Multivoxel Pattern Analysis Reveals Increased Memory Targeting and Reduced Use of Retrieved Details during Single-Agenda Source Monitoring

Susan G. R. McDuff,¹ Hillary C. Frankel,³ and Kenneth A. Norman^{1,2}

¹Department of Psychology, and ²Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08540, and ³Harvard Medical School, Boston, Massachusetts 02115

We used multivoxel pattern analysis (MVPA) of functional MRI (fMRI) data to gain insight into how subjects' retrieval agendas influence source memory judgments (was item X studied using source Y?). In Experiment 1, we used a single-agenda test where subjects judged whether items were studied with the targeted source or not. In Experiment 2, we used a multiagenda test where subjects judged whether items were studied using the targeted source, studied using a different source, or nonstudied. To evaluate the differences between single- and multiagenda source monitoring, we trained a classifier to detect source-specific fMRI activity at study, and then we applied the classifier to data from the test phase. We focused on trials where the targeted source and the actual source differed, so we could use MVPA to track neural activity associated with both the targeted source and the actual source. Our results indicate that single-agenda monitoring was associated with increased focus on the targeted source (as evidenced by increased targeted-source activity, relative to baseline) and reduced use of information relating to the actual, nontarget source. In the multiagenda experiment, high levels of actual-source activity were associated with increased correct rejections, suggesting that subjects were using recollection of actual-source information to avoid source memory errors. In the single-agenda experiment, there were comparable levels of actual-source activity (suggesting that recollection was taking place), but the relationship between actual-source activity and behavior was absent (suggesting that subjects were failing to make proper use of this information).

Key words: memory; long-term memory; fMRI; source memory; retrieval strategies; pattern classification

Introduction

Research on source memory, the ability to recall the conditions under which a memory was acquired, has increasingly come to focus on how agendas at the time of retrieval can influence which mnemonic features are retrieved and used to make source judgments (Johnson et al., 1993; Mitchell et al., 2008). One factor that has been shown to influence source memory judgments is the number of sources mentioned in the test instructions: Subjects are more likely to falsely attribute items to a source when that source is the only one mentioned at test (single-agenda source monitoring) compared to when multiple sources are mentioned at test (multiagenda source monitoring) (Lindsay and Johnson, 1989; Zaragoza and Koshmider, 1989; Dodson and Johnson, 1993; Henkel et al., 2000).

The goal of our study is to map out differences in how subjects make source judgments on single-agenda versus multiagenda tests. One possible difference relates to how subjects cue memory.

Results from single-agenda tests suggest that, on these tests, subjects attempt to constrain retrieval to the targeted source by activating source-specific information from the study phase (Herron and Rugg, 2003a; see also Rugg, 2004; Jacoby et al., 2005a,b); according to the encoding specificity principle (Tulving and Thomson, 1973), activating targeted-source information should boost recall of the targeted source and reduce retrieval of nontarget source information. Subjects may be more prone to apply this kind of constraint on single-agenda versus multiagenda tests. Another possible difference relates to decision-making: Johnson and Raye (2000) make a distinction between activation (retrieval) of information and how strongly subjects "weight" (use) retrieved information. If subjects focus on the target source during single-agenda tests, they may retrieve information pertaining to nontarget sources but fail to properly use this information when making their source judgments.

To evaluate these hypotheses, we ran functional MRI (fMRI) studies where we manipulated use of single-agenda versus multiagenda instructions. In these studies, we used multivoxel pattern analysis (MVPA) (Norman et al., 2006) to measure activation (at test) of source-specific patterns of brain activity from the study phase. We focused on trials where the targeted source and the actual source differed (i.e., subjects were asked "Was this word studied with source X?" when it was actually studied with source Y). We hypothesized that different approaches to memory

Received July 30, 2008; revised Nov. 22, 2008; accepted Dec. 3, 2008.

This work was supported by National Institute of Mental Health Grant P50 MH062196 to K.A.N. and a National Science Foundation Graduate Research Fellowship to S.G.R.M. We thank Marcia Johnson and an anonymous reviewer for their comments on this manuscript.

Correspondence should be addressed to Kenneth A. Norman, Department of Psychology, Princeton University, Green Hall, Washington Road, Princeton, NJ 08540. E-mail: knorman@princeton.edu.

DOI:10.1523/JNEUROSCI.3587-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/290508-09\$15.00/0

Figure 1. Illustration of the experimental procedure used during a single scanning run, where each run consisted of a study phase and a test phase (separated by a 1-min serial response task). The left side of the figure illustrates the sequence of events during the study phase. The study phase consisted of miniblocks (lasting 3 trials) during which subjects performed a particular encoding task (artist, function, or read). The lower-left part of the figure shows the trial sequence during an artist-task miniblock. On each study trial, subjects were instructed to perform the specified encoding task for 2 s (without overtly responding); after 2 s, an asterisk appeared, at which point subjects had 2 s to enter their 1-to-5 response. The right side of the figure illustrates the sequence of events during the test phase. The test phase consisted of miniblocks (lasting 3 trials) during which subjects were asked to target a particular encoding task. The lower-right part of the figure shows two example test trials from a read-task miniblock. Subjects were presented with a mixture of old items from the study phase and new items. On each test trial, subjects were given a task cue, followed by a test word; upon presentation of the test word, subjects had to indicate whether they had studied that word with the targeted task. The test screens shown here are from Experiment 1 (the test instructions differed slightly for Experiment 2).

cueing and decision-making would be associated with distinct patterns of activation. If subjects are more prone to constrain retrieval (by activating targeted-source information) in the single-agenda case, this should show up as an increase in neural activation of the targeted-source pattern in the single-agenda condition. Also, we can evaluate how well subjects are using recollection of the actual (nontarget) source by looking at the relationship between neural activation of the actual-source pattern and behavior. If subjects are using actual-source information to make their source decisions, high levels of actual-source activation (indicating recollection of the actual, nontarget source) should be associated with increased correct rejections. We predicted that actual-source activation would be more closely tied to behavior in the multi-agenda (vs the single-agenda) experiment.

Materials and Methods

Subjects. Eleven people participated in Experiment 1 (four female, ages 20–26). Twelve people participated in Experiment 2 (six female, ages 19–28). Subjects were drawn from the graduate and undergraduate student community at Princeton University and received financial compensation for their participation. The experiments were run sequentially (first Experiment 1, then Experiment 2).

Materials. Experimental stimuli consisted of 216 noun words drawn from the MRC database (Coltheart, 1981; Wilson, 1988). The words that we used in the experiment were all between 4 and 9 letters in length ($M = 5.33$) and had a Kucera and Francis frequency rating of between 1 and 50 ($M = 17.52$). The familiarity rating of the words was between 500 and 620 ($M = 541.84$), the concreteness rating was between 550 and 670 ($M = 592.22$), and the imagery rating was between 490 and 659 ($M = 585.48$).

The words were presented to subjects on the computer via a projection system that reflected the images onto a mirror above subjects' eyes in the bore of the magnet. Subjects studied a total of 162 words. All 162 of these words were also presented on the source memory test, mixed in with 54 new words. The E-Prime software package (Psychology Software Tools) was used to present stimuli and collect responses.

Overview of experiments. The behavioral paradigm that we used was an exclusion test (Jacoby, 1991). Subjects were asked to study nouns; each word was encoded using either the artist encoding task, the function encoding task, or the read encoding task (the tasks are described below). During the test phase, subjects viewed all studied items in addition to new, unstudied items. On each trial, subjects were given a task cue ("Artist?", "Function?", or "Read?"), followed by a blank screen (lasting 3–7 s), followed by the test word. When the test word appeared, subjects had to indicate whether that word was studied using the targeted task; Experiment 1 and Experiment 2 used slightly different test instructions (described below). Test trials in this paradigm be divided up into three types: congruent trials, where the test word was studied using the targeted source; incongruent trials, where the test word was studied using a nontarget source; and new-item trials, where the test word did not appear at study.

Experiment 1 used single-agenda instructions. In this experiment, subjects were instructed to press one button to indicate with high confidence that the test word was studied using the targeted task, a second button to indicate with low confidence that the test word was studied using the targeted task, and a third button to indicate that the test word was not studied using the targeted task (subjects rarely used the "low-confidence yes" response, so we collapsed together "high-confidence yes" and low-confidence yes responses into a single "yes" response category when analyzing the data). Experiment 2 was identical to Experiment 1, except it used multi-agenda instructions: For each test word, subjects were instructed to press one button to indicate that the test word was studied using the targeted task, a second button to indicate that the test word was studied using a different task, and a third button to indicate that the test word was new (nonstudied). Note that, here, subjects had to discriminate between three classes of items (targeted task, different task, and new), whereas in Experiment 1 subjects only had to discriminate between two classes of items (targeted and not targeted, where "not targeted" encompassed items studied using a different task and new items).

Detailed procedure. Both Experiment 1 and Experiment 2 consisted of six runs, where each run consisted of a study phase and then a test phase that probed subjects' memory for the immediately preceding study phase. The sequence of events during a run is illustrated in Figure 1 and described below (the test instructions shown in the figure are from Experiment 1).

During the study phase, words were presented, one at a time. For each

word, subjects encoded that word using the artist task, the function task, or the read task. The artist and function tasks were adapted from Dzulki-fli and Wilding (2005) (see also Johnson et al., 1997a), and the read task was adapted from Davachi et al. (2003). For the artist task, subjects were asked to rate how easy it would be to draw each object, on a scale of 1–5 where 1 is easy and 5 is hard. For the function task, subjects were asked to think of functions for each object, and then to press a key corresponding to the number of functions they were able to generate (1–5). For the read task, subjects silently read words backwards to themselves (the words were displayed forwards, not backwards) and rated how difficult it was to do so (where 1 is easy and 5 is hard). For all tasks, subjects entered their responses on a keypad while lying in the scanner.

Each study phase consisted of 27 trials that were evenly split across the 3 encoding tasks (i.e., subjects studied 9 words using the artist task, 9 words using the function task, and 9 words using the read task). Trials were arranged in miniblocks of 3 trials, where all trials in the miniblock used the same encoding task. For each miniblock, subjects were first given a task cue (lasting 6 s) that notified them which task they would be performing. The task cue was followed by 3 words presented for 4 s each. For each word, the 4-s presentation period was broken into “study” and “response” phases as follows. For the first two seconds, subjects saw only the word and the rating scale that they would use. During the last two seconds, a small star appeared above the word, and subjects had to enter their numerical (1–5) rating using the keypad. After each study phase, subjects spent 1 min completing a basic serial response task, where numbers appeared on the screen and subjects were required to press buttons corresponding to these numbers. The test phase immediately followed the end of the serial response task.

During the test phase, subjects were presented with a mixture of all 27 items they had seen at study and 9 new words. For each word, subjects were first given a task cue that specified one of the three encoding tasks (e.g., Artist?). We will refer to this task as the targeted task. The task cue was presented for 1 s, and then a blank screen was presented for variable-length delay period, lasting 3, 5, or 7 s (for each test phase, 24 test trials had a 3-s delay, 9 trials had a 5-s delay, and 3 trials had a 7-s delay). After the delay, the test word was presented.

As mentioned above, the instructions for responding to the test word were different for Experiment 1 and Experiment 2. In Experiment 1, subjects were required to indicate whether the item was studied with the targeted task or not. In Experiment 2, subjects were required to indicate whether the item was studied using the targeted task, a different task, or was new. In both experiments, subjects had two seconds to enter their response; if they did not respond in time, the trial timed out (and “no response” was recorded in the data file). After each test trial, there was a variable-length delay (24 test trials had a 2-s delay, 9 trials had a 4-s delay, and 3 trials had a 6-s delay). During the test period, subjects switched retrieval orientations every 3 trials. That is, they received 3 trials in a row where they were asked to target one encoding task; then, for the next 3 trials, they were asked to target a different encoding task, and so on. Exactly half of the test trials were incongruent, 25% were congruent trials, and 25% were new trials. Each of the three encoding tasks served equally often as the targeted task.

fMRI data acquisition. The fMRI data were acquired on a Siemens Allegra 3 Tesla scanner at the Center for the Study of Brain, Mind, and Behavior at Princeton University. Anatomical brain images were acquired with an MP-RAGE sequence consisting of the following parameters: 176 sagittally oriented slices, repetition time (TR) = 2500 ms; echo time (TE) = 4.38 ms; voxel size = $1.0 \times 1.0 \times 1.0$ mm; flip angle = 78°; field of view (FOV) = 256 mm. Functional images were acquired with an EPI sequence where 34 sagittal slices covering the whole brain were collected every 2 s (TR length). TE = 30 ms; voxel size = $3.0 \times 3.0 \times 3.9$ mm; flip angle = 75°; FOV = 192 mm. Anatomical images were acquired at the start of the session. The main part of the experiment consisted of six functional runs. As described earlier, each run encompassed a single study phase and test phase; 292 images were collected per run.

fMRI data analysis: Preprocessing of fMRI data. We used the Analysis of Functional NeuroImages (AFNI) fMRI data analysis software package (Cox, 1996) to preprocess the data. First, all functional images were

coregistered with the first functional scan, and signal spikes were removed. A motion correction algorithm was applied to the data to remove any artifacts associated with head motion. Within each run, linear and quadratic trends were removed to remove the effects of scanner drift. No spatial smoothing was applied to the dataset.

Multivoxel pattern analysis: Overview. We used MVPA to measure activation (at test) of source-specific patterns of fMRI activity from the study phase. The MVPA approach to analyzing fMRI data involves training a pattern classifier to detect multivoxel patterns of fMRI data corresponding to particular cognitive states (for reviews, see Haynes and Rees, 2006; Norman et al., 2006). By aggregating the information that is present in multiple voxels' responses, MVPA achieves a higher level of sensitivity to the subject's cognitive state than standard mass-univariate approaches. In our study, the increased sensitivity of MVPA makes it possible to track fluctuations in activation of the targeted source and also fluctuations in source-specific recollection across single brain scans (where each scan was acquired over a 2 s period). These fluctuations, in turn, can be related to subjects' behavior.

Our MVPA analysis procedure closely resembled the procedure that we used previously by Polyn et al. (2005). The procedure was composed of three steps: First, we selected voxels to use in the classification analysis. Second, we trained a classifier to discriminate between study-phase brain patterns corresponding to subjects performing the artist, function, and read encoding tasks. Third, to measure activation of study-phase patterns, we applied the trained classifier to single (2-s) brain scans from the test phase. The output of the classifier gives us a graded index of how well the test pattern matches the artist, function, and read brain patterns from the study phase. As is typical for MVPA analyses, the analysis procedure was performed within individual subjects (i.e., voxel selection, classifier training, and classifier testing were all performed on data from the same subject).

Multivoxel pattern analysis: Details. To select voxels for the classification analysis, we ran a mass-univariate General Linear Model analysis in AFNI, and we found the 1000 voxels (across the whole brain) that were most strongly affected by the encoding task manipulation (artist vs function vs read) at study. After completing voxel selection, the functional data from these 1000 voxels were loaded into MATLAB (Mathworks) using the Princeton MVPA Toolkit (<http://www.cs.bmb.princeton.edu/mvpa>). All of the subsequent classification steps used the MVPA Toolkit. First, we z-scored the functional data separately for each voxel and each run, to ensure that we had a normalized activation value across runs. Next, we trained a simple neural network classifier (looking just at the selected voxels) to discriminate between single brain scans acquired while subjects were performing the artist, function, and read encoding tasks at study. The neural network consisted of two layers: an input layer with 1000 units (corresponding to each of the 1000 selected voxels), and an output layer with 3 units (one per encoding task). Each input unit was connected in a feedforward manner to all 3 output units; these connection weights define a function that maps between voxel activity values and encoding task. Neural network training was implemented using the backpropagation algorithm, which iteratively adjusts connection weights to minimize prediction error when mapping between inputs and outputs (Rumelhart et al., 1996). After this training process, we used the classifier to evaluate a series of test patterns (single brain scans) that had not been presented during classifier training. For most of the analyses reported here, we selected voxels and trained the classifier using all 6 runs of study-phase data, and then we applied the classifier to all 6 runs of test-phase data. We also present the results of analyses where we selected voxels and trained the classifier using data from 5 of 6 study-phase runs, and then we applied the classifier to data from the sixth (“left out”) study-phase run. Additional details regarding our MVPA analysis methods are provided in the supplemental materials, available at www.jneurosci.org, including classifier “importance maps” that graphically depict which voxels the classifier used to distinguish between the study phase task conditions.

Logic of MVPA analyses. We focused our MVPA analyses on new-item and incongruent test trials. For these analyses, we binned the task-specific classifier outputs for each trial according to the role each task played on that trial. For new-item trials, the classifier outputs were

Table 1. Summary of predictions for single- and multiagenda source monitoring

Test type	Targeted-task activation (TT – OT)	Constraint (correlation between TT and AT)	Relationship between AT and behavior
Single-agenda (Experiment 1)	High	Yes, negative correlation between TT and AT	None
Multiagenda (Experiment 2)	Low	Negative correlation between TT and AT may be curtailed by restriction in the range of TT	Positive (high levels of AT predict correct rejections)

binned according to whether a particular task was the targeted task (TT) on that trial or one of the other tasks (OT) on that trial. For incongruent-item trials, the classifier outputs were binned according to whether a particular task was the TT on that trial, the actual task (AT) that was performed on that item at study, or the other task (OT). For example, if subjects were asked Artist? but the item was originally studied using the Function task, then TT = artist, AT = function, and OT = read. As described below, the OT can be used as a baseline when measuring activity related to the targeted task and the actual task. Note that the mapping of tasks (artist, function, read) onto conditions (TT, AT, and OT) varies from trial to trial.

Our analysis procedure is founded on two key claims. The first claim is that we can measure memory targeting by looking at the difference between TT and OT activity. We hypothesized that subjects would attempt to constrain retrieval to the targeted task by performing the targeted task on the test word (Jacoby et al., 2005a, 2005b), and that subjects would be more prone to do this in the single-agenda (vs multiagenda) condition. If this is the case, we should see a selective increase in TT activity relative to OT activity, and this increase should be larger in Experiment 1 than in Experiment 2.

The second claim is that, on incongruent trials, recollection of the actual task should lead to an increase in AT activity relative to OT activity. Prior research has established that recollection of memories from a particular source is associated with activation of source-specific patterns of activity from the study phase (Nyberg et al., 2000; Wheeler et al., 2000; Vaidya et al., 2002; Wheeler and Buckner, 2003; Kahn et al., 2004; Smith et al., 2004; Johnson and Rugg, 2004, 2007; Woodruff et al., 2005). Thus, in our exclusion paradigm, we would expect strong recollection of the actual source on incongruent trials to be associated with strong AT activity (relative to OT activity). Likewise, weak recollection of the actual source should be associated with weak AT activity (relative to OT activity). Note that our MVPA approach to dissociating targeted-task versus actual-task activity only works for incongruent trials. On congruent trials, the targeted task and the actual task are the same, so there is no way to tease apart activity relating to targeting versus recollection of the actual source; as such, we did not include congruent trials in our analysis. Also, we acknowledge that, in principle, other factors besides actual-task recollection could affect AT activity. For example, subjects might enact a strategy of performing all three tasks at test, to see which one fits best with the test word. A key point in this regard is that nonselectively performing all three tasks will affect TT, AT, and OT activity equally, so this strategy cannot be used to explain differences between TT and AT activity (on the one hand) and OT activity (on the other).

Importantly, the idea that actual-task activity indexes recollection makes it possible to assess the relationship between targeted-task activity and recollection of the actual task. If targeting of one task reduces recollection of other tasks, then – across incongruent trials – high levels of TT activity (indicating strong constraint) should be associated with low levels of AT activity (indicating low recollection of the actual task), resulting in a negative correlation between TT and AT activity. We expected that this negative correlation would be easier to observe in Experiment 1 than in Experiment 2, insofar as we expected TT activity to be lower overall in Experiment 2; this (anticipated) restriction in the range of TT should reduce the size of the correlation.

The link between actual-task activity and recollection also makes it possible to determine whether recollection of the actual (nontarget) task is differentially used in single- and multiagenda source monitoring. The key idea here is that AT recollection is not, on its own, sufficient to trigger a correct rejection. AT recollection needs to happen early in the trial (otherwise, subjects will not be able to respond before the 2-s deadline)

and, more importantly, subjects need to attend to this early trial recollection to benefit from it (Johnson and Raye, 2000). We can measure how strongly subjects are attending to AT recollection by measuring the relationship between early trial AT activity and behavioral accuracy. If subjects are attending to AT recollection, high levels of AT activity early in the trial (indicating that AT recollection occurred, and that it occurred early enough to influence responding) should be associated with increased correct rejections. If subjects are not attending to AT recollection, this relationship between early trial AT activity and correct rejections should be absent. We expected that subjects would devote more scrutiny to AT recollection in the multiagenda test than in the single-agenda test; as such, we predicted that the relationship between early trial AT activity and behavior would be stronger in Experiment 2 than in Experiment 1. Table 1 presents a summary of the predictions described above.

Results

Classification of task-related activity during the study phase

All of our MVPA analyses depend on the idea that we can train a classifier to successfully detect patterns of brain activity associated with performing the artist, function, and read encoding tasks. To assess our ability to discriminate between task-specific encoding states, we trained the classifier on study-phase data from 5 of 6 scanner runs, and measured the classifier's ability to correctly predict (for each individual scan) which encoding task the subject was performing on the remaining study run. The correspondence between the classifier's predictions and the actual encoding task (as indexed by correlation) was significantly above chance for each individual subject. The average percentage correct classification of individual brain volumes (chance = 33.3%) was 79.1% for Experiment 1 (SEM = 3.7%) and 73.9% for Experiment 2 (SEM = 3.0%); classification accuracy was not significantly different across experiments, $t_{(21)} = 1.10$, $p > 0.05$. For further details on study-phase classification, see the supplemental materials, available at www.jneurosci.org.

Classification of task-related activity at test

For all subsequent analyses, the classifier was trained on study-phase data from all 6 scanner runs, and was applied to test-phase data from all 6 scanner runs. For each new-item and incongruent test trial, we measured the activity of each classifier output unit (artist, function, and read) for 7 successive scans (lasting 2 s each), starting with the scan when the test word was presented. As discussed earlier, classifier outputs from new-item and incongruent-item trials were binned according to whether that task was the targeted task, the actual task performed on the item at study, or the other task.

For all of the results presented below, dependent measures (e.g., classifier output for a particular condition) were computed separately for each individual subject. Figures and tables show the mean and SEM (across subjects) of these per-subject measures. We used two-tailed t tests (applied to these per-subject measures) to assess whether effects were reliable across subjects.

Analysis of targeted-task activity

To evaluate our prediction that the amount of targeted task activity would be higher in Experiment 1 versus Experiment 2, we

computed the average amount of TT activity in both studies. Figure 2*A* shows the event-related classifier output averages for both experiments. The left side of the figure plots average classifier output for TT and OT on new-item trials, and the right side of the figure plots average classifier output for TT, AT, and OT on incongruent trials. Figure 2*B* plots baseline-corrected targeted-task activity (TT–OT) for new-item trials and incongruent trials, as a function of Experiment (1 vs 2). For all of the subplots in Figure 2, classifier output is shown for 7 successive scans, starting with the scan when the test word was presented.

In Experiment 1, for both new-item and incongruent trials, TT activity was significantly higher than OT activity at multiple time points. This provides strong evidence that subjects were activating the TT representation at test. The supplemental materials, available at www.jneurosci.org, contain further analyses exploring the timing of TT activity. These additional analyses, which control for “spill-over” of TT activity from the preceding trial, demonstrate that TT activity was triggered by the test word, as opposed to the task cue that preceded the test word; these timing results fit with the idea, mentioned earlier, that TT activity reflects subjects performing the targeted task on the test word. In Experiment 2, the TT–OT difference was also significant for some time points, but numerically the TT–OT difference scores were smaller in Experiment 2 than in Experiment 1. When we directly compared TT–OT difference scores across experiments, we found that the difference was significant for incongruent trials at time points 3 and 5, and the difference was significant for new-item trials at time point 4; when we combined new-item and incongruent trials, the cross-experiment difference in TT–OT was significant at time points 3, 4, and 5.

Analysis of the relationship between targeted-task and actual-task activity

According to the encoding specificity principle, TT activity should lead to reduced recollection of AT information on incongruent trials. As a first-pass measure, we compared the overall level of baseline-corrected AT activity (i.e., AT–OT) in the two experiments; as discussed above, this difference score provides an index of the degree of AT recollection that is taking place on incongruent trials. If TT activity suppresses AT recollection, then AT activity should be lower in Experiment 1 (where TT activity was relatively high) than in Experiment 2 (where TT activity was relatively low). Contrary to this prediction, we found that AT activity rose significantly above baseline in both experiments, and that the level of baseline-corrected AT activity was virtually identical across experiments (Fig. 2*B*, bottom).

Figure 2. Event-related averages of classifier output in Experiment 1 and Experiment 2. Classifier output is shown for 7 successive scans, starting with the scan when the test word was presented. The left side of the figure shows classifier output for new-item trials, and the right side of the figure shows classifier output for incongruent-item trials. The graphs in *A* show raw classifier outputs, and the graphs in *B* show difference scores (TT–OT and AT–OT). Error bars in *B* indicate the SEM (across subjects) of the difference score. Individual bars marked with asterisks are significantly different from zero; pairs of bars marked with asterisks are significantly different from each other. * $p < 0.05$; ** $p < 0.01$.

To further investigate the relationship between TT and AT activity, we ran a more sensitive within-subjects analysis where we correlated (across incongruent trials) the level of TT activity with the level of AT activity. If TT activity suppresses AT recollection, then we would expect to see a negative correlation within individual subjects (assuming that there is adequate across-trial variability in TT activity). The TT–AT correlation was computed separately for each time point (scan) in the trial, starting with the scan when the test word was presented (e.g., we correlated TT activity at time point 1 with AT activity at time point 1; we correlated TT activity at time point 2 with AT activity at time point 2; and so on).

Figure 3. Correlation between TT and AT classifier activity (across incongruent trials) in Experiment 1 and Experiment 2. Correlation values were computed separately for each time point (scan) in the trial, starting with the scan when the test word was presented. TT–AT correlation values were corrected by subtracting out the corresponding TT–OT correlation value; this corrected score indicates whether TT activity was more strongly correlated with AT activity than OT activity. Error bars indicate the SEM (across subjects) of the corrected correlation measure. Individual bars marked with asterisks are significantly different from zero. * $p < 0.05$; ** $p < 0.01$.

One complicating factor in this analysis is that classifiers can register a negative correlation between cognitive states even if (at a “process” level) the cognitive states are not related to each other. Intuitively, the more that one pattern is present in the brain, the less any other pattern will be present. With the neural network classifiers that we are using, the process is not completely zero-sum (i.e., it is possible to increase one classifier output without reducing other outputs), but we commonly observe some degree of negative correlation. To deal with this issue, we also computed the correlation between the TT and OT classifier outputs; then we subtracted out the TT–OT correlation from the TT–AT correlation. This measure factors out the “baseline” level of negative correlation (which should apply equally to TT–OT and TT–AT) and makes it possible to test whether TT activity is more negatively correlated with AT activity than with activity of the third (irrelevant) task.

Figure 3 shows the average “corrected correlation value” (across subjects) for each time point, for both Experiment 1 and Experiment 2. After correcting for the TT–OT correlation, there was a significant negative correlation between TT and AT activity at multiple time points (4, 5, and 6) in Experiment 1. In contrast, the correlation between TT and AT activity was not significant for any time point in Experiment 2. These results fit with our prediction that, when subjects attempt to constrain recall by activating the targeted task, TT activity will reduce recall of memories from other sources. The lack of a significant correlation for Experiment 2 can be explained in terms of the lower overall level of (baseline-corrected) TT activity in that experiment, which effectively restricts the range of TT and squelches the correlation. Importantly, the correlational nature of these results prevents us from making strong causal inferences. The observed negative correlation in Experiment 1 could be caused either by TT activity blocking AT recollection, or by AT recollection displacing TT activity (intuitively, strong recollection of the actual task will make it difficult to focus on the targeted task). The two possibilities are not mutually exclusive, and it seems likely that both of these situations occur to some extent.

Table 2. Experiment 1: Mean proportions of trials where subjects responded “yes,” responded “no,” or failed to respond in time as a function of trial type (congruent, incongruent, and new)

	Yes	No	Failed to respond
Congruent	0.69 (0.06)	0.15 (0.03)	0.16 (0.06)
Incongruent	0.09 (0.03)	0.77 (0.05)	0.14 (0.05)
New	0.06 (0.02)	0.83 (0.04)	0.12 (0.04)

Numbers in parentheses indicate SEM.

Table 3. Experiment 2: Mean proportion of trials where subjects responded “same task,” responded “different task,” responded “new,” or failed to respond in time as a function of trial type (congruent, incongruent, and new)

	Same task	Different task	New	Failed to respond
Congruent	0.50 (0.05)	0.11 (0.04)	0.05 (0.02)	0.35 (0.07)
Incongruent	0.07 (0.02)	0.48 (0.05)	0.08 (0.02)	0.37 (0.08)
New	0.01 (0.01)	0.07 (0.04)	0.57 (0.07)	0.35 (0.07)

Numbers in parentheses indicate SEM.

Analysis of the relationship between actual-task activity and behavior

The finding that AT activity was above-baseline in both experiments allows us to look at how AT recollection affected behavior in Experiment 1 versus Experiment 2; as discussed earlier, we can use the relationship between AT activity and behavior on incongruent trials to assess the weight that subjects are giving to AT recollection when making source memory decisions. Behavioral results for Experiments 1 and 2 are presented in Tables 2 and 3, respectively (see the supplemental materials, available at www.jneurosci.org, for additional behavioral analyses). As in previous comparisons of single-agenda versus multiagenda tests (Lindsay and Johnson, 1989), false alarms were higher in the single-agenda experiment; this trend was significant for new items, $t_{(21)} = 2.28$, $p = 0.03$, but not for incongruent items, $t_{(21)} = 0.56$, $p > 0.05$. To get an overall sense of the relationship between AT activity and behavior, we plotted event-related averages of baseline-corrected AT scores (i.e., AT–OT) for correct rejections and errors, in both Experiment 1 and Experiment 2 (note that both incorrect responses and failures to respond in time were counted as errors). The results of these analyses are shown in Figure 4A.

We also used an area-under-the-receiver-operating-characteristic-curve (AUC) measure (Fawcett, 2006) to sensitively index how well AT activity discriminates between correct rejection and error trials. Specifically, the AUC measure indexes the overlap between the observed distributions of AT activity scores associated with correct rejections versus errors; AUC provides extra information (beyond looking at means and SDs) because it factors in the entire shape of the distribution. The AUC analysis was run separately for each time point (scan) in the trial, starting with the scan when the test word was presented. AUC scores range from 0 to 1, where 0.5 indicates chance discrimination. AUC scores >0.5 indicate that AT activity was associated with increased correct rejections, and AUC scores <0.5 indicate that AT activity was associated with increased errors. If subjects are using AT recollection when making their source memory judgments, AUC scores should be >0.5 . Note that we used AT (alone) instead of AT–OT as our trial-by-trial measure of recollection when computing AUC. Subtracting out OT is valuable for demonstrating that (on average) the actual task is activated more strongly than the other task. However, when measuring recol-

lection on a trial-by-trial basis, subtracting out OT adds noise relative to using AT alone (since the proportion of OT variance that is not shared with AT is large, relative to the proportion of variance that is shared with AT). The AUC scores for Experiments 1 and 2 are shown in Figure 4B.

The results of these analyses show clear differences between Experiment 1 and Experiment 2. In Experiment 1, there was an overall trend for AT activity to be negatively associated with behavioral accuracy: For all but one time point, AT–OT was numerically higher for errors than for correct rejections; this trend was significant at the end of the trial, at time point 6. In contrast, in Experiment 2, early trial AT activity was positively associated with behavioral accuracy: AT–OT was significantly higher for correct rejections than errors at time point 2, and the AUC measure was significantly >0.5 at this time point. When we directly compared the AUC scores from the two experiments, we found that AUC scores were significantly higher for Experiment 2 versus 1 (indicating a more positive relationship between AT activity and correct rejections) at time points 2, 6, and 7.

To summarize, our key prediction regarding the relationship between AT activity and behavior was confirmed: Early trial AT activity was associated with correct rejections in Experiment 2 but not Experiment 1. The other main finding from this analysis, the relationship between late-trial AT and errors in Experiment 1, was unexpected, and merits further discussion. One possible interpretation of this result is that subjects in Experiment 1 were treating recollection of any information (even AT information) as evidence for the targeted task; for an example of how subjects can misattribute retrieved information from one source as evidence for another source, see Henkel et al. (2000). However, the timing of the effect in Experiment 1 argues against this interpretation: If AT recollection were actually causing errors, then we would expect to see this effect early in the trial, but the association between AT activity and errors was only present late in the trial (~ 10 – 12 s after stimulus onset). The timing of this effect suggests that increased AT activity on error trials in Experiment 1 may reflect postdecisional processing (i.e., subjects recalling and thinking about the actual task after they made an error) as opposed to predecisional processing; informally, subjects often reported that they would respond yes to an item on an incongruent trial and then, immediately afterward, they would realize that the item was from the wrong source (Van Zandt and Maldonado-Molina, 2004). Errors in both experiments may be attributable to factors that are “invisible” to the classifier (e.g., item familiarity) as opposed to task-specific activity (for additional discussion of this point, see the Limitations of our analysis procedure section below).

Figure 4. Analyses of the relationship between actual-task activity and behavior on incongruent trials in Experiments 1 and 2. **A**, shows event-related averages of baseline-corrected AT activity (AT–OT) for Experiments 1 and 2, split by whether subjects made a correct rejection (CR) or an error (ERR). In **B**, we used AUC measure (Fawcett, 2006) to index how well AT activity discriminated between correct rejection and error trials. AUC scores range from 0 to 1, where 0.5 indicates chance discrimination. AUC scores >0.5 indicate that AT activity was associated with increased correct rejections, and AUC scores <0.5 indicate that AT activity was associated with increased errors. Error bars in **A** indicate the SEM (across subjects) of the CR versus ERR difference for that time point (note that CR and ERR bars were not individually tested vs zero). Error bars in **B** indicate the SEM (across subjects) of the AUC score. In **A**, pairs of bars marked with asterisks are significantly different from each other. In **B**, individual bars marked with asterisks are significantly different from chance (0.5), and pairs of bars marked with asterisks are significantly different from each other. $*p < 0.05$; $**p < 0.01$.

Discussion

The goal of this study was to use neural data to gain psychological insight into how subjects make source memory judgments when they are asked to consider one source (single-agenda) versus when they are asked to consider multiple sources (multiagenda). Our first prediction was that subjects would be more likely to perform the targeted encoding task at test given single-agenda versus multiagenda instructions. Our MVPA results support this prediction: Targeted-task activation was significantly higher (relative to baseline) in Experiment 1 than in Experiment 2. We also hypothesized that activation of the targeted task would be associated with reduced recollection of the actual task on incongruent trials. Support for this claim was mixed: The level of actual-task activation (relative to baseline) was similar across experiments despite the difference in targeted-task activation. However, a more sensitive within-subjects analysis revealed that TT and AT activity were negatively correlated within individual subjects (Fig. 3). Our final prediction was that subjects would make better use of AT recollection on a multiagenda task, compared with a single-agenda task. Our results support this prediction: In Experiment 2, high levels of AT activation were associated with increased correct rejections on incongruent trials, but this relationship was not present in Experiment 1 (despite similar overall

levels of AT activation). To our knowledge, this is the first demonstration that subjects can retrieve diagnostic source information but nonetheless fail to use this information when making their source judgments.

The key methodological innovation underlying these findings was our use of pattern classifiers to track the appearance of task-specific activity during retrieval. As discussed earlier, MVPA increases sensitivity to the comings and goings of cognitive states by aggregating the information that is present in multiple voxels. This increase in sensitivity allowed us to derive meaningful measures of TT, AT and OT activity for each incongruent trial. Importantly, these MVPA analyses should be viewed as complementing (not replacing) standard voxel-based General Linear Model analyses. While MVPA is useful for addressing questions about what information is present in the subject's head at a particular point in time, whole-brain MVPA analyses are less useful for mapping out which brain regions are involved in particular cognitive processes (for a discussion of pitfalls associated with using whole-brain MVPA for brain mapping, see Norman et al., 2006). In the supplemental materials, available at www.jneurosci.org, we present voxel-based analyses exploring which brain regions discriminate between different tasks at study and which brain regions discriminate between correct versus incorrect responses to incongruent items at test.

Relationship to previous neural studies of agenda-dependent memory

Our results add to the growing body of neural evidence supporting agenda-dependent memory, the idea that subjects' goals at the time of retrieval can impact what information comes to mind and how subjects use this information (Mitchell et al., 2008; for examples of relevant neuroimaging studies, see Johnson et al., 1997b; Ranganath et al., 2000; Rugg and Wilding, 2000; Dobbins and Wagner, 2005; Dobbins and Han, 2006; for reviews of relevant studies, see Rugg, 2004; Simons, 2009). To our knowledge, there has only been one previous neuroimaging study that directly compared single-agenda to multiagenda source monitoring: Raye et al. (2000) (Experiment 1C) compared a single-agenda test ("Was the item studied as a picture?") to a multiagenda test ("Was the item studied as a picture or studied auditorily?") and found differences in frontal activity on the two types of tests. This finding indicates that processing is different in the two conditions but it does not indicate whether the difference relates to memory cuing or to the evaluation of retrieved information (or both).

While direct comparisons of single-agenda and multiagenda tests are scarce, there have been numerous imaging studies that speak to our hypotheses (set forth in the Introduction) about how subjects approach single-agenda tests. For example, a recent study by Woodruff et al. (2006) used a single-agenda source memory test and, like our study, found that subjects activate information relating to the targeted source. In Woodruff et al. (2006), subjects studied picture and word stimuli mixed together. At test, subjects were asked to target items from a particular source (e.g., pictures). For each test item, subjects were asked to say "yes" to items that were studied using the targeted source, and to say "no" otherwise. Woodruff et al. (2006) focused their fMRI analysis on new-item trials. They found that brain activity on new-item trials differed as a function of whether subjects were targeting picture versus word memories. Furthermore, brain activity patterns associated with targeting picture versus word memories were similar to brain activity patterns associated with studying pictures versus words, respectively (for a similar result, see Hornberger et al., 2006).

Additional relevant evidence comes from single-agenda studies that have compared ERPs on congruent and incongruent trials. Spe-

cifically, these studies have looked at the effect of retrieval orientation on the parietal old/new ERP effect, an ERP correlate of recollection (for discussion of this ERP effect, see Rugg et al., 2000; Rugg and Curran, 2007). Many of these studies have found that the parietal old/new effect is larger for congruent than for incongruent trials, suggesting that subjects have some ability to prevent recollection of information that mismatches the targeted source (Herron and Rugg, 2003a,b; Džulkifli and Wilding, 2005; Herron and Wilding, 2005; Džulkifli et al., 2006; for similar results from a slightly different paradigm, see Dywan et al., 1998, 2001, 2002).

The fMRI and ERP studies reviewed in this section provide some support for the idea that (on single-agenda tests) subjects attempt to constrain retrieval to the targeted source: The fMRI studies found activation of the targeted source, and the ERP studies found reduced recollection of nontarget memories. However, these studies did not address the relationship between activation of the targeted source and recollection of nontarget memories. In our study, we were able to address this relationship by simultaneously measuring targeted-task activation and actual-task activation, and then correlating these measures across trials. Also, the fMRI and ERP studies reviewed above did not measure the relationship between recollection of nontarget memories and behavior. In our study, we demonstrated a significant link between actual-task activity and behavioral accuracy in Experiment 2 (but not in Experiment 1), and we used this link to argue that subjects make better use of nontarget recollection in Experiment 2 versus Experiment 1.

Limitations of our analysis procedure

Importantly, the classifier was trained to detect patterns of activity that discriminate between the three tasks. This training procedure gives the classifier the ability to detect recollection of task-specific details, but it does not give the classifier the ability to detect recollection of nondiagnostic details (i.e., details shared by all three tasks) or feelings of familiarity. For evidence that subjects are influenced by nondiagnostic forms of memory on exclusion tests, see Dobbins and McCarthy (2008). Another limitation of our analysis procedure is that it focuses on activation (at test) of patterns from the study phase. As such, the analysis procedure will not detect processes that are engaged only at test (not at study).

Future directions

Our long-term goal is to exploit the sensitivity of MVPA to examine how memory cuing and decision-making processes vary (across subject populations, and as a function of situational factors). For example, recent results from Jacoby et al. (2005b) and Velanova et al. (2007) suggest that (on single-agenda tests with tasks as sources) elderly adults are less likely than young adults to perform the targeted task at test. Also, ERP studies have identified several manipulations that affect how strongly subjects orient to the targeted task, e.g., reducing the memorability of the targeted source (Herron and Rugg, 2003a; Džulkifli et al., 2006; but see Herron and Wilding, 2005) and varying the targeted source unpredictably from trial to trial at test (vs using a blocked design) (Johnson and Rugg, 2006). We plan to explore these (and other) factors using variants of the design used here.

Conclusions

Our MVPA approach provides a new kind of evidence regarding how information is processed during memory retrieval. Using this technique, we compared retrieval processing during single-agenda (in Experiment 1) and multiagenda source monitoring (in Experiment 2). We observed that single-agenda source monitoring is associated with increased memory targeting and reduced use of retrieved

diagnostic details. Going forward, the ability to separately track targeted-task and actual-task activity should help us to develop more nuanced theories of how subjects cue memory, how cues interact with stored memory traces, how subjects make memory decisions, and how these processes go awry in subjects with memory disorders.

References

- Coltheart M (1981) The MRC psycholinguistic database. *Q J Exp Psychol A* 33A:497–505.
- Cox R (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.
- Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci U S A* 100:2157–2162.
- Dobbins IG, Han S (2006) Cue- versus probe-dependent prefrontal cortex activity during contextual remembering. *J Cogn Neurosci* 18:1439–1452.
- Dobbins IG, McCarthy D (2008) Cue framing effects in source remembering: A memory misattribution model. *Mem Cognit* 36:104–118.
- Dobbins IG, Wagner AD (2005) Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cereb Cortex* 15:1768–1778.
- Dodson CS, Johnson MK (1993) Rate of false source attributions depends on how questions are asked. *Am J Psychol* 106:541–557.
- Dywan J, Segalowitz SJ, Webster L (1998) Source monitoring: ERP evidence for greater reactivity to nontarget information in older adults. *Brain Cogn* 36:390–430.
- Dywan J, Segalowitz SJ, Webster L, Hendry K, Harding J (2001) Event-related potential evidence for age-related differences in attentional allocation during a source monitoring task. *Dev Neuropsychol* 19:99–120.
- Dywan J, Segalowitz S, Arsenault A (2002) Electrophysiological response during source memory decisions in older and younger adults. *Brain Cogn* 49:322–340.
- Dzulkifli MA, Wilding EL (2005) Electrophysiological indices of strategic episodic retrieval processing. *Neuropsychologia* 43:1152–1162.
- Dzulkifli MA, Herron JE, Wilding EL (2006) Memory retrieval processing: neural indices of processes supporting episodic retrieval. *Neuropsychologia* 44:1120–1130.
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874.
- Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7:523–534.
- Henkel LA, Franklin N, Johnson MK (2000) Cross-modal source monitoring confusions between perceived and imagined events. *J Exp Psychol Learn Mem Cogn* 26:321–335.
- Herron JE, Rugg MD (2003a) Strategic influences on recollection in the exclusion task: electrophysiological evidence. *Psychon Bull Rev* 10:703–710.
- Herron JE, Rugg MD (2003b) Retrieval orientation and the control of recollection. *J Cogn Neurosci* 15:843–854.
- Herron JE, Wilding EL (2005) An electrophysiological investigation of factors facilitating strategic recollection. *J Cogn Neurosci* 17:777–787.
- Hornberger M, Rugg MD, Henson RN (2006) fMRI correlates of retrieval orientation. *Neuropsychologia* 44:1425–1436.
- Jacoby LL (1991) A process dissociation framework: separating automatic from intentional uses of memory. *J Mem Lang* 30:513–541.
- Jacoby LL, Shimizu Y, Daniels KA, Rhodes MG (2005a) Modes of cognitive control in recognition and source memory: depth of retrieval. *Psychon Bull Rev* 12:852–857.
- Jacoby LL, Shimizu Y, Velanova K, Rhodes MG (2005b) Age differences in depth of retrieval: memory for foils. *J Mem Lang* 52:493–504.
- Johnson JD, Rugg MD (2006) Electrophysiological correlates of retrieval processing: effects of consistent versus inconsistent retrieval demands. *J Cogn Neurosci* 18:1531–1544.
- Johnson JD, Rugg MD (2007) Recollection and the reinstatement of encoding-related cortical activity. *Cereb Cortex* 17:2507–2515.
- Johnson MK, Raye CL (2000) Cognitive and brain mechanisms of false memories and beliefs. In: *Memory, brain, and belief* (Schacter DL, Scarry E, eds), pp 35–86. Cambridge, MA: Harvard UP.
- Johnson MK, Hashtroudi S, Lindsay DS (1993) Source monitoring. *Psychol Bull* 114:3–28.
- Johnson MK, Kounios J, Nölde SF (1997a) Electrophysiological brain activity and memory source monitoring. *Neuroreport* 8:1317–1320.
- Johnson MK, Nölde SF, Mather M, Kounios J, Schacter DL, Curran T (1997b) The similarity of brain activity associated with true and false recognition memory depends on test format. *Psychol Sci* 8:250–257.
- Kahn I, Davachi L, Wagner AD (2004) Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *J Neurosci* 24:4172–4180.
- Lindsay DS, Johnson MK (1989) The eyewitness suggestibility effect and memory for source. *Mem Cognit* 17:349–358.
- Mitchell KJ, Raye CL, McGuire JT, Frankel H, Greene EJ, Johnson MK (2008) Neuroimaging evidence for agenda-dependent monitoring of different features during short-term source memory tests. *J Exp Psychol Learn Mem Cogn* 34:780–790.
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430.
- Nyberg L, Habib R, McIntosh AR, Tulving E (2000) Reactivation of encoding-related brain activity during memory retrieval. *Proc Natl Acad Sci U S A* 97:11120–11124.
- Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966.
- Ranganath C, Johnson MK, D'Esposito M (2000) Left anterior prefrontal activation increases with demands to recall specific perceptual information. *J Neurosci* 20:RC108.
- Raye CL, Johnson MK, Mitchell KJ, Nölde SF, D'Esposito M (2000) fMRI investigations of left and right PFC contributions to episodic remembering. *Psychobiology* 28:197–206.
- Rugg MD (2004) Retrieval processing in human memory: electrophysiological and fMRI evidence. In: *The cognitive neurosciences*, 3rd Ed (Gazzaniga M, ed), Cambridge, MA: MIT.
- Rugg MD, Curran T (2007) Event-related potentials and recognition memory. *Trends Cogn Sci* 11:251–257.
- Rugg MD, Wilding EL (2000) Retrieval processing and episodic memory. *Trends Cogn Sci* 4:108–115.
- Rugg MD, Allan K, Birch CS (2000) Electrophysiological evidence for the modulation of retrieval orientation by depth of study processing. *J Cogn Neurosci* 12:664–678.
- Rumelhart D, Durbin R, Golden R, and Chauvin Y (1996) Backpropagation: the basic theory. In: *Backpropagation: theory, architectures, and applications* (Chauvin Y, Rumelhart D, eds), pp 1–34. Mahwah, NJ: Erlbaum.
- Simons JS (2009) Constraints on cognitive theory from neuroimaging studies of source memory. In: *Neuroimaging of human memory: linking cognitive process to neural systems* (Roesler F, Ranganath C, Roder B, Kluge RH, eds), pp 403–424. New York: Oxford UP, in press.
- Smith AP, Henson RN, Dolan RJ, Rugg MD (2004) fMRI correlates of the episodic retrieval of emotional contexts. *Neuroimage* 22:868–878.
- Tulving E, Thomson DM (1973) Encoding specificity and retrieval processes in episodic memory. *Psychol Rev* 80:352–373.
- Vaidya CJ, Zhao M, Desmond JE, Gabrieli JD (2002) Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. *Neuropsychologia* 40:2136–2143.
- Van Zandt T, Maldonado-Molina MM (2004) Response reversals in recognition memory. *J Exp Psychol Learn Mem Cogn* 30:1147–1166.
- Velanova K, Lustig C, Jacoby LL, Buckner RL (2007) Evidence for frontally mediated controlled processing differences in older adults. *Cereb Cortex* 17:1033–1046.
- Wheeler ME, Buckner RL (2003) Functional dissociation among components of remembering: control, perceived oldness, and content. *J Neurosci* 23:3869–3880.
- Wheeler ME, Buckner RL (2004) Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337–1349.
- Wheeler ME, Petersen SE, Buckner RL (2000) Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc Natl Acad Sci U S A* 97:11125–11129.
- Wilson M (1988) The MRC psycholinguistic database: machine readable dictionary, Version 20. *Behav Res Methods Instrum Comput* 20:6–11.
- Woodruff CC, Johnson JD, Uncapher MR, Rugg MD (2005) Content-specificity of the neural correlates of recollection. *Neuropsychologia* 43:1022–1032.
- Woodruff CC, Uncapher MR, Rugg MD (2006) Neural correlates of differential retrieval orientation: Sustained and item-related components. *Neuropsychologia* 44:3000–3010.
- Zaragoza MS, Koshmider JW 3rd (1989) Misled subjects may know more than their performance implies. *J Exp Psychol Learn Mem Cogn* 15:246–255.