**The Impact of Depression on Brain Activity During Source Memory Retrieval**

Elyssa M. Barrick and Daniel G. Dillon

Number of words:

Number of figures: 6

Number of tables: 2

Supplemental information: stimulus list, 1 table

Keywords:

Corresponding author:

Daniel G. Dillon, Ph.D.

Center for Depression, Anxiety and Stress Research

McLean Hospital

115 Mill Street, Belmont, MA 02478

Email: ddillon@mclean.harvard.edu

Phone: 617-855-4233

Fax: 617-855-4231

**Abstract**

**Background:** Recollection is disrupted in Major Depressive Disorder (MDD), but this disruption can be minimized by focused attention at encoding and retrieval. The neural mechanisms responsible for these clinically important phenomena are unclear. Thus, we used event-related potentials (ERPs) to examine recollection in MDD.

**Methods:** Twenty-four adults with MDD and 24 controls encoded words shown on the left or right (perceptual source) by making animacy or mobility judgments (conceptual source). ERPs were recorded during cued source retrieval, which depends on recollection.

**Results:** Mobility judgments prompted deeper encoding than animacy judgments, and memory was characterized by a *Group* x *Cue* x *Encoding Task* interaction: depressed adults were generally less accurate and less confident than controls, but they showed excellent conceptual source memory following deeper encoding. In parallel, a positive parietal ERP deflection that tracks recollection was globally reduced in depression, but sustained left parietal activation was seen during conceptual source judgments for deeply encoded words in depressed adults.

**Conclusions:** This study links two reliable effects of depression on recollection to electrophysiological activity over parietal cortex. First, accuracy and confidence were reduced in MDD, and the most reliable ERP correlate of recollection—a positive parietal deflection from 400-800 ms—was blunted. Second, depressed adults showed excellent memory when both the encoding and retrieval tasks demanded sustained attention, and this combination elicited sustained left parietal activity. These results link the impact of depression on recollection to left parietal activity, and they highlight the need for further work on this important topic.

**Introduction**

Memory retrieval plays a key role in Major Depressive Disorder (MDD) and, increasingly, in its treatment. Retrieval in depression is “overgeneral” (1): cued to recall specific episodes from their lives, depressed adults tend to offer categorical accounts, summaries that convey gist but few details. This lack of precision has consequences, as overgeneral retrieval predicts a longer course of illness (2–4). Moreover, enhancing retrieval specificity can reduce hopelessness and brooding rumination while enhancing problem solving (5), and it can lead to lasting changes in depressive symptoms (6). In short, memory retrieval is impaired in depression and targeting it for improvement can bring lasting relief.

Consequently, the depth of our ignorance regarding the neurobiology of memory retrieval in depression is astonishing, particularly since the cognitive neuroscience of episodic retrieval in healthy adults has been studied extensively (e.g., 7; 8; 9). This does not reflect lack of desire; a decade ago, the National Institutes of Mental Health, Aging, and Neurological Disorders and Stroke called for integrated research on depression and memory (10). Furthermore, the nature of the problem is clear. As one might expect from the overgeneral memory literature, depression appears to impair recollection—the retrieval of contextual details specifying the spatiotemporal source of a memory (11–14). However, despite dozens of event-related potential (ERP) and functional magnetic resonance imaging (fMRI) studies of recollection in healthy adults, a parallel literature in MDD has not emerged.

This study takes a step towards addressing this gap in the literature by using ERPs to study source memory in MDD. We adapted a design that can tease apart neural systems engaged by conceptual versus perceptual source retrieval (15–17), using neutral stimuli to avoid confounds associated with mood congruent encoding (18–20). At study, participants viewed words presented on the left or right above a question that specified one of two encoding tasks—an animacy judgment or a mobility judgment. At test, they were cued to retrieve the presentation side (perceptual source) and the encoding task (conceptual source) on separate trials.

A recent fMRI/ERP study (15) found that both conceptual and perceptual retrieval elicited the most well-studied ERP marker of recollection—a positive deflection over parietal sites that extends from about 400-800 ms post-stimulus, often with a left hemisphere maximum, and that is thought to reflect the transmission of information between the hippocampus and parietal lobes (7). Furthermore, retrieval of both sources also recruited the precuneus and elicited a negative polarity ERP maximal over posterior electrodes and commonly referred to as the late posterior negativity, or LPN (21–23). Critically, however, the LPN extended over left frontal cortex during conceptual retrieval, and this was mirrored by dorsolateral PFC activation in the fMRI session. These data suggests that retrieval attempts first activate parieto-hippocampal circuits, bringing candidate memories to mind and generating the parietal ERP effect. Next, those candidates are reviewed until one is selected and endorsed. The review and selection of perceptual memories strongly engages posterior cortical regions, but conceptual retrieval in particular activates left PFC regions that support semantic encoding, elaboration, and selection (24). Because MDD is associated with volumetric loses in hippocampus and PFC (25), we expected globally reduced source accuracy in MDD. In addition, because rumination may occupy left PFC circuits, we expected especially strong disruption of conceptual source memory.

However, in the course of our analysis it became clear that we had overlooked a key factor. Specifically, several studies report that if attention is sustained at encoding or retrieval, depressed adults can perform well (26–30). As detailed below, one of our tasks promoted deeper encoding than the other, and when words from that task were targeted for conceptual source retrieval, the MDD group was quite accurate. Thus, this study highlights neural mechanisms linked to disrupted source memory in MDD, but also activity that supports memory when encoding and retrieval conditions are salubrious.

**Materials and Methods**

**Participants and self-report measures**

Participants (18-62 years old, right-handed, no neurological or unstable medical conditions) were recruited from the community and compensated ($25.00/hour). Consent was obtained with a protocol approved by the Partners HealthCare Human Research Committee. Prospective participants were screened by phone or online, and at the ERP session we assessed psychiatric history with the MINI International Neuropsychiatric Interview, version 6.0 (31) and administered the Beck Depression Inventory II (BDI-II; 32). Controls had to report no current or past psychiatric conditions. Depressed adults had to report current depression, no history of other DSM-IV Axis I diagnosis (except secondary generalized anxiety, social anxiety, and specific phobia due to high comorbidity with MDD), no medication use in the past 2 weeks (6 weeks for fluoxetine, 6 months for neuroleptics), and a BDI-II score ≥ 14. Thirty-four controls and 26 depressed adults met these criteria and completed the ERP session. Data from 10 controls and 2 depressed adults were excluded due to EEG artifacts (see below), leaving final samples of 24 individuals per group.

In addition to the BDI-II, we administered the Mood and Anxiety Symptom Questionnaire(MASQ; 33), the Ruminative Responses Scale (RRS; 34), and the Pittsburgh Sleep Quality Index (PSQI; 35). The Wechsler Test of Adult Reading (WTAR; 36) was used to assess IQ. One depressed participant did not complete the PSQI.

**Task**

The task was programmed in PsychoPy (37). Due to a hardware change, RT data were not acquired for one control and one depressed participant.

**Stimuli.** We used the MRC Psycholinguistic Database (38) to select 25 words from four categories: “living/immobile” (e.g., *oak*), “non-living/immobile” (e.g., *shed*), “living/mobile” (e.g., *dog*), and “non-living/mobile” (e.g., *kite*). ANOVA yielded no between-list differences for number of letters (mean±S.D.; 5.27±1.29), number of syllables (1.52±0.50), frequency of occurrence (35.58±79.02), concreteness (598.87±20.18), or imageability (596.80±25.31), all *ps* > 0.064. Words are listed in the Supplement.

**Encoding.** The task included six encoding-retrieval cycles. Each encoding block included 16 trials (Figure 1, *left*) in which a word appeared on the left or right above one of two questions: “living/non-living?” or “mobile/immobile?” Participants responded by pressing a button. A jittered interval (500-2000 ms) separated the trials.

**Counting.** Immediately after encoding, a 3-digit number (e.g., 931) was shown and participants counted backwards from that number in steps of three for 30 s. This task was used to disrupt rehearsal and clear working memory (39).

PLEASE INSERT FIGURE 1 ABOUT HERE

**Retrieval.** Each block comprised 48 trials that included a cue, word, and response screen (Figure 1, *right*). On 16 trials each, the cue was “Side” or “Question” and the word came from the preceding encoding block; these cues prompted retrieval of the perceptual and conceptual source, respectively. On the remaining trials the cue was “Odd/Even” the word was a numeral between “one” and “ninety-six”, and the participant judged parity. On Odd/Even trials participants had to read a cue, interpret it, and retrieve information, but retrieval was directed at semantic rather than episodic memory. Thus, comparing ERP data from Side or Question trials versus Odd/Even trials should isolate activity mediating episodic retrieval. Presentation order of words and cues was random. The response screen consisted of ‘RESPOND’ printed above the word with the numbers 1-5 printed below and corresponding to a choice and level of confidence in that choice (Figure 1, *right*). A jittered interval (500-2000 ms) separated the trials.

**EEG Recording**

The EEG was recorded during retrieval with a 128-sensor HydroCel GSN Electrical Geodesics Inc (EGI) net (sample rate = 1000 Hz, 0.02–100 Hz). Data were referenced to the vertex and impedances were kept below 45 kΩ when possible (maximum: 75 kΩ).

