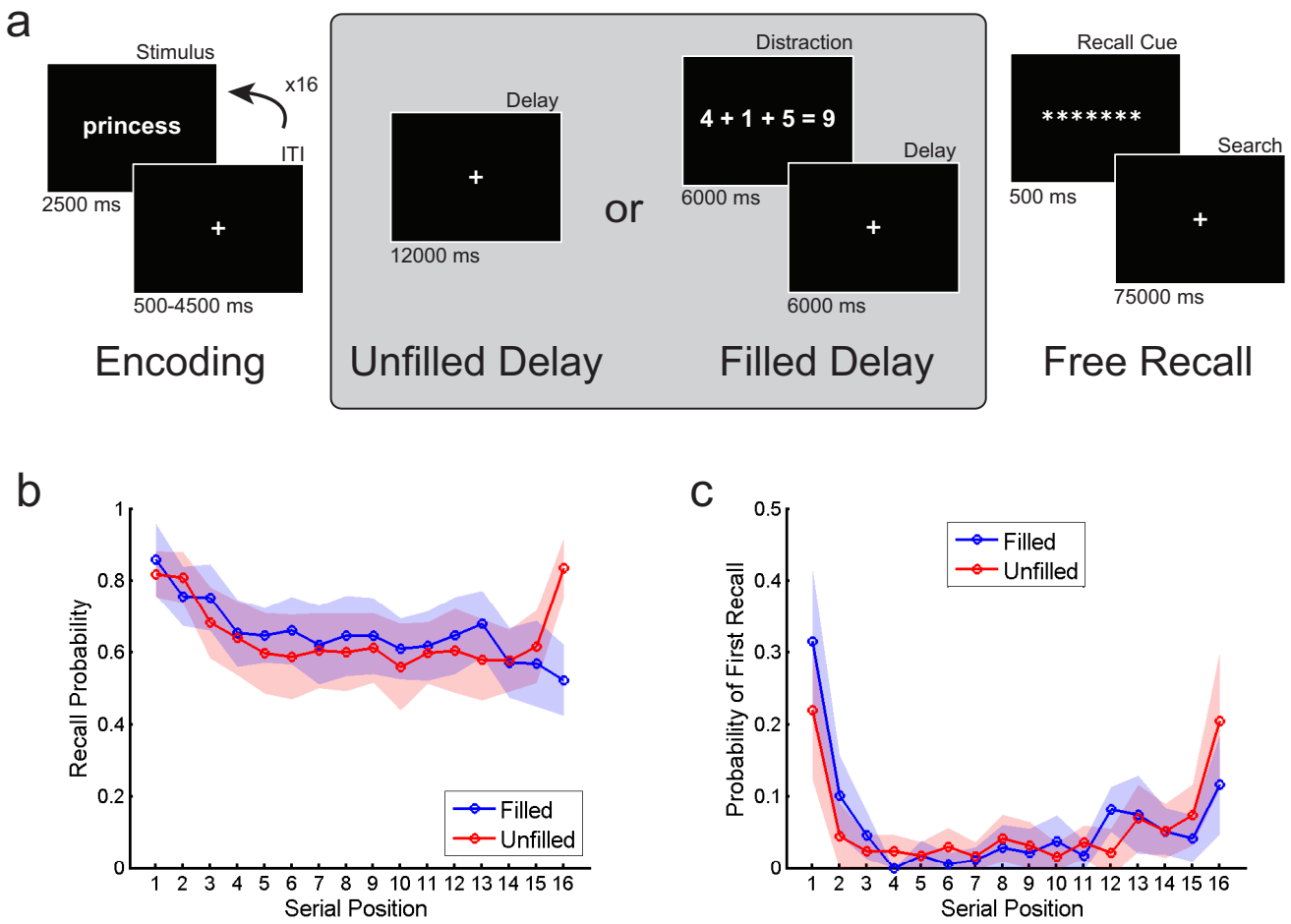

Figure 1. Using free recall to investigate context-based mechanisms. **(a)** Free recall paradigm. **(b)** Mean proportion of items recalled and **(c)** probability of first recall for each delay condition. Shaded regions represent SEM.

