





Integrating Region- and Network-level Contributions to Episodic Recollection Using Multilevel Structural Equation Modeling

Kyle A. Kurkela¹, Rose A. Cooper², Ehri Ryu¹, and Maureen Ritchey¹

Abstract

■ The brain is composed of networks of interacting brain regions that support higher-order cognition. Among these, a core network of regions has been associated with recollection and other forms of episodic construction. Past research has focused largely on the roles of individual brain regions in recollection or on their mutual engagement as part of an integrated network. However, the relationship between these region- and network-level contributions remains poorly understood. Here, we applied multilevel structural equation modeling to examine the functional organization of the posterior medial (PM) network and its relationship to episodic memory outcomes. We evaluated two aspects of functional

heterogeneity in the PM network: first, the organization of individual regions into subnetworks, and second, the presence of regionally specific contributions while accounting for network-level effects. Our results suggest that the PM network is composed of ventral and dorsal subnetworks, with the ventral subnetwork making a unique contribution to recollection, especially to recollection of spatial information, and that memory-related activity in individual regions is well accounted for by these network-level effects. These findings highlight the importance of considering the functions of individual brain regions within the context of their affiliated networks. ■

INTRODUCTION

The brain is composed of networks of interacting brain regions, and the functioning of these networks is crucial for higher-level cognition. Within the domain of episodic memory, researchers have identified a set of cortical and hippocampal regions that has consistently been linked to our ability to mentally construct events, including recollecting past events, imagining future events, and constructing fictional scenarios (Ritchey & Cooper, 2020; Buckner & DiNicola, 2019; Rugg & Vilberg, 2013; Ranganath & Ritchey, 2012; Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010). These regions include the hippocampus, angular gyrus (AG), precuneus, retrosplenial cortex (RSC), posterior cingulate cortex (PCC), parahippocampal cortex (PHC), and medial pFC (MPFC), which are all key components of the default mode network, particularly its medial-temporal and core subnetworks (Andrews-Hanna et al., 2010). In recognition of their special role in recollection-based memory processes, these regions have been collectively referred to as the “core recollection network” (Rugg & Vilberg, 2013). More recently, the functions of this network have been situated within the context of two cortico-hippocampal brain networks—an anterior temporal network and a posterior medial (PM) network—that differ on the basis of their

structural and functional connectivity with the medial-temporal lobes and their relation to distinct aspects of memory-guided behavior (Ranganath & Ritchey, 2012). In this framework, the PM network corresponds to the recollection network described above but is assigned a more general role in representing contextual information and event models that support episodic construction, including but not limited to recollection. Exactly how this network supports episodic construction, however, remains unclear. Past research suggests that there are both regional- and network-level contributions of the PM network to episodic recollection. For instance, the hippocampus has long been known to be essential for episodic memory (e.g., Corkin, 2002; Riedel et al., 1999). Using neuroimaging to look beyond the hippocampus, however, it is apparent that the rest of the PM network is also reliably engaged during recollection (Kim, 2013; Rugg & Vilberg, 2013; Spaniol et al., 2009). These regions are robustly structurally and functionally connected with the hippocampus, supporting the idea that they constitute an integrated functional network, yet how this network-level involvement relates to their individual functions remains an open question. Here, we use multilevel structural equation modeling (SEM) to examine heterogeneity in the function of the PM network during an episodic retrieval task. Specifically, we investigated the subnetwork architecture of the PM network, as well as the contributions of individual PM regions to predicting memory outcomes.

¹Boston College, Chestnut Hill, MA, ²Northeastern University, Boston, MA

A great deal of research has focused on the roles of individual brain regions in supporting episodic construction, delineating specific roles for the hippocampus, AG, and other regions of the PM network (Ritchey & Cooper, 2020). The hippocampus, for example, is posited to support the binding together of contextual details in memory (Diana, Yonelinas, & Ranganath, 2007; Eichenbaum, Yonelinas, & Ranganath, 2007; Davachi, 2006) and is thought to perform a pattern completion function in which partial representations evoked by memory cues are “completed” by reinstating related information stored in memory (Horner, Bisby, Bush, Lin, & Burgess, 2015; Norman & O’Reilly, 2003; Marr, 1971). The AG, on the other hand, is thought to support the representation of multimodal episodic details brought to mind during recollection (Humphreys, Lambon Ralph, & Simons, 2021; Ramanan, Piguet, & Irish, 2018; Rugg & King, 2018). Some fMRI studies have directly tested for cognitive and temporal dissociations among the regions of the PM network, finding evidence for functional specialization in the context of both episodic memory (Richter, Cooper, Bays, & Simons, 2016; Vilberg & Rugg, 2012, 2014) and imagination (Thakral, Madore, & Schacter, 2020). For example, Richter et al. (2016) used fMRI to identify brain activity that tracked the success, precision, and subjective vividness of episodic recollection. The authors modeled these measures jointly and found that the hippocampus uniquely tracked whether or not retrieval was successful, the AG uniquely tracked the precision of remembered information, and the precuneus uniquely tracked subjective memory vividness. These findings suggest that individual regions of the PM network make distinct contributions to the recollection process.

There is other evidence pointing to the importance of mutual engagement of these regions as part of an integrated network. fMRI studies have consistently observed increased neural activity across the PM network that is related to successful, vivid episodic recollection (Kim, 2013; Rugg & Vilberg, 2013; Spaniol et al., 2009). Moreover, functional connectivity within the PM network scales with the quality of recollection, including its subjective vividness or the recovery of source details (Cooper & Ritchey, 2019; Geib, Stanley, Dennis, Woldorff, & Cabeza, 2017; Geib, Stanley, Wing, Laurienti, & Cabeza, 2017; King, de Chastelaine, Elward, Wang, & Rugg, 2015; Schedlbauer, Copara, Watrous, & Ekstrom, 2014; Watrous, Tandon, Conner, Pieters, & Ekstrom, 2013). Interestingly, another line of research suggests that the PM network may not act as a single homogeneous network but is instead composed of at least two, highly related subnetworks (Barnett et al., 2021; Cooper, Kurkela, Davis, & Ritchey, 2021; Buckner & DiNicola, 2019; Andrews-Hanna et al., 2010). For example, Cooper et al. (2021) examined the functional connectivity of the PM network regions while participants watched a short movie and, using data-driven community detection, showed that the PM network fractured into ventral and dorsal subnetworks that were differentially

engaged during event perception. Barnett et al. (2021) used a similar data-driven approach to delineate multiple subnetworks in the PM network using resting-state data while also showing evidence that the regions of the different PM subnetworks represented similar information during a memory-guided decision-making task. Although these studies provide support for a subnetwork view of the PM network, it currently remains unclear whether these subnetworks are dissociable in their contributions to episodic recollection.

Prior studies have made it clear that activity in regions of the PM network are related to memory and to one another, but it remains unclear whether their contributions to memory are regionally specific or shared across the network. SEM is a well-suited tool for delineating regional and network-level contributions to behavior (Bolt et al., 2018), allowing for the capturing of common, distributed, network-level contributions by estimating latent variables that captures the covariance among regions of a network. Structural models can then estimate the statistical dependency between these network latent variables and some behavioral variable while also estimating the regional-specific effect of each of the regions, statistically controlling for their membership within larger networks. For instance, Bolt et al. (2018) used SEM to parse the unique contributions of the right dorsolateral pFC to cognitive control from those of the larger frontoparietal control network, showing that the unit of behavioral significance for many common cognitive control tasks was not the right dorsolateral pFC but the shared contributions of the frontoparietal control network. This approach differs from common applications of SEM to study functional interactions supporting cognition (see McIntosh & Protzner, 2012, for a review), which in the context of episodic memory have largely focused on building models of effective connectivity among brain regions (McCormick, St-Laurent, Ty, Valiante, & McAndrews, 2015; Addis, Leclerc, Muscatell, & Kensinger, 2010; McCormick, Moscovitch, Protzner, Huber, & McAndrews, 2010; Rosenbaum, Furey, Horwitz, & Grady, 2010; Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006; Rajah & McIntosh, 2005). For instance, past work taking this approach has shown that episodic retrieval involves increased communication among left frontal and parietal regions (Iidaka et al., 2006) as well as between the hippocampus and regions in the frontal lobes (McCormick et al., 2010, 2015) and sensory cortex (McCormick et al., 2015). Here, rather than focusing on interactions among brain regions, we apply SEM to estimate the specific regional and common network contributions to episodic remembering within a single statistical model.

Taken together, the literature suggests that the regions of the PM network perform dissociable yet interrelated functions and, as a result, make separable contributions to the recollection process. It remains unclear, however, exactly how to combine the findings from experiments taking region-focused approaches and network-focused

approaches—highlighting the need for an approach that can simultaneously take into account network-wide and region-specific contributions to episodic retrieval. The present study uses SEM to model heterogeneity of function of the PM network. We sought to model two key aspects of functional heterogeneity within the network: that larger networks fracture into related subnetworks and that regions of the network make extranetwork contributions to cognition. To this end, we first compared a single network model to a two related subnetworks model, motivated by previous evidence for dissociable ventral and dorsal PM subnetworks that exhibit distinct patterns of functional connectivity during movie watching (Cooper et al., 2021). Next, we modeled region-specific contributions to behavior, controlling for network-level effects (cf. Bolt et al., 2018), to determine whether any regions acted outside their networks in support of episodic recollection.

METHODS

Experiment

Participants

Twenty-eight participants from Cooper and Ritchey (2019) were included in the final set of analyses after excluding individuals who did not complete the study or who had inadequate memory performance (see Cooper & Ritchey, 2019). Participants were selected such that they were between the ages of 18 and 35 years ($M = 21.82$ years, $SD = 3.57$ years; 16 women, 12 men) and had no history of neurological or psychiatric illness. Participants' self-reported ethnicity was as follows: not Hispanic or Latino ($n = 22$) and Hispanic or Latino ($n = 6$). Race was self-reported as White ($n = 18$), Asian ($n = 3$), more than one race ($n = 3$), Black ($n = 2$), and other ($n = 1$), with one participant electing not to report their race ($n = 1$). Participants reported an average of $M = 15.2$ ($SD = 1.67$) years of education. Informed consent was obtained from all participants before the experiment, and participants were reimbursed for their time. All procedures were approved by the Boston College institutional review board.

