

Information Processing in the Mental Workspace Is Fundamentally Distributed

Alexander Schlegel, Prescott Alexander, and Peter U. Tse

Abstract

■ The brain is a complex, interconnected information processing network. In humans, this network supports a mental workspace that enables high-level abilities such as scientific and artistic creativity. Do the component processes underlying these abilities occur in discrete anatomical modules, or are they distributed widely throughout the brain? How does the flow of information within this network support specific cognitive functions? Current approaches have limited ability to answer such questions. Here, we report novel multivariate methods to ana-

lyze information flow within the mental workspace during visual imagery manipulation. We find that mental imagery entails distributed information flow and shared representations throughout the cortex. These findings challenge existing, anatomically modular models of the neural basis of higher-order mental functions, suggesting that such processes may occur at least in part at a fundamentally distributed level of organization. The novel methods we report may be useful in studying other similarly complex, high-level informational processes. ■

INTRODUCTION

A hallmark of human cognition is the ability to volitionally construct and flexibly manipulate mental representations. Such abilities have been studied using several overlapping psychological constructs including working memory (Baddeley, 2003), mental imagery (Tong, 2013; Kosslyn, Behrmann, & Jeannerod, 1995), visuospatial ability (Uttal et al., 2013), mental models (Hegarty, 2004), analogical reasoning (Bassok, Dunbar, & Holyoak, 2012), and mental workspace (Logie, 2003). In general, these terms denote the ability to work volitionally and flexibly with mental representations, a skill that underlies much of human life from mundane tasks, such as planning seating arrangements at family get-togethers, to our species' greatest artistic and scientific achievements. For instance, Albert Einstein wrote that his scientific thought process consisted primarily of "certain signs and more or less clear images which can be 'voluntarily' reproduced and combined" (Hadamard, 1954). Here, we will use Logie's term "mental workspace" to refer to the mental space in which these flexible cognitive processes occur.

How does the human brain support the mental workspace underlying flexible and creative mental phenomena such as mathematical, scientific, and artistic thought (Schlegel et al., 2015; Logie, 2003)? Understanding how the brain enables the imaginative abilities of the mental workspace is an important goal for many fields (Insel, Landis, & Collins, 2013; Markram, 2012), and several models have proposed potential mechanisms (Graham &

Rockmore, 2011; Tononi, 2008; Postle, 2006; Baddeley, 2003; Logie, 2003; Rumelhart & McClelland, 1986). A key aspect of these behaviors and the models that attempt to explain them is the ability to both represent and manipulate mental images. Previous research has shown that manipulating visual imagery in the mental workspace recruits a neural network extending throughout the cerebral cortex and associated structures (Schlegel et al., 2013). An important question to answer concerning this network is whether the component processes underlying the network's function (e.g., executive or representational subsystems) occur in localized anatomical modules or whether these component processes occur at a more fundamentally distributed level of organization that transcends anatomical boundaries. However, our ability to measure and analyze complex informational processes that are distributed widely in the human brain remains underdeveloped, and thus, such questions are currently difficult to answer (Barnett & Seth, 2014; Crowe et al., 2013; Lizier, Heinze, Horstmann, Haynes, & Prokopenko, 2011).

Manipulation of visual imagery requires multiple component processes including (a) forming a mental representation of an image and (b) performing an operation to manipulate that representation. Standard models of working memory propose that each of these component processes is mediated by an anatomically localized neural "module." As an example, the "central executive" in Baddeley's model of working memory has been proposed to reside in dorsal lateral pFC (DLPFC) and direct the maintenance of mental representations that are stored in modality-specific regions such as visual cortex for the "visuospatial sketchpad" or auditory cortex for

the “phonological loop” (Crowe et al., 2013; Lee, Kravitz, & Baker, 2013; Baddeley, 2003; Kane & Engle, 2002; Ishai, Ungerleider, & Haxby, 2000). Similarly, Postle argues that pFC is not involved in the representation of working memory contents; instead, his model proposes that mental representations are processed exclusively by domain-specific sensory- or action-related regions (Postle, 2006). Thus, although these models hold that working memory and related abilities may recruit a “distributed” neural network in the sense that the complex functions of the network are mediated collectively by anatomically widespread regions, the component processes that underlie those complex functions are relegated to anatomically distinct modules. In many cases, arguments for anatomically modular models are based on a failure to find (i.e., acceptance of the null hypothesis) or often even look for relevant information in regions outside those that the models propose (Sreenivasan, Vytlačil, & D’Esposito, 2014; Lee et al., 2013; Postle, 2006; Baddeley, 2003; Ishai et al., 2000). For instance, both Lee and colleagues (2013) and Ishai and colleagues (2000) found information pertaining to the visual but not the nonvisual aspects of working memory representations in extrastriate visual cortex and found the opposite for lateral pFC. Both groups interpreted their results to suggest that extrastriate visual cortex processes the visual aspects of working memory tasks but not the nonvisual aspects and that lateral pFC processes the nonvisual but not the visual aspects. Although such conclusions are a common practice in the field, they amount to acceptance of null results regarding the information that was not detected in each respective area; they thus run the risk of failing to account for information that may have been present but that was not detected by their methods. Baddeley’s anatomically localized model of working memory similarly relies on studies that either did not find or did not look for relevant information outside the hypothesized regions (Baddeley, 2003).

There is reason to believe that anatomically modular models of working memory may be missing a piece of the puzzle. Mounting empirical evidence derived from new, network- and information-based analytical techniques paints a more complex picture of high-level cognitive processing, suggesting that it may, in many cases, occur at a level of organization that transcends any single neural structure (Sporns, 2014; Turk-Browne, 2013; Bassett et al., 2010; Ester, Serences, & Awh, 2009; Van den Heuvel, Stam, Kahn, & Hulshoff Pol, 2009; Tononi, 2008; Rumelhart & McClelland, 1986). However, traditional techniques based on univariate differences in brain activity and/or anatomically localized analyses are insensitive to informational processes that occur via complex patterns of interaction between regions. We therefore hypothesized that novel, more-sensitive analytical methods that target the complex informational structure of high-level cognition and that consider information carried by patterns of connectivity between regions would reveal that the mental workspace emerges out of a fundamen-

tally distributed sharing of informational processes throughout the cortex. This hypothesis runs contrary to traditional modular accounts, which claim that information is segregated to specific anatomical regions, such as visual information occurring only in visual cortex or executive processing occurring only in pFC (Lee et al., 2013; Postle, 2006; Baddeley, 2003; Ishai et al., 2000).

To evaluate our hypothesis and investigate how the mental workspace network implements both the representation and manipulation of visual imagery, we used fMRI to record cortical activity as participants completed a series of trials involving the mental manipulation of shapes maintained in working memory. Our analyses investigated the distribution of neural information related to two component processes of this task: the maintenance of shapes in imagery and the mental manipulation of those shapes. For each process, we asked where in the cortex information related to that process occurred and whether the nature of this information was consistent with a fundamentally distributed or a modular processing model.

During each trial, participants recalled one of four abstract shapes memorized previously (Figure 1A) and performed one of four mental operations on that shape (90° clockwise rotation, 90° counterclockwise rotation, horizontal flip, or vertical flip; Figure 1B). To enable the functional analyses described below, the shapes were related in a two-level hierarchy of similarity (see Figure 1A). The operations shared an analogous relationship (see Figure 1B). To ensure that neural activity associated with the shapes and operations was because of visual imagery rather than the presented visual stimuli, we constructed a unique mapping for each participant from shapes to letters and from operations to numbers. Each trial occurred as follows: At the start of a trial, four letter–number pairs (e.g., “C3”) appeared for 2 sec, with an arrow pointing to a single pair to indicate the shape and operation for the current trial. The other three pairs were shown as a visual control to ensure that any successful classification analyses were because of mental imagery rather than the visual stimuli. After a 6-sec period during which the participant performed the indicated mental operation, four shapes at various orientations appeared on the screen for 2 sec. One of these was the shape indicated previously, whereas the other three shapes again served as a visual control. The participant indicated whether the displayed shape was at the orientation that would result from the indicated operation and was then given feedback regarding whether the response was correct or incorrect (Figure 1C shows a trial schematic).

Our analyses of the task-related fMRI data used a combination of existing and novel multivariate methods to investigate the informational structure of the network underlying the mental workspace. First, we performed ROI classification analyses with trials labeled based on either shape or operation. These analyses investigated which regions supported information about mental representations and/or

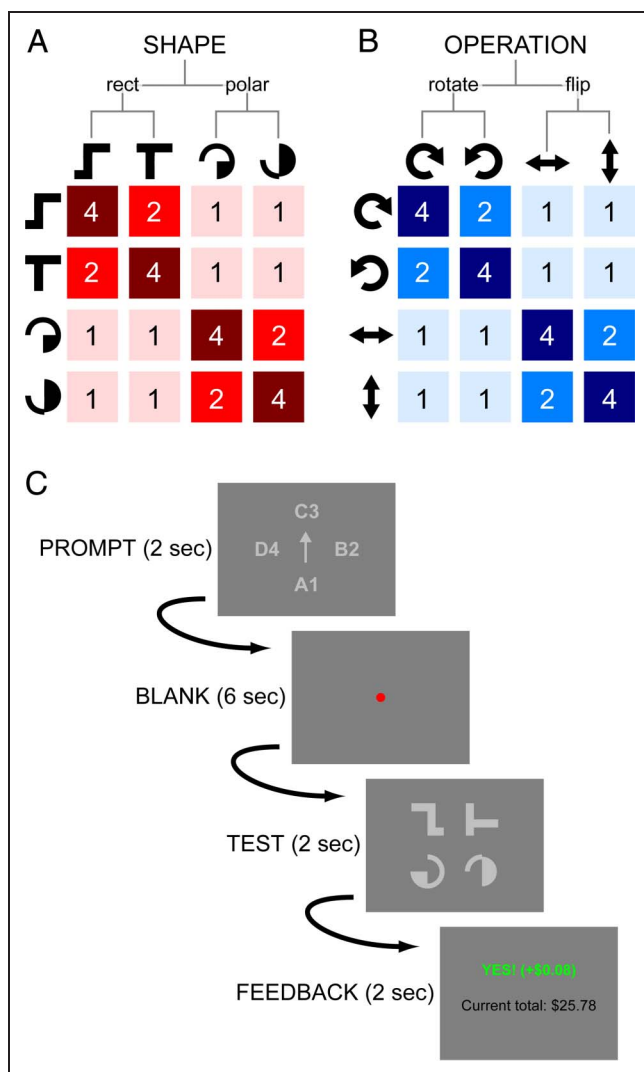


Figure 1. Experimental design. (A) The four shapes used in the experiment, related in a two-level hierarchical structure. Two shapes were derived from a 4×4 rectangular grid, and two shapes were derived from an analogous polar grid. At the bottom is a similarity structure that represents the matrix form of the hierarchy. Each shape is more similar to itself than it is to any other shape, and each rectangular shape is more similar to the other rectangular shape than it is to either polar shape (and vice versa). See Schlegel et al. (2013) for details on the particular values used in the similarity structure. (B) The four mental operations used in the experiment, also related in a two-level hierarchy: 90° clockwise rotation, 90° counterclockwise rotation, horizontal flip, and vertical flip. The similarity structure for operations is constructed in the same manner as for shapes. (C) Trial schematic. A 2-sec prompt screen indicated the shape and operation for the current trial. This was followed by a 6-sec blank screen during which the participant performed the indicated operation on the indicated shape. Next, a 2-sec test screen appeared, during which the participant indicated whether a displayed shape matched the output of the indicated operation. Finally, the participant was given response feedback.

mental manipulations. Second, we developed a novel ROI cross-classification analysis to investigate whether this information shared common characteristics between regions. Third, we developed a novel classification analysis on patterns of information flow between cortical regions

to determine how information related to the task was transferred between regions. Each of these three types of analysis could potentially provide evidence for either a modular or distributed processing model. In combination, our analyses reveal that information about both mental representations and mental manipulations is supported by many regions across the cortex, that information about at least mental representations shares common characteristics across these regions, and that this information becomes distributed via complex, bidirectional patterns of information flow between regions. Together, these findings lend support to a fundamentally distributed model of processing in the neural network underlying the mental workspace.

METHODS

Participants

Nineteen participants (six women, aged 18–51 years) with normal or corrected-to-normal vision gave informed written consent according to the guidelines of the Committee for the Protection of Human Subjects at Dartmouth College before participating. All experimental protocols were approved by the Committee for the Protection of Human Subjects (institutional review board #15822). Participation consisted of two experimental sessions: one behavioral session in which participants practiced the task until they reached criterion (described below) and a subsequent 1.75-hr fMRI scanning session.

Experimental Design

During each of a series of trials, participants performed one of four mental operations on one of four abstract visual shapes. The four mental operations were 90° clockwise rotation, 90° counterclockwise rotation, horizontal flip, and vertical flip. The four abstract shapes are shown in Figure 1: Two shapes were constructed from a 4×4 rectangular grid, and two were constructed from an analogous polar grid. All shapes were matched for area. To equate the visual presentation between conditions, we did not display the shape or operation to use in a given trial. Instead, each shape was mapped to one of the letters A, B, C, or D, and each operation was mapped to one of the numbers 1, 2, 3, or 4. Each participant was assigned a unique mapping and spent the practice session committing the shapes, operations, and mapping to memory. The practice session concluded once the participant responded correctly on 10 consecutive trials. At the start of each trial, a 2-sec-long prompt screen displayed four letter–number pairs (e.g., “C3”). An arrow pointed to one of these pairs to indicate the shape and operation to use for the current trial. This screen was replaced by a fixation dot for 6 sec during which the participant performed the indicated mental operation on the indicated shape. After this period, a 2-sec-long test screen

displayed each of the four shapes at various orientations relative to the starting orientations learned by the participants. The participant was instructed to identify the current trial's shape on the screen and indicate via a button press within that 2-sec period whether it was in the orientation that would result from the trial's indicated operation. In half of the trials, the shape was in the correct orientation, and in the other half, it was in a random, incorrect orientation. During the fMRI session, the operations and shapes were counterbalanced across all trials, and correct/incorrect trials and display positions were randomized. To encourage attentiveness, participants were paid based on their performance (receiving money for correct responses and losing money for incorrect responses, with a minimum base rate of reimbursement). Participants completed 15 fMRI runs, each of which consisted of 16 trials interleaved with 8 sec of rest to ensure that the BOLD response for a given trial was not influenced by activity from the previous trial (5 min 28 sec per run). Thus, each stimulus and operation occurred four times per run (60 times in total during the experiment), and 240 trials were administered over the scanning session.

MRI Acquisition

MRI data were collected using a 3.0-T Philips (Amsterdam, The Netherlands) Achieva Intera scanner with a 32-channel sense head coil located at the Dartmouth Brain Imaging Center. One T1-weighted structural image was collected using a magnetization-prepared rapid acquisition gradient-echo sequence (repetition time [TR] = 8.176 msec, echo time = 3.72 msec, flip angle = 8°, field of view = 240 × 220 mm, 188 sagittal slices, voxel size = 0.9375 × 0.9375 × 1 mm, acquisition time = 3.12 min). T2*-weighted gradient EPI scans were used to acquire functional images covering the whole brain (TR = 2000 msec, echo time = 20 msec, flip angle = 90°, field of view = 240 × 240 mm, voxel size = 3 × 3 × 3.5 mm, slice gap = 0 mm, 35 slices).

MRI Data Preprocessing

High-resolution anatomical images were processed using the FreeSurfer image analysis suite (Dale, Fischl, & Sereno, 1999). Standard preprocessing of fMRI data was carried out: Data were motion and slice-time corrected, high-pass filtered temporally with a 100-sec cutoff, and smoothed spatially with a 6-mm FWHM Gaussian kernel, all using FMRIB Software Library (FSL; Smith et al., 2004). Data from each run were concatenated temporally for each participant after aligning each run using FSL's FLIRT tool and demeaning each voxel's time course. For the ROI classification (described below), data were prewhitened using FSL's MELODIC tool (i.e., principal components were extracted using MELODIC's default dimensionality estimation method with a minimum of 10 components per ROI).

ROI Classification Analysis

Each trial could be labeled based on either the shape that was represented in visual imagery or the operation that was performed to manipulate that representation. For each of these two labeling schemes, we used PyMVPA (Hanke et al., 2009) to perform a four-way spatiotemporal multivariate classification analysis in each of the six ROIs that showed information pertaining to manipulation of visual imagery in a previous study (see Figure 2A;

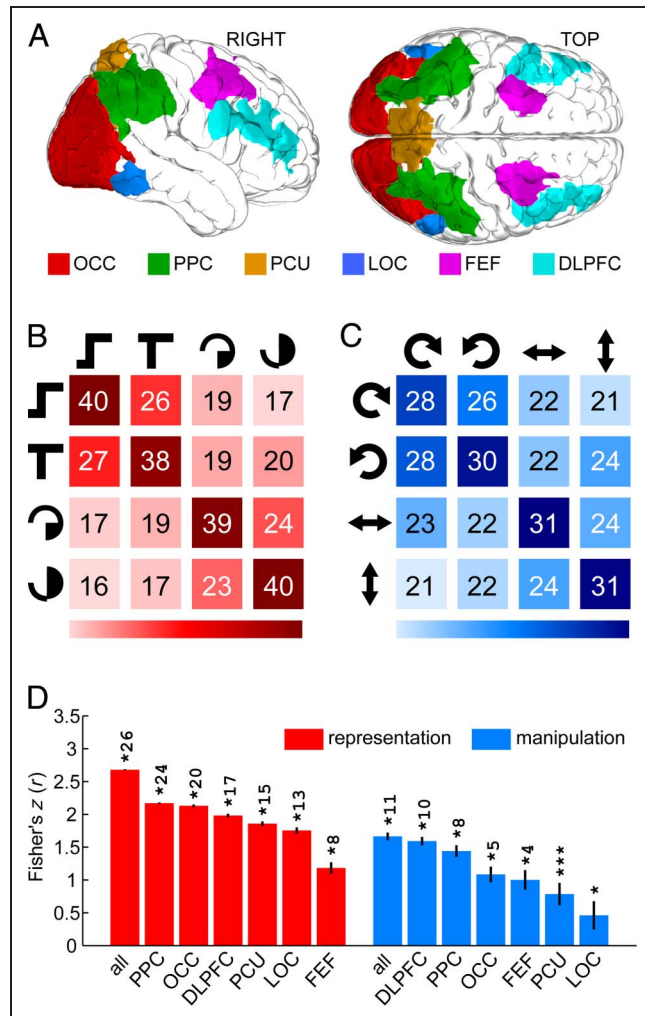


Figure 2. ROI classification results. (A) The six bilateral ROIs used in the current experiment, derived from the results of a previous study (see Methods). (B) Mean confusion matrix from a four-way classification among mental representations across the entire mental workspace network (compare with Figure 1A). (C) Analogous confusion matrix from a classification among mental manipulations (compare with Figure 1B). (D) Results of four-way classification analyses in each ROI. Correlations between resulting confusion matrices and similarity structures in Figure 1 are Fisher's Z transformed. Error bars are jackknife-corrected standard errors of the mean (see Methods). Asterisks indicate significance in a one-tailed jackknifed *t* test comparing Fisher's Z-transformed correlations with zero across participants (**p* ≤ .05, ****p* ≤ .001; *(*n*): *p* ≤ 1 × 10⁻ⁿ). Results are FDR corrected for multiple comparisons across the seven ROIs.

Schlegel et al., 2013). Five of these (lateral occipital cortex [LOC], posterior parietal cortex [PPC], precuneus [PCU], DLPFC, and FEF) were bilateral ROIs that showed greater activity during visual imagery manipulation than visual imagery maintenance in a whole-brain, group-level general linear model analysis. These ROIs were transformed separately for each participant from Montreal Neurological Institute space to that participant's native functional space for use in the current study. The remaining mask (occipital cortex [OCC]) was defined anatomically in each participant's native anatomical space using the following labels from FreeSurfer's cortical parcellation: inferior occipital gyrus and sulcus, middle occipital gyrus and sulci, superior occipital gyrus, cuneus, occipital pole, superior occipital and transverse occipital sulci, and anterior occipital sulcus (all bilateral). For the control ROI analysis, the thalamus was defined functionally as above, and the ventricle mask was defined anatomically from the following FreeSurfer cortical parcellation masks: left and right lateral ventricles, left and right inferior lateral ventricles, third ventricle, fourth ventricle, and fifth ventricle.

For the spatiotemporal multivariate classification, we used a linear support vector machine classifier and leave-one-trial-out cross-validation. Because we only considered correct response trials, a nonuniform number of trials existed for each condition and participant (57.4 trials per condition on average [$SEM = 0.203$]; see Table S1 for details). Although the difference in number of trials was small, we ensured that they could not affect the classification results by including a target balancing step in our cross-validation procedure. In this step, each classification fold was performed 10 times using random, balanced samples of the data, and the results for that fold were averaged across the 10 bootstrapped folds. For each classification, we used the spatiotemporal pattern of prewhitened BOLD data from the first 3 TRs of each correct response trial, shifted by 1 TR to account for the hemodynamic response function (HRF) delay inherent in fMRI data. We shifted by 1 TR only to include as much trial data as possible without also including data that could have been influenced by the test display. Prewhitening reduced each ROI's voxel-based pattern to an average of 93.6 data features ($SEM = 4.83$). Thus, each classification used spatiotemporal patterns of, on average, 280.8 dimensions ($SEM = 14.5$). Each feature dimension was z scored by run before classification to reduce between-run differences in signal that may have occurred because of scanner or physiological noise.

Our measure of classifier performance was the correlation between the confusion matrix resulting from the classification and the matrix form of either the shape or operation similarity structure (see Figure 1A and B). This measure is more sensitive than classification accuracy because it also takes into account confusions between conditions that result from the hierarchical relationship between the shapes and between the operations. We used a jackknife procedure to perform random effects

analyses evaluating the significance of the correlations (Miller, Patterson, & Ulrich, 1998). In the case of noisy estimates such as individual subject confusion matrices, jackknifed analyses can provide cleaner results without biasing statistical significance (see Miller et al., 1998, for more details on this method). In a jackknifed analysis with N participants, N grand means of the data (in this case, confusion matrices) are calculated, each with one participant left out. The correlation between each of these grand mean confusion matrices and the model similarity structure was then calculated, and a one-tailed t test evaluated whether the Fisher's Z -transformed correlations were positive (i.e., whether there was a significant correlation between confusion matrices and the model similarity structure across participants). Because the jackknife procedure reduces the variance between participants artificially, a correction must be applied to the t statistic calculation; specifically, the sample standard deviation between correlations is multiplied by the square root of $(N - 1)$.

ROI Cross-classification Analysis

To assess whether information about mental representations or mental manipulations was shared between areas, we performed a cross-classification analysis in which a classifier was trained on data from one ROI and tested on data from a second ROI. This analysis used the same procedures as the ROI classification analysis described above. However, because the voxel-based feature space of each ROI differed, data from pairs of ROIs needed to be transformed into a common feature space before classification. To do this, we first used FSL's MELODIC tool to transform each ROI's data from voxel space to 50 principal component signals using PCA. After this step, each ROI's pattern had the same dimensionality, but those patterns' features would be unlikely to correspond across ROIs. Therefore, for each pair of ROIs, these component signals were matched pairwise as follows to maximize the total similarity between component signals. First, the correlation distance ($1 - |r|$) between each pair of components was calculated, yielding a 50×50 correlation distance matrix. Next, the rows and columns of this matrix were reordered using the Hungarian algorithm to minimize the matrix trace (Kuhn, 1955). The components meeting along the diagonal of this reordered, trace-minimized matrix defined the pairwise matching. If two components were matched by this procedure but were anticorrelated, one component was negated to produce positively correlated component pairs. We performed this matching procedure for each fold of the cross-validation independently, excluding test data to avoid inflating the similarity between training and testing patterns artificially. Once this procedure was complete, data from the two ROIs shared a common feature space, that is, the two feature spaces had the same dimensionality, and corresponding features in the two spaces were

maximally similar. Cross-classification could then proceed by training the classifier on data from one ROI and testing it on data from the other ROI. Each ROI served both as the training set and as the testing set, with results averaged between the two cases. Figure 3 provides a visual schematic of the cross-classification analysis procedure.

Information Flow Classification Analysis

The goal of this analysis was to determine whether patterns of directed connectivity between processes occurring in pairs of ROIs could be used to classify either mental representations or mental manipulations. To this end, we transformed the functional data using PCA as above, but with dimensionality fixed at 10 components. For each participant, task condition (i.e., unique combination of shape and operation), and directed pair of areas (e.g., from PPC to DLPFC), we then calculated the Granger causality with a lag of 1 TR between each directed pair of principal component signals (e.g., between component i of PPC and component j of DLPFC). As input data for each component, we used the temporal concatenation of data from the first 5 TRs of each correct response trial of that condition, shifted by 1 TR to account for the HRF delay. We used 5 TRs in this analysis rather than the 3 TRs used in previous analyses to maximize the amount of data for the Granger causality calculations, which require signals of long duration to reveal influence. However, we did not test the optimum number of TRs to include in any of these analyses. For each participant and directed pair of ROIs, this procedure yielded sixteen 10×10 Granger-causal (GC) graphs, which were used as the patterns for classification. Each pattern was labeled based on either shape or operation and analyzed using a multivariate classification as in the ROI classifications described above. Because these patterns were defined for each task condition rather than for each trial, we used leave-one-operation-out cross-validation for the representation analysis and leave-one-shape-out cross-validation for the manipulation analysis. Directed connections with classification results that passed false discovery rate (FDR) correction for multiple comparisons across the 30 directed pairs in each analysis were used to construct directed graphs, which were then sorted topologically (see Figure 6B and D). Figure 5 provides a visual schematic of the information flow classification analysis procedure.

RESULTS

Performance accuracy was high after an initial training session during which participants memorized the shapes, operations, and corresponding letter and number mappings (responses were correct in 95.8% of trials across participants and conditions). One-way between-participant analyses of variance showed no significant differences in accuracy across conditions, confirming that the difficulties

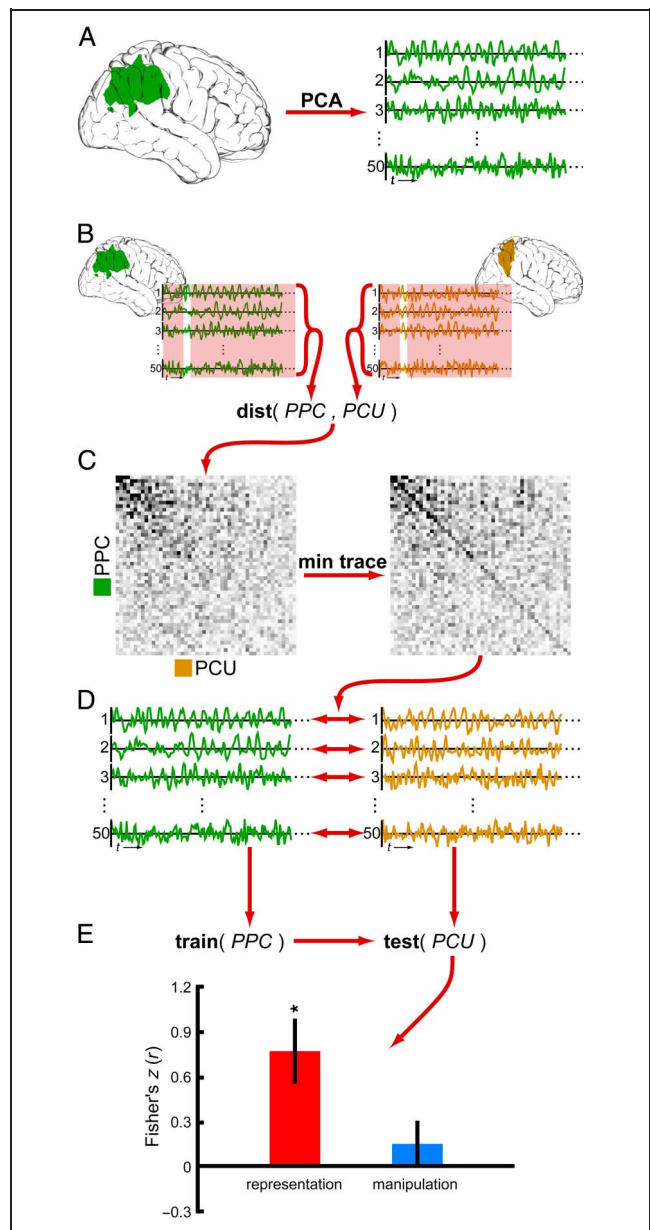


Figure 3. Visual schematic of cross-classification analysis procedure. (A) Functional data from each ROI (PPC shown here) were transformed from voxel space to 50 principal component signals using PCA. (B) For a single cross-classification analysis between two ROIs (PPC and PCU here), the correlation distance between each pair of principal component signals was calculated. This calculation was performed independently for each classification fold, leaving out the test data from that fold (visualized here as a gap in the data that was used to calculate distances). This resulted in a 50×50 correlation distance matrix. (C) The trace of this correlation distance matrix was minimized using the Hungarian algorithm to compute a matching of component signals between the two ROIs that maximized their pairwise similarity (i.e., minimized their correlation distance). (D) This procedure resulted in a common 50-dimensional feature space shared between the two ROIs. Matched principal component signals between ROIs were maximally similar to each other. (E) A cross-classification analysis was performed using these transformed functional data. The classifier was trained on data from one ROI (PPC in this case) and tested on data from the other ROI (PCU in this case). Other than the difference between training and testing data, the classification was carried out identically to that of the ROI-based analyses in Figure 2.

of shapes and operations were well matched (for shapes: $F(3, 72) = 1.65, p = .185$; for operations: $F(3, 72) = 0.369, p = .775$; see Table S1 for behavioral results). In addition, participants showed no difference in response times (RTs) between the operation conditions ($F(3, 72) = 0.0509, p = .985$; Table S2). Participants did show a significant difference in RTs between the shape conditions ($F(3, 72) = 6.107, p = 9.24 \times 10^{-4}$), but an ROI classification analysis for shape (see next section) with RT-matched subsets of trials confirmed that our results were not because of this RT difference (Table S3; all corrected p values $< .002$).

ROI Classification Analysis

Our ROIs for analysis of the fMRI data were the six bilateral cortical regions that contained information pertaining to the transformation of visual imagery in a previous study that used data independent from those of the current study (Figure 2A; see Methods for details on how these ROIs were defined; Schlegel et al., 2013). Each area has been shown to play a role in neural processing related to the current task (Crowe et al., 2013; Harrison & Tong, 2009; Margulies et al., 2009; Zacks, 2008; Schall, 2004; Tanaka, 1996). We used multivariate decoding methods to determine whether each ROI supported information about mental representations and/or mental manipulations of visual imagery, that is, whether patterns of neural activity in each ROI could be used to classify either the shape that was represented in visual imagery during each trial or the operation that was used to manipulate that representation.

Because of the hierarchical relationship that we introduced among shapes and operations, we measured classifier performance using a representational similarity analysis in which we correlated the confusion matrix resulting from each four-way classification with the matrix form of this hierarchical similarity structure (Figure 1A and B; Schlegel et al., 2013; Kriegeskorte, Mur, & Bandettini, 2008). This measure allowed us to use information from both correct classifications (classification “hits”) and specific patterns of confusion (classification “misses”) between conditions that resulted from the relationships among shapes and among operations. Thus, classification was only “successful” if the classifier performed according to our hypothesized pattern of correct classification and confusion, allowing us to achieve greater sensitivity than purely “accuracy”-based classification methods and to verify that our results were not because of task-irrelevant factors such as the letters or numbers used in the task mapping.

Initial classifications using the union of all ROIs confirmed that the information processing structure of this network matched precisely the similarity structures of both shape and operation sets (Figure 2B and C; for shapes: $t(18) = 10.6, p = 8.59 \times 10^{-26}$; for operations: $t(18) = 16.0, p = 4.54 \times 10^{-12}$; results are FDR corrected for multiple comparisons). This result also held true for

classification analyses performed on each ROI separately (Figure 2D; FDR corrected for multiple comparisons across the seven total analyses for each classification scheme). Because all of our results were significant, we verified the specificity of our analysis by conducting control classifications using two additional masks. The first was a functionally defined, bilateral thalamus ROI from our previous study that showed increased but not task-specific activity during mental manipulation of imagery compared with maintenance of imagery; the second was an anatomically defined ventricle mask. None of the four control classification analyses using these masks reached significance, confirming that our original analyses detected information about the shapes and operations specifically within our six ROIs (see Table S4 for ROI control analysis results). As a further control to confirm that our analysis was valid and unbiased, we shuffled the labels randomly in each classification and found that the correlations between confusion matrices and model similarity structures were no longer significant (Table S5). Thus, neural activity in each ROI supported information about both representation and manipulation of visual imagery. This result provides evidence that processing of both representations and manipulations is distributed throughout the mental workspace network, running counter to models such as Baddeley’s or Postle’s that propose that its component processes are segregated to particular cortical regions (Sreenivasan et al., 2014; Lee et al., 2013; Postle, 2006; Baddeley, 2003; Ishai et al., 2000). The large effect sizes and specificity of our results underscore the sensitivity of our experimental design and representational similarity-based analysis for uncovering information that other techniques such as univariate analyses or two-way classifications may have missed.

ROI Cross-classification Analysis

Our previous analysis suggests that information about both representations and manipulations is distributed throughout the mental workspace network, but what is the nature of the information carried by each region? Our hypothesis implies that information is shared commonly throughout the network. Alternatively, however, each network node could process a unique informational aspect of representation and manipulation. For instance, Lee and colleagues (2013) suggest that, whereas visual cortex represents image-level information (e.g., edges, corners, contours), information in pFC is conceptual in nature (e.g., “the T shape” or “the tadpole shape”). In this alternative scenario, we would expect our previous classification analysis to succeed in both visual cortex and pFC, although the classifier would have picked up on different information in each region. To resolve between these possibilities, we developed a novel multivariate cross-classification analysis to investigate whether information is shared among the nodes of the mental workspace network. In a cross-classification analysis, a

classifier is trained on one data set (in this case, one ROI) and tested on a different data set (in this case, a different ROI). A successful cross-classification provides evidence that information is shared between the two data sets. In this case, it would provide evidence that information is shared between the two ROIs, rather than the alternative possibility that both ROIs support information about the task but in separate formats. However, we initially face a technical hurdle to cross-classifying between ROIs because cross-classification requires the two data sets to share the same feature space. In other words, cross-classification would require the feature space of each ROI to have identical dimensionality (e.g., same number of voxels) and each feature of one ROI to carry the same meaning as the corresponding feature in the other ROI. Voxel-based ROIs do not meet either of these criteria, so we first needed to transform each ROI's data into a common feature space before we could perform the cross-classification analysis.

Conceptually, we hypothesized that the functional data for a given ROI were a set of signals in voxel space that represented a mixture of a number of underlying informational subprocesses that were shared in a distributed manner between the ROIs. If this characterization is valid, then PCA would allow us to transform our voxel-based data independently for each ROI to recover a set of principal component signals that represented those underlying subprocesses that were mixed between the voxel-space signals that we actually measured. We therefore used PCA to convert the voxel-based data from each ROI into 50 principal component signals. We chose the number 50 to construct classification patterns of sufficient size while remaining smaller than the size of our smallest ROIs; however, we did not test whether this was the optimum dimensionality to use. This step allowed us to establish feature spaces for the ROIs that had uniform dimensionality. The second step required to construct a common feature space for cross-classification was to rearrange the dimensions of these feature spaces such that corresponding features carried the same informational meaning across ROIs. To achieve this, for each cross-classification between two ROIs, we performed a pairwise matching of component signals between the two ROIs to maximize the total correlation between matched component signal pairs (i.e., so that each component signal from the first ROI was matched to a maximally similar signal from the second ROI). We performed this matching step independently for each fold of the cross-classification, leaving out data from the testing set to avoid artificially inflating the similarity of test patterns across the two ROIs.

This two-step process yielded a common 50-dimensional feature space for each fold of each cross-classification analysis (see Figure 3 for a visual schematic of the procedure). Classification then proceeded exactly as in the previous ROI classification analysis. We cross-classified between each pair of ROIs, with results presented in Figure 4 (all results FDR corrected across the 15 ROI

pairs). We could successfully cross-classify mental representations between most pairs of ROIs, providing evidence that information about mental visual representations is shared widely throughout the mental workspace network. The cross-classification of mental manipulation was significant only between DLPFC and PPC ($t(18) = 1.93$, $p = .0346$ [uncorrected]). However, this result did not hold after FDR correction. This result suggests that information about manipulations of visual imagery is distributed but may be more compartmentalized in the network, with DLPFC and PPC possibly sharing some information. As in the ROI classification above, we confirmed the validity of the analysis by performing control analyses in which labels were shuffled randomly. In this case, the cross-classifications were no longer significant, ruling out the possibility that our cross-classification results occurred because of unknown biases introduced by our analysis pipeline (Table S6). Thus, information about mental representations is not only distributed throughout the network but also shared between many network nodes. The trend in our data suggests that information about mental manipulations may be shared between DLPFC and PPC, but we did not find evidence of information sharing related to manipulations between any of the other ROI pairs.

Information Flow Classification Analysis

To investigate how this information becomes shared, we developed a new method to analyze whether information

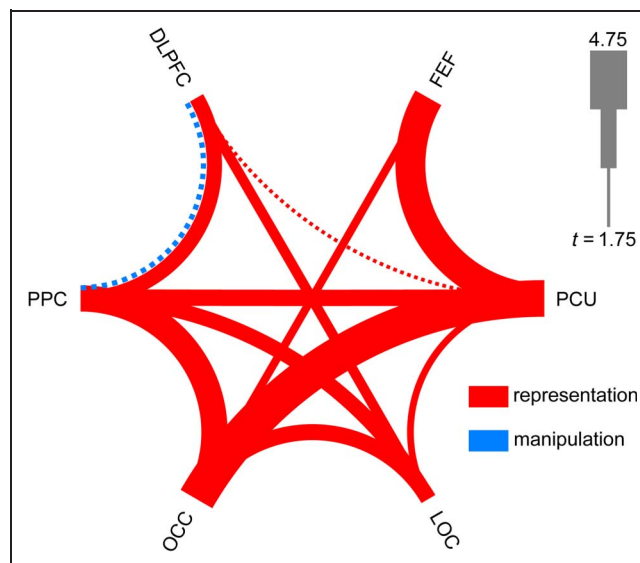


Figure 4. ROI cross-classification results. Arcs indicate pairs of ROIs in which cross-classification was successful. Dotted arcs indicate classifications that were significant when uncorrected, but which did not pass FDR correction across the 15 ROI pairs. All other displayed classifications passed FDR correction. Arc thickness indicates t statistic values in a one-tailed, jackknifed t test of Fisher's Z -transformed correlations between confusion matrices and model similarity structures, compared with zero (see text). Abbreviations are as in Figure 2.

is carried in patterns of directed connectivity between pairs of network nodes. This analysis abstracted away from information contained in patterns of activity within neural regions, seeking instead to probe the informational content of patterns of information flow between pairs of neural regions. Established methods for assessing directed connectivity produce a single value that characterizes the degree to which processing in one region is predictive of later processing in another region (Barnett & Seth, 2014; Friston, 2011; Lizier et al., 2011). These methods can detect increases or decreases in directed connectivity but are insensitive to information that may be carried via patterns of such connectivity. Because of this limitation, two processes (e.g., clockwise and counterclockwise mental rotation) may entail distinct patterns of directed connectivity without involving different overall magnitudes of directed connectivity and would thus be indistinguishable by current methods. Furthermore, in the present analysis, we were not concerned directly with whether information flowed between nodes, because in a densely connected, distributed network, each node will likely exert a complex pattern of control over all other nodes. Rather, we wanted to know whether the condition-specific patterns of directed connectivity between the underlying informational processes that were distributed among these nodes supported information about specific mental representations and manipulations. If so, then the current analysis would provide further evidence for the findings of the previous two analyses that the information processing underlying the mental workspace occurs at a fundamentally distributed level of organization in the cortex.

As directed connectivity patterns, we used GC graphs constructed independently for each unique task condition (Barnett & Seth, 2014). Granger causality is a statistical method for evaluating the ability of a source signal to predict the future of a destination signal beyond the predictive power provided by the destination signal's own past. Although the validity of Granger causality for fMRI data has come under scrutiny, computational and empirical work has shown that it is a viable technique when proper precautions such as those used in this study are taken (Barnett & Seth, 2014; Friston, Moran, & Seth, 2013; Wen, Rangarajan, & Ding, 2013). Specifically, we investigated differences in patterns of Granger causality between conditions rather than attempting to establish "ground-truth" connectivity between regions. Our GC graphs were constructed as follows: First, voxel-based data from each ROI were transformed individually using PCA into 10 principal component signals, with the same rationale as described above for the cross-classification analysis. We used 10 components here instead of 50 so that our resulting GC graphs would have a reasonable dimensionality for classification, but we again did not evaluate the optimal dimensionality to use. Next, we constructed a 10×10 GC graph for each of the 16 unique task conditions (e.g., Shape 1 + clockwise rotation), each

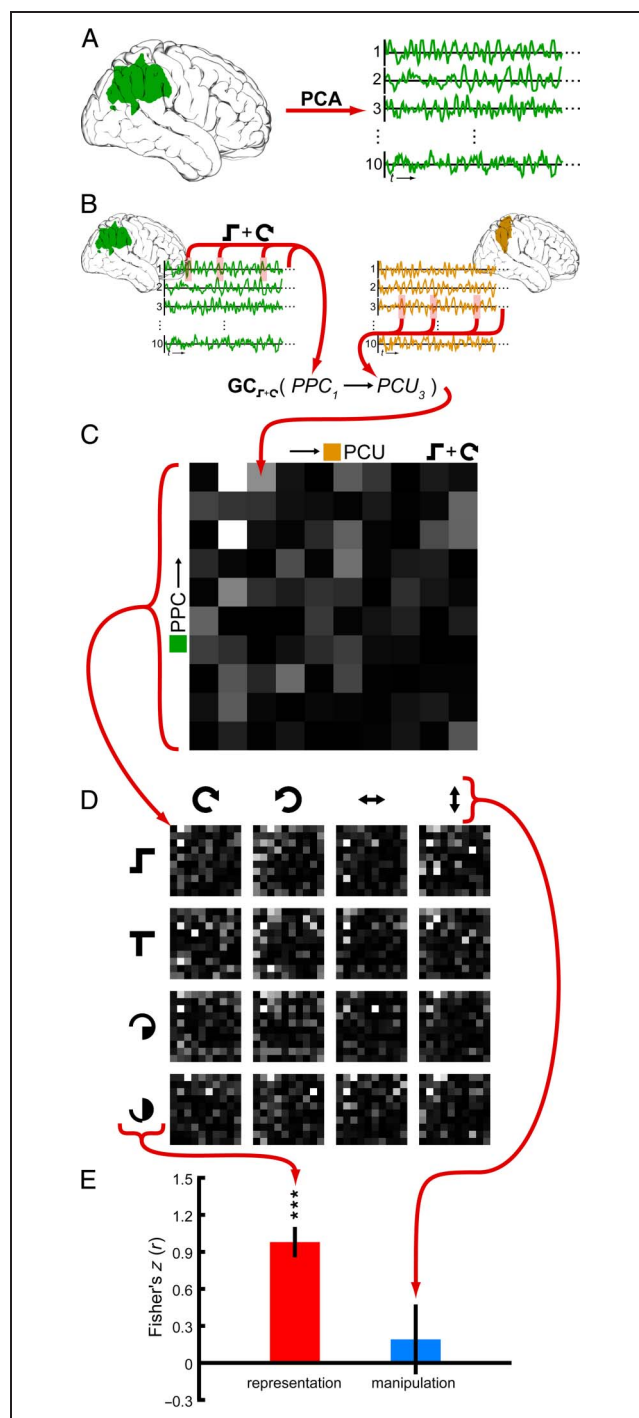


Figure 5. Visual schematic of information flow classification analysis procedure. (A) Functional data from each ROI (PPC shown here) were transformed from voxel space to 10 principal component signals using PCA. (B) For a given directed ROI pair (PPC to PCU here) and condition (Shape 1 and clockwise rotation here), the Granger causality from the source ROI to the destination ROI was calculated for each pair of principal component signals (PPC component 1 and PCU component 3 here), using only data from trials of that condition. (C) This resulted in a 10×10 GC graph for each participant, directed ROI pair, and condition. (D) The resulting 16 GC graphs for a given participant and directed ROI pair could be labeled based on either shape or operation. (E) A classification analysis then proceeded as in the other analyses, except that either leave-one-shape-out or leave-one-operation-out cross-validation was performed.

participant, and each directed pair of ROIs (e.g., from PPC to DLPFC). Each GC graph was constructed by computing the Granger causality from each of the 10 principal components in the source ROI to each of the 10 principal components in the destination ROI, using only data from one task condition.

For each participant, task condition, and directed pair of ROIs, this process yielded a pattern of directed connectivity (the GC graph) that represented a task-specific, directed pattern of information flow between regions. We then used these GC graphs as inputs to classification

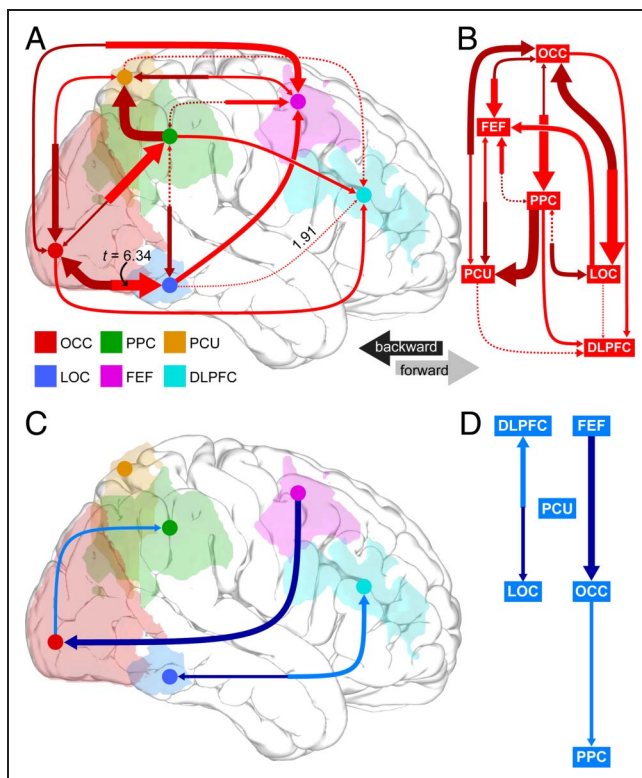


Figure 6. Information flow classification results. (A) Graph indicating directed ROI pairs in which patterns of information flow could be used successfully to classify mental representation. Dotted arrows indicate classifications that were significant when uncorrected, but which did not pass FDR correction across the 30 directed ROI pairs. All other displayed classifications passed FDR correction. Arrow thickness indicates t statistic values in a one-tailed, jackknifed t test of Fisher's Z -transformed correlations between confusion matrices and model similarity structures, compared with zero (see text). Light red arrows indicate posterior-to-anterior connections, and dark red arrows indicate anterior-to-posterior connections. The greatest effect occurred for the OCC-to-LOC connection, and the smallest significant effect occurred for the LOC-to-DLPFC connection. Both effect sizes for these two connections are indicated on the graph for reference. Abbreviations are as in Figure 2. (B) A topological sorting of the graph from A reveals that the OCC resides at the top of a bottom-up hierarchy of information flow for mental visual representations. (C) Graph indicating directed ROI pairs in which patterns of information flow could be used successfully to classify mental manipulation. Arrow properties are as in A. (D) A topological sorting of the graph from C reveals that the DLPFC and FEF reside at the top of a top-down hierarchy of information flow for mental manipulations of visual imagery.

analyses as described above, with results presented in Figure 6 (Figure 5 provides a visual schematic of this analysis). Complementing our ROI classification and ROI cross-classification results, we found that frequently bidirectional patterns of directed information flow between many nodes of the mental workspace network could be used to classify the mental shape representations. A topological sorting of the resulting directed graph of significant classification results revealed a posterior-to-anterior hierarchy for mental representations, with the OCC at the top and connectivity cascading down to the DLPFC (i.e., a bottom-up hierarchy; Figure 6B). The pattern of results for the manipulation classification shows a sparser graph, with the DLPFC and FEF at the top of an anterior-to-posterior hierarchy (i.e., a top-down hierarchy; Figure 6D). Here, being placed at the top of the hierarchy indicates dominance in the sense that a higher node supports more information in outward flowing rather than inward flowing directed connectivity patterns. As in the previous two analyses, we performed control classifications with shuffled labels, confirming the validity of the analysis (Table S7).

DISCUSSION

The mental workspace is a cognitive system that enables the volitional, flexible mental operations underlying the mathematical, scientific, and artistic creativity that distinguish humans as a species (Logie, 2003; Dehaene & Naccache, 2001). Here, we applied novel network-level pattern analysis methods to reveal the structure of information flow in the neural network that supports the mental workspace. We find that the component processes of representing and manipulating visual imagery entail a level of informational organization that transcends the anatomical structures that standard models of working memory regard as functionally encapsulated modules. Instead, our data imply that such processes emerge out of the fundamentally distributed sharing and flow of information between the nodes of a cortex-wide network. We found that representations entail the sharing and flow of information between all of the ROIs we tested. Mental manipulations showed patterns of information flow between all but one of our ROIs, but we did not find significant sharing of information at the scale of our fMRI data after correcting for multiple comparisons. It is important to note, however, that further information sharing and flow could have occurred at spatial or temporal scales or levels of information processing to which our data or analyses were insensitive. Because fMRI data are temporally low-pass filtered by the HRF, our data can only address information flow that occurs on the scale of seconds. Nonetheless, our findings call into question "textbook" anatomically modular models of the neural basis of working memory and other higher-order mental functions (Sreenivasan et al., 2014; Lee et al., 2013; Postle, 2006; Baddeley, 2003; Kane & Engle, 2002).

Existing neural models of working memory and related processes could be described as “distributed” in the sense that they assign the component functions of working memory to anatomical modules that are distributed across the brain. However, a key advance in this study is to suggest that even these component processes that underlie the more complex functions we studied are distributed in the brain. Thus, contrary to models such as Baddeley’s that localize executive functions to lateral pFC and the storage of visual representations to OCC, our data suggest that informational processing in the mental workspace is more fundamentally distributed. Thus, anatomy may be incidental for at least some aspects of the high-level mental functions studied here, with the actual functional separation of processes occurring at a higher level of informational organization.

Our work advances recently developed analytical techniques that approach the brain as an information processing network. Multivariate classification and representational similarity analyses allow the informational structure of processes at many levels of organization to be probed (Haxby, Connolly, & Guntupalli, 2014; Kriegeskorte et al., 2008). Directed connectivity measures enable the investigation of effective functional coupling between network nodes (Barnett & Seth, 2014; Seghier & Friston, 2013; Lizier et al., 2011; Schurger, Pereira, Treisman, & Cohen, 2010; Xue et al., 2010). Here, we adapted these techniques to answer two new kinds of question. First, our ROI cross-classification analysis was able to evaluate whether information is shared between multiple network nodes. Note that traditional representational similarity analyses as proposed by Kriegeskorte and colleagues (2008) are not able to answer this question generally. For instance, it could have been the case that visual cortex represents mental images only at a stimulus level (e.g., edges, corners, contours) whereas pFC represents those images only at a conceptual level (e.g., “the T shape” or “the tadpole shape”). In this case, the dissimilarity structures derived from each ROI could still be highly correlated with each other (e.g., Shape 1 and Shape 2 are similar at the stimulus level because they are both derived from a rectilinear grid and are also “conceptually” similar because they both look like letters). However, these matching dissimilarity structures would have derived from very different underlying informational spaces, and thus, it would be erroneous to conclude that the correlation between those dissimilarity structures indicates sharing of information between the ROIs. The second question that our new techniques allowed us to address was whether patterns of information flow between network nodes carry information about the functional significance of the connections between those nodes. These questions and the techniques described here to investigate them are generally applicable across a range of topics both within neuroscience—for instance, learning (Bassett et al., 2010), intelligence (Jung & Haier, 2007), language (Schlegel, Rudelson, & Tse, 2012), and attention (Baldauf & Desimone, 2014)—and in other fields that study similar

informational networks in biology and beyond (Bassett & Gazzaniga, 2011).

It should be noted that using fMRI restricted our sensitivity to functional interactions occurring at millimeter or larger spatial scales and on the order of seconds. It is likely that we missed the sharing and flow of information occurring in more local small-scale neural circuits and on shorter timescales than we could measure. For instance, the reduced sharing of information and connectivity we found for manipulations of visual imagery may not be an indication that such sharing and connectivity do not occur in the brain, if such processes occur at finer spatial or temporal scales than fMRI can measure. In addition, focusing on the six ROIs that had previously shown information pertaining to visual imagery increased the power of our analyses within this restricted network. However, this statistical power was gained at the expense of potentially missing a larger scope for the mental workspace network. Indeed, we previously found six additional bilateral neural regions in the cerebellum, thalamus, medial temporal lobe, supplementary eye field, frontal operculum, and medial frontal cortex with activity that differed depending on whether visual imagery was manipulated or maintained, but we could not classify between different mental operations in these regions and thus have yet to determine the contribution of these additional nodes to the types of behaviors studied here (Schlegel et al., 2013). Future work should investigate whether the mental workspace network is even larger and more distributed than we report here.

Although we found shared information pertaining to representations in each of the ROIs we studied, an alternative explanation for this finding could be that information about representations merely spreads passively from a single area such as visual cortex that actually processes that information. However, our finding of widespread bidirectional information flow between many network nodes suggests that this is an unlikely possibility. The bidirectionality, density, and hierarchical nature of the connectivity between these nodes lead more parsimoniously to an interpretation that the brain processes mental visual representations in a fundamentally distributed manner.

Connectivity analyses such as those presented here are vulnerable to the lurking variable problem, in which two network nodes appear to support a direct informational connection when in fact each supports independent yet parallel processes or is mutually driven by a third unknown process. Our information flow results may be affected by this situation, because our network showed a dense pattern of connectivity and we did not test each connection for mediating variables. Because of this, we suggest that these findings be interpreted more holistically as providing evidence for fundamentally distributed information processing in the brain, rather than as having deduced a precise wiring diagram of the mental workspace network.

Finally, our results should not be interpreted to mean that anatomically modular processing does not occur in

the cortex. Indeed, a rich body of evidence from lesion and other studies suggests that there are many cognitive functions, working memory included, for which specific regions of the cortex are necessary (Damasio & Damasio, 1989). However, our data do suggest that anatomical modularity cannot provide a complete explanation of the neural processes underlying the mental workspace. To contrast exclusively between either “distributed” or “modular” processing would be too simplistic, because even a relatively simple conceptual construct like a “mental representation” likely has a complex implementation in the brain that plays out over multiple regions and at multiple levels of informational abstraction. Thus, although our data speak against models that make anatomically modular claims such as “Visual cortex is the representational module of visual working memory,” they do not inform other statements of modularity such as “Visual cortex and DLPFC both mediate mental representation, but they play unique roles in that process.” Our data even leave open the possibility that cognitive processes such as mental representation require both anatomically modular and fundamentally distributed processing, but at different temporal scales. In support of this idea, a recent study by Siegel, Buschman, and Miller (2015) suggests that both localized and distributed processing may occur during different stages of working memory tasks. The unique contribution of each type of processing remains to be determined.

Our results provide new evidence that high-level cognitive processes such as the representation and manipulation of visual imagery are mediated via the complex, fundamentally distributed sharing and flow of information throughout the cerebral cortex. Although much work in cognitive neuroscience has been concerned with reducing the brain’s functions to discrete, localized regions, our results provide evidence that the component processes of at least some forms of high-level cognition transcend anatomically segregated structures, emerging fundamentally from the interaction between several levels of organization (Bassett & Gazzaniga, 2011; Bressler & Menon, 2010). The field has found studying such interactions vital yet difficult (Insel et al., 2013; Markram, 2012; Bressler & Menon, 2010; Bullmore & Sporns, 2009), and the new methods reported here to investigate the structure, sharing, and flow of information in the brain may prove useful in understanding many other complex cognitive processes (Schlegel et al., 2012, 2015; Baldauf & Desimone, 2014; Bassett et al., 2010; Jung & Haier, 2007). Future work should investigate how precisely the distributed flow of information in the cortex supports high-level cognitive abilities and whether this mode of information processing is unique to certain forms of cognition or common across many cortical functions.

Acknowledgments

A. S. would like to thank Patrick Cavanagh, David Kraemer, and Frank Tong for serving on his PhD dissertation committee, of

which this study was a part. This study was funded by NSF Graduate Research Fellowship 2012095475 to A. S. and Templeton Foundation Grant 14316 to P. U. T. All data and code are available online at www.dartmouth.edu/~petertse.

Reprint requests should be sent to Alexander Schlegel, Department of Psychological and Brain Sciences, H. B. 6207, Moore Hall, Dartmouth College, Hanover, NH 03755, or via e-mail: schlegel@gmail.com.

REFERENCES

- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–839.
- Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science*, 344, 424–427.
- Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, 223, 50–68.
- Bassett, D. S., & Gazzaniga, M. S. (2011). Understanding complexity in the human brain. *Trends in Cognitive Sciences*, 15, 200–209.
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2010). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences, U.S.A.*, 108, 7641–7646.
- Bassok, M., Dunbar, K. N., & Holyoak, K. J. (2012). Introduction to the special section on the neural substrate of analogical reasoning and metaphor comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 261–263.
- Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, 14, 277–290.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10, 186–198.
- Crowe, D. A., Goodwin, S. J., Blackman, R. K., Sakellaridi, S., Sponheim, S. R., Macdonald, A. W., et al. (2013). Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition. *Nature Neuroscience*, 16, 1484–1491.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9, 179–194.
- Damasio, H., & Damasio, A. R. (1989). *Lesion analysis in neuropsychology*. Oxford: Oxford University Press.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79, 1–37.
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience*, 29, 15258–15265.
- Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Current Opinion in Neurobiology*, 23, 172–178.
- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, 1, 13–36.
- Graham, D. J., & Rockmore, D. (2011). The packet switching brain. *Journal of Cognitive Neuroscience*, 23, 267–276.
- Hadamard, J. (1954). *The psychology of invention in the mathematical field* (p. 142). New York: Dover Publications, Inc.
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python

- toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7, 37–53.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458, 632–635.
- Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, 37, 435–456.
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, 8, 280–285.
- Insel, T. R., Landis, S. C., & Collins, F. S. (2013). The NIH brain initiative. *Science*, 340, 687–688.
- Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, 28, 979–990.
- Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, 30, 135–187.
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9, 637–671.
- Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia*, 33, 1335–1344.
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 1–28.
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2, 83–97.
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, 16, 997–999.
- Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J.-D., & Prokopenko, M. (2011). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of Computational Neuroscience*, 30, 85–107.
- Logie, R. H. (2003). Spatial and visual working memory: A mental workspace. *Psychology of Learning and Motivation*, 42, 37–78.
- Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., et al. (2009). Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences, U.S.A.*, 106, 20069–20074.
- Markram, H. (2012). The human brain project. *Scientific American*, 306, 50–55.
- Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, 35, 99–115.
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23–38.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Schall, J. D. (2004). On the role of frontal eye field in guiding attention and saccades. *Vision Research*, 44, 1453–1467.
- Schlegel, A., Alexander, P., Fogelson, S. V., Li, X., Lu, Z., Kohler, P. J., et al. (2015). The artist emerges: Visual art learning alters neural structure and function. *Neuroimage*, 105, 440–451.
- Schlegel, A., Kohler, P. J., Fogelson, S. V., Alexander, P., Konuthula, D., & Tse, P. U. (2013). Network structure and dynamics of the mental workspace. *Proceedings of the National Academy of Sciences, U.S.A.*, 110, 16277–16282.
- Schlegel, A., Rudelson, J. J., & Tse, P. U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience*, 24, 1664–1670.
- Schurger, A., Pereira, F., Treisman, A., & Cohen, J. D. (2010). Reproducibility distinguishes conscious from nonconscious neural representations. *Science*, 327, 97–99.
- Seghier, M. L., & Friston, K. J. (2013). Network discovery with large DCMs. *Neuroimage*, 68, 181–191.
- Siegel, M., Buschman, T. J., & Miller, E. K. (2015). Cortical information flow during flexible sensorimotor decisions. *Science*, 348, 1352–1355.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23(Suppl. 1), S208–S219.
- Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, 17, 652–660.
- Sreenivasan, K. K., Vytlačil, J., & D’Esposito, M. (2014). Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. *Journal of Cognitive Neuroscience*, 26, 1141–1153.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19, 109–139.
- Tong, F. (2013). Imagery and visual working memory: One and the same? *Trends in Cognitive Sciences*, 17, 489–490.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215, 216–242.
- Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139, 352–402.
- Van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, 29, 7619–7624.
- Wen, X., Rangarajan, G., & Ding, M. (2013). Is Granger causality a viable technique for analyzing fMRI data? *PLoS One*, 8, e67428.
- Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater neural pattern similarity across repetitions is associated with better memory. *Science*, 330, 97–101.
- Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience*, 20, 1–19.