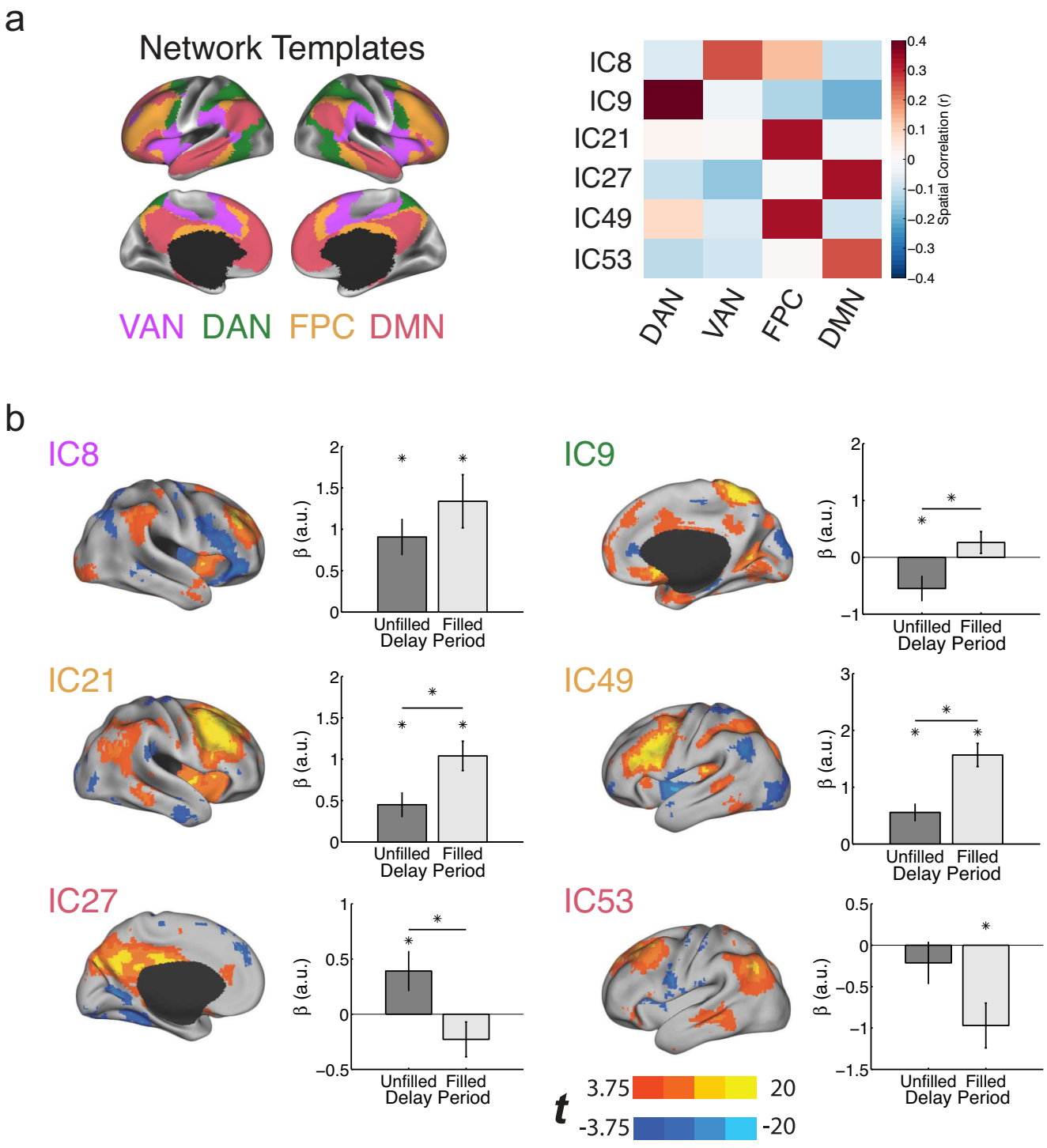

Figure 2. Estimation of large-scale networks using ICA. **(a)** Spatial ICA was applied to temporally concatenated group fMRI data, resulting in identification of 61 ICs. A spatial correlation analysis was used to identify ICs corresponding to parcellation maps of the cortical surface (Yeo et al., 2011). Six ICs exhibited positive spatial correlations with networks of interest. DAN, dorsal attention network; VAN ventral attention network; FPCN, frontoparietal control network; DMN default mode network. **(b)** Large-scale networks and their response to distraction. Left, representative t-maps of group-level spatial IC loadings ($p < 0.001$, uncorrected; see Figure S1 for complete visualization). Right, average β parameters estimating IC activity during the filled and unfilled delay periods. $*$, $p < 0.05$ FDR corrected. a.u., arbitrary units.

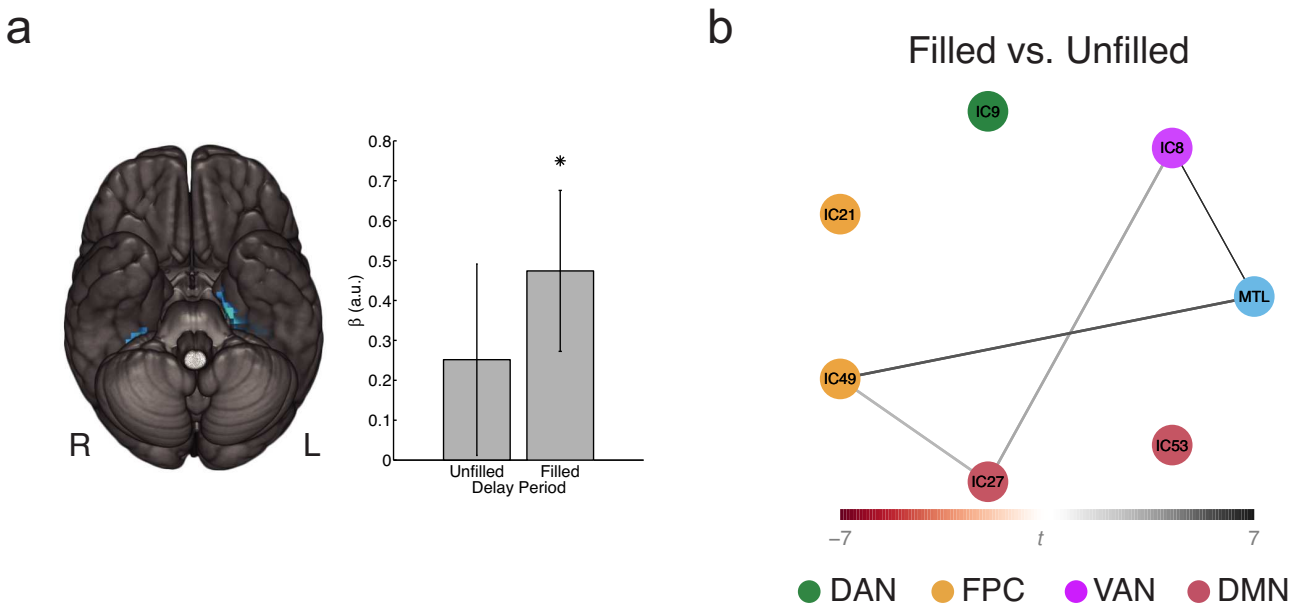

Figure 3. Distraction causes engagement of MTL and shifts in functional connectivity. **(a)** A functional ROI within bilateral MTL exhibits increased activity in the presence of distraction. Mean β estimates of sustained delay-period activity within the MTL are depicted, with error bars depicting standard error. $*$, $p < 0.05$. **(b)** Reorganized functional connectivity in the presence of distraction. Significant differences ($p < 0.05$, FDR corrected) in functional connectivity are denoted by weighted lines. Black lines indicate increased functional connectivity during the filled, relative to the unfilled, delay intervals.

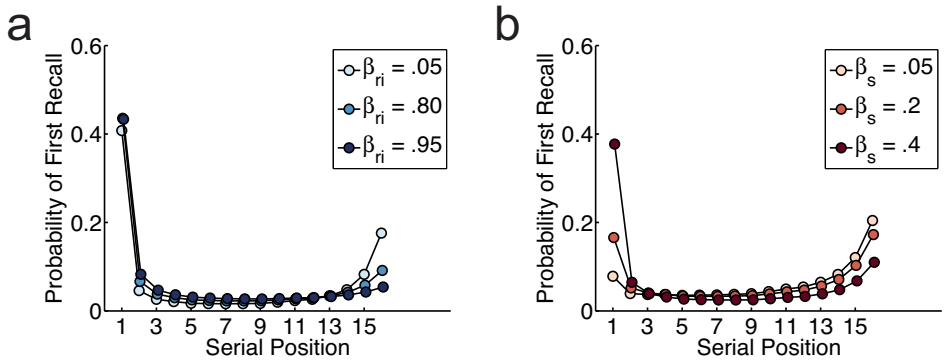

Figure 4. Relating large-scale network activity to cognitive mechanisms. **(a)** Behavioral correlates of contextual disruption (CD). A low value of the β_{ri} parameter indicates that a contextual retrieval cue is successfully maintained in the face of distraction, leading to an enhanced likelihood of initiating recall with an item from the end of the list. In contrast, a high value of β_{ri} indicates that the contextual retrieval cue is disrupted, leading to a more even spread of recall initiation across the list items. **(b)** Behavioral correlates of contextual retrieval (CR). A high value of the β_s parameter indicates the successful retrieval of the context representation associated with the beginning of the study list, leading to an enhanced likelihood of initiating recall with an item from the start of the list. In contrast, a low β_s value indicates failure to retrieve start-list context, leading to a more even spread of recall initiation across the list items.

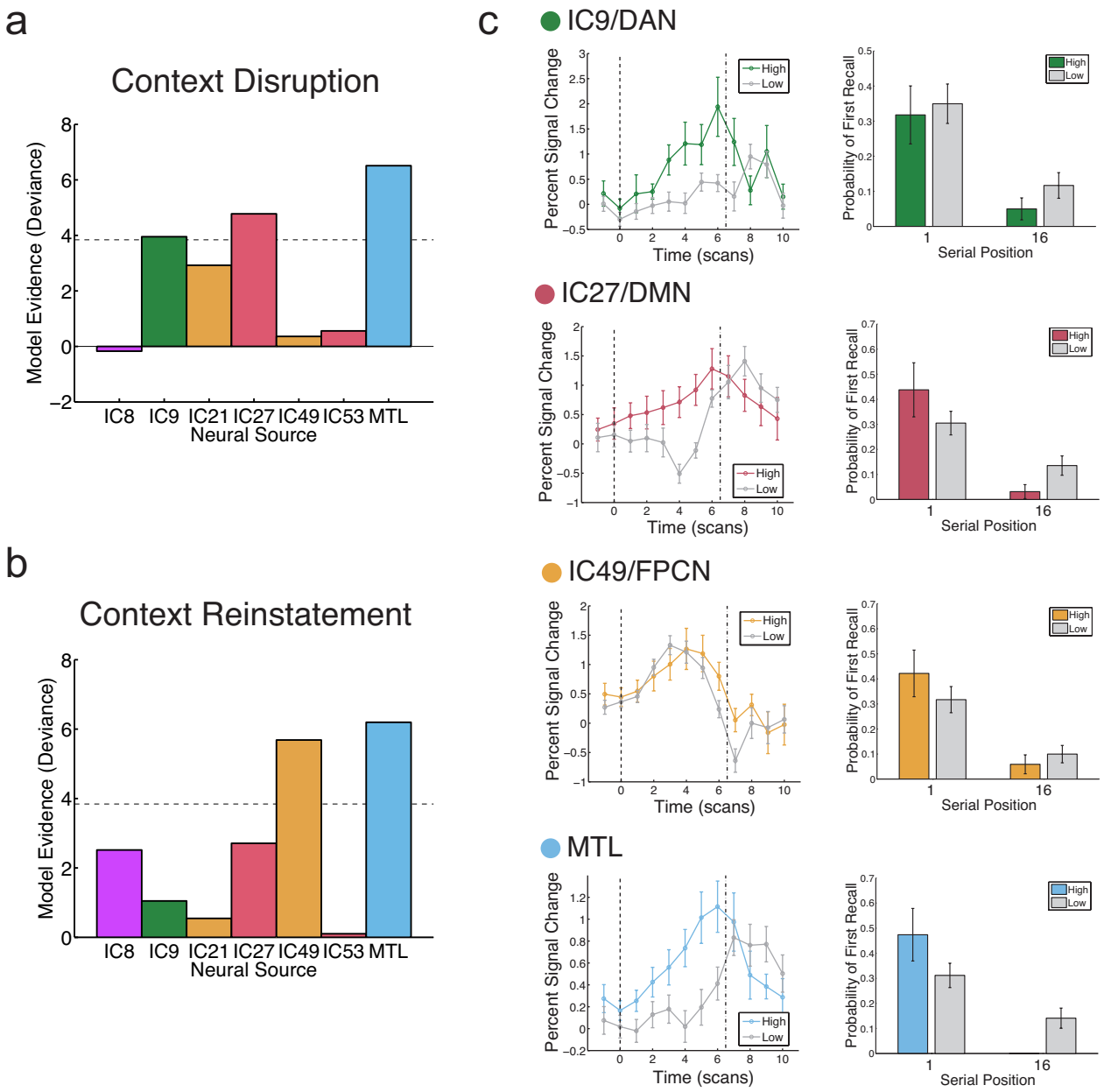

Figure 5. Large-scale network and MTL activity reflect distinct computational mechanisms during the filled-delay condition. **(a)** Evidence for neurally informed context disruption models. **(b)** Evidence for context reinstatement models. Deviance statistics above the critical value (denoted in the dashed line) are significant at $p < 0.05$. **(c)** Changes in recall initiation associated with delay-period activation. Left, average activity during the filled delay. The dashed and dotted vertical lines denote the onset of the delay and recall periods, respectively. Right, average probability of recall initiation. Data are partitioned at the 80% quantile. Error bars depict SEM.

Supplemental Data

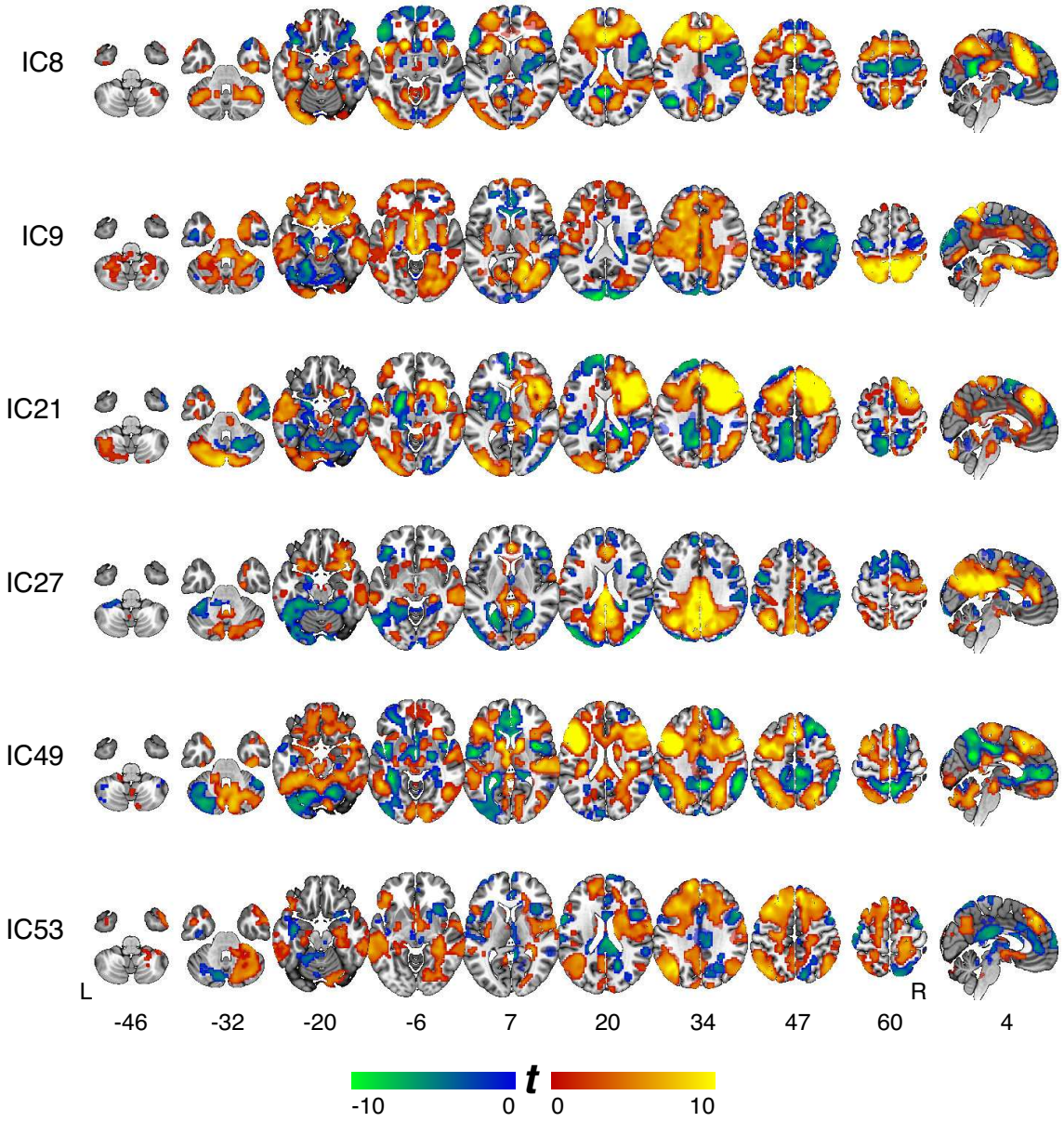

Figure S1. Spatial consistency of IC maps. *t*-maps of group-level spatial IC loadings ($p < 0.001$, uncorrected). Slice coordinates are given in MNI space. L, left; R, right.

Table S1. Positive spatial loadings of independent components

Network	Gyrus (Hemisphere)	BA	*k*	Coordinates (x,y,z)	*t*
IC8	Middle Frontal Gyrus (R)	8	13000	36, 35, 40	16.89
	Superior Frontal Gyrus (L)	9		-24, 50, 31	15.62
	Superior Frontal Gyrus (R)	6		0, 11, 52	14.87
	Precuneus (R)	31		15, -70, 31	13.37
	Fusiform Gyrus (L)	19	1585	-48, -73, -8	12.36
	Fusiform Gyrus (R)	18	575	27, -97, -11	8.44
	Middle Occipital Gyrus (R)	19		51, -73, 7	3.82
	Middle Frontal Gyrus (R)	6	55	48, 5, 52	7.48
	Anterior Cingulate (R)	32	87	15, 44, -5	6.60
	Superior Temporal Gyrus (L)	38	244	-57, 17, -20	6.25
	Inferior Parietal Lobule (L)	40	277	-63, -34, 46	5.45
IC9	Precuneus (R)	7	18667	24, -55, 64	18.34
	Lingual Gyrus (R)	18		12, -70, 1	14.41
	Superior Parietal Lobule (L)	7		-15, -61, 67	13.53
	Caudate (R)			6, 17, -8	13.52
	Declive (L)		209	-12, -94, -17	6.08
	Superior Frontal Gyrus (R)	6	159	27, 8, 67	5.54
IC21	Cingulate Gyrus (R)	32	15590	9, 26, 37	19.29
	Middle Frontal Gyrus (R)	6		36, -4, 43	16.34
	Middle Frontal Gyrus (R)	9		45, 29, 25	15.65
	Middle Occipital Gyrus (L)	19		-39, -88, 10	14.17
	Middle Temporal Gyrus (L)	21	1285	-60, -4, -8	11.92
	Superior Temporal Gyrus (L)	22		-78, -28, 16	6.08
	Uncus (L)	28		-21, 2, -32	5.01
	Inferior Frontal Gyrus (L)	9		-60, 14, 25	4.43
	Paracentral Lobule (R)	5	418	0, -28, 61	9.29
	Middle Frontal Gyrus (L)	46	146	-42, 50, -2	7.91
	Precuneus (R)	31	64	9, -46, 34	7.12
	Inferior Parietal Lobule (L)	40	150	-51, -28, 46	6.57
	Culmen (R)		57	21, -31, -17	5.11
IC27	Cingulate Gyrus (R)	31	7257	9, -40, 31	19.34
	Precuneus (L)	19		-36, -70, 40	14.81
	Precuneus (R)	7		33, -67, 40	9.43
	Middle Occipital Gyrus (R)	18		30, -88, -2	9.33
	Anterior Cingulate (R)	32	423	3, 41, 10	10.49
	Extra-Nuclear (R)	13	616	39, 14, -14	8.74
	Parahippocampal Gyrus (R)	35		24, -13, -32	4.71
	Middle Temporal Gyrus (R)	21	399	60, -19, -11	8.19
	Inferior Frontal Gyrus (L)	45	340	-63, 32, 4	7.88
	Precentral Gyrus (L)	4		-66, -4, 25	5.66
	Extra-Nuclear (L)	13	217	-30, 17, -17	7.29
	Inferior Frontal Gyrus (R)	45	235	69, 29, 13	6.84
	Precentral Gyrus (L)	9	206	-39, 20, 31	6.82
	Uvula (L)		385	-9, -76, -32	6.41
	Cerebellar Tonsil (R)			45, -70, -35	5.75
	Middle Temporal Gyrus (L)	21	108	-66, -31, -8	6.08
	Middle Frontal Gyrus (R)	8	162	57, 29, 37	5.72
	Superior Frontal Gyrus (R)	8	67	30, 50, 43	5.67
	Inferior Occipital Gyrus (L)	19	80	-33, -79, 4	5.66
IC49	Precentral Gyrus (L)	6	16590	-36, 8, 28	17.94
	Inferior Frontal Gyrus (R)	9		54, 11, 28	16.69
	Medial Frontal Gyrus (R)	6		0, 23, 46	15.68
	Precuneus (R)	7		33, -61, 40	12.88
	Angular Gyrus (L)	39	1794	-39, -55, 43	10.86
	Precuneus (L)	19		-15, -79, 52	8.13
IC53	Superior Frontal Gyrus (L)	8	11595	-15, 50, 37	13.77
	Angular Gyrus (L)	39		-45, -67, 40	11.14
	Middle Frontal Gyrus (L)	6		-36, 17, 43	10.47
	Postcentral Gyrus (R)	3		27, -25, 58	9.93
	Inferior Parietal Lobule (R)	40	453	45, -52, 46	8.38
	Middle Temporal Gyrus (L)	21	784	-66, -40, -8	7.93
	Caudate (L)		71	-12, -13, 22	7.60
	Lentiform Nucleus (L)		85	-30, -10, 1	6.96
	Cuneus (R)	18	99	18, -94, 28	6.06
	Parahippocampal Gyrus (L)	28	85	-21, -28, -11	6.06
	Caudate (L)		147	-9, 20, 1	5.97
	Middle Temporal Gyrus (L)	21	70	-48, -1, -32	4.82

Note: BA = Brodmann area; L = left; R = right. *k* = cluster extent, in voxels. Brodmann areas and anatomical locations are approximate. Coordinates are given in Talairach space. The top 5 peak t-values per cluster separated by a minimum distance of 10 mm are reported. Peaks localized within grey matter are reported.

Table S2. Maximum likelihood parameter estimates and credible intervals for neurally informed model variants.

	Model														
	Baseline	Context Retrieval							Context Disruption						
Parameter		IC8	IC9	IC21	IC27	IC49	IC53	MTL	IC8	IC9	IC21	IC27	IC49	IC53	MTL
β_{enc}	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.69 (0.66 0.75)	0.69 (0.66 0.75)	0.70 (0.66 0.75)	0.70 (0.66 0.75)	0.71 (0.66 0.75)
β_{rec}	0.88 (0.85 0.91)	0.87 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.87 (0.85 0.91)	0.87 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.88 (0.85 0.91)	0.87 (0.85 0.91)	0.88 (0.85 0.91)	0.87 (0.85 0.91)
γ	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.50 (0.44 0.55)	0.49 (0.44 0.55)	0.49 (0.44 0.55)	0.48 (0.44 0.55)
$\xi_{d,u}$	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)	0.34 (0.32 0.34)
$\xi_{d,f}$	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)	0.34 (0.33 0.35)
s	0.98 (0.81 1.14)	0.98 (0.81 1.14)	0.99 (0.81 1.14)	0.99 (0.81 1.14)	0.98 (0.81 1.14)	0.98 (0.81 1.14)	0.98 (0.81 1.14)	0.99 (0.81 1.14)	0.99 (0.81 1.14)	0.99 (0.81 1.14)	1.00 (0.81 1.14)	1.01 (0.81 1.14)	0.97 (0.81 1.14)	0.99 (0.81 1.14)	0.98 (0.81 1.14)
ϕ_s	1.88 (1.62 2.15)	1.89 (1.62 2.15)	1.85 (1.62 2.15)	1.87 (1.62 2.15)	1.87 (1.62 2.15)	1.89 (1.62 2.15)	1.89 (1.62 2.15)	1.84 (1.62 2.15)	1.85 (1.62 2.15)	1.87 (1.62 2.15)	1.84 (1.62 2.15)	1.83 (1.62 2.15)	1.85 (1.62 2.15)	1.87 (1.62 2.15)	1.89 (1.62 2.15)
ϕ_d	0.76 (0.68 0.82)	0.76 (0.68 0.82)	0.75 (0.68 0.82)	0.75 (0.68 0.82)	0.75 (0.68 0.82)	0.76 (0.68 0.82)	0.76 (0.68 0.82)	0.74 (0.68 0.82)	0.75 (0.68 0.82)	0.75 (0.68 0.82)	0.73 (0.68 0.82)	0.72 (0.68 0.82)	0.77 (0.68 0.82)	0.75 (0.68 0.82)	0.78 (0.68 0.82)
τ	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.69 (0.66 0.72)	0.70 (0.66 0.72)	0.71 (0.66 0.72)	0.68 (0.66 0.72)	0.69 (0.66 0.72)	0.68 (0.66 0.72)
λ	0.51 (0.39 0.63)	0.50 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.51 (0.39 0.63)	0.52 (0.39 0.63)	0.51 (0.39 0.63)	0.53 (0.39 0.63)	0.49 (0.39 0.63)	0.51 (0.39 0.63)	0.52 (0.39 0.63)
$\beta_{ri,u}$	0.81 (0.73 0.87)	0.81 (0.73 0.87)	0.80 (0.73 0.87)	0.81 (0.73 0.87)	0.81 (0.73 0.87)	0.81 (0.73 0.87)	0.81 (0.73 0.87)	0.80 (0.73 0.87)	0.81 (0.78 0.84)	0.81 (0.73 0.86)	0.80 (0.76 0.85)	0.79 (0.73 0.84)	0.80 (0.74 0.85)	0.81 (0.78 0.82)	0.82 (0.79 0.84)
$\beta_{ri,f}$	0.78 (0.70 0.85)	0.79 (0.70 0.85)	0.77 (0.70 0.85)	0.78 (0.70 0.85)	0.78 (0.70 0.85)	0.80 (0.70 0.85)	0.79 (0.70 0.85)	0.77 (0.70 0.85)	0.78 (0.68 0.85)	0.66 (0.67 0.76)	0.69 (0.63 0.80)	0.37 (0.37 0.38)	0.82 (0.64 0.90)	0.76 (0.72 0.79)	0.81 (0.77 0.83)
$\beta_{s,u}$	0.37 (0.32 0.42)	0.37 (0.32 0.42)	0.37 (0.31 0.43)	0.37 (0.32 0.44)	0.37 (0.32 0.40)	0.36 (0.31 0.40)	0.36 (0.30 0.42)	0.37 (0.31 0.43)	0.37 (0.32 0.41)	0.37 (0.30 0.46)	0.38 (0.32 0.42)	0.38 (0.33 0.44)	0.37 (0.32 0.43)	0.37 (0.32 0.41)	0.36 (0.31 0.42)
$\beta_{s,f}$	0.44 (0.39 0.51)	0.44 (0.39 0.50)	0.45 (0.41 0.48)	0.45 (0.40 0.49)	0.44 (0.39 0.51)	0.44 (0.38 0.50)	0.44 (0.39 0.49)	0.45 (0.40 0.50)	0.45 (0.38 0.51)	0.45 (0.39 0.50)	0.46 (0.41 0.51)	0.49 (0.43 0.55)	0.43 (0.36 0.54)	0.45 (0.41 0.49)	0.42 (0.37 0.49)
ν_u	–	0.01 (-0.02 0.04)	0.05 (0.01 0.10)	-0.00 (-0.05 0.04)	-0.00 (-0.04 0.03)	0.04 (0.00 0.09)	-0.04 (-0.09 0.00)	0.04 (0.00 0.08)	-0.00 (-0.03 0.02)	0.04 (-0.00 0.09)	-0.02 (-0.07 0.02)	-0.01 (-0.07 0.02)	0.01 (-0.02 0.05)	-0.00 (-0.05 0.03)	0.03 (0.00 0.06)
ν_f	–	0.02 (0.00 0.05)	0.02 (-0.01 0.05)	-0.02 (-0.06 0.02)	0.04 (-0.00 0.08)	0.05 (0.03 0.08)	0.01 (-0.03 0.05)	0.05 (0.01 0.09)	-0.00 (-0.07 0.04)	0.43 (0.21 0.46)	-0.15 (-0.22 -0.05)	0.98 (0.97 0.98)	0.15 (0.01 0.26)	-0.07 (-0.12 0.04)	0.12 (0.05 0.16)

Note: 95% credible intervals are denoted in parentheses.