An ERP Study of Multidimensional Source Retrieval in Depression

Elyssa M. Barrick and **Daniel G. Dillon**

Center for Depression, Anxiety and Stress Research, McLean Hospital

Abstract

We collected event-related potentials (ERPs) from 24 unmedicated adults with Major Depressive Disorder (MDD) and 24 controls during source memory retrieval. Words were encoded on the left or right during animacy and mobility judgments. Mobility judgments were slower than animacy judgments, suggesting deeper encoding. Participants then recalled the encoding judgment (Question cue) and position (Side cue) for each word. Depressed adults, but not controls, showed better accuracy for words from the mobility task presented under the Question vs. Side Cue. Furthermore, depressed adults showed larger left parietal ERPs to words from the mobility task presented under the Question vs. the Side Cue from 400–800 ms and 800–1400 ms. This ERP effect was negatively correlated with sleep quality. Thus, deep encoding followed by retrieval of the encoding judgment supported memory in MDD and augmented left parietal ERPs that have been linked to recollection and that appear sensitive to sleep disturbance.

Graphical abstract
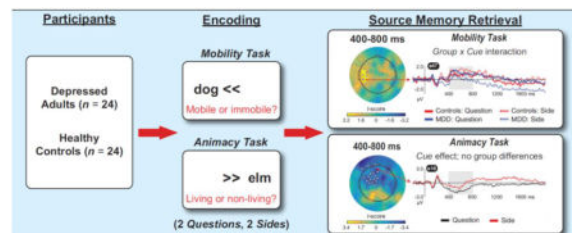
Keywords

source memory; retrieval; ERP; recollection; depression

Corresponding author: Daniel G. Dillon, Ph.D. Center for Depression, Anxiety and Stress Research McLean Hospital 115 Mill Street, Belmont, MA 02478, ddillon@mclean.harvard.edu, Phone: 617 -855-4233 Fax: 617-855-4231.

1. Introduction

Major Depressive Disorder (MDD) is associated with poor episodic memory (Airaksinen et al., 2004; Burt et al., 1995; Rock et al., 2014; Zakzanis et al., 1998). Importantly, the extent of the memory deficit appears to vary with the degree of support provided at encoding. For example, Zakzanis et al. (1998) found that depressed adults performed worse when the encoding tasks provided less structure (e.g., memorization of uncategorized vs. categorized word lists). The cognitive initiative framework can account for such results (Hertel, 1997; Hertel and Hardin, 1990). The framework's core hypothesis is that depressed individuals can control attention and use strategies to enhance encoding and improve memory but that, in the absence of external support or emotionally compelling material, they often fail to do so.

To test this hypothesis, Hertel and Rude (1991) had depressed and healthy adults encode neutral words in a task with focused and unfocused conditions. Specifically, participants were asked to judge whether single words (e.g., "artist") fit well into sentence frames (e.g., "The young man's portrait was painted by the ___"). In the focused condition, each word was shown for one second and disappeared when the sentence was presented (for eight seconds), such that the participant had to keep the word in working memory in order to respond accurately. Furthermore, participants in the focused condition could only respond when prompted, and they did so by repeating the word and then verbally indicating whether or not it fit into the frame. By contrast, in the unfocused condition the word remained onscreen while the sentence frame was presented, the participant could respond at any time, and the response consisted only of indicating whether or not the word was a quality fit. Thus, the focused condition made more demands on attention and working memory than the unfocused condition. This manipulation was designed to test the hypothesis that if depressed participants had to devote sufficient resources to encoding each word, they would be less likely to ruminate or otherwise engage in off-task thinking and memory would be enhanced. Consistent with this prediction, a *Group* x *Task* interaction emerged for free recall: relative to controls, depressed adults recalled fewer words from the unfocused condition, but there was no difference for words from the focused condition. Thus, depressed adults performed well when the encoding task demanded sustained engagement.

The cognitive initiative framework has also been applied to retrieval. Depressed adults typically show larger deficits for recall than for recognition (Burt et al., 1995), and when recognition is analyzed to estimate contributions made by recollection vs. familiarity, depression impairs the former more than the latter (Hertel and Milan, 1994; MacQueen et al., 2002). The cognitive initiative framework explains these data by pointing to the greater need for controlled attention, effortful searching, and post-retrieval monitoring during recall vs. recognition, and in support of recollection vs. familiarity. Importantly, the results of Hertel and Rude (1991) suggest that depressed adults should show improved recollection if the task used to probe retrieval helps them focus their attention appropriately.