**Behavioral Analysis**

We analyzed trial-level data using linear mixed models implemented with the R (40) library *lme4* (41), as this method easily accommodates covariates that might influence source memory or depression, such as age and gender (42; 43). Specific models are described below, but in all cases we computed a first model with task elements and covariates as fixed effects but with *Group* omitted. We added *Group* in a second model and then used likelihood ratio tests to compare model fits by chi-square. If the second model was a significant improvement, we report its parameters; otherwise, we report parameters from the first model. All models used *Word* and *Subject* as random effects. When modeling encoding accuracy (coded 0 or 1), we used glmer with the logit link function. Finally, we extracted *p*-values via the R library *lmerTest*.

**Encoding**. We dropped trials (< 1%) with no response or where RT exceeded the participant’s mean±3SD. Our first accuracy model included *Encoding Task*, *Side*, *Block*, *Gender*, and *Age*. Our first RT model included the same factors plus *Accuracy*. We compared these to models that included all the same factors plus *Group*.

**Retrieval**. We dropped trials with no response or where RT exceeded the participant’s mean±3SD (< 2%). Next we analyzed the Odd/Even trials.Accuracy (percent correct) on these trials was at ceiling (controls: 98.43±0.12; MDD: 99.13±0.09), and RT (in ms) was similar between groups (controls: 862.58±51; MDD: 779.00±48). Adding *Group* did not improve models that included *Block, Age*, and *Gender*, χ2s < 2.1, *p*s > 0.14. Thus, the Odd/Even trials elicited similar behavior across groups and are thus suitable as a control condition.

For Side and Question trials, accuracy was coded: incorrect, high confidence = 1; incorrect, low confidence = 2; guess = 3; correct, low confidence = 4; correct, high confidence = 5. Confidence was coded guess = 1, low confidence = 2, high confidence = 3. We computed three models for accuracy and confidence. The first included *Block*, *Cue*, *Encoding Task*, *Encoding Side*, *Age*, and *Gender*. The second added a *Cue* x *Encoding Task* interaction, and the third added a *Group* x *Cue* x *Encoding Task* interaction, plus the main effect and two-way interactions involving *Group*. We used similar models to analyze correct RT.

**ERP Analysis**

**Pre-processing.** Pre-processing was conducted with EEGLAB (44) and ERPLAB (45) toolboxes for MATLAB (MathWorks, Natick). EEG data were merged, re-referenced to the average of all electrodes, and filtered (0.1-30 Hz). Bad channels were interpolated, independent component analysis was used to extract components capturing blinks, HEOG, and EKG, and the cleaned data were time-locked to word onsets and segmented (-200 to 2000 ms). The pre-stimulus interval was used for baseline correction, and segments where any raw value or the maximum-minimum voltage difference (200 ms intervals, 100 ms sliding window) exceeded 100 μV were rejected due to artifact. We used *a priori* criteria of > 18 bad channels or more than 50% of trials rejected to exclude datasets (10 controls, 2 MDD). The mean number of clean segments in each bin defined by *Group* x *Cue* x *Task* ranged from 21-28 for source hits.Guesses were excluded from ERP analysis and there were too few segments per bin for analysis of misses. Finally, data were averaged for group-level analyses.

**Group-level analyses**. We first inspected the ERP waveforms associated with all correct responses to the Question, Side, and Odd/Even cues, regardless of encoding task. There was an obvious group difference in the left parietal ERP from 400-800 ms (see Results), thus we extracted mean amplitude data over this window from parietal electrodes in the left (P1, P3, P5, P7) and right (P2, P4, P6, P8) hemispheres for a *Group* x *Cue* x *Hemisphere* ANOVA. This focus on a few electrodes and one time window constitutes a traditional ERP analysis and has been used in many studies of recollection (7; 46; 47)

Additional analyses used a newer approach and were intended to parallel the memory accuracy data, which were characterized by a *Group* x *Cue* x *Encoding Task* interaction. To identify activity mediating this interaction, we computed Question minus Side difference waves separately for words from the mobility and animacy tasks in each group. We then submitted the difference waves to mass univariate analysis (48), focusing on mean amplitudes from 400-800 ms, 800-1400 ms, and 1400-2000 ms. Mass univariate analysis is widely used in fMRI research (49) and here entails a one-sample *t*-test (within-group analysis) or a two-sample *t*-test (between-group analysis) at each electrode. By examining every electrode and multiple time windows, this makes much better use of the spatiotemporal richness of ERP data than the traditional method. To correct for multiple comparisons, we used cluster-based permutation (50). All electrodes within 4 cm of each other were considered neighbors, and neighboring electrodes significant at *p* < 0.05 (uncorrected) were considered clusters. The sum of all *p­*-values in a cluster constituted its mass. We then performed 2500 permutations, selecting the most extreme cluster mass score from each permutation to generate a distribution (51) that was used to judge the probability of observing clusters of various sizes. Only clusters significant at *p* < 0.05 were considered reliable.

**Results**

**Demographics**

There were no group differences in gender, age, education, or estimated IQ (Table 1). The MDD group endorsed poorer sleep plus more depression, anxiety, and rumination than the controls, with the mean BDI-II score indicating moderate depression.

PLEASE INSERT TABLE 1 ABOUT HERE

**Behavior**

**Encoding**. The mobility task was harder than the animacy task, judging by percent correct (mobility: 92.42±0.26; animacy: 95.85±0.20; *Z* = -4.91, *p* < 0.001) and RT (mobility: 1,801±552 ms; animacy: 1,664±535 ms; *Z* = 10.54, *p* < 0.001). Participants were faster when correct (*Accuracy*: *Z* = -3.46, *p* < 0.001), and RT decreased over the session (*Block*: *Z* = -6.34, *p* < 0.001). *Group* did not improve the accuracy or RT models, χ2s < 1.93, *p*s > 0.16, thus depressed and healthy adults performed similarly.

**Source accuracy**. Source accuracy was influenced by depression and adding *Group* improved the model, χ2 = 26.40, *p* < 0.001. There was a *Group* x *Cue* x *Encoding Task* interaction, *Z* = -2.13, *p* = 0.033, which subsumed significant *Cue* x *Encoding Task* and *Group* x *Encoding Task* interactions and main effects of *Cue* and *Encoding Task* (*Z*s > 2.7, *p*s < 0.006). Figure 2A (left panel) shows that the triple interaction emerged because accuracy under the Side cue did not vary by encoding task, but accuracy under the Question cue was better following mobility versus animacy judgments, and this effect was larger in MDD. In the MDD group, a Question minus Side subtraction (Figure 2A, right) was positive for the mobility task but negative for the animacy task, and this difference was significant, *t*(23) = 4.47, *p* < 0.001, *d* = 0.91. In controls the Question minus Side subtraction was negative for both tasks but less negative for the mobility task, *t*(23) = 2.04, *p* = 0.053, *d* = 0.42. A between-groups test on Question minus Side differences for the animacy task was not significant, *t* < 1, but difference scores for the mobility task were more positive in MDD, *t*(46) = 3.04, *p* = 0.004, *d* = 0.88.

Because the ERP analysis focused on hits, we repeated this analysis with hit rates (Table 2). A *Group* x *Cue* x *Encoding Task* ANOVA on these data did not yield a triple interaction, *F* < 1. However, for depressed adults the Question minus Side accuracy subtraction was again more positive for the mobility task (5.03±11.95) than the animacy task (-8.76±9.82), *t*(23) = 3.82, *p* = 0.001, *d* = 1.26. The same was true for controls (mobility: -2.97±9.74; animacy: -12.27±14.11; *t*(23) = 2.83, *p* = 0.010, *d* = 0.76). As before, there was no group effect for the animacy task, *t*(46) = 1.00, *p* = 0.32, *d* = 0.29, but the Question minus Side difference score was again larger in depressed adults for the mobility task, *t*(46) = 2.54, *p* = 0.015, *d* = 0.73.

Finally, simple pairwise comparisons showed accuracy for depressed versus healthy adults only in the Question/mobility cell of the design, *Z* = 1.98, *p* = 0.048; in all other cells accuracy was non-significantly higher in controls (*Z*s < 1.63, *p*s > 0.10). Thus, depressed adults generally performed worse than controls, except for words from the mobility task resented under the Question cue. Finally, the best-fitting model included effects of *Age*, *Z* = -3.31, *p* < 0.002, *Gender*, *Z* = 3.32, *p* < 0.002, and *Block, Z* = 3.23, *p* = 0.001, reflecting higher accuracy in younger adults, in men, and in later blocks.

PLEASE INSERT FIGURE 2 AND TABLE 2 ABOUT HERE

**Source confidence**. Figure 2B (top panel) shows that depressed adults were less confident than controls. Accordingly, the model was improved by adding *Group*, χ2 = 18.46, *p* = 0.001, and included a trending *Group* x *Cue* interaction, *Z* = -1.65, *p* = 0.098, as the difference in confidence was stronger under the Side cue, *Z* = 2.42, *p* = 0.016, than the Question cue, *Z* = 1.14, *p* = 0.255. The model also revealed effects of *Cue*, *Z* = -5.33, *p* < 0.001, *Encoding Task*, *Z* = 2.91, *p* = 0.004, and *Block*, *Z* = 3.69, *p* < 0.001, as participants were more confident when responding to the Question cue, to words from the mobility task, and in later blocks.

**Source RT**. Correct RT was similar across groups, with everyone slower in response to Question vs. Side cues (Figure 2B, bottom). Thus, the model was not improved by adding *Group*, *p =* 0.08, but it included a strong effect of *Cue*, *Z* = -45.51, *p* < 0.001. There were also effects of *Run*, *Z* = -18.82, *p* < 0.001, *Confidence*, *Z* = 21.61, *p* < 0.001, and *Gender*, *Z* = -3.09, *p* = 0.003, reflecting shorter correct RTs in later blocks, for high confidence responses, and in males.

**ERPs**

Figures 3 displays waveforms elicited by correct responses to the Question, Side, and Odd/Even cues. There is robust activity over parietal electrodes from 400-800 ms in controls that is markedly reduced in MDD. A *Group* x *Cue* x *Hemisphere* ANOVA confirmed this impression, yielding an effect of *Group*, *F*(1,46) = 4.35, *p* = 0.043. There was also an effect of *Cue*, *F*(2,92) = 10.37, *p* < 0.001, with follow-up tests yielding reliable differences between all conditions (REGWQ; Question > Side > Odd/Even, *ps* < 0.043). These data suggest that recollection was strongest under the Question cue and reduced in MDD. In both groups, the LPN is apparent at Oz from 800-2000 ms on Question and Side (but not Odd/Even) trials, and it extends over left frontal cortex on Question trials, replicating prior findings (15).

PLEASE INSERT FIGURE 3 ABOUT HERE

To test our *a priori* hypothesis, we subtracted activity on Odd/Even trials from Question and Side trials. Figure 4 shows that the Question condition drove activity over left parietal electrodes from 400-800 ms, although this effect was only reliable in controls (a negative difference over right frontal cortex was only significant in the MDD group). From 1400-2000 ms, this subtraction revealed a negative difference over left PFC in both groups. By contrast, the Side condition elicited an LPN over midline posterior sites from 800-1400 ms. Contrary to expectations, between-groups tests on these difference waves yielded no reliable findings.

PLEASE INSERT FIGURE 4 ABOUT HERE

***Group x Cue x Encoding Task***. Our planned ERP analysis did not reveal group differences but it was not designed to capture the *Group* x *Cue* x *Encoding Task* interaction that characterized source accuracy. To address this limitation, we computed Question minus Side difference waves for words from each encoding task and submitted them to within and between-groups analysis, duplicating our approach to the accuracy data. Figure 5 that in depressed adults, the Question minus Side difference varied drastically by encoding task (see Table 3 for details). Words from the mobility task elicited sustained activation over left parietal cortex, leading to significant effects in all time windows. By contrast, words from the animacy task elicited a negativity over fronto-central sites that was stronger over the left hemisphere; again, significant differences were observed in all time windows. In controls, no significant effects were seen in any time window. Between-groups comparisons revealed reliable differences for the mobility task over left centro-parietal electrodes from 400-800 and 800-1400 ms, with stronger activation in the MDD group. Thus, the Question (minus Side) cue had a strong, task-dependent effect on brain activity in MDD that was muted in controls.

PLEASE INSERT FIGURE 5 AND TABLE 3 ABOUT HERE

**Individual Differences**

We computed Pearson correlations in the MDD group to determine if variation in depressive severity, brooding rumination, or sleep disruption affected source accuracy or ERP amplitudes. We found no relationship with source accuracy, left parietal activity averaged over encoding tasks, or left frontal activity on Question/animacy trials (|*p*s| < 0.31, *p*s > 0.15). However, as shown in Figure 6, the magnitude of parietal activity isolated by the Question minus Side subtraction for words from the mobility task was negatively correlated with PSQI scores from 400-800 ms and 800-1400 ms (*r*s < -0.47, *ps* < 0.02). To confirm that these results were not simply a correlate of depressive severity, we computed hierarchical regressions using ERP amplitude as the criterion and entering BDI-II and PSQI scores a predictors in steps 1 and 2, respectively. PSQI significantly predicted ERP amplitude after accounting for BDI-II in both windows (400-800 ms; β = -0.45, *p* = 0.03; 800-1400 ms; β = -0.49, *p* = 0.03), and adding PSQI improved both models (Δ*R*2s > 0.16, Δ*F*s > 4.5, *p*s < 0.05).

PLEASE INSERT FIGURE 6 ABOUT HERE

**Discussion**

This study yielded two sets of parallel behavioral and ERP findings. First, relative to controls the depressed adults were generally less accurate and less confident in their source memories, and this was mirrored by a significant reduction in parietal ERP amplitude from 400-800 ms (Figure 3). The negative effect of depression on memory was relatively modest, but in addition to generating lower confidence ratings than controls in all four cells of the design, the depressed adults generated lower mean accuracy scores in three cells of the design. Worse performance by one group in 7/8 cells is improbable under the null-hypothesis (binomial test, *p* = 0.035 one-tailed), thus the data are consistent with a negative effect of MDD on source memory accompanied by weak recruitment of brain activity linked to recollection.

However, when words from the mobility task were presented under the Question cue, depressed adults showed better memory accuracy than the controls (Figure 2A, right). This result was echoed by robust left parietal activity (Figure 5, top row). Considered alongside the first set of findings, it appears that the depressed adults under-recruited left parietal circuits except on Question/mobility trials, when strong recruitment was associated with excellent memory.

The first set of findings conformed to our expectations, which were informed by the prior literature. Depressed adults consistently show recollection deficits, thus we expected them to perform worse than controls and to show weaker parietal activity from ~400-800 ms, as this activity is a marker of recollection (7). The second set of findings was unexpected, but perhaps it should have been anticipated. Hertel has consistently documented negative effects of depression on memory when encoding tasks are loosely structured and leave time for mind-wandering, but when encoding tasks demand sustained attention and thus admit little mind-wandering, she has found excellent memory in depression (Hertel & Rude, 1991). Both of our encoding tasks offered structure, and both are conventionally considered “deep” tasks. However, both groups took longer to complete the mobility task versus the animacy task, and this was accompanied by a difference in accuracy—probably because judging animacy is straightforward, but judging mobility is difficult for some words (e.g., *oak* would be considered immobile in the context of instructions provided to participants, but oak trees sway in the breeze). Therefore, we argue that the mobility task required more sustained attention and prompted deeper encoding than the judgment task, thus offering richer targets for retrieval under the Question cue. Moreover, responses to the Question cue were slower than responses to the Side cue, indicating that participants labored over conceptual retrieval more than perceptual retrieval. Consequently, Question/mobility trials demand sustained attention at encoding and retrieval, and Hertel’s work predicts that this is when depressed adults should do best—as indeed they do.

Until now it has not been possible to do more than speculate about the brain systems that support good source memory in depression when this is observed, because (to our knowledge) there is only one prior functional study of source memory in depression, and that study found no group differences in behavior (REF). Here we show that the same activity implicated in source memory failure is implicated in source memory success following deep encoding—namely, left parietal cortex. Correct responses on Question/mobility trials elicited strong activity over left parietal cortex extending over all time windows analyzed in depressed adults, and this activation exceeded that seen in controls from 400-800 and 800-1400 ms. Because of the aforementioned link to recollection, these ERP effects suggest that the combination of deep encoding and conceptual retrieval supports the sustained retrieval of rich detail in depressed adults. By contrast, when conceptual retrieval was directed at words from the shallower encoding condition (animacy task), the depressed adults generated a lasting negativity over left PFC. As outlined in the Introduction, this likely reflects some combination of cue elaboration or selection amongst competing memory representations, which are probably more necessary in the context of this more difficult cell of the design.

These results provide valuable initial insights into source memory retrieval in depression, but some limitations should be mentioned. First, the negative effect of MDD on memory was relatively modest, and it may be useful to modify the design and/or the recruitment strategy in order to identify more severe disruptions. In fact, our use of unmedicated outpatients and neutral stimuli, in the context of a fully anticipated memory test with no stress manipulation, is probably the

**References**

1. Williams JMG, Barnhofer T, Crane C, Herman D, Raes F, Watkins E, Dalgleish T (2007): Autobiographical memory specificity and emotional disorder. *Psychol Bull*. 133: 122–48.

2. Brittlebank AD, Scott J, Williams JM, Ferrier IN (1993): Autobiographical memory in depression: state or trait marker? *Br J Psychiatry*. 162: 118–21.

3. Peeters F, Wessel I, Merckelbach H, Boon-Vermeeren M (2002): Autobiographical memory specificity and the course of major depressive disorder. *Compr Psychiatry*. 43: 344–350.

4. Sumner JA, Griffith JW, Mineka S (2010): Overgeneral autobiographical memory as a predictor of the course of depression: A meta-analysis. *Behav Res Ther*. 48: 614–625.

5. Raes F, Williams JMG, Hermans D (2009): Reducing cognitive vulnerability to depression: A preliminary investigation of MEmory Specificity Training (MEST) in inpatients with depressive symptomatology. *J Behav Ther Exp Psychiatry*. 40: 24–38.

6. Neshat-Doost HT, Dalgleish T, Yule W, Kalantari M, Ahmadi SJ, Dyregrov a., Jobson L (2012): Enhancing Autobiographical Memory Specificity Through Cognitive Training: An Intervention for Depression Translated From Basic Science. *Clin Psychol Sci*. . doi: 10.1177/2167702612454613.

7. Rugg MD, Curran T (2007): Event-related potentials and recognition memory. *Trends Cogn Sci*. 11: 251–257.

8. Eichenbaum H, Yonelinas a P, Ranganath C (2007): The medial temporal lobe and recognition memory. *Annu Rev Neurosci*. 30: 123–152.

9. Rugg MD, Vilberg KL (2013): Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol*. 23: 255–260.

10. Steffens DC, Otey E, Alexopoulos GS, Butters MA, Cuthbert B, Ganguli M, *et al.* (2006): Perspectives on Depression, Mild Cognitive Impairment, and Cognitive Decline. *Arch Gen Psychiatry*. 63: 130.

11. Raes F, Hermans D, Williams JMG, Demyttenaere K, Sabbe B, Pieters G, Eelen P (2006): Is overgeneral autobiographical memory an isolated memory phenomenon in major depression? *Memory*. 14: 584–594.

12. MacQueen GM, Galway TM, Hay J, Young LT, Joffe RT (2002): Recollection memory deficits in patients with major depressive disorder predicted by past depressions but not current mood state or treatment status. *Psychol Med*. 32: 251–258.

13. MacQueen GM, Campbell S, McEwen BS, Macdonald K, Amano S, Joffe RT, *et al.* (2003): Course of illness, hippocampal function, and hippocampal volume in major depression. *Proc Natl Acad Sci U S A*. 100: 1387–1392.

14. Ramponi C, Barnard P, Nimmo‐Smith I (2004): Recollection deficits in dysphoric mood: An effect of schematic models and executive mode? *Memory*. 12: 655–670.

15. Bergström ZM, Henson RN, Taylor JR, Simons JS (2013): Multimodal imaging reveals the spatiotemporal dynamics of recollection. *Neuroimage*. 68: 141–153.

16. Simons JS, Gilbert SJ, Owen AM, Fletcher PC, Burgess PW, Jon S, *et al.* (2005): Distinct Roles for Lateral and Medial Anterior Prefrontal Cortex in Contextual Recollection. 813–820.

17. Dobbins IG, Wagner AD (2005): Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cereb Cortex*. 15: 1768–1778.

18. Bower GH (1981): Mood and memory. *Am Psychol*. 36: 129–148.

19. Bower GH (1987): Commentary on mood and memory. *Behav Res Ther*. 25: 443–455.

20. Dillon DG, Dobbins IG, Pizzagalli DA (2014): Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. *Soc Cogn Affect Neurosci*. 9: 1576–83.

21. Johansson M, Mecklinger A (2003): The late posterior negativity in ERP studies of episodic memory: action monitoring and retrieval of attribute conjunctions. *Biol Psychol*. 64: 91–117.

22. Mecklinger A, Johansson M, Parra M, Hanslmayr S (2007): Source-retrieval requirements influence late ERP and EEG memory effects. *Brain Res*. 1172: 110–123.

23. Cycowicz YM, Friedman D, Snodgrass JG (2001): Remembering the color of objects: an ERP investigation of source memory. *Cereb Cortex*. 11: 322–334.

24. Badre D, Wagner AD (2007): Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. 45: 2883–2901.

25. Treadway MT, Waskom ML, Dillon DG, Holmes AJ, Park MTM, Chakravarty MM, *et al.* (2015): Illness progression, recent stress, and morphometry of hippocampal subfields and medial prefrontal cortex in major depression. *Biol Psychiatry*. 77: 285–294.

26. Hertel PT (1997): On the Contributions of Deficent Cognitive Control to Memory Impairments in Depression. *Cogn Emot*. 11: 569–583.

27. Hertel PT, Brozovich F (2010): Cognitive Habits and Memory Distortions in Anxiety and Depression. *Curr Dir Psychol Sci*. 19: 155–160.

28. Hertel PT, Benbow AA, Geraerts E (2012): Brooding deficits in memory: Focusing attention improves subsequent recall. *Cogn Emot*. 26: 1516–1525.

29. Hertel PT, Hardin TS (1990): Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. *J Exp Psychol Gen*. 119: 45–59.

30. Hertel PT, Rude SS (1991): Depressive deficits in memory: focusing attention improves subsequent recall [see comments]. *J Exp Psychol Gen*. 120: 301–309.

31. Sheehan D V., Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, *et al.* (1998): The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. (Vol. 59), pp 22–33.

32. Beck AT, Steer RA, Brown GK (1996): Manual for the Beck depression inventory-II. *San Antonio, TX Psychol Corp*. 1–82.

33. Watson D, Weber K, Assenheimer JS, Clark LA, Strauss ME, McCormick RA (1995): Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *J Abnorm Psychol*. 104: 3–14.

34. Treynor W, Gonzalez R, Nolen-hoeksema S (2003): Rumination Reconsidered : A Psychometric Analysis. 27: 247–259.

35. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ, III CFR, *et al.* (1989): The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 28: 193–213.

36. Holdnack HA (2001): *Wechsler Test of Adult Reading: WTAR*. San Antonio, Texas: The Psychological Corporation.

37. Peirce JW (2008): Generating Stimuli for Neuroscience Using PsychoPy. *Front Neuroinform*. 2: 10.

38. Coltheart M (1981): The MRC psycholinguistic database. *Q J Exp Psychol Sect A*. 33: 497–505.

39. Reitman JS, Higman B, Lifson A, Rosenblum J (1974): Without Surreptitious Rehearsal, Information in Short-Term Memory Decays I. *J Verbal Learning Verbal Behav*. 13: 365–377.

40. R Developement Core Team (2015): R: A Language and Environment for Statistical Computing. *R Found Stat Comput*. 1: 409.

41. Bates D, Maechler M, Bolker B, Walker S (2015): lme4: linear mixed-effects models using S4 classes. R package version 1.1-8. *R*. . doi: http://CRAN.R-project.org/package=lme4.

42. Mark RE, Rugg MD (1998): Age effects on brain activity associated with episodic memory retrieval. An electrophysiological study. *Brain*. 121: 861–873.

43. Nolen-Hoeksema S (2001): Gender Differences in Depression. *Curr Dir Psychol Sci*. 10: 173–176.

44. Delorme A, Makeig S (2004): EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 134: 9–21.

45. Lopez-Calderon J, Luck SJ (2014): ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci*. 8: 213.

46. Hayama HR, Johnson JD, Rugg MD (2008): The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*. 46: 1211–1223.

47. Wilding EL, Rugg MD (1996): An event-related potential study of recognition memory with and without retrieval of source. *Brain*. 119: 889–905.

48. Groppe DM, Urbach TP, Kutas M (2011): Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*. 48.

49. Friston KJ, Holmes a. P, 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.

50. Groppe DM, Urbach TP, Kutas M (2011): Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*. 48: 1726–37.

51. Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ (1999): Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans Med Imaging*. 18: 32–42.

**Figure Captions**

*Figure 1*. Encoding (left) and recognition (right) trial structures. Encoding trials began with three centrally presented arrows pointing to the side on which the word would appear. The encoding task was presented next, either “living or non-living?” (animacy judgment) or “mobile or immobile?” (mobility judgment, not shown). Finally, a word was presented directly above the encoding question; participants had 3500 ms to respond. Retrieval trials began with presentation of one of three cues (“Side”, “Question”, or “Odd/Even”). After a 1000 ms delay, a word was presented. On Side and Question trials, the word came from the immediately preceding encoding block, while on Odd/Even trials the word was a numeral (e.g., “seventy-seven”). Finally, a response screen was presented and persisted until the participant responded or 10 seconds had elapsed. The response options for a Side trial are displayed. On Question trials, “left” and “right” were replaced with “living/non-living” and “mobile/immobile”, respectively; on Odd/Even trials they were replaced with “odd” and “even”.

*Figure 2*. Source memory (A) accuracy, (B, *top*) confidence, and (B*, bottom*) correct RT. Bar heights correspond to the mean, error bars = SEM. Asterisks denote *p* < 0.05.

*Figure 3*. Waveforms elicited by correct responses to the Question (black), Side (red), and Odd/Even (blue) cues. Representative electrodes from the left and right hemisphere are depicted for frontal and parietal scalp; a midline occipital electrode is displayed as this is where the late posterior negativity (LPN) was maximal. Gray shading demarcates the parietal ERP associated with recollection. Asterisks indicate the reduction in parietal activity in depressed adults.

*Figure 4*. Topographies showing activity elicited by Question and Side hits, with activity on correct Odd/Even trials subtracted out. Columns correspond to the three time windows analyzed (400-800, 800-1400, 1400-2000 ms). Electrodes in clusters associated with significant effects (within-group) are shown in white. Between-group comparisons revealed no differences.

*Figure 5*. Results of the mass univariate analysis conducted on Question minus Side difference waves, sorted by group and encoding task. Columns correspond to the three time windows analyzed (400-800, 800-1400, 1400-2000 ms). Electrodes in clusters associated with significant effects are shown in white. Paralleling the behavioral analyses, there were MDD > control differences in response to words from the mobility but not the animacy task.

*Figure 6.* ERP amplitudes are related to sleep quality in depressed adults. There were significant negative correlations between sleep disturbance as measured by the PSQI (*x*-axis) and ERP amplitudes captured by the Question minus Side difference wave for words from the mobility task (*y­*-axis) in the 400-800 (left) and 800-1400 (right) ms time windows.