**Motivated Memories: Effects of reward and recollection in the core recollection network and beyond**

Rachael L. Elward, Kaia L. Vilberg, & Michael D. Rugg

Center for Vital Longevity and School of Behavioral and Brain Sciences,

1600 Viceroy Drive, Suite 800, Dallas, TX 75235

Telephone: 972-883-3781.

Fax: 972-883-3250.

Running Title: Recollection and monetary reward

Keywords: Retrieval, fMRI, Salience, Angular Gyrus, Hippocampus, Striatum

Correspondence: Rachael L. Elward, Center for Vital Longevity, 1600 Viceroy Drive, Suite 800, Dallas, TX 75235

Email: Rachael.elward@utdallas.edu

**Abstract**

fMRI was employed to assess whether the neural correlates of accurate source memory are modulated by the reward value of recollected information. Study items comprised pictures of objects, each paired with a depiction of one of two coins. The reward value of the coins ($2.00 vs. $0.02) was disclosed after study. At test, a source memory procedure was employed in which subjects discriminated between studied and unstudied objects and, for objects judged studied, indicated the identity of the coin paired with the object at study. Correct judgments earned a reward corresponding to the value of the coin, whereas incorrect judgments were penalized. No regions were identified where the magnitude of recollection effects was modulated by reward. Exclusive effects of source accuracy were evident in the hippocampus. Different striatal sub-regions demonstrated exclusive recollection effects, exclusive reward effects and overlap between the two effects. The left angular gyrus and medial prefrontal cortex were additively responsive to source accuracy and the reward. The findings suggest that reward value and recollection success are conjointly but independently represented in at least two cortical regions, and that striatal retrieval success effects cannot be accounted for in terms of a single construct, such as goal satisfaction.

Episodic retrieval is frequently engaged in service of a specific retrieval goal. You may remember reading that rapper 'Vanilla Ice' is from Dallas, Texas, but this information probably has little motivational significance for your daily life. However, if you should take part in a quiz with a large prize, and the quiz master asks a question about Vanilla Ice, then any information that is associated with him (i.e. that he was born in Dallas) may become significant to you. How do motivational factors, such as the prospect of a large monetary reward, impact the processes supporting memory retrieval?

To date, studies investigating the effects of motivational significance on the neural correlates of long term memory have almost invariably employed designs in which study items were associated with different levels of reward at the time they were encoded (e.g. Adcock et al. 2006; Wittman 2008; Wolosin et al. 2012; Murty and Adcock 2013). These studies have demonstrated that when study items are associated with the prospect of a large rather than a small monetary reward at retrieval, encoding-related activity in the hippocampus is facilitated, and subsequent memory performance improves. While shedding valuable light on the question of how reward modulates neural activity at the time of encoding, the influence of reward on encoding efficacy in such studies makes it difficult to cleanly assess reward effects at retrieval. Following the example given in the introductory paragraph above, in the present study we investigated how the reward value associated with an accurate episodic memory judgment impacts the neural correlates of successful retrieval when the value of the reward was not known at the time of encoding.

A recent review has highlighted the ubiquity with which the striatum, a brain region strongly associated with reward, demonstrates enhanced activity when memory retrieval is successful, even in the absence of reward contingencies (Scimeca and Badre 2012; see also

Spaniol et al. 2009; Kim 2011; Han et al. 2010). Scimeca and Badre (2012) outlined a number of different, and not mutually exclusive, hypotheses to account for the involvement of the striatum in retrieval. Among these are the possibilities that the striatum modulates the re-encoding of retrieved information in accordance with its likely future significance, and that the region plays a role in control processes that regulate the entry of retrieved information into working memory.

To our knowledge, the sole prior study to investigate the effect of reward on the neural correlates of successful retrieval from long-term memory is that of Han et al. (2010). They associated monetary reward either with accurate 'old' (hit) or 'new' (correction rejection) recognition memory judgments. When hits were rewarded, activity in bilateral caudate was greater for items judged old relative to items judged new. Crucially, the reverse pattern was obtained when correct rejections were rewarded. In both cases, the effects of reward were independent of the accuracy of the selected response. Thus, the striatal 'retrieval success' effects observed in the former condition were actually tracking the reward value of the associated memory judgment. In light of these findings, Han et al. (2010) proposed that retrieval-related activity in the striatum is sensitive to what they referred to as 'perceived goal attainment' rather than to successful memory retrieval. They further suggested that the striatal retrieval success effects often evident when memory judgments are unrewarded reflect the tendency of participants to treat the identification of previously studied items as the goal of the memory test.

Like Han et al. (2010), the present study investigated how monetary reward impacts retrieval-related neural activity. Unlike in the prior study, however, here reward value was signaled not by study status (old vs. new), but by a contextual feature that had been paired with items at study. Unlike in the previous study, therefore, here reward is associated not with successful versus unsuccessful retrieval, but with the content of retrieved information. Thus,

successful retrieval could be associated with either high or low reward magnitude, the level of the reward depending on the recollected content. Additionally, retrieval of contextual features of a study episode depends much more heavily upon recollection (one of the two putative bases of recognition memory; see e.g. Eichenbaum et al. 2007; Yonelinas 2002; Wixted et al. 2010) than it does upon familiarity (the other putative basis). Thus, any effects of reward on the neural correlates of retrieval success in the present study can reasonably securely be ascribed to its effect on neural activity selectively associated with recollection. Much of this activity is found within a set of brain regions that are sometimes referred to as the 'core' or 'general' recollection network (Johnson and Rugg 2007; Hayama et al. 2012; for reviews see Kim 2010 and Rugg and Vilberg 2013). These regions include hippocampus and parahippocampal cortex, medial prefrontal cortex (mPFC), and posterior midline and lateral parietal cortex (angular gyrus).

The present study affords the opportunity to assess whether recollection-related activity in any of the components of the core recollection network is modulated by the reward value of retrieved information. This is an especially pertinent question in the case of the mPFC which, along with its sensitivity to successful recollection, is also highly sensitive to reward anticipation and delivery (e.g. Bzdok et al. 2013). Therefore, as has been suggested in the case of the striatum (Han et al. 2010, see above), it is of interest to ask whether retrieval-related activity in the mPFC is a proxy for some more general construct such as goal attainment. If this is the case recollection effects should be supplanted, or at least overshadowed, by those of reward. The study also permits assessment of the proposal that hippocampal activity will be potentiated during retrieval of high versus low value memories by virtue of a striatally-mediated feedback loop that facilitates the re-encoding of highly adaptive memories (Scimeca and Badre 2012; see also Shohamy and Adcock 2010).

**Materials and Methods**

Participants. Thirty right-handed English-speaking adults aged between 18-29 years participated in the experiment. All were free from neurological and psychiatric disorder according to self-report. Before participating, informed consent was obtained and participants were made aware that, in addition to receiving $30 per hour for participation in the experiment, they could receive an additional $10 depending on the accuracy of their memory performance. The experiment was approved by the UTD Institutional Review Board. Six participants were excluded from the analyses described below because they demonstrated chance or near-chance source memory (accuracy $< 0.54$)[1], three participants were excluded because of insufficient trial numbers (fewer than 8) in one or more of the critical response categories, and one subject was excluded because of excessive movement.

Experimental materials and procedure. Study and test items were colored pictures of objects presented within a grey frame. Three lists of 80 pictures were created so that within each list an equal number of pictures denoted objects typically encountered indoors (e.g. a bed) or outdoors (e.g. a hot-air balloon). During study, pictures from one of the three lists were presented in association with the depiction of a Lira coin, whereas pictures from another list were presented in association with the depiction a Deutschmark (DM; see Figure 1). Pictures from the third list served as foils (new items) during the test phase. Item order in the study and test lists was pseudo-randomized so that no more than three items from the same stimulus category were presented consecutively. Each list began with two filler items that were not included in the analyses. Across participants, pictures were fully counterbalanced such that every object was presented equally frequently in association with each of the two coins at study and as a foil at test.

Study and test phases were conducted inside the MRI scanner. Functional data were acquired at test only. Each study trial began with the presentation of a red fixation cross for 500ms. This was followed by a picture paired with one of the two coins (see Figure 1, study); the coin and item images were presented together for 4s, and followed by a white fixation cross that was presented for 500ms. The study task was to indicate by button press whether the object would typically be encountered indoors or outdoors. Participants were made aware before encoding that there would be a subsequent test on which their memories would be tested both for the objects, and for which of the two coins was paired with each object. They were instructed to attempt to associate the image of study object with the associated coin so as to be able maximize performance on the later test.

The test phase followed the study phase after an interval of approximately 2 minutes. During the interval (see Figure 1, Interval), participants were informed that a DM was worth about $1, whereas a Lira was worth considerably less (approximately 2000 Lira to $1). To reinforce this point, participants were informed that 3000 Lira, but only 3 DM were needed to buy a loaf of bread. The purpose of providing this information was to associate the image of the DM with high monetary value and the Lira with low value. Participants were then informed that the amount of money that they would be awarded for a correct memory judgment would differ in accordance with the value of the coins that they remember. They were informed they would receive $2.00 for each correct identification of an object paired with a DM, but only $0.02 for a correct identification of an object previously paired with a Lira. They were further informed that incorrect judgments would incur losses equivalent to the reward value had the response been correct (the employment of a penalty for incorrect responses was to encourage the adoption of a neutral response bias. An analogous pay-off structure was employed by Han et al. 2010). The

cumulative reward available per participant was capped at $10. There was no reward or loss associated with 'New' responses.

Each test trial commenced with a red fixation cross for 500ms. This was followed by the presentation of the test item for 400ms. This in turn was followed by a white fixation cross that was displayed for 3100ms, during which time participants responded. The test trials were intermixed with 129 fixation only trials. On these trials, participants were asked not to respond but to wait for the next trial to begin. The offset of the cross was followed by 250ms blank interval, leading to a stimulus-onset-asynchrony of 4.25s. Participants made a three-way response judgment, to discriminate between items that were previously presented with a DM (High Reward Response), items previously presented with a Lira (Low Reward Response), and items that were not previously presented (New). Responses were made on a scanner compatible button box with the index, middle and ring fingers of the right hand. For half of the participants, 'new' response were made with the index finger, and for the remainder, the new response was assigned to the ring finger. In each case, the mapping between DM and Lira responses was alternated over participants between the two other fingers. The test phase was divided into three blocks of approximately ten minutes duration. During fixation only trials, a red fixation cross was presented for 500ms followed by a white fixation cross for 3500ms and a 250ms blank interval. Prior to entering the scanner, participants undertook a practice study-test cycle. Crucially, no mention was made during practice of the relative value of the two coins, or that the coins would be associated with differential monetary rewards and losses. In addition, participants were asked at the outset of the experiment if they had seen the two coins before and if they knew anything about them. Two participants reported that they were familiar with the coins but neither was able to say which coin was more valuable than the other. It is worth noting that one of these

subjects was in any case excluded from our primary data analyses because of insufficient trial numbers.

MRI data acquisition. Blood oxygenation level-dependent (BOLD), T2*-weighted echoplanar functional images (SENSE factor 1.5, flip angle 70º, 80 x 80 matrix, FOV = 24 cm, TR = 2000 ms, TE = 30 ms) and T1-weighted anatomical images (MPRAGE sequence, 240 x 240 matrix, 1 mm isotropic voxels) were acquired using a 3T Philips Achieva MRI (Philips Medical Systems, Andover, MA, USA) scanner equipped with an 32 channel receiver head coil. Two hundred and seventy-three functional volumes were acquired during each of the three test blocks. Each volume comprised 34 slices, acquired in an ascending sequence, oriented parallel to the AC-PC line (thickness 3mm, 1mm inter-slice gap, 3mm isotropic voxels). The first 5 volumes of each scanning session were discarded to allow equilibration of tissue magnetization. For each session, mean across-volume signal intensity across was normalized across volumes to a nominal value of 100.

Data analysis. Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK), run under Matlab R2010a (The Mathworks Inc. USA) was used for fMRI data analysis. Functional images were subjected to realignment (to the mean image), slice timing correction (using the $17^{th}$ slice as the reference), reorientation, spatial normalization to a standard EPI template (based on the Montreal Neurological Institute (MNI) reference brain; Cocosco et al. 1997) and smoothing with an 8mm full-width half maximum Gaussian kernel. Each participant's structural volume was normalized to the MNI T1 template prior to averaging to create an across-subjects (N = 20) mean image. Functional analysis was performed using a General Linear Model (GLM) in which a delta (or stick) function was used to model neural activity at item onset. This function was convolved with a canonical hemodynamic

response function (HRF) to model the predicted BOLD response (Friston et al. 1995). Six event types were modeled: i) High Reward Response-Source Correct (an item previously presented with a DM, given a high reward response), ii) Low Reward Response-Source Correct (an item previously presented with a Lira, given a low reward response), iii) Low Reward Response-Source Incorrect (an item previously presented with a DM, given a low reward response), iv) High Reward Response-Source Incorrect (an item previously presented with a Lira, given a high reward response), v) correct rejections, and vi) events of no interest such as false alarms, item misses, and trials associated with multiple or omitted responses. The model also included as covariates six regressors representing motion-related variance (three for rigid-body translation and three for rotation), regressors modeling the separate scan sessions, and the across-scan mean. An AR(1) model was used to estimate and correct for non-sphericity of the error covariance (Friston et al. 2002). Subject-specific parameter estimates for the four event types of interest were taken forward to a repeated measures, 2 (source correct vs. source incorrect) x 2 (high reward response vs. low reward response) ANOVA as implemented in SPM. T and F contrasts derived from the ANOVA model were thresholded at $p < 0.001$, uncorrected, with a 23 voxel extent threshold. This cluster extent threshold gives a corrected whole-brain cluster-wise significance level of $p < .05$ as estimated using Monte-Carlo simulations implemented with the 3dClustSim function in AFNI (Van Essen et al. 1998). Coordinates of significant effects are reported in MNI space. Effects of interest are displayed on sections of the participants' mean normalized structural image or rendered onto a fiducial brain (Van Essen et al. 1998). Because of our *a priori* interest in possible reward and memory effects within the striatum (see Introduction), additional targeted analyses were conducted in this region.

**Results**

Behavioral Results

As was expected on the basis of preliminary findings, no participant reported any prior experience or knowledge of DM or Lira coins, including their historical values. The item hit rate (proportion of studied items given an old response – either Lira or DM) and the correct rejection rate were both 0.87 (SDs of 0.09 and 0.1 respectively). Source memory accuracy (calculated as the proportion of item hits associated with a correct source response) was 0.67. Participants made significantly more source correct responses to low reward items than to high reward items, 0.73 vs. 0.60 respectively, $t(19) = 3.92$, $p < 0.01$. Participants also made more false alarms to new items in association with the low reward response (Lira) than with the high reward response (DM), 0.09 vs. 0.03, respectively, $t(19) = 4.17$, $p < 0.01$.

Reaction times (RT) are illustrated in Figure 2. The data were entered into a 2 x 2 ANOVA with factors of source accuracy (correct vs. incorrect source) and reward response (high vs. low). The main effect of reward and its interaction with source accuracy were both non-significant, $F(1, 19) = 1.12$, and $< 1$, respectively. There was, however, a main effect of accuracy, $F(1, 19) = 15.14$, $p < .001$, reflecting slower RTs for incorrect source memory judgments regardless of their reward value.

fMRI Results

Regions demonstrating main effects of source accuracy and reward are illustrated in Figure 3A and listed in Table 1. Source accuracy effects were identified in, among other regions, ventral parietal and posterior cingulate cortex, mPFC, and bilateral hippocampus and adjacent MTL cortex. Reward effects were identified in bilateral insula, anterior cingulate, mPFC, posterior cingulate, left ventral parietal cortex and bilateral orbitofrontal cortex. As is evident

from Figure 3A, even at the pre-experimentally defined threshold of $p < .001$ for each contrast, accuracy and reward effects overlapped substantially in left lateral parietal cortex and mPFC. We also used the same ANOVA model to search for regions demonstrating an interaction between source accuracy and reward effects (see Methods). No regions were identified where these factors interacted.

We next addressed the question of whether effects of source accuracy and reward could be regionally dissociated. This was accomplished by exclusively masking each main effect by the alternate effect, thresholding the mask at $p < .05$ one-sided (note that the more liberal the threshold of an exclusive mask, the more conservative is the outcome of the masking procedure). Regions exclusively sensitive to source accuracy included bilateral hippocampus and amygdala, and a region of the lateral striatum in the vicinity of the putamen (See Table 2, Figure 3B). Effects exclusive to reward were identified in a region of left ventral striatum, as well as in bilateral insula, anterior cingulate/dorsal mPFC, and lateral orbitofrontal cortex bilaterally (See Table 3, Figure 3B).

Striatal source memory and reward effects: The foregoing analyses yielded clear evidence of exclusive effects of both source memory (left putamen) and reward (left nucleus accumbens) in the striatum (Figure 3B). At the thresholds that identified these effects ($p < .001$ and 21 voxels), there was however no evidence of any striatal region where the two effects overlapped. When the threshold of each contrast was lowered to $p < .01$, a cluster of 148 voxels was identified that extended from the thalamus into the right ventral striatum (Figure 4). A smaller cluster (34 voxels) was also identified in the left thalamus/ventral striatum. The peaks of the conjoint right ventral striatal effects were within a few mm of each other (source accuracy: 9, 5, -8, peak $Z = 4.0$, reward: 9, 8, -5, peak $Z = 3.56$; center of mass 9, 7, -7). Each effect survived small volume

correction (family-wise error rate, estimated according to Gaussian Random-Field theory; Worsley et al. 1996) within a 3 mm sphere centered on the ventral striatal retrieval success effect reported in the meta-analysis of Spaniol et al. (2009, co-ordinates: 10, 10, -4; see also Scimeca and Badre 2012).

<u>Response Bias:</u> Participants demonstrated a significant bias toward low reward responses. Thus, as a group they were loss-averse, preferring to opt for a potential loss of $0.02 rather than $2.00 when retrieved information was not diagnostic of the associated coin. The question thus arises whether any of the fMRI findings described above are attributable to, or modulated by, response bias. For example it might be that, regardless of accuracy, more mnemonic evidence was required to support a high than a low reward response, such that effects attributed to reward should in fact be interpreted as retrieval effects, reflecting the greater amount of retrieved information that was required before selecting a high reward response.

Because of the substantial variation that existed across participants in the strength of their response bias, we were able to address this question. We quantified bias as the difference in the proportion of accurate source judgments associated with high versus low reward. We acknowledge that a purer measure of bias would be the difference between the source attributions made to false alarms. False alarm rates were however too low to provide stable bias estimates, with several subjects never endorsing a false alarm with a high reward response. The correlation between the bias estimated from accurate source judgments and false alarms was high, however (r (18) = .588, p < .01) suggesting that the two estimates index a common process. Mean bias was 0.12, with a range of -0.06 to 0.43. We created two sub-groups by splitting the subject sample at the median (.11). The mean score for the low bias group was not significantly greater than zero (mean = 0.01, SD = 0.05), t(9)= 0.9), whereas the mean score for the high bias

group was 0.22 (SD = 0.08). The two groups did not differ with respect to item (t(18) = 1.7, n.s.) or source accuracy (t(18) = 0.9, n.s.). A mixed-model ANOVA was conducted on the RT data with factors of reward, source memory accuracy and group. No interactions with group were significant (maximum F(1, 18) = 1.6).

We used a whole-brain, between-subjects ANOVA to assess whether any of the effects of source accuracy or reward were modified by bias. In each case, we inclusively masked the relevant main effect (source accuracy or reward, thresholded at p < .001) with the interaction of the effect with group (high vs. low bias, liberally thresholded at p < .01). The analyses identified a single 24-voxel cluster in the right insula (33, 23, -11, peak Z = 4.04) where reward effects were greater in magnitude in the high-bias group. Thus, these findings provide little evidence that fMRI reward effects were a proxy for retrieval success. This impression is supported by two further observations. First, there is substantial evidence that RT co-varies with quality of mnemonic evidence and the confidence of the associated judgment (e.g. Rotello & Zeng 2008). Consistent with this evidence, inaccurate source judgments were made more slowly than accurate judgments. RTs were however equivalent for high- and low-reward responses. Second, reward effects were not identified in all recollection-sensitive regions. Notably, no effects of reward were observed in the hippocampus, where recollection-related activity has consistently been reported to co-vary with the amount and quality of retrieved contextual (source-specifying) information (Rugg et al. 2012) or, alternatively, with the correlated variable of 'memory strength' (e.g. Wais et al. 2010).


**Discussion**

Relative to inaccurate source memory judgments, accurate source judgments were associated with enhanced activity in the core recollection network[2] and the striatum. In addition, irrespective of source accuracy, test items attracting a high- rather than a low-reward response elicited enhanced activity in several regions previously implicated in reward processing, including the ventral striatum. Significantly, there was overlap between recollection- and reward-related effects in medial prefrontal cortex, angular gyrus and striatum, but no evidence of an interaction between the effects. Thus, in these regions, source accuracy and response reward value were seemingly independently represented.

As just noted, source recollection was associated with enhanced activity in the striatum, in agreement with prior findings (Scimeca and Badre 2012; Kim 2011). The recollection effects partially overlapped reward effects that were also evident in the ventral striatum. Han et al. (2010) proposed that striatal retrieval success effects reflect perceived goal attainment and conjectured that, without an external incentive, identification of studied items is motivationally more satisfying than is identification of unstudied items. Hence 'old' judgments elicit greater activity than 'new' judgments in regions responsive to goal satisfaction, such as the striatum. Han et al. (2010) further argued that this difference in the motivational significance of the two classes of memory judgment can be modulated by external reward contingencies, to such an extent in their study that retrieval success effects were entirely supplanted by reward. Whereas this account can explain the exclusive reward effects that we identified in the left ventral striatum, recollection effects in other striatal regions, including a cluster where the effects overlapped those of reward, were unaffected by reward value. Thus, the account may not generalize from the simple recognition memory judgments investigated by Han et al. (2010) to circumstances where reward is based on the identity of recollected content.

As just noted, in addition to striatal regions where reward and recollection effects overlapped, we also identified regions in the striatum where the respective effects were exclusive to only one of these two factors. A ventral striatal region in the vicinity of the left nucleus accumbens was exclusively sensitive to reward, whereas a more lateral region in the vicinity of the putamen demonstrated an exclusive recollection effect (Figure 3B). These findings suggest that striatal recollection and reward effects reflect the engagement of dissociable processes, and thus that the effects cannot be accounted for in terms of a single construct such as goal attainment (cf. Han et al. 2010; Kim 2010, 2011). Of course, this is not to say that striatal recollection effects represent a memory signal that supports mnemonic judgments (although see Scimeca and Badre 2012 for a review of neuropsychological evidence favoring this possibility). The effects could instead reflect non-mnemonic processes that are contingent on successful recollection. For example, the effects might, in part, reflect differences in the confidence of accurate and inaccurate source judgments (Cabeza et al. 2012; Schwarze et al. 2013).

In agreement with numerous prior reports, activity in the left angular gyrus was enhanced for recollected test items (for reviews see Vilberg and Rugg 2008; Cabeza at al. 2008). Angular gyrus activity was, however, also independently sensitive to reward value (Figure 3A). These findings are seemingly inconsistent with the proposal that recollection effects in the left angular gyrus reflect the region's role in 'bottom-up' attentional re-orienting (Cabeza et al. 2008, 2012). According to this proposal the left angular gyrus supports the re-orienting of attention towards recollected information in a manner analogous to that originally proposed for the right temporo-parietal junction with respect to the detection of salient sensory events (Corbetta and Shulman 2002; Corbetta et al. 2008). Attentional re-orienting is held to co-vary with the salience of recollected information (Cabeza et al. 2008), thereby accounting for the finding (Vilberg and

Rugg 2007, 2009) that rather than having an 'all-or-none character' (as was reported for the left temporo-parietal junction; Yu et al. 2012), angular gyrus recollection effects are graded according to the amount of information retrieved. In the present case, it seems reasonable to assume that the salience of information recollected on high and low reward trials differed by virtue of its association with differing levels of reward (see Ciaramelli 2011 for a similar assumption). Therefore, according to an attentional re-orienting account, angular gyrus recollection effects should have been greater for recollected items associated with the high than the low reward study context.

Whereas the present findings are arguably difficult to reconcile with the attentional re-orienting account of retrieval-related angular gyrus activity as it is currently articulated (Cabeza et al. 2012), they leave open the questions of the functional significance of recollection-related activity in the region, and why it should be sensitive to the reward value of the selected response. With respect to the latter issue, it is noteworthy that reward anticipation effects in the angular gyrus have been reported with sufficient consistency to be identified in a recent meta-analysis (Lui et al. 2011). Thus, the sensitivity of the angular gyrus to reward extends beyond the context of the present source memory test. One possibility is that the region does indeed respond to salient internal events, as suggested by Cabeza et al. (2012), but that the saliencies of different classes of event (here, successful recollection and the selection of a high reward response) are separately represented. Regardless of the present findings, however, there are other grounds for questioning the validity of an attentional re-orienting account of angular gyrus function (Vilberg and Rugg 2012; Hutchinson et al. 2012; Nelson et al. 2012). An alternative account of reward effects in this region is that they reflect engagement of some domain-specific process, such as representation of magnitude or quantity, that is supported by the region (e.g. Gobel et al. 2004,

Ischebeck et al. 2006; see also Lui et al. 2011). How such an account might be reconciled with current ideas about the role of the angular gyrus in successful recollection (e.g. Vilberg and Rugg 2008; Shimamura 2011; Binder 2009) is an important question for future work.

Unlike the angular gyrus, the mPFC has been consistently recognized as supporting reward processing, and has been assigned a variety of roles in this regard, including value representation and value-based decision-making (e.g. Roy et al. 2012; O'Doherty 2011). The mPFC has also been implicated in linking mnemonic information to response selection (Euston et al. 2012), self-directed cognition (Bzdok et al. 2013), and the representation of future events (along with other components of the 'default mode' network: Schacter et al. 2012). Whereas it has been suggested on the basis of a recent meta-analysis that reward and mnemonic processing engage ventral and dorsal aspects of the mPFC respectively (Bzdok et al. 2013; although see Andrews-Hannah et al. 2012), there is little evidence for this dissociation in the present data (see Figure 3). Indeed, as is evident in Figure 3, the mPFC region demonstrating the strongest evidence of an exclusive reward effect is in the vicinity of the dorsal anterior cingulate. As in the case of the angular gyrus, therefore, it appears that common regions of the mPFC independently represent the outcome of a recollection attempt and the reward value of the ensuing response choice. Thus, as in the case of the striatum, there was no evidence to suggest that mPFC recollection effects reflect a confound between successful retrieval and a non-mnemonic factor such as goal attainment or satisfaction.

There were some regions that demonstrated recollection effects exclusively, including the hippocampus and adjacent MTL cortex. The finding of enhanced MTL activity in association with accurate source memory replicates numerous prior reports (for reviews see Diana et al. 2007; Rugg and Vilberg 2013). At first sight the present findings are in disagreement with those

of Wolosin et al. (2012), who reported that retrieval-related hippocampal activity was modulated by the reward value of the test items. In that study, though, value was assigned to the items at the time of encoding. Thus, unlike in the present case, encoding of items was modulated by the prospect of future reward. Therefore, retrieval-related reward effects reported by Wolosin et al. (2012) likely reflect retrieval of memory representations that differed in both their strength and encoded content. The present findings suggest that when memory representations do not differ in these respects, the hippocampus can be insensitive to the reward value associated with a recollected episode. Thus, there was no evidence here for the differential engagement of hippocampally-mediated 'adaptive encoding' according to the value of the retrieved information (cf. Shohamy and Adcock 2010; Scimeca and Badre 2012).

In addition to regions demonstrating exclusive recollection effects, other regions were exclusively sensitive to reward. These included the anterior insula, dorsal mPFC/anterior cingulate, lateral orbitofrontal cortex and, as already discussed, a sub-region of the ventral striatum. All of these regions have been implicated in reward processing in numerous prior studies (see Lui 2011 for review). Additionally, the insula has been associated with loss aversion more specifically (e.g. Rudorf et al. 2012). On the assumption that participants strongly biased towards low-reward responses were more loss-aversive, this offers a possible explanation for the finding that this high-bias sub-group demonstrated larger reward-related effects than low-bias participants in the right insula. The finding that there were no regions where the effects of reward varied with response accuracy, and hence with the quality of mnemonic evidence, suggests that the present reward effects reflect processes downstream from those responsible for representing or evaluating the products of a retrieval attempt.

In conclusion, the present findings suggest that the neural correlates of successful source recollection are largely impervious to reward value, and hence to the motivational significance and salience of recollected information. Nonetheless, activity in two cortical regions sensitive to successful recollection also discriminated between responses associated with high versus low reward. Thus, recollection success and the reward value of mnemonic judgments appear to be independently represented in some components of the putative core recollection network.

**References**

Adcock A, Thangavel A, Whitfeild-Gabrieli S, Knutson B, Gabrieli JDE. 2006. Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron 50:507–517

Andrews-Hanna JP. 2012. The brain's default network and its adaptive role in internal mentation. Neuroscientist 18:251–270.

Binder JR. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767-2796

Bzdok D, Langner R, Schilbach L, Engemann DA, Laird AR, Fox PT, Eickhoff SB. 2013. Segregation of the human medial prefrontal cortex in social cognition. Front Hum Neurosci. 7:1-17

Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. 2008. The parietal cortex and episodic memory and attentional account. Nat. Rev. Neurosci. 9:613-622

Cabeza R, Ciaramelli E, Moscovitch M. 2012. Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. TICS. 16:338-352

Ciaramelli E. 2011. Testing the AtoM model: Cueing and payoff paradigms. Paper presented at 5th International Conference on Memory. York, UK (Abstract retrieved from http://www.york.ac.uk/conferences/icom5/Abstract%20Booklet_ICOM5220711.pdf , page 48).

Cocosco C, Kollokian V, Kwan R, Evans A. 1997. Brainweb Online interface to a 3D MRI simulated brain database. In: NeuroImage (Proc. of HBM'97). Vol. 5 (4, part 2/4). Academic Press, New York, p. S425.

Corbetta M & Shulman GL (2002) Control of goal-directed and stimulus driven attention in the brain. Nat Rev Neurosci, 3, 201–215.

Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. TICS, 11:378-386

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. Annu Rev Neurosci. 30:123-52.

Euston DR, Gruber AJ, McNaughton BL. 2012. The role of medial prefrontal cortex in memory and decision making. Neuron. 76:1057-1070

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp. 2:189–210.

Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J. 2002. Classical and Bayesian inference in neuroimaging: applications. Neuroimage 16:484–512.

Gobel SM, Johansen-Berg H, Berhens T, Rushworth MFS 2004. Response-selection-related parietal activation during number comparison. J. Cogn. Neurosci. 16, 1536–1551.

Han S, Huettel SA, Raposo A, Adcock A, Dobbins I. 2010. Functional significance of striatal responses during episodic decisions: Recovery or goal attainment? J Neurosci 30:4767-4775

Hayama HR, Vilberg KL, Rugg, MD. 2012. Overlap between the neural correlates of cued recall and source memory: evidence for a generic recollection network? J Cogn Neurosci 24:1127-37

Hutchinson JB, Uncapher MR, Wagner AD. 2012. Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. Learn Mem 6:343-56

Ischebeck A, Zamarian L, Siedentopf C, Koppelstätter F, Benke T, Felber S, and Delazera M. 2006. How specifically do we learn? Imaging the learning of multiplication and subtraction Neuroimage. 30:1365

Johnson JD, Rugg MD. 2007. Recollection and the reinstatement of encoding-related cortical activity. Cereb. Cortex 17:2507–2515

Kim H. 2010. Dissociating the roles of the default-mode dorsal and ventral networks in episodic memory retrieval. Neuroimage 50:1648–1657.

Kim H. 2011. Differential neural activity in the recognition of old versus new events: An activation likelihood estimation meta-analysis. Hum Brain Mapp. 34:814–836

Liu X, Hairston J, Schrier M Fan J. 2011. Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev. 35:1219-36

Murty VP, Adcock A. 2013. Enriched Encoding: Reward Motivation Organizes Cortical Networks for Hippocampal Detection of Unexpected Events. Cereb Cortex.

Nelson SM, McDermott KB, Peterson SE. 2012. In favor of a 'fractionation' view of ventral parietal cortex: comment on Cabeza et al. Trends Cogn Sci 16:399-400

O'Doherty J. 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. Ann. N.Y. Acad. Sci. 1239:118–129

Rotello CM, Zeng M. 2008. Analysis of RT distributions in the remember–know paradigm. Psychon B & Review 15:825-832

Roy M, Shohamy D, Wagner TD. 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning TICS 16:147-156

Rudorf S, Preuschoff K, Weber B. 2012. Neural Correlates of Anticipation Risk Reflect Risk Preferences. J. Neurosci. 32:16683-16692

Rugg MD, Vilberg, KL. 2013. Brain networks underlying episodic memory retrieval. Curr Opin in Neurobiol. 23:255-60

Rugg MD, Vilberg KL, Mattson JT, Yu SS, Johnson JD, Suzuki M. 2012. Item memory, context memory and the hippocampus: fMRI evidence. Neuropsychologia 50:3070-9

Schacter DL, Addis DR, Hassabia D, Martin VC, Spreng RN, Szpunar KK. 2012. The future of memory: Remembering, imagining, and the brain. Neuron. 76:677-694

Schwarze U, Bingel U, Badre D, Sommer T. 2013. Ventral striatal activity correlates with memory confidence for old- and new-responses in a difficult recognition test. PLoS ONE 8(3):e54324

Sciemeca J, Badre D. 2012. Striatal contributions to declarative memory retrieval. Neuron. 75:380-392

Shohamy D, Adcock A. 2010. Dopamine and adaptive memory. TICS 14:464-472

Shimamura AP. 2011. Episodic retrieval and the cortical binding of relational activity. Cogn Affect Behav Neurosci. 11:277-291

Spaniol J. Davidson PSR. Kim ASN. Han H. Moscoviych M. Gray CL. 2009. Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. Neuropsychologia. 47: 765–1779.

Van Essen DC, Drury HA, Joshi S, Miller MI. 1998. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. Proc. Natl. Acad. Sci. 95:788–795.

Vilberg KL, Rugg MD. 2007. Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. Neuropsychologia 45:2216–2225

Vilberg KL, Rugg MD. 2008. Memory retrieval and the parietal cortex: a review of evidence from event related fMRI. Neuropsychologia 46:1787–1799.

Vilberg KL, Rugg MD. 2009. Functional significance of retrieval-related activity in lateral parietal cortex: Evidence from fMRI and ERPs. Hum Brain Mapp. 30:1490–1501

Vilberg KL, Rugg MD. 2012. The neural correlates of recollection: Transient versus sustained fMRI effects. J Neurosci, 32:15679 –15687

Wais PE, Squire LR, Wixted JT. 2010. In search of recollection and familiarity signals in the hippocampus. J Cogn Neurosci. 22: 109-23

Wittmann BC, Schiltz K, Boehler CN, Duzel E. 2008. Mesolimbic interaction of emotional valence and reward improves memory formation. Neuropsychologia 46:1000–1008

Wixted J, Mickes L, Squire LR. 2010. Measuring Recollection and Familiarity in the Medial Temporal Lobe. Hippocampus. 20:1195-1205

Wolosin SM, Zeitamova D, Preston AR. 2012. Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. J Cognitive Neurosci 24:1532–1547

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. 1996. A unified statistical approach for determining significant signals in images of cerebral activation Hum Brain Mapp. 4:58-73

Yonelinas AP. 2002. The Nature of Recollection and Familiarity: A Review of 30 Years of Research. Journal of Memory and Language. 46:441-517

Yu SS, Johnson JD, Rugg MD. 2012. Dissociation of recollection-related neural activity in ventral lateral parietal cortex. Cognitive Neuroscience, 3:142-149

## Footnotes

1. The outcome of an analysis of the fMRI data that included these subjects is reported in supplementary results. The results are qualitatively very similar to those reported in the principal analyses although, as would be expected, memory effects are somewhat weaker. Thus, there is no reason to think that the analyses reported below were biased by the exclusion of these low-performing subjects.

2. We adopt the widely held assumption that successful source memory depends largely upon recollection of the relevant details about the study episode (see Introduction). It has been argued that source memory can also be supported by an acontextual sense of familiarity, especially if item and source information become unitized during study (Diana, Yonelinas, Ranganath, 2008). We cannot reject the possibility that source judgments in the present experiment received some contribution from familiarity. However, the findings show that successful source judgments were accompanied by robust effects in regions, such as the hippocampus, demonstrated to be recollection- but not familiarity-sensitive in studies employing several different operational definitions of recollection and familiarity (Rugg and Vilberg, 2013). Additionally, there were no source memory effects in regions, such as the intra-parietal sulcus, previously implicated in familiarity-based retrieval (Vilberg and Rugg, 2008; Kim, 2010).  Together, this pattern of results suggests that any contribution of familiarity to successful source judgments was minimal.

Table 1: <u>Outcome of the contrasts between activity associated with accurate vs. inaccurate source judgments (Memory), and high vs. low reward responses (Reward).</u>

| Contrast | Region | Peak Z (# vox) | Location (x, y, z) |
|---|---|---|---|
| Memory | Medial Prefrontal Cortex | 5.11 (800) | 6, 50, -5 |
| | L Inferior Frontal Gyrus | 3.68 (21) | -48, 29, 1 |
| | R Inferior Frontal Gyrus | 4.67 (53) | 51, 5, 22 |
| | L Middle Temporal Gyrus | 4 (24) | -57, -13, -20 |
| | L Hippocampus | 4.59 (307) | -30, -13, -11 |
| | R Hippocampus/amygdala | 3.96 (70) | 15, -4, -17 |
| | R Thalamus | 4.28 (91) | 6, -10, 1 |
| | Cingulate Cortex | 4.23 (155) | 0, -34, 37 |
| | Supramarginal/Angular Gyrus | 4.51 (683) | -54, -49, 22 |
| | R Superior Occipital Gyrus | 3.76 (37) | 30, -79, 46 |
| Reward | R Orbital Gyrus | 3.84 (24) | 33, 47, -11 |
| | Anterior Cingulate Cortex | 4.88 (996) | 3, 44, 10 |
| | Ventral striatum | 4.3 (317) | -9, 5, -5 |
| | L Orbital Gyrus | | -30, 44, -8 |
| | L Insula | | -36, 20, -11 |
| | R Insula | 4.58 (121) | 36, 20, -11 |
| | Middle Cingulate Cortex | 4.12 (50) | 3, -25, 34 |
| | L Angular Gyrus | 3.45 (60) | -42, -61, 37 |

| | | |
|---|---|---|
| R Inferior Parietal Lobule | 3.63 (77) | 48, -55, 46 |
| L Precentral Gyrus | 3.76 (43) | -24, -13, 70 |

Z-values refer to the peak of each activated cluster. Cluster sub-peak coordinates are also listed.

Left (L), right (R).

Table 2: <u>Exclusive effects of accurate vs. inaccurate source judgments, revealed by masking the Memory contrast (p<0.001, 21 cluster extent) exclusively with the Reward contrast (p < .05)</u>

| Region | Peak Z (# vox) | Location (x, y, z) |
|---|---|---|
| L Superior Medial Gyrus | 3.48 (21) | -9, 65, 13 |
| R Olfactory cortex | 3.82 (25) | 9, 14, -14 |
| R Inferior Frontal Gyrus | 4.67 (53) | 51, 5, 22 |
| L Middle Cingulate Cortex | 4.12 (50) | 0, -4, 34 |
| R Parahippocampal Gyrus | 3.87 (44) | 21, -4, -23 |
| L Hippocampus | 4.59 (250) | -30, -13, -11 |
| L. Putamen | 3.41 | -21, 2, -2 |
| L Supramarginal Gyrus | 4.05 (27) | -60, -46, 25 |
| L Middle Temporal Gyrus | 4.00 (132) | -54, -49, 16 |
| R Superior Occipital Gyrus | 3.72 (24) | 30, -79, 43 |

Table 3: Exclusive effects of high vs. low reward responses, revealed by masking the Reward contrast (p<0.001, 21 cluster extent) exclusively with the Memory contrast (p < .05)

| Region | Peak Z (# vox) | Location (x, y, z) |
| --- | --- | --- |
| L Precentral Gyrus | 29 (3.67) | -24, -13, 73 |
| Ventral striatum | 136 (4.3) | -9, 5, -5 |
| R Insula | 107 (4.58) | 36, 20, -11 |
| L Superior Medial Gyrus | 106 (4.37) | 3, 38, 34 |
| L Anterior Cingulate Cortex | 23 (3.83) | -15, 41, 7 |
| L Middle Orbital Gyrus | 42 (3.87) | -30, 44, -8 |
| R Middle Orbital Gyrus | 4 (3.84) | 33, 47, -11 |

Figure 1: Schematic of the experimental paradigm and critical contrasts employed in the fMRI analyses.
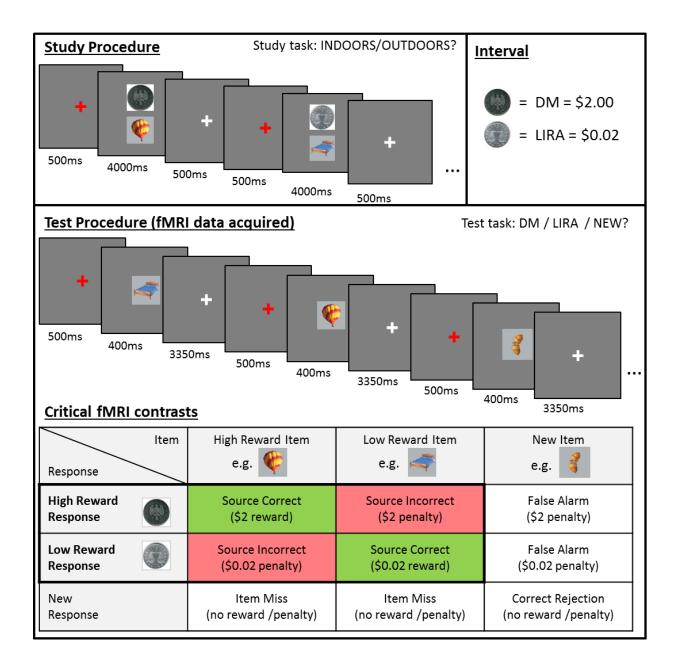
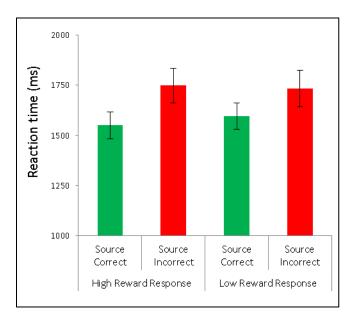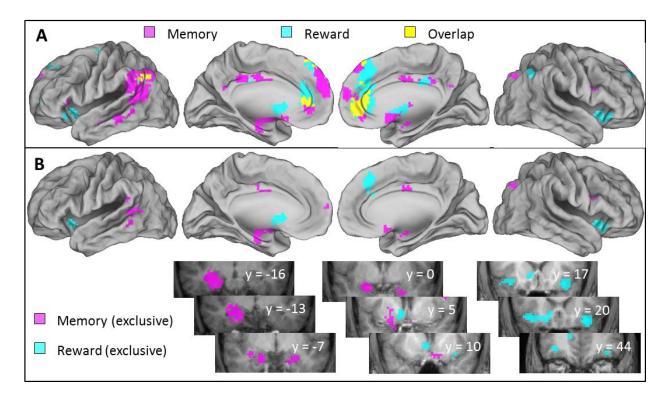Figure 2. Mean RTs to the four critical classes of test item (+/- 1 sem).
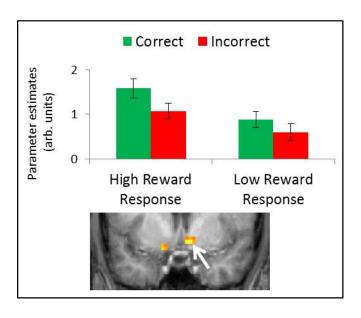
Figure 3. A) Outcome of the contrasts between high and low reward responses (reward), and correct and incorrect source judgments (memory), thresholded at p < .001. B) Exclusive effects of memory and reward in the hippocampus (left), striatum (middle) and insula and lateral orbitofrontal cortex (right).

Figure 4. Overlapping effects of reward response and source memory accuracy in the striatum (3mm sphere centered on 9, 5, -8).

<u>Supplementary Data</u>

Supplementary analyses were conducted on all twenty-six participants with sufficient trial numbers (minimum of 8 per condition) regardless of their behavioral performance. Mean item memory (hit rate – false alarm rate) was 0.74, and the probability of a correct source judgment on recognized items was 0.63. The fMRI data were analyzed in a manner identical to that described in 'fMRI methods'.

<u>fMRI results:</u>

The findings from the subsidiary analysis are illustrated in figure S1 and detailed in tables S1-S3. Memory effects were evident in much of the recollection network, although they were somewhat less extensive than in the primary analyses reported in the main text. Reward effects were, if anything, more extensive than in the primary analyses. The two classes of effect overlapped in left angular gyrus and medial PFC. Crucially, as in the principal analyses, no clusters demonstrating a memory X reward interaction could be identified.

Table S 1: Regions showing reliable effects of memory (source correct > source incorrect), reward (high reward > low reward) and regions where both contrasts are reliable (overlap) for all 26 subjects.

| Contrast | Region | Cluster | Peak Z | x | y | z |
|----------|--------|---------|--------|-----|-----|-----|
| Memory | Left Angular Gyrus | 368 | 4.57 | -45 | -70 | 43 |
| | Left Hippocampus | 124 | 4.13 | -30 | -13 | -14 |
| | Right Nucleus Accumbens | 38 | 4.12 | 9 | 5 | -8 |
| | Left Middle Temporal Gyrus | 46 | 3.78 | -63 | -31 | -8 |
| | Medial Prefrontal Cortex | 31 | 3.71 | 3 | 50 | -8 |
| | Right Hippocampus | 25 | 3.70 | 21 | -7 | -17 |
| | Left Caudate | 24 | 3.77 | -9 | 14 | -17 |
| Reward | Left Superior Medial Gyrus | 1891 | 4.94 | 3 | 41 | 37 |
| | Ventral striatum | | 4.72 | -9 | 5 | -5 |
| | Left Insula | | 4.63 | -36 | 20 | -11 |
| | Right Insula | 129 | 4.63 | 36 | 20 | -11 |
| | Middle Cingulate Cortex | 46 | 4.26 | 0 | -25 | 31 |
| | Right Inferior Parietal Lobule | 120 | 4.20 | 48 | -55 | 46 |
| | Left Angular Gyrus | 178 | 4.14 | -39 | -64 | 40 |
| | Right Middle Orbital Gyrus | 49 | 3.92 | 33 | 47 | -11 |

Table S2: Exclusive effects of accurate vs. inaccurate source judgments, revealed by masking the Memory contrast (p<0.001, 21 cluster extent) exclusively with the Reward contrast (p < .05) for all 26 subjects.

| Region | Cluster | Peak Z | x | y | z |
|---|---|---|---|---|---|
| Left Hippocampus | 94 | 4.08 | -30 | -16 | -17 |
| Left Middle Temporal Gyrus | 24 | 3.91 | -54 | -49 | 19 |
| Left Middle Temporal Gyrus | 22 | 3.43 | -48 | -37 | -5 |

Table S3: Exclusive effects of high vs. low reward responses, revealed by masking the Reward contrast (p<0.001, 21 cluster extent) exclusively with the Memory contrast (p < .05)

| Region | Cluster | Peak Z | x | y | z |
|---|---|---|---|---|---|
| Left Superior Medial Gyrus | 838 | 4.94 | 3 | 41 | 37 |
| Right Insula Lobe | 124 | 4.63 | 36 | 20 | -11 |
| Right Inferior Parietal Lobule | 35 | 4.19 | 48 | -55 | 46 |
| Right Middle Orbital Gyrus | 49 | 3.92 | 33 | 47 | -11 |
| Left Cerebellum | 32 | 3.89 | -36 | -67 | -38 |
| Right Caudate | 32 | 3.88 | 21 | 23 | 16 |

Figure S1: A) Outcome of the contrasts between high and low reward responses (reward), and correct and incorrect source judgments (memory), thresholded at p < .001, along with regions where these two contrasts overlap in all 26 subjects. B) Exclusive effects of memory and reward in the hippocampus (left), striatum (middle) and insula and lateral orbitofrontal cortex (right) in all 26 subjects.