Given the elegant behavioral work on these issues, the paucity of relevant neuroscientific data is surprising. In particular, although there are many studies of hippocampal volume in depression (for review, see MacQueen and Frodl, 2011) and some functional imaging investigations of encoding (Bremner et al., 2004; Dillon et al., 2014; Dillon and Pizzagalli,

2013; Hamilton and Gotlib, 2008), there are remarkably few neuroscientific studies of memory retrieval in MDD. Over a decade ago, the National Institutes of Mental Health, Aging, and Neurological Disorders and Stroke called for neuroscientific research on depression and memory (Steffens et al., 2006). However, despite dozens of event-related potential (ERP) and functional magnetic resonance imaging (fMRI) studies of episodic retrieval in healthy adults (e.g., Eichenbaum et al., 2007; Rugg and Curran, 2007; Rugg and Vilberg, 2013), no similar literature has emerged in MDD.

The current study was designed to address this gap. Because depression affects recollection more than familiarity—and given the difficulties associated with imaging free recall—we elected to conduct an ERP investigation of source memory in MDD. Source memory refers to conscious retrieval of the spatiotemporal details that define an encoding episode (Johnson et al., 1993). It depends heavily on recollection, and there is evidence that source memory is disrupted in depressed adults (Degl'Innocenti and Bäckman, 1999). We used a design that recruits neural systems engaged during conceptual and perceptual retrieval (Bergström et al., 2013; Dobbins and Wagner, 2005; Simons et al., 2005a). At study, participants viewed neutral words shown on the left or right above a question specifying an animacy or mobility judgment. At test, they were cued to recall the presentation side (perceptual source, "Side" cue) and encoding task (conceptual source, "Question" cue). We used neutral words rather than emotional words to limit mood-congruency effects (Watkins et al., 1992), as these might obscure a more fundamental impact of depression on the neurocognitive processes that mediate source memory.

A recent fMRI/ERP study in healthy adults (Bergström et al., 2013) found that both conceptual and perceptual retrieval elicited the most well-studied ERP marker of recollection, which is a positive deflection over parietal scalp that extends from about 400–800 ms post-stimulus, typically with a left hemisphere maximum (Rugg and Curran, 2007). Both forms of retrieval also activated the precuneus and elicited a negative polarity ERP maximal over posterior electrodes that is referred to as the late posterior negativity, or LPN (Cycowicz et al., 2001; Johansson and Mecklinger, 2003; Mecklinger et al., 2007). Intriguingly, the LPN extended over left frontal scalp during conceptual but not perceptual retrieval, and this was mirrored by fMRI activation in left dorsolateral PFC (DLPFC). Related fMRI studies confirmed that left and medial PFC regions were more strongly activated during conceptual vs. perceptual source retrieval (Simons et al., 2005a,b).

In light of the prior literature linking depression to poor performance on cognitively demanding retrieval tasks, we expected reduced source memory accuracy in MDD. In particular, because depression has been consistently linked to DLPFC hypofunction (Koenigs & Grafman, 2009) and diminished left PFC activation at rest (Davidson et al., 2002), and because brooding rumination—a common problem in depression (Treynor et al., 2003)—can recruit DLPFC neurons that would otherwise support conceptual memory (Cooney et al., 2010), we predicted that MDD would have an especially strong negative impact on conceptual source memory. To test these hypotheses, we computed between-group contrasts of ERPs elicited during conceptual and perceptual source retrieval, collapsed over the encoding tasks. In a second analysis intended to more closely track the behavioral results described below, we examined group differences in conceptual and perceptual

retrieval for words from each encoding task considered separately. This analysis provided an opportunity to determine whether the different degrees of support provided at encoding and retrieval affected memory in a manner consistent with the cognitive initiative framework (Hertel, 1997).

Finally, we computed correlations that related behavior and ERP amplitudes to individual differences in depressive severity, brooding rumination, and sleep disturbance in the MDD group. Our decision to investigate relationships with depressive severity and brooding rumination was based on the literature reviewed above—we expected that more severe depression and a greater tendency to ruminate would be associated with poorer memory. We examined sleep disturbance because it affects processes relevant to episodic retrieval, including executive function and the activation of parietal regions implicated in recollection (Chee et al., 2006; Durmer and Dinges, 2005; McEwen, 2006), and because of substantial evidence of disrupted sleep in MDD and other psychiatric disorders (Deldin et al., 2006; Tsuno et al., 2005; Wulff et al., 2010). We anticipated that negative relationships between memory, left parietal ERPs associated with recollection, and depression would be strongest in those individuals who reported the worst sleep.