With respect to statistical power, our analyses focused on trial-to-trial variability in brain activity and memory outcomes. Our data set had a total of 3888 trials nested within 28 participants (22 individuals contributing 144 trials; six individuals contributing 120 trials—see Cooper & Ritchey, 2019). Wolf, Harrington, Clark, and Miller (2013) ran a series of Monte Carlo simulations to determine the minimum sample size required to achieve SEMs that had at least 80% power to detect nonzero parameters with an alpha of .05. To determine if our models were sufficiently powered, we compared the number of observations that we had at the within-subject, across-trial level (i.e., 3888 observations) to the most conservative recommendations made for achieving adequate power in the Wolf simulations, which argue for 460 observations. However, we note

that, unlike our study, the Wolf simulations did not use a multilevel model, and they did not use categorical variables. There does not currently exist any published recommendations for the required n to achieve sufficient power using the type of multilevel SEM that we use in this article. In the absence of alternative recommendations, we argue that our models are likely to be sufficiently powered given that our number of observations at the within-subject, across-trial level far exceeds the most conservative recommendations made by Wolf et al. (2013). Although we have sufficient power to model trial-to-trial variability in this data set, we remain underpowered to model between-subject variability, and thus, we refrain from interpreting any existing subject-to-subject variability in our analyses.

Materials

Memoranda consisted of 144 unique events that were constructed using a combination of 144 episode-unique object stimuli from Brady, Konkle, Gill, Oliva, and Alvarez (2013), six panoramic scenes from the SUN 360 database (Xiao, Ehinger, Oliva, & Torralba, 2012), and 12 sounds from the International Affective Digitized Sounds database (Bradley & Lang, 2007). The grayscale object images were altered such that they took on 1 of 120 unique colors taken from the equally spaced positions in CIELAB color space. Importantly, the object images were selected from the Brady et al. (2013) database such that the objects did not have a characteristic color associated with them (e.g., a picture of the fruit “orange” being characteristically related to color “orange”). In a similar manner, the 360° panoramic background images were transformed into 120 equally spaced 100° field of view images taken at regular intervals around the panorama. Events consisted of the simultaneous presentation of the colored object on top of a randomly chosen view from the panorama, coupled with the presentation of one of the affective sounds. Participants were encouraged to integrate the three features together into a single meaningful event. For example, a red radio could be placed on top of a beach scene with a view of the ocean while the sound of a woman screaming played in the background. See Cooper and Ritchey (2019) for further information.

Procedure

Participants completed six interleaved encoding and retrieval phases while undergoing MRI scanning (see Cooper & Ritchey, 2019). To summarize, during scanned encoding phases, participants were told that they would encounter 24 events consisting of a foreground object, a background scene, and an emotionally evocative sound. Participants were asked to remember each of the events in as much detail as possible, with explicit instructions to try and remember the color of the foreground object, the position of the object within the background scene, and the emotional valence of the sound. During scanned

retrieval phases, participants were tested on their ability to reconstruct episode features from memory. At the beginning of each retrieval trial, participants were shown gray-scale versions of the object stimuli from the previous encoding phase. During this remember period, participants were asked to bring to mind the cued episode in as much detail as possible. Immediately after the remember period, participants were asked to report the emotional valence of the episode's sound using a confidence scale. The confidence scale asked participants to identify their response to the emotional valence question as either with high confidence or with low confidence. After reporting the sound's valence, participants were asked to report the quality of their memory for the remaining two features in a counterbalanced order. Specifically, participants were presented with the object image in a random color on top of a random view from the correct background scene. Participants were instructed to reconstruct the color of the target image using an interactive 360° color wheel and to position the object within the background scene using a similar interactive 360° panoramic scale.

fMRI Data Acquisition

MRI scanning was performed at the Harvard Center for Brain Science using a 3-T Siemens Prisma MRI scanner with a 32-channel head coil. Structural MRI images were collected using a T-1 weighted multiecho MPRAGE protocol with a field of view of 256 mm, 1 mm isotropic voxels, 176 sagittal slices with an interleaved acquisition, a repetition time (TR) of 2530 msec, an echo time of 1.69/3.55/5.41/7.27 msec, a flip angle of 7°, phase encoding from anterior–posterior, parallel imaging with GRAPPA, and an acceleration factor of 2. Functional images were acquired using a whole-brain multiband EPI sequence with a field of view of 208 mm, 2 mm isotropic voxels, 69 slices at T > C –25.0 with interleaved acquisition, a TR of 1500 msec, an echo time of 28 msec, a flip angle of 75°, anterior–posterior phase encoding, parallel imaging with GRAPPA, and an acceleration factor of 2. A total of six scan runs were collected, each of which consisted of 466 TRs.

Analyses

Behavioral Data

Behavioral data from the data set consisted of trialwise error values measured in degrees for the object color and scene position features and of binary data (i.e., 1 = correct; 0 = incorrect) for the sound valence feature (i.e., collapsed across confidence). For consistency across behavioral measures, we transformed the object color and scene position measures into binary measures representing whether a response was correct (1) or incorrect (0), similar to what we have done previously (Cooper & Ritchey, 2019). This was done by taking any trial with an

error smaller than the accuracy threshold (see below) and giving it a score of 1 and taking any trial with an error greater than the accuracy threshold and giving it a score of 0. Descriptive statistics of our behavioral variables are detailed in Table 1.

The accuracy threshold was determined by fitting two probability density functions to group aggregate data within a mixture modeling framework and is described in detail in Cooper and Ritchey (2019). In brief, Cooper and Ritchey (2019) estimated (for the color and scene features separately) the probability that each error resulted from the von Mises as opposed to the uniform distribution. Errors that had less than a 50% chance of fitting the von Mises distribution were labeled as incorrect. This analysis resulted in a threshold of $\pm 57^\circ$ for the object color feature and $\pm 30^\circ$ for the scene position feature and served as our threshold for labeling a response as “correct” or “incorrect.”

MRI Preprocessing

All preprocessing of the MRI data was performed using fMRIPrep v1.0.3 (Esteban et al., 2019). Data were preprocessed using the same steps as in Cooper and Ritchey (2019). First, each T1w volume was corrected for intensity nonuniformity and skull-stripped. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template Version 2009c was performed through nonlinear registration, using brain-extracted versions of both the T1w volume and template. All analyses reported here use structural and functional data in Montreal Neurological Institute space. Brain tissue segmentation of cerebrospinal fluid, white matter, and gray matter was performed on the brain-extracted T1w image. Functional data were slice time-corrected, motion-corrected, and corrected for field distortion. This was followed by coregistration to the corresponding T1w using boundary-based registration with 9 *df*. Physiological noise regressors were extracted using CompCor. A mask to exclude signals with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six aCompCor components were calculated within the intersection of the subcortical mask and the union of cerebrospinal fluid and white matter masks calculated in T1w space, after their projection to the native space of each functional run. Framewise displacement was also calculated for each functional run. No smoothing of the data was performed. For further details of the pipeline please refer to the online documentation: <https://fmriprep.readthedocs.io/en/1.0.3/index.html>. Entire scan runs were excluded from further analysis if more than 20% of frames had a framewise displacement exceeding 0.3 mm. Spike regressors were additionally created and added to our trialwise models (detailed below) by flagging all frames that had a framewise displacement greater than 0.6 mm. In total, six participants had a single scan run excluded from further analysis because of excessive motion.

Table 1. Descriptive Statistics for Variables of Interest

	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>ICC</i>	<i>Correlations</i>										
						<i>RSC</i>	<i>PHC</i>	<i>pAG</i>	<i>aAG</i>	<i>pHipp</i>	<i>Prec</i>	<i>PCC</i>	<i>MPFC</i>	<i>SCENE</i>	<i>COLOR</i>	
RSC	0.049	0.170	-0.567	0.808	0.138											
PHC	-0.028	0.181	-0.763	1.048	0.155	0.324										
pAG	-0.007	0.315	-1.292	1.232	0.138	0.296	0.330									
aAG	0.188	0.352	-1.346	1.648	0.184	0.296	0.119	0.295								
pHipp	0.077	0.156	-0.614	1.182	0.022	0.191	0.184	0.135	0.276							
Prec	0.116	0.215	-0.923	0.971	0.123	0.326	0.265	0.247	0.356	0.285						
PCC	0.144	0.218	-0.614	1.044	0.142	0.282	0.173	0.189	0.418	0.283	0.410					
MPFC	0.135	0.279	-1.176	1.308	0.036	0.162	0.101	0.066	0.342	0.285	0.282	0.402				
SCENE	0.675	0.468	0.000	1.000	0.205	0.194	0.130	0.086	0.050	0.051	0.094	0.055	-0.026			
COLOR	0.724	0.447	0.000	1.000	0.135	0.070	0.046	0.024	0.046	0.032	0.081	0.061	-0.033	0.251		
SOUND	0.758	0.429	0.000	1.000	0.118	0.091	0.070	0.034	0.039	0.066	0.099	0.078	0.021	0.229	0.170	

Neural measures (RSC–MPFC) are the mean *t* value across all voxels within that ROI from the single-trial estimation step. Behavioral measures (SCENE, COLOR, SOUND) are binary, coded such that 1 = correct and 0 = incorrect. Correlations between neural measures are Pearson’s correlation coefficients. Correlations between the behavioral and neural variables are point-biserial correlations coefficients. All descriptive statistics, excluding the ICCs, were calculated ignoring the nested structure. Prec = precuneus; SCENE = scene position correct; COLOR = object color correct; SOUND = sound valence correct.

Trialwise Response Estimates

To estimate trialwise response estimates, we used a multimodel approach proposed by Mumford, Turner, Ashby, and Poldrack (2012) and implemented in SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/>) using in-house MATLAB (<https://www.mathworks.com/products/matlab.html>) scripts. In this approach, a separate general linear model (GLM) is built to estimate the amplitude of the BOLD response for each trial by modeling the response for each trial using its own dedicated regressor and modeling all other trials as a separate regressor (including the following nuisance regressors: translation in the x , y , and z dimensions; rotation in pitch, roll, and yaw; the first five principal components from aCompCorr, framewise displacement, and spike regressors censoring high motion frames). In total, 3888 GLMs were constructed—one for each trial in our data set. Encoding and retrieval phases for our experiment were interleaved, such that each scan run had 24 encoding trials followed by 24 retrieval trials. The GLMs for the current study were restricted to TRs encompassing the retrieval trials from Cooper and Ritchey (2019), that is, excluding the encoding trials. Each retrieval trial was modeled by convolving SPM12's hemodynamic response function with a stick function placed at the onset of the remember period of each retrieval trial. The statistic used as our estimate of the BOLD response was the t -statistic for the regressor corresponding to each individual trial. The t -statistic provides a more sensitive measure than beta values when searching for information within the brain (Misaki, Kim, Bandettini, & Kriegeskorte, 2010) and downweights noisy voxels, allowing them to have a smaller influence on results. Trialwise response estimates were extracted from our ROIs averaged within each ROI and submitted to further analysis.

ROIs

The ROIs for the present analysis are the same ones used in a previous study from our lab investigating interactions among PM regions (Cooper et al., 2021). The ROIs were created using a combination of cortical ROIs from the Default A and Default C subnetworks from the Schaefer Atlas (Schaefer et al., 2018) and a hippocampal ROI from a probabilistic parcellation (Ritchey, Montchal, Yonelinas, & Ranganath, 2015). These anatomical ROIs were combined with a meta-analytic map generated using Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) using the search term “episodic.” Functional peaks from this map within regions of the PM network were selected, and ROIs were drawn around these peaks such that they were of equal size and each contained 100 contiguous voxels. These ROIs were additionally constrained to the left hemisphere because cortical memory retrieval effects are often found to be strongest in the left hemisphere of the brain, which was also evident in the meta-

analytic map used to create the ROIs. The final set of ROIs included the posterior hippocampus (pHipp), PHC, RSC, precuneus, PCC, posterior AG (pAG), anterior AG (aAG), and MPFC. We previously examined the functional connectivity of these ROIs in an independent data set (Cooper et al., 2021). In a community detection analysis, we found evidence for the subnetwork organization of the PM network, including a ventral PM subnetwork including RSC, PHC, and pAG and a dorsal PM subnetwork including the MPFC, pHipp, precuneus, PCC, and aAG. Although there exist multiple possible subnetwork organizations of the PM network, we were motivated by these previous findings to examine a ventral-dorsal subnetwork architecture. We additionally note that this subnetwork grouping appears to align with correlations among the ROIs observed in the current data set (see Table 1).

Multilevel SEM

The present study took a multilevel SEM approach to investigate functional heterogeneity in the PM network. Multilevel SEM allows for the estimation of latent constructs and for modeling of structural paths among those latent constructs in data sets that have a nested structure. This approach is optimal for the current data set that contains observations of ROI activity across trials that are nested within participants. In our data, trials are the Level 1 units and participants are the Level 2 units. Because of the nested structure, the data have two sources of variation: one because of the difference between trials within participants and the other because of the difference between participants. For the neural data, the former represents where the BOLD response estimate (i.e., the t value) is relative to that participant's own average across all trials and the latter represents where each participant's average t value compared with other participant's average t values. Our primary interest was in modeling within-subject, trial-to-trial variability. The between-subject variability in the neural data could represent meaningful differences in individual characteristics, but we did not have a priori hypotheses about these individual differences in the present sample. Therefore, in the multilevel model for the neural data (see below), the between-subject model is specified only so that this source of variability is accounted for and therefore the statistical validity of the within-subject model is not compromised. For the behavioral data, the two sources of variation represent differences in overall memory quality on a trial-to-trial basis and differences in overall accuracy across trials on a subject-to-subject basis.

All modeling was performed using Mplus software Version 8.2 (Muthén & Muthén, 1997–2017). Models were determined to have acceptable levels of model fit if they displayed the following fit indices: root-mean-square error of approximation (RMSEA) $< .06$, comparative fit index (CFI) $> .95$, and a standardized root-mean-square residual (SRMR¹) $< .08$ (Hu & Bentler, 1999). For SRMR,

a separate index was calculated for the within-cluster (i.e., within-subject) and between-cluster (i.e., between-subject) levels. The models were estimated using the MLR (maximum likelihood with robust standard errors) and the WLSMV (weighted least squares means and variances adjusted) estimators in Mplus. The cutoff values cited above are ones that are commonly applied in the SEM literature. We note that these values were originally determined based on simulations of SEMs on continuous variables estimated using a maximum likelihood estimator, whereas our behavioral data contain binary variables and all of our models that contain the behavioral variables use the WLSMV estimator. A recently published report by Xia and Yang (2019) suggests that models fit on categorical data using an estimator similar to Mplus's WLSMV estimator may be overly optimistic when using traditional fit index cutoffs. In the absence of alternative model fit thresholds, we interpret our results with respect to traditional model fit index thresholds. We do, however, use caution and carefully examine all statistics available to make a judgment with respect to model fit. For a summary of all of the models fit in the current article, see Appendix A.

Preliminary analyses. Before performing our multilevel SEM analysis, we verified the necessity of a multilevel analysis by calculating intraclass correlations (ICCs) for each of our variables of interest and by fitting a “null” model designed to test if there is any structure in the between-subject covariance matrix (see Jak, Oort, & Dolan, 2013). The ICC is a statistic that reflects the proportion of variance of a variable that can be attributed to individual differences among our participants. Data sets that contain variables that have ICCs close to 0 will not see an additional benefit from multilevel modeling. In the null model, a saturated (i.e., a model that estimates parameters for all possible variances and covariances among variables) was specified for the within-subject covariance structure, and a null model in which all of the variances and covariances are constrained to be zero was specified for the between-subject covariance structure. If this model fails to satisfactorily fit the data, it suggests that the between-subject variances need to be allowed in the model and therefore calls for a two-level model. If the ICCs are greater than .1 or the null model fits the data poorly, we will conclude that multilevel modeling is required for our data set.

Measurement models. After verifying the necessity of multilevel modeling, we examined the measurement structure underlying our eight ROIs. We tested two measurement models: a single-factor model for the integrated PM network hypothesis and a two-factor model for the two subnetworks hypothesis. In the single-factor model, the eight PM network regions loaded onto a single latent factor at the within-subject level. We did not impose any restriction at the between-subject level

because we did not have a priori hypotheses about the nature of the between-subject variability in neural data. In the two-factor model, RSC, PHC, and pAG loaded on one factor representing the ventral PM network (vPMN), and MPFC, pHipp, precuneus, PCC, and aAG loaded on the other factor representing the dorsal PM network (dPMN; see Figure 1). Again, we did not impose any restriction at the between-subject level. For the measurement models for neural data, the MLR estimator was used.

We next fit a two-level categorical factor model to the behavioral data. This model contained a single latent factor that loaded onto each of our memory measures (i.e., scene position memory, object color memory, and sound valence memory). We placed restrictions on this model such that it had cross-level metric invariance—the factor loadings for Level 1 (i.e., within-subjects, across trials) of the model were set equal to the corresponding factor loadings for Level 2 (i.e., between-subjects, across-subjects) of the model. These restrictions were placed on the model for two reasons. First, the cross-level metric invariance model facilitates the interpretation of the latent construct at both levels as being the within-subject and between-subject components of the same underlying construct. In this context, the Level 1 latent variable represents overall memory for each episode, whereas the Level 2 latent variable represents participant's overall memory ability. Second, the cross-level metric invariance model limits the number of free parameters in the model, avoiding possible estimation problems common to overparameterized models (see Jak, 2019). Because the behavioral variables were binary, the WLSMV estimator in Mplus was used to estimate this model. This behavioral model with cross-level metric invariance was then stitched together with the neural model to form our final measurement model.

Structural models. After establishing good-fitting measurement models, the neural and behavioral models were stitched together to form a single model (see Figure 1). We subsequently fit a series of models to the data to quantify the contribution of the PM network (or PM subnetworks) to memory quality and whether any of the regions made a region-specific contribution to memory quality over and above their network (or subnetwork) contribution. For these models, the WLSMV estimator was used. The baseline structural model contained a structural path from the PM network (or PM subnetworks) latent variable(s) to the memory quality latent variable at the within-subject level. After fitting the baseline model, a series of models were fit to test for region-specific contributions (i.e., one at a time). Each of these models included an additional structural path from the region to memory quality. This direct path reflects the predictive effect of the region after accounting for its participation in the network (or subnetwork). In a secondary set of analyses, we examined paths between the neural

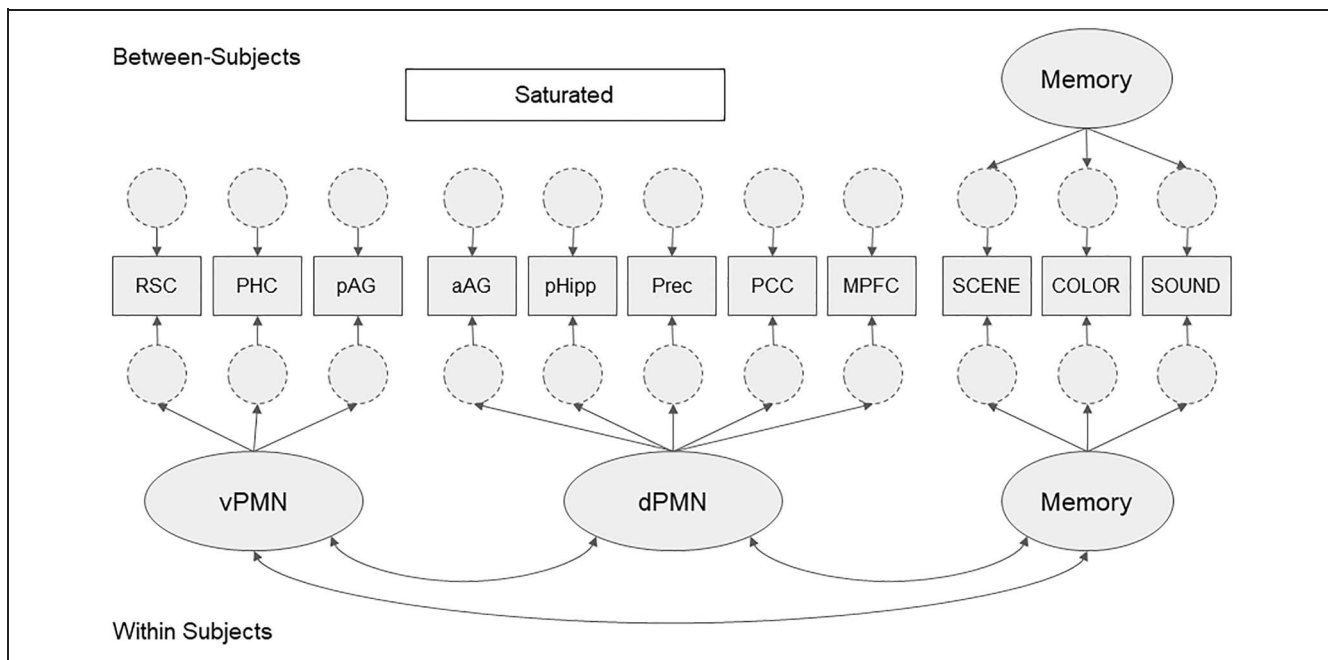


Figure 1. Measurement model. A graphical representation of our combined measurement model including both neural and behavioral variables (Model 6), following the graphing conventions of Kim, Dedrick, Cao, and Ferron (2016). Our measurement model contained two latent variables for the neural data at the within-subject level, a single latent variable for the behavioral data at the within-subject level, and a single latent variable representing the behavioral data at the between-subject level. The factor loadings for the memory latent variable were set equal across the levels. At the between-subject level, the eight neural variables were allowed to covary with one another and with the memory factor. See Table 3 for standardized parameters of the within-subject part of the model. Memory = overall memory quality; Prec = precuneus; Scene = scene position feature correct; Color = object color feature correct; Sound = sound valence feature correct.

variables and memory outcomes for each individual event feature, allowing memory features to covary but removing the latent variable for overall memory quality.

RESULTS

Preliminary Analyses

The ICCs were larger than .1 for the neural measures, except for pHipp whose ICC was .022 and MPFC whose ICC was .036. The null model for the neural data did not fit the data well (Model 1: $\chi^2 = 2951.578$, $df = 36$, $p < .001$, RMSEA = 0.144, CFI = .000, SRMR_{within} = 0.047, SRMR_{between} = 0.352, AIC = -10496.752, BIC = -10221.064). Thus, we concluded that multilevel modeling was appropriate for our neural data. For the behavioral measures, the ICCs ranged from .118 to .205, indicating that about 10%–20% of the variance in the memory measures are due to between-subject differences. The null model resulted in adequate fit to the data (Model 2: $\chi^2 = 20.882$, $df = 6$, $p < .01$, RMSEA = 0.025, CFI = .958, SRMR_{within} = 0.000, SRMR_{between} = 0.634). However, the fit statistics for the behavioral null model were obtained with the WLSMV estimator (because the behavioral memory measures were binary), and applying the same criteria for the WLSMV fit statistics have been shown to be less sensitive to discover model–data misfit (Xia & Yang, 2019). Thus, considering the large ICC values and

the limitation in the performance of the WLSMV fit statistics, we concluded that adopting a multilevel model was also appropriate for the behavioral data.

Measurement Models

The one-factor model for the neural data resulted in the following fit statistics (Model 3: $\chi^2 = 290.442$, $df = 20$, $p < .001$, RMSEA = 0.059, CFI = .905, SRMR_{within} = 0.063, SRMR_{between} = 0.005, AIC = -12584.469, BIC = -12208.530). The two-factor model fit the data well and better than the one-factor model (Model 4: $\chi^2 = 90.096$, $df = 19$, $p < .001$, RMSEA = 0.031, CFI = .975, SRMR_{within} = 0.035, SRMR_{between} = 0.003, AIC = -13145.343, BIC = -12763.138). To further compare model fits, we examined the estimated correlation between the two latent factors in the two-factor model and compared the estimated communality values of the two models. We reasoned that additional evidence in favor of the two-factor model would be seen if the correlation between the latent factors was estimated to be low–moderate and if the estimated communality values were all equivalent or higher for the two-factor model relative to the one-factor model. We note that the correlation between the vPMN and dPMN latent variables was high but not perfect ($r = .630$ or ~39.7% of variance shared), and the communality values in the two-factor model were

all equivalent or higher compared with the one-factor model (see Table 2). Taken together, these results suggest that a two-factor model was a better model for the neural data.

For the behavioral data, the two-level categorical single factor model with equality constraints on the factor loadings across levels fit the data well (Model 5: $\chi^2 = 0.163$, $df = 2$, $p < .9218$, RMSEA = 0.000, CFI = 1.00, SRMR_{within} = 0.001, SRMR_{between} = 0.022). This model has cross-level metric invariance, meaning that the latent variables at each level can be interpreted as the within-subject and between-subject components, respectively, of the same construct “overall memory quality.” Cross-level metric invariance additionally allowed us to calculate the proportion of variance in the overall memory quality factor that is attributable to individual differences and trial-to-trial differences. The memory quality factor had an ICC of .329, meaning that 32.9% of the variability in memory quality comes from individual differences and 67.1% of the variability comes from trial-to-trial differences. This is advantageous for our purposes, because our primary interest was explaining trial-to-trial variability in memory quality.

After finding good fitting neural and behavioral measurement models, we proceeded to fit a joint measurement model by stitching the two-factor neural model and the cross-level metric invariance behavioral model together (see Figure 1, Model 6). At the between-subject level, the regions were allowed to covary with one another and also allowed to covary with the between-subject memory quality factor. This joint measurement model fit the data adequately (Model 6: $\chi^2 = 534.782$, $df = 59$, $p < .001$, RMSEA = 0.046, CFI = .974, SRMR_{within} = 0.034, SRMR_{between} = 0.043). The key parameter estimates for the within-subject part of this model are reported in

Table 2. Community Values

<i>param</i>	<i>One Factor</i>			<i>Two Factor</i>		
	<i>est</i>	<i>se</i>	<i>pval</i>	<i>est</i>	<i>se</i>	<i>pval</i>
pHipp	0.188	0.024	<.001	0.193	0.025	<.001
Prec	0.397	0.039	<.001	0.395	0.039	<.001
PCC	0.437	0.029	<.001	0.479	0.024	<.001
MPFC	0.251	0.030	<.001	0.286	0.032	<.001
PHC	0.177	0.027	<.001	0.344	0.030	<.001
RSC	0.231	0.036	<.001	0.391	0.035	<.001
aAG	0.472	0.030	<.001	0.485	0.032	<.001
pAG	0.215	0.038	<.001	0.363	0.036	<.001

Communality value estimates from the one-factor and two-factor measurement models. All of the estimated communality values are equivalent or higher in the two-factor model compared with the one-factor model. param = parameter; est = estimate; se = standard error; pval = *p* value; Prec = precuneus.

Table 3. Measurement Model Standardized Parameter Estimates

<i>paramHeader</i>	<i>param</i>	<i>est</i>	<i>se</i>	<i>pval</i>
vPMN.BY	RSC	0.627	0.006	<.001
	PHC	0.567	0.008	<.001
	pAG	0.608	0.007	<.001
dPMN.BY	MPFC	0.516	0.006	<.001
	pHipp	0.437	0.008	<.001
	Prec	0.647	0.007	<.001
	PCC	0.685	0.004	<.001
Memory.BY	aAG	0.711	0.005	<.001
	SCENE	0.707	0.057	<.001
	COLOR	0.465	0.044	<.001
dPMN.WITH	SOUND	0.460	0.030	<.001
	vPMN	0.634	0.008	<.001
Memory.WITH	vPMN	0.200	0.025	<.001
	dPMN	0.102	0.026	<.001

Select standardized parameter estimates in the within-subject level model (Model 6). This table was created using the R package MplusAutomation (Hallquist & Wiley, 2018). Parameter headers (*paramHeader*) follow standard Mplus syntax, where the BY key word indicates a loading parameter (λ) and the WITH key word indicates a covariance parameter (θ). param = parameter; est = estimate; se = standard error; pval = *p* value. See Figure 1 caption for abbreviations.

Table 3. This is the model that we then incorporated into our SEM, linking the neural and behavioral variables.

Structural Models

Overall Memory Quality Models

We next estimated a series of structural models to tease apart network and region-specific contributions to overall memory. In the first model (Model 7), each of the two sub-networks was allowed to have a structural path to overall memory quality. In this baseline model, the vPMN uniquely (i.e., when statistically controlling for the dPMN) predicted the overall quality with which events were remembered whereas the dPMN did not (see Figure 2). When modeled separately, however, both the vPMN ($\beta = 0.190$, $SE = 0.021$, $p < .001$) and the dPMN ($\beta = 0.170$, $SE = 0.022$, $p < .001$) predicted memory quality (i.e., in models that included only one of the two paths). Models estimating region-specific contributions to overall memory quality are depicted in Figure 2, and structural path parameter estimates for these models are reported in Table 4. Of the PM network regions, only the MPFC displayed a statistically significant region-specific ability to predict memory quality when controlling for its participation in its PM subnetwork (see Table 4; $\alpha = .05$, family-wise error corrected for multiple comparisons).

Figure 2. Path diagram. Path diagram representing the within-subject level of our two-level baseline model (i.e., Model 7) with standardized parameter estimates (standard error in parentheses). The dotted line depicts the additional region-specific contribution path added in the region-specific family of models detailed in Table 4. See Figure 1 caption for abbreviations. * $p < .05$, ** $p < .01$, *** $p < .001$.

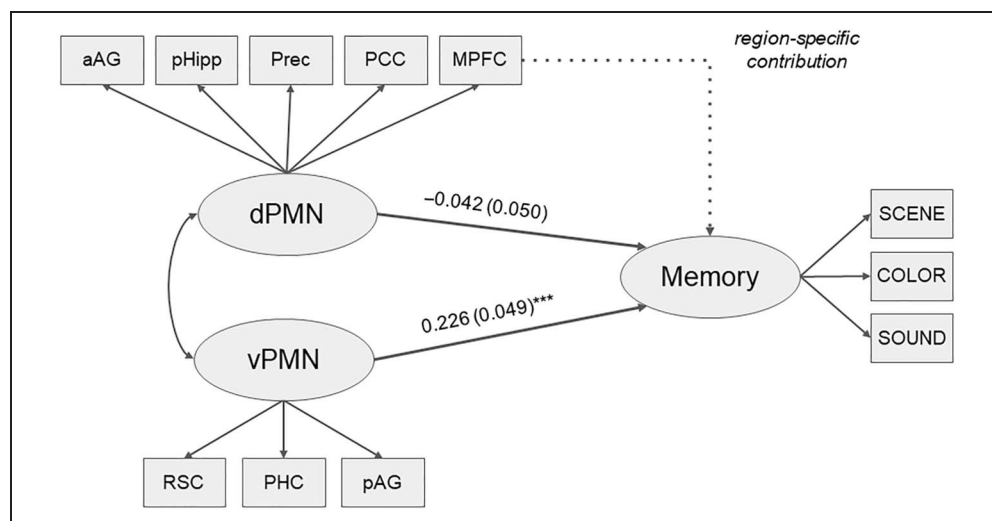
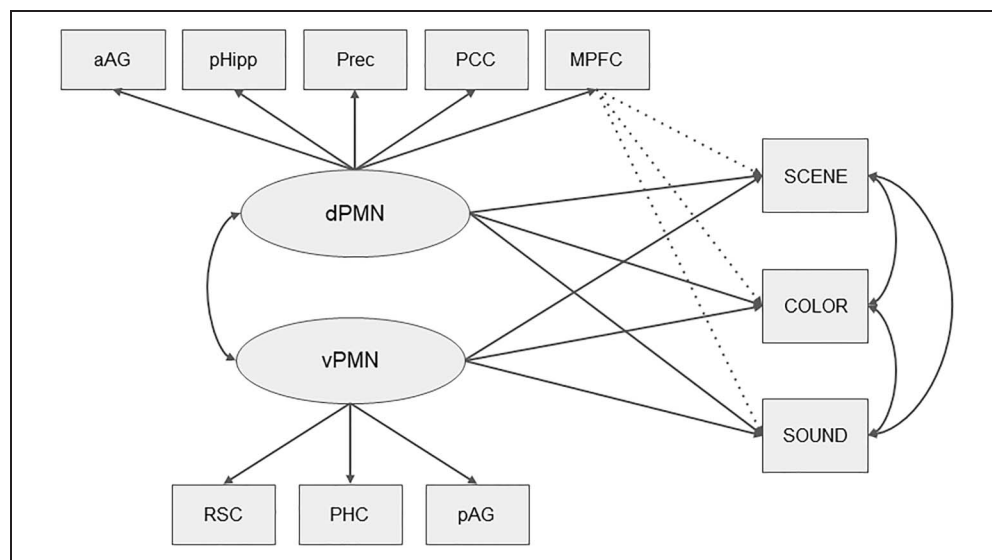


Table 4. Key Parameter Estimates from Region-specific Models

<i>Title</i>	<i>paramHeader</i>	<i>param</i>	<i>est</i>	<i>se</i>	<i>pval</i>
Model 07 _{pHipp}	MEMQ.ON	PHIPP	0.059	0.030	.048
Model 07 _{Prec}	MEMQ.ON	PREC	0.110	0.045	.016
Model 07 _{PCC}	MEMQ.ON	PCC	0.066	0.081	.420
Model 07 _{MPFC}	MEMQ.ON	MPFC	-0.158	0.037	<.001
Model 07 _{PHC}	MEMQ.ON	PHC	0.097	0.090	.282
Model 07 _{RSC}	MEMQ.ON	RSC	0.026	0.045	.575
Model 07 _{aAG}	MEMQ.ON	AAG	0.047	0.057	.415
Model 07 _{pAG}	MEMQ.ON	PAG	-0.063	0.051	.218

This table reports the key parameter estimates for the family of models delineating region-specific contributions. This table was created using the R package MplusAutomation (Hallquist & Wiley, 2018). Parameter headers (paramHeader) follow standard Mplus syntax, where the ON key word indicates a path parameter (β). param = parameter; est = estimate; se = standard error; pval = p value. See Figure 1 caption for abbreviations. Listed p values are uncorrected for multiple comparisons.

Figure 3. Memory feature model path diagram. Path diagram representing our feature specific memory model (Model 9) capturing network-level and region-specific contributions to feature memory. The baseline model contained paths from each subnetwork latent variable to each memory feature (i.e., six in total, solid lines). The region-specific models contained all of the paths from the baseline model with the addition of paths from a single region to each memory feature (i.e., three additional paths, dotted lines), iterated across the entire set of regions. See Table 5 for parameter estimates.



Inspection of the parameter estimates from this alternate model (Model 07_{MPFC}) suggests that the MPFC had a negative relationship with memory quality when controlling for its participation in the dPMN. The absence of other region-specific effects suggests that the contributions of the other PM regions were well described by the subnetwork level effects.

Memory Feature Models

Our primary aim was to examine the region-specific and network-level contributions of PM regions to overall memory quality during retrieval. Our experimental design, however, also afforded us the opportunity to examine their contributions to the retrieval of different memory features

Table 5. Memory Feature Models: Parameter Estimates

<i>Model</i>	<i>paramHeader</i>	<i>param</i>	<i>est</i>	<i>se</i>	<i>pval</i>
Model 9	COLOR.ON	VPMN	-0.004	0.047	.933
	COLOR.ON	DPMN	0.041	0.050	.417
	SOUND.ON	VPMN	0.006	0.043	.882
	SOUND.ON	DPMN	0.114	0.059	.052
	SCENE.ON	VPMN	0.277	0.043	.000
	SCENE.ON	DPMN	-0.147	0.040	.000
Model 9 _{pHIPP}	COLOR.ON	PHIPP	0.015	0.024	.544
	SOUND.ON	PHIPP	0.038	0.032	.234
	SCENE.ON	PHIPP	0.045	0.033	.163
Model 9 _{PREC}	COLOR.ON	PREC	0.073	0.049	.135
	SOUND.ON	PREC	-0.001	0.040	.971
	SCENE.ON	PREC	0.061	0.087	.485
Model 9 _{PCC}	COLOR.ON	PCC	0.026	0.050	.600
	SOUND.ON	PCC	0.014	0.078	.863
	SCENE.ON	PCC	0.083	0.126	.511
Model 9 _{MPFC}	COLOR.ON	MPFC	-0.096	0.027	.000*
	SOUND.ON	MPFC	-0.040	0.036	.272
	SCENE.ON	MPFC	-0.089	0.038	.020
Model 9 _{PHC}	COLOR.ON	PHC	0.012	0.105	.907
	SOUND.ON	PHC	-0.004	0.094	.966
	SCENE.ON	PHC	0.052	0.046	.261
Model 9 _{RSC}	COLOR.ON	RSC	-0.027	0.055	.616
	SOUND.ON	RSC	0.008	0.043	.847
	SCENE.ON	RSC	0.062	0.046	.177
Model 9 _{PAG}	COLOR.ON	PAG	0.025	0.060	.674
	SOUND.ON	PAG	-0.004	0.037	.923
	SCENE.ON	PAG	-0.079	0.033	.016
Model 9 _{AAG}	COLOR.ON	AAG	0.013	0.066	.845
	SOUND.ON	AAG	-0.056	0.053	.285
	SCENE.ON	AAG	0.047	0.089	.596

Key parameter estimates from a family of memory feature specific models. This table was created using the R package MplusAutomation (Hallquist & Wiley, 2018). Parameter headers (paramHeader) follow standard Mplus syntax, where the ON key word indicates a path parameter from the variable listed in the “param” column to variable listed in the “paramHeader” column. param = parameter; est = standardized estimate; se = standard error; pval = *p* value. See Figure 1 caption for abbreviations. See Figure 3 for path diagram. Listed *p* values are uncorrected for multiple comparisons.

* Statistically significant after a Bonferroni correction for multiple comparisons.

(i.e., scene perspective, object color, sound valence). To examine this, we updated our joint measurement model so that the behavioral measures simply covaried with one another instead of loading onto a common factor. This updated measurement model fit the data well (Model 8: $\chi^2 = 683.198$, $df = 37$, $p < .001$, RMSEA = 0.067, CFI = .965, SRMR_{within} = 0.030, SRMR_{between} = 0.000). Using this measurement model, we then fit a series of structural models to examine the network-level and region-specific contributions to each of the features of our events (see Figure 3). Key parameter estimates from this family of models can be found in Table 5. The baseline model (Model 9; see Figure 3) suggests that there were statistically significant network-level contributions of the vPMN and the dPMN to scene feature memory, such that the vPMN contributed positively to scene memory whereas the dPMN contributed negatively. No other network-level effects were statistically significant, although it is worth noting that, in contrast to its negative relationship with scene memory, the dPMN trended toward positive relationships with sound memory. Interestingly, the parameter estimates for the covariances among the residuals of the behavioral variables suggest that there remains a joint “holistic” remembering property that is not explained by PM network activity (see Appendix B for a full table of model parameters). The results from a family of models containing region-specific paths from each region to each memory feature (see Figure 3 and Table 5) suggest that the MPFC made a region-specific negative contribution to object color memory. No other regions made a region-specific contribution after controlling for family-wise error using a Bonferroni correction.

DISCUSSION

In the current study, we examined heterogeneity in the function of the PM network using a multilevel SEM framework. Our measurement models indicated that a two-factor model with latent factors representing ventral and dorsal subnetworks was the best model for our neural data compared with a single-factor model grouping all PM regions together. Our structural models indicated that the contributions of individual regions of the PM network to memory quality are largely subsumed by subnetwork-level contributions, apart from the MPFC, which made a unique, region-specific contribution to memory quality. Interestingly, the region-specific contribution of the MPFC was found to be negative, such that less MPFC activation (when controlling for subnetwork membership) was associated with more accurate recollections. Feature-specific analyses revealed that the dissociation between vPMN and dPMN was driven largely by their distinct contributions to memory for scene information, compared with object color or sound valence information. Together, these results reveal new insights into how memory outcomes can be explained by a combination of network-level and region-specific factors.

Our results support the presence of dissociable subnetworks within the PM network (Barnett et al., 2021; Cooper et al., 2021; Buckner & DiNicola, 2019; Andrews-Hanna et al., 2010). Previous studies have shown evidence for highly related subnetworks during rest (Barnett et al., 2021; Andrews-Hanna et al., 2010) and during movie watching (Cooper et al., 2021). Our results extend these findings, showing evidence that a similar subnetwork organization explains the trialwise involvement of PM regions during retrieval of multifeature events. Our models also showed that the coactivation of the vPMN makes contributions to memory quality that go above and beyond those made by coactivation of the dPMN (see Figure 2). The vPMN has previously been shown to modulate its connectivity in response to event transitions, and individual differences in episodic memory ability have been linked to dynamic changes in vPMN connectivity during movie watching (Cooper et al., 2021). The vPMN regions have also been shown to represent similar information during a memory-guided decision-making task (Barnett et al., 2021). The vPMN is strongly related to episodic retrieval and autobiographical remembering, whereas portions of the dPMN have been linked to mentalizing about the mental states of others (Andrews-Hanna, Saxe, & Yarkoni, 2014; Andrews-Hanna, Smallwood, & Spreng, 2014). Additional evidence suggests that the vPMN may be particularly responsive to remembering and orienting toward visual-spatial information and the dPMN toward people (Silson, Steel, Kidder, Gilmore, & Baker, 2019; Peer, Salomon, Goldberg, Blanke, & Arzy, 2015) and the thematic elements of autobiographical remembrances (Gurguryan & Sheldon, 2019).

The fact that the vPMN in our data set was uniquely related to overall memory quality could be reflective of our experimental design, which required the recollection of fine-grained visual-spatial details. At least two other aspects of our results seem to support this conclusion. First, memory for the scene feature—which in our experimental design requires the recollection of the fine-grained visual-spatial details—loaded most strongly onto our overall memory quality factor (see Table 3). Second, the vPMN significantly contributed only to scene feature memory in our memory feature models (see Table 5). The specific role of the vPMN in supporting scene memory is consistent with recent frameworks proposing that the anterior hippocampus and anterior regions of the neocortex support memory for coarse, gist-level, schematic details in memory, whereas posterior regions of the hippocampus and the neocortex—including PHC, RSC, and pAG—support memory for fine-grained perceptual details, especially spatial details (Sheldon, Fenerci, & Gurguryan, 2019; Sekeres, Winocur, & Moscovitch, 2018; Robin & Moscovitch, 2017). In contrast, the dPMN was negatively correlated with scene memory and tended to be positively related to sound memory, which may have been mediated by relatively coarse representations of

the sound valence that were sufficient to drive memory for this feature.

When taking into consideration the covariance among PM network regions, we did not find much evidence for independent, region-specific contributions, suggesting that the network-level effects could adequately account for their roles in predicting memory outcomes. Nevertheless, we had expected that there might be more region-specific effects, based on evidence that many of these regions play specialized roles in recollection. There are several reasons for why we did not see the region-specific effects that we had hypothesized. One possibility is that there is something unique about our experimental paradigm that did not allow us to observe region-specific contributions. For example, the hippocampus may have emerged as making a region-specific contribution if we had operationalized our measure of memory success to more specifically target the hippocampus' proposed function. The hippocampus' contribution to predicting overall memory success may be subsumed by the network-level contribution, but this may not be the case if the measure was more specific to successful pattern completion, for instance. Another possibility lies in how we modeled the neural response. In the current report, we modeled the neural response by assuming that it was transient, starting at the presentation of the memory cue during our "remember" periods. Previous research suggests that the memory-related neural response in the AG is not transient with respect to the onset of recall but is instead sustained throughout the duration of the recall period (Vilberg & Rugg, 2012, 2014). It is possible that modeling a sustained response throughout the recall period would allow for the identification of region-specific contributions of the AG. Our experimental design only allowed for 4 sec for recall, so the responses captured here are likely to be similar to the transient responses seen in Vilberg and Rugg (2012, 2014). As an additional test of this possibility, all of our models were rerun using single-trial estimates modeling the entire 4-sec retrieval period. The key results of the current report remained unchanged. Another possible explanation is that the identification of region-specific contributions within our framework assumes that the operations and representations of individual regions can be decoupled. However, in a typically functioning brain, the activity of two brain regions may be highly correlated if the involvement of one brain region depends on the output of the other, even if they are performing otherwise separate functions. Thus, although the current results suggest strong evidence for network-level effects in the context of the typically functioning brain, the roles of individual brain regions may be better revealed in studies documenting the consequences of region-specific disruptions, such as studies of patients with focal brain damage (Corkin, 2002; Moscovitch & Winocur, 1992), or in electrophysiological studies that can resolve fine temporal differences in information processing among regions in the same

network (Treder et al., 2021, Fox, Foster, Kucyi, Daitch, & Parvizi, 2018).

The one region in which we found a region-specific effect was the MPFC. The MPFC has been commonly described as part of the PM network (Ritchey & Cooper, 2020; Rugg & Vilberg, 2013) and is thought to support the formation and retrieval of gist-level, schema-based representations (Sekeres et al., 2018; Robin & Moscovitch, 2017; Schlichting & Preston, 2015; van Kesteren, Ruitter, Fernández, & Henson, 2012). Our results indicated that, after accounting for MPFC's participation in the dPMN subnetwork, the MPFC had a region-specific negative relationship with memory quality. The negative relationship between MPFC and memory success is not without precedent, with fMRI experiments of memory encoding often finding that less MPFC activation is associated with greater subsequent memory, particularly for objective compared with subjective memory judgments (Maillet & Rajah, 2014). This MPFC activation is thought to be associated with mind wandering or off-task thoughts (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009), which interferes with the formation of a lasting memory trace. The current experiment, however, was primarily focused on retrieval where previous reports have indicated a positive relation between MPFC activity and measures of subjective memory success (Kim, 2016; McDermott, Szpunar, & Christ, 2009; Spaniol et al., 2009). One possible explanation of this surprising result is that it reflects the role of the MPFC in schema based memory. Our experimental design relies on participants arbitrarily associating event elements at a fine level of detail. If participants were relying on a schema to meet our task demands, this could potentially lead to decreased performance on the fine-grained memory measures in our experiment. However, in the absence of any independent measures of schema use in our experiment, another plausible interpretation is that the observed negative relationship may be the result of a statistical artifact. In the current study, the MPFC was only weakly correlated with the quality with which events were remembered but was still positively correlated with other regions of the network (see Table 1). It was only after controlling for its subnetwork participation that we saw a strong negative contribution to memory. Thus, the result seen here could be the result of a conditioning-on-a-collider bias, also known as Berkson's paradox (Lübke, Gehrke, Horst, & Szepannek, 2020; Berkson, 1946). In this paradox, two variables that, in reality, do not have a statistical association are induced to have a negative association by statistically controlling for a variable that they both cause. In the current scenario, it could be the case that MPFC activation and memory quality are (at least in part) correlated with increases in PM network coactivation, but memory quality and MPFC activation are not related to one another.

The SEM methodology applied in the current report has a number of distinct advantages. First, the current SEM approach has an advantage over previous reports of

brain–behavior correlations in that it can simply and simultaneously capture the network-level and region-specific contributions of brain regions to behavioral phenomena. Second, the current report expanded upon previous deployments of this methodology (Bolt et al., 2018) by applying a multilevel SEM to simultaneously model within-subject and between-subject variation in the BOLD response, seeking to relate trial-by-trial, within-subject variability in BOLD response to trial-by-trial variability in memory while controlling for individual differences. Third, our data set has a distinct advantage over previous studies of episodic remembering because it incorporates multiple measures of the quality of retrieval of an episode. This allowed us to model overall memory quality as a latent variable loading onto our measures of memory for three different features of each episode. By operationalizing memory success in this way, we were able to capture trial-to-trial variability in the joint remembering of event features. This is key because holistic recollection is thought to be a key characteristic separating episodic remembering from other forms of memory (Tulving, 1983).

Our SEM approach is related to, but distinct from, other methods for relating regions and networks to episodic remembering. For example, previous studies have used data-driven, hierarchical clustering methods to parcellate PMN subnetworks (Barnett et al., 2021; Cooper et al., 2021; Andrews-Hanna et al., 2010) but did not relate trial-wise coactivation within those subnetworks to episodic remembering. Another set of related methodological approaches is effective connectivity approaches. Specifically, some effective connectivity approaches also use SEM, but they use SEM to attempt to test hypothetical models of the underlying causal relations among ROIs (e.g., McIntosh & Gonzalez-Lima, 1994; see McIntosh & Protzner, 2012, for a review). The latent variable modeling approach applied here, in contrast, does not attempt to make such causal inferences. Instead, our approach uses a latent variable to capture the coactivation seen within a network and relates this coactivation to a behavioral variable of interest. Lastly, the current approach is conceptually similar to partial least squares (PLS) analyses (Krishnan, Williams, McIntosh, & Abdi, 2011; McIntosh & Lobaugh, 2004; McIntosh, Bookstein, Haxby, & Grady, 1996). PLS involves maximizing the covariation between behavior and the signals extracted from voxels of the brain, extracting latent variables reflecting distributed coactivation across the brain that explains variance in some

behavior of interest. The SEM approach used in the current report is similar to PLS in that it also estimates a latent variable using the covariation of regional activation profiles but has the advantage of being exclusively hypothesis driven and computationally and conceptually simpler. Many PLS applications (but not all Krishnan et al., 2011), in contrast, are data driven in nature. Additionally, PLS typically operates on all of the voxels collected during the course of an experiment, whereas the current approach operates on a set of hypothesized ROIs.

The current report makes an important contribution to the literature on the role of the PM network in episodic remembering. It does, however, have its limitations. Our multilevel approach allowed us to model trialwise neural activation and behavioral profiles while controlling for individual differences. Multilevel SEM, however, also allows researchers to build models of individual differences in neural activation and behavior beyond simply controlling for this important source of variability. We did not attempt to model individual differences in the current report in large part because our data set would be underpowered to do so. Future research could utilize larger sample sizes to model individual differences related to particular participant characteristics (see Bolt et al., 2018, for an SEM application to individual differences). Additionally, the current analysis was focused on a set of a priori ROIs that were the same across individuals. Although this is a good starting point and is a strategy often adopted by researchers, recent research in high-precision functional mapping suggests that individually defined ROIs may provide more accurate insights into network organization and function (Gilmore, Nelson, & McDermott, 2021; Buckner & DiNicola, 2019). Finally, although our memory measures captured multiple aspects of each episode (specifically, memory for multiple episodic features), they may not have adequately captured the functioning of core alliances within the PM network (Ritchey & Cooper, 2020).

In conclusion, the brain is simultaneously composed of large-scale brain networks and individual regions composing those networks. Here, we demonstrate the importance of considering both network and regional levels of analysis when studying brain–behavior relationships, finding evidence in favor of a specific subnetwork organization of the PM network in its relation to episodic memory outcomes. Future work should continue to characterize the PM network by examining how these levels of analysis differentially support the various subprocesses and representations underlying episodic recollection.

APPENDIX A

Table A1. Summaries of All Models Included in the Article along with Their Model Fit Indices

<i>Model</i>	<i>Description</i>	<i>Estimator</i>	<i>Chi-square (df)</i>	<i>RMSEA</i>	<i>CFI</i>	<i>SRMR_W</i>	<i>SRMR_B</i>
Preliminary							
1	Neural preliminary	MLR	2951.578 (36), $p < .001$.144	.000	.047	.352
2	Behavioral preliminary	WLSMV	20.882 (6), $p = .002$.025	.958	.000	.634
Measurement models							
3	One-factor neural	MLR	290.442 (20), $p < .001$.059	.905	.063	.005
4	Two-factor neural	MLR	90.096 (19), $p < .001$.031	.975	.035	.003
5	Latent variable behavioral	WLSMV	0.163 (2), $p = .922$.000	1.000	.001	.022
6	Stitched model	WLSMV	534.782 (59), $p < .001$.046	.974	.034	.043
8	Stitched model without behavioral latent variable	WLSMV	683.198 (37), $p < .001$.067	.965	.030	.000
Structural models							
7	Behavioral latent variable	WLSMV	See Table 4 and Figure 2				
9	Without behavioral latent variable	WLSMV	See Table 5 and Figure 3				

The preliminary models were used to establish the appropriateness of the multilevel modeling approach for the neural (Model 1) and behavioral (Model 2) data. The measurement models characterized the loading of individual measures onto their corresponding latent variables, separately for the neural (Models 3 and 4) and behavioral (Model 5) data. Measurement models were stitched together into combined models with (Model 6) and without (Model 8) the latent variable for overall memory quality. Finally, structural models were used to test the paths connecting the neural and behavioral latent variables with each other and with the individual measures, for both the overall memory quality (Model 7) and memory feature (Model 9) analyses.

APPENDIX B

Table B1. Parameter Estimates for the Within-subject Part of Our Feature Specific Memory Model (Model 9; see Figure 3)

<i>paramHeader</i>	<i>param</i>	<i>est</i>	<i>se</i>	<i>pval</i>
VPMN.BY	RSC	0.627	0.006	<.001
	PHC	0.567	0.008	<.001
	PAG	0.608	0.007	<.001
DPMN.BY	MPFC	0.516	0.006	<.001
	PHIPP	0.438	0.008	<.001
	PREC	0.647	0.007	<.001
	PCC	0.685	0.004	<.001
SCENE.ON	AAG	0.711	0.005	<.001
	VPMN	0.277	0.043	<.001
	DPMN	-0.147	0.040	<.001
COLOR.ON	VPMN	-0.004	0.047	.933
	DPMN	0.041	0.050	.417
SOUND.ON	VPMN	0.006	0.043	.882
	DPMN	0.114	0.059	.052
DPMN.WITH	VPMN	0.634	0.008	<.001
SCENE.WITH	COLOR	0.340	0.025	<.001
	SOUND	0.313	0.026	<.001
COLOR.WITH	SOUND	0.232	0.026	<.001
Variances	VPMN	1	0	.999
	DPMN	1	0	.999
	PHIPP	0.809	0.007	<.001
	PREC	0.581	0.009	<.001
	PCC	0.531	0.006	<.001
	MPFC	0.734	0.006	<.001
	PHC	0.679	0.009	<.001
	RSC	0.607	0.008	<.001
	AAG	0.495	0.007	<.001
	PAG	0.630	0.008	<.001

This table was created using the R package *MplusAutomation* (Hallquist & Wiley, 2018). Parameter headers (*paramHeader*) follow standard Mplus syntax, where the ON key word indicates a path parameter from the variable listed in the “*param*” column to variable listed in the “*paramHeader*” column, the BY key word indicates a loading parameter (λ), and the WITH key word indicates a covariance parameter (θ). *param* = parameter; *est* = standardized estimate; *se* = standard error; *pval* = *p* value. See Figure 1 caption for abbreviations. See Figure 3 for path diagram.

Acknowledgments

This work was supported by National Institutes of Health (NIH) grant R00MH103401 (M. R.) and National Science Foundation grant BCS-2047415 (M. R.). The data were collected at the Harvard Center for Brain Science, involving the use of instrumentation supported by the NIH Shared Instrumentation Grant Program (grant no. S10OD020039).

Reprint requests should be sent to Kyle A. Kurkela or Maureen Ritchey, Department of Psychology and Neuroscience, Boston College, Chestnut Hill, MA 02467, or via e-mail: kyle.kurkela@bc.edu or maureen.ritchey@bc.edu.

Author Contributions

Kyle A. Kurkela: Conceptualization; Formal analysis; Methodology; Writing—Original draft. Rose A. Cooper: Conceptualization; Methodology; Writing—Review & editing. Ehri Ryu: Formal analysis; Methodology; Supervision; Writing—Review & editing. Maureen Ritchey: Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing—Review & editing.

Funding Information

National Institute of Mental Health (<https://dx.doi.org/10.13039/100000025>), grant number: R00MH103401; National Science Foundation (<https://dx.doi.org/10.13039/100000001>), grant number: BCS-2047415; National Institutes of Health (<https://dx.doi.org/10.13039/100000002>), grant number: S10OD020039.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were $M(\text{an})/M = .407$, $W(\text{oman})/M = .32$, $M/W = .115$, and $W/W = .159$, the comparable proportions for the articles that these authorship teams cited were $M/M = .549$, $W/M = .257$, $M/W = .109$, and $W/W = .085$ (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: $M/M = .356$; $W/M = .305$; $M/W = .153$; $W/W = .186$.

Note

1. In multilevel models, SRMR at the between-cluster level may be large because the observed variances are small, not necessarily because there is a large extent of misfit.

REFERENCES

- Addis, D. R., Leclerc, C. M., Muscatell, K. A., & Kensinger, E. A. (2010). There are age-related changes in neural connectivity during the encoding of positive, but not negative, information. *Cortex*, *46*, 425–433. <https://doi.org/10.1016/j.cortex.2009.04.011>, PubMed: 19555933
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, *65*, 550–562. <https://doi.org/10.1016/j.neuron.2010.02.005>, PubMed: 20188659
- Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2014). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *Neuroimage*, *91*, 324–335. <https://doi.org/10.1016/j.neuroimage.2014.01.032>, PubMed: 24486981
- Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*, 29–52. <https://doi.org/10.1111/nyas.12360>, PubMed: 24502540
- Barnett, A. J., Reilly, W., Dimsdale-Zucker, H. R., Mizrak, E., Reagh, Z., & Ranganath, C. (2021). Intrinsic connectivity reveals functionally distinct cortico-hippocampal networks in the human brain. *PLoS Biology*, *19*, e3001275. <https://doi.org/10.1371/journal.pbio.3001275>, PubMed: 34077415
- Berkson, J. (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics*, *2*, 47–53. <https://doi.org/10.2307/3002000>, PubMed: 21001024
- Bolt, T., Prince, E. B., Nomi, J. S., Messinger, D., Llabre, M. M., & Uddin, L. Q. (2018). Combining region- and network-level brain-behavior relationships in a structural equation model. *Neuroimage*, *165*, 158–169. <https://doi.org/10.1016/j.neuroimage.2017.10.007>, PubMed: 29030103
- Bradley, M. M., & Lang, P. J. (2007). *The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual*. Tech. Rep. B-3. Gainesville, FL: University of Florida.
- Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*, 981–990. <https://doi.org/10.1177/0956797612465439>, PubMed: 23630219
- Buckner, R. L., & DiNicola, L. M. (2019). The brain's default network: Updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, *20*, 593–608. <https://doi.org/10.1038/s41583-019-0212-7>, PubMed: 31492945
- Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 8719–8724. <https://doi.org/10.1073/pnas.0900234106>, PubMed: 19433790
- Cooper, R. A., Kurkela, K. A., Davis, S. W., & Ritchey, M. (2021). Mapping the organization and dynamics of the posterior medial network during movie watching. *Neuroimage*, *236*, 118075. <https://doi.org/10.1016/j.neuroimage.2021.118075>, PubMed: 33910099
- Cooper, R. A., & Ritchey, M. (2019). Cortico-hippocampal network connections support the multidimensional quality of episodic memory. *eLife*, *8*, e45591. <https://doi.org/10.7554/eLife.45591>, PubMed: 30900990
- Corkin, S. (2002). What's new with the amnesic patient H.M.? *Nature Reviews Neuroscience*, *3*, 153–160. <https://doi.org/10.1038/nrn726>, PubMed: 11836523
- Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*, 693–700. <https://doi.org/10.1016/j.conb.2006.10.012>, PubMed: 17097284
- Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*, 379–386. <https://doi.org/10.1016/j.tics.2007.08.001>, PubMed: 17707683
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. <https://doi.org/10.1146/annurev.neuro.30.051606.094328>, PubMed: 17417939
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., et al. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*, 111–116. <https://doi.org/10.1038/s41592-018-0235-4>, PubMed: 30532080
- Fox, K. C. R., Foster, B. L., Kucyi, A., Daitch, A. L., & Parvizi, J. (2018). Intracranial electrophysiology of the human default network. *Trends in Cognitive Sciences*, *22*, 307–324. <https://doi.org/10.1016/j.tics.2018.02.002>, PubMed: 29525387
- Geib, B. R., Stanley, M. L., Dennis, N. A., Woldorff, M. G., & Cabeza, R. (2017). From hippocampus to whole-brain: The role of integrative processing in episodic memory retrieval. *Human Brain Mapping*, *38*, 2242–2259. <https://doi.org/10.1002/hbm.23518>, PubMed: 28112460
- Geib, B. R., Stanley, M. L., Wing, E. A., Laurienti, P. J., & Cabeza, R. (2017). Hippocampal contributions to the large-scale episodic memory network predict vivid visual memories. *Cerebral Cortex*, *27*, 680–693. <https://doi.org/10.1093/cercor/bhv272>, PubMed: 26523034
- Gilmore, A. W., Nelson, S. M., & McDermott, K. B. (2021). Precision functional mapping of human memory systems. *Current Opinion in Behavioral Sciences*, *40*, 52–57. <https://doi.org/10.1016/j.cobeha.2020.12.013>
- Gurguryan, L., & Sheldon, S. (2019). Retrieval orientation alters neural activity during autobiographical memory recollection. *Neuroimage*, *199*, 534–544. <https://doi.org/10.1016/j.neuroimage.2019.05.077>, PubMed: 31152842
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling*, *25*, 621–638. <https://doi.org/10.1080/10705511.2017.1402334>, PubMed: 30083048
- Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, *6*, 7462. <https://doi.org/10.1038/ncomms8462>, PubMed: 26136141
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. <https://doi.org/10.1080/10705519909540118>
- Humphreys, G. F., Lambon Ralph, M. A., & Simons, J. S. (2021). A unifying account of angular gyrus contributions to episodic and semantic cognition. *Trends in Neurosciences*, *44*, 452–463. <https://doi.org/10.1016/j.tics.2021.01.006>, PubMed: 33612312
- Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., & Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebral Cortex*, *16*, 1349–1360. <https://doi.org/10.1093/cercor/bhl040>, PubMed: 16861334
- Jak, S. (2019). Cross-level invariance in multilevel factor models. *Structural Equation Modeling*, *26*, 607–622. <https://doi.org/10.1080/10705511.2018.1534205>
- Jak, S., Oort, F. J., & Dolan, C. V. (2013). A test for cluster bias: Detecting violations of measurement invariance across clusters in multilevel data. *Structural Equation Modeling*, *20*, 265–282. <https://doi.org/10.1080/10705511.2013.769392>

- Kim, H. (2013). Differential neural activity in the recognition of old versus new events: An activation likelihood estimation meta-analysis. *Human Brain Mapping, 34*, 814–836. <https://doi.org/10.1002/hbm.21474>, PubMed: 22110008
- Kim, H. (2016). Default network activation during episodic and semantic memory retrieval: A selective meta-analytic comparison. *Neuropsychologia, 80*, 35–46. <https://doi.org/10.1016/j.neuropsychologia.2015.11.006>, PubMed: 26562053
- Kim, E. S., Dedrick, R. F., Cao, C., & Ferron, J. M. (2016). Multilevel factor analysis: Reporting guidelines and a review of reporting practices. *Multivariate Behavioral Research, 51*, 881–898. <https://doi.org/10.1080/00273171.2016.1228042>, PubMed: 27754708
- King, D. R., de Chastelaine, M., Elward, R. L., Wang, T. H., & Rugg, M. D. (2015). Recollection-related increases in functional connectivity predict individual differences in memory accuracy. *Journal of Neuroscience, 35*, 1763–1772. <https://doi.org/10.1523/JNEUROSCI.3219-14.2015>, PubMed: 25632149
- Krishnan, A., Williams, L. J., McIntosh, A. R., & Abdi, H. (2011). Partial least squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage, 56*, 455–475. <https://doi.org/10.1016/j.neuroimage.2010.07.034>, PubMed: 20656037
- Lübke, K., Gehrke, M., Horst, J., & Szepannek, G. (2020). Why we should teach causal inference: Examples in linear regression with simulated data. *Journal of Statistics Education, 28*, 133–139. <https://doi.org/10.1080/10691898.2020.1752859>
- Maillet, D., & Rajah, M. N. (2014). Dissociable roles of default-mode regions during episodic encoding. *Neuroimage, 89*, 244–255. <https://doi.org/10.1016/j.neuroimage.2013.11.050>, PubMed: 24315838
- Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society of London, 262*, 23–81. <https://doi.org/10.1098/rstb.1971.0078>, PubMed: 4399412
- McCormick, C., Moscovitch, M., Protzner, A. B., Huber, C. G., & McAndrews, M. P. (2010). Hippocampal-neocortical networks differ during encoding and retrieval of relational memory: Functional and effective connectivity analyses. *Neuropsychologia, 48*, 3272–3281. <https://doi.org/10.1016/j.neuropsychologia.2010.07.010>, PubMed: 20637787
- McCormick, C., St-Laurent, M., Ty, A., Valiante, T. A., & McAndrews, M. P. (2015). Functional and effective hippocampal-neocortical connectivity during construction and elaboration of autobiographical memory retrieval. *Cerebral Cortex, 25*, 1297–1305. <https://doi.org/10.1093/cercor/bht324>, PubMed: 24275829
- McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia, 47*, 2290–2298. <https://doi.org/10.1016/j.neuropsychologia.2008.12.025>, PubMed: 19159634
- McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. (1996). Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage, 3*, 143–157. <https://doi.org/10.1006/nimg.1996.0016>, PubMed: 9345485
- McIntosh, A. R., & Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping, 2*, 2–22. <https://doi.org/10.1002/hbm.460020104>
- McIntosh, A. R., & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: Applications and advances. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2004.07.020>, PubMed: 15501095
- McIntosh, A. R., & Protzner, A. B. (2012). Structural equation models of imaging data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (Vol. 740, pp. 636–649). Guilford Press.
- Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage, 53*, 103–118. <https://doi.org/10.1016/j.neuroimage.2010.05.051>, PubMed: 20580933
- Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik (Ed.), *The handbook of aging and cognition* (Vol. 586, pp. 315–372). Erlbaum.
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage, 59*, 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>, PubMed: 21924359
- Muthén, L. K., & Muthén, B. O. (1997–2017). *Mplus user's guide* (8th ed.). Los Angeles: Muthén & Muthén.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110*, 611–646. <https://doi.org/10.1037/0033-295X.110.4.611>, PubMed: 14599236
- Peer, M., Salomon, R., Goldberg, I., Blanke, O., & Arzy, S. (2015). Brain system for mental orientation in space, time, and person. *Proceedings of the National Academy of Sciences, U.S.A., 112*, 11072–11077. <https://doi.org/10.1073/pnas.1504242112>, PubMed: 26283353
- Rajah, M. N., & McIntosh, A. R. (2005). Overlap in the functional neural systems involved in semantic and episodic memory retrieval. *Journal of Cognitive Neuroscience, 17*, 470–482. <https://doi.org/10.1162/0898929053279478>, PubMed: 15814006
- Ramanan, S., Pigué, O., & Irish, M. (2018). Rethinking the role of the angular gyrus in remembering the past and imagining the future: The contextual integration model. *Neuroscientist, 24*, 342–352. <https://doi.org/10.1177/1073858417735514>, PubMed: 29283042
- Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature reviews Neuroscience, 13*, 713–726. <https://doi.org/10.1038/nrn3338>, PubMed: 22992647
- Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *eLife, 5*, e18260. <https://doi.org/10.7554/eLife.18260>, PubMed: 27776631
- Riedel, G., Micheau, J., Lam, A. G., Roloff, E. L., Martin, S. J., Bridge, H., et al. (1999). Reversible neural inactivation reveals hippocampal participation in several memory processes. *Nature Neuroscience, 2*, 898–905. <https://doi.org/10.1038/13202>, PubMed: 10491611
- Ritchey, M., & Cooper, R. A. (2020). Deconstructing the posterior medial episodic network. *Trends in Cognitive Sciences, 24*, 451–465. <https://doi.org/10.1016/j.tics.2020.03.006>, PubMed: 32340798
- Ritchey, M., Montchal, M. E., Yonelinas, A. P., & Ranganath, C. (2015). Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *eLife, 4*, e05025. <https://doi.org/10.7554/eLife.05025>, PubMed: 25584461
- Robin, J., & Moscovitch, M. (2017). Details, gist and schema: Hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences, 17*, 114–123. <https://doi.org/10.1016/j.cobeha.2017.07.016>
- Rosenbaum, R. S., Furey, M. L., Horwitz, B., & Grady, C. L. (2010). Altered connectivity among emotion-related brain regions during short-term memory in Alzheimer's disease. *Neurobiology of Aging, 31*, 780–786. <https://doi.org/10.1016/j.neurobiolaging.2008.06.002>, PubMed: 18639365
- Rugg, M. D., & King, D. R. (2018). Ventral lateral parietal cortex and episodic memory retrieval. *Cortex, 107*, 238–250. <https://doi.org/10.1016/j.cortex.2017.07.012>, PubMed: 28802589

- Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology*, 23, 255–260. <https://doi.org/10.1016/j.conb.2012.11.005>, PubMed: 23206590
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., et al. (2018). Local–global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, 28, 3095–3114. <https://doi.org/10.1093/cercor/bhx179>, PubMed: 28981612
- Schedlbauer, A. M., Copara, M. S., Watrous, A. J., & Ekstrom, A. D. (2014). Multiple interacting brain areas underlie successful spatiotemporal memory retrieval in humans. *Scientific Reports*, 4, 6431. <https://doi.org/10.1038/srep06431>, PubMed: 25234342
- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8. <https://doi.org/10.1016/j.cobeha.2014.07.005>, PubMed: 25750931
- Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters*, 680, 39–53. <https://doi.org/10.1016/j.neulet.2018.05.006>, PubMed: 29733974
- Sheldon, S., Fenerci, C., & Gurguryan, L. (2019). A neurocognitive perspective on the forms and functions of autobiographical memory retrieval. *Frontiers in Systems Neuroscience*, 13, 4. <https://doi.org/10.3389/fnsys.2019.00004>, PubMed: 30760984
- Silson, E. H., Steel, A., Kidder, A., Gilmore, A. W., & Baker, C. I. (2019). Distinct subdivisions of human medial parietal cortex support recollection of people and places. *eLife*, 8, e47391. <https://doi.org/10.7554/eLife.47391>, PubMed: 31305238
- Spaniol, J., Davidson, P. S. R., Kim, A. S. N., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47, 1765–1779. <https://doi.org/10.1016/j.neuropsychologia.2009.02.028>, PubMed: 19428409
- Thakral, P. P., Madore, K. P., & Schacter, D. L. (2020). The core episodic simulation network dissociates as a function of subjective experience and objective content. *Neuropsychologia*, 136, 107263. <https://doi.org/10.1016/j.neuropsychologia.2019.107263>, PubMed: 31743681
- Treder, M. S., Charest, I., Michelmann, S., Martín-Buro, M. C., Roux, F., Carceller-Benito, F., et al. (2021). The hippocampus as the switchboard between perception and memory. *Proceedings of the National Academy of Sciences, U.S.A.*, 118, e2114171118. <https://doi.org/10.1073/pnas.2114171118>, PubMed: 34880133
- Tulving, E. (1983). *Elements of episodic memory*. Oxford, United Kingdom: Oxford University Press.
- van Kesteren, M. T. R., Ruiters, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35, 211–219. <https://doi.org/10.1016/j.tins.2012.02.001>, PubMed: 22398180
- Vilberg, K. L., & Rugg, M. D. (2012). The neural correlates of recollection: Transient versus sustained fMRI effects. *Journal of Neuroscience*, 32, 15679–15687. <https://doi.org/10.1523/JNEUROSCI.3065-12.2012>, PubMed: 23136408
- Vilberg, K. L., & Rugg, M. D. (2014). Temporal dissociations within the core recollection network. *Cognitive Neuroscience*, 5, 77–84. <https://doi.org/10.1080/17588928.2013.860088>, PubMed: 24283400
- Watrous, A. J., Tandon, N., Conner, C. R., Pieters, T., & Ekstrom, A. D. (2013). Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nature Neuroscience*, 16, 349–356. <https://doi.org/10.1038/nn.3315>, PubMed: 23354333
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 76, 913–934. <https://doi.org/10.1177/0013164413495237>, PubMed: 25705052
- Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, 51, 409–428. <https://doi.org/10.3758/s13428-018-1055-2>, PubMed: 29869222
- Xiao, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2012). Recognizing scene viewpoint using panoramic place representation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2695–2702). <https://doi.org/10.1109/CVPR.2012.6247991>
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8, 665–670. <https://doi.org/10.1038/nmeth.1635>, PubMed: 21706013