2. Materials and Methods

2.1. Participants and self-report

Adults (18–62 years old, right-handed, no neurological or unstable medical conditions) were recruited from the community and compensated $25/hour, following a protocol approved by the Partners HealthCare Human Research Committee. They were screened by phone or online, at which time the Beck Depression Inventory II (BDI-II; Beck et al., 1996) was administered. Individuals were invited to participate in the MDD group if they endorsed symptoms consistent with a current Major Depressive Episode, had a BDI-II score   14 (the cut-off for mild depression; Beck et al., 1996), and reported no other Axis I psychopathology with the exception of generalized anxiety, social anxiety, and/or specific phobia. Controls had to report no current or past Axis I psychopathology. On the day of the ERP session, we assessed psychiatric status with the MINI International Neuropsychiatric Interview, version 6.0 (Sheehan et al., 1998). Depressed adults had to again report current depression, no history of other DSM-IV Axis I diagnosis (except generalized anxiety, social anxiety, or specific phobia), and no medication use in the past two weeks (six weeks for fluoxetine, six months for neuroleptics). Thirty-four controls and 26 depressed adults completed the session. Data from 10 controls and 2 depressed adults were excluded due to excessive artifacts (see below), leaving 24 individuals per group.

Following the EEG session, we administered the BDI-II again, along with the Mood and Anxiety Symptom Questionnaire (MASQ; Watson et al., 1995), the Ruminative Responses Scale (RRS; Treynor et al., 2003), and the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989). These probe symptoms of depression and anxiety, trait rumination, and sleep quality over the last month, respectively. Finally, the Wechsler Test of Adult Reading (WTAR; Holdnack, 2001) was used to estimate IQ. One control did not complete the MASQ and one depressed participant did not complete the PSQI. WTAR data from non-native

English speakers (2 controls, 2 MDD) were not analyzed, as WTAR results may be invalid in this population. The entire protocol took between 2.5 and 3.0 hours to complete.

As shown in Table 1, there were no group differences in gender, age, education, or WTAR scores. Relative to controls, the MDD group endorsed poorer sleep, more rumination, and greaterdepression and anxiety. The mean BDI-II score indicatedmoderate depression. Regarding comorbid anxiety, two depressed participants met criteria for generalized anxiety disorderin the past six months, several reported sub -threshold symptoms in the past month (social anxiety, $n = 2$; panic attacks, $n = 2$; agoraphobia, $n = 2$), and seven reported at least one lifetime panic attack.

2.2. Task

The task was programmed in PsychoPy (Peirce, 2009). Due to a hardware change, RT data were not recorded for one control and one depressed participant.

2.2.1. Stimuli—We used the MRC Psycholinguistic Database (Coltheart, 1981) 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 significant differences for number of letters (mean±S.D.; 5.27±1.29) or syllables (1.52±0.50), frequency (35.58±79.02), concreteness (598.87±20.18), or imageability (596.80±25.31), $ps > 0.064$. Words are listed in the Supplement.

2.2.2. 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?" (animacy task) or "mobile/immobile?" (mobility task) (see Figure 1 for stimulus durations). Participants responded by pressing one of two buttons ("c" or "m") on a standard computer keyboard with the index fingers of their left and right hands. A jittered interval (500–2000 ms) separated the trials. The computer randomly selected a new set of six encoding lists for each participant, with the only constraint being that each 16-word list had to include four words from each category (living/immobile, non-living/immobile, living/mobile, non-living/mobile). Once the lists were created, the word order in each list was randomized.

2.2.3. Counting—Immediately after encoding, a 3-digit number (e.g., 931) was shown and participants counted backwards, out loud, from that number in steps of three for 30 s. Backward counting served to disrupt rehearsal and clear working memory (Reitman et al., 1974). The experimenter watched the participant to ensure that the counting was completed, but no data were recorded during this phase of the experiment.

2.2.4. Retrieval—Each block comprised 48 trials that included a cue, a word, and a 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 trials probed perceptual and conceptual source retrieval, respectively. On the remaining 16 trials the cue was "Odd/ Even", the word was a numeral between "one" and "ninety-six", and the participant judged parity. All trials involved reading a cue, interpreting it, and retrieving information, but on Odd/Even trials retrieval was directed at semantic rather than episodic memory. Thus, by

comparing ERP data from Side or Question trials to Odd/Even trials, we intended to isolate activity mediating episodic retrieval (see Bergström et al., 2013, for a similar approach). The presentation order of words and cues was randomized for each participant.

To avoid confounding ERPs associated with memory from those associated with response preparation and execution, participants could not respond until a response screen was shown, 3000 ms after each word was presented. 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 (Figure 1, *right*). As in other studies of multidimensional source memory (Starns and Hicks, 2005), "guess" was included as a response option. Participants were asked to select this when they could not recover any information favoring one source over the other. A jittered interval (500–2000 ms) separated the trials. Participants responded by pressing the "c" and "v" keys with the first two fingers of their left hands, and the "b", "n", and "m" keys with the first three fingers of their right hands.

2.3. 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 vertex and impedances were kept below 45 kΩ when possible (maximum: 75 kΩ).

2.4. Behavioral Analysis

The behavioral data were cleaned by dropping trials with no response or where *log*(RT) exceeded the participant's mean±3SD(< 4% of encoding trials, < 1% of retrieval trials). Analysis involved *t*-tests andmixed -model Type III ANOVAs implementedin the R software (R Core Team, 2017)library *afex*(Singmann et al., 2016) . For both encoding and retrieval, accuracy was computed as percent correct. For RT, ANOVAs were computed on *log*transformed data because of improved fit to normality, but results are described foruntransformed RT data to enhanceinterpretability . At encoding, *Group* x *Task* (mobility, animacy) ANOVAs were run for accuracy and correct RT. At retrieval, between-groups *t*-tests were first used to compare accuracy, confidence, and correct RT on Odd/Even trials, to verify that MDD did not affect performance in this control condition. Next, a *Group* x *Cue* x *Task*ANOVA was run on the number of guesses. Finally, accuracy, confidence, and correct RT on Question and Side trials were analyzedwith *Group* x *Cue* x *Task*ANOVAs . Alpha was set at 0.05.

2.5. ERP Analysis

2.5.1. Pre-processing—Pre-processing wasconducted with the EEGLAB (Delorme and Makeig, 2004)and ERP LAB (Lopez-Calderon and Luck, 2014)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 components analysis was used to remove activity reflectingblinks, 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. We used *a priori* criteria of > 18 bad channels or more than 50% of trials

rejected (Luck, 2014) to exclude excessively noisy 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 and there were too few clean segments for analyzing misses . Thus, the analysis was focused on correct responses, a common approach in this literature (Bergström et al., 2013; Dobbins and Wagner, 2005; Han et al., 2012; Simons et al., 2005a).

2.5.2. Group-level analyses—In the first ERP analysis , we computed "Question minus Odd/Even" and "Side minus Odd/Even" difference waves , collapsed over the encoding tasks for each participant, and then compared the MDD and control groups. This approach has been used previously (Bergström et al., 2013, Simons et al., 2005a,b), and it allowed us to test our *a priori* hypothesis that depression would disrupt source retrieval in general and conceptual retrieval in particular. The second ERP analysis was designed t o isolate cue effects while holding encoding processes constant, and involved (a) computing" Question minus Side" difference waves separately for words from the mobility and animacy tasks in each participant , and then (b) computing between-group contrasts on the difference waves. This analysis was designed to parallel the analysis of the source memory accuracy data.

All within and between-group analyse s were statistically analyzed by submitting difference waves to mass univariate analysis (Groppe et al., 2011a, b) and focusing on mean amplitudes from 400–800 ms, 800–1400 ms, and 1400–2000 ms. The 400–800 ms interval was selected to capture left parietal ERP effects associated with recollection (Rugg and Curran, 2007), with the latter two windows intended to encompass the LPN seen during source recollection (Bergström et al., 2013; Johansson and Mecklinger, 2003). Mass univariate analysis is widely used in fMRI studies (Friston et al., 1995) and increasingly in ERP research (Groppe et al., 2011a,b; Luck and Gaspelin, 2017; Maris and Oostenveld, 2007; Pernet et al., 2015). Here it entailed one-sample (within-group analysis) or two-sample (between-group analysis) *t*-tests at each electrode. To correct for m ultiple comparisons, we used cluster-based permutation (Groppe et al., 2011a). All electrodes within 4 cm of each other were considered neighbors, and neighboring electrodes significant at $p < 0.05$ (uncorrected) were considered clusters. The use of a 4 cm inter-electrode distance is appropriate for detecting neighboring electrodes in regions where they are relatively widely spaced (e.g., over parietal scalp) on the 128 channel EGI net (Song et al., 2015). The sum of all *p*-values in a cluster constituted its mass. We then performed 2500 between-participant permutations of the data, each time randomly assigning participants to the two groups without replacement, computing the group difference, and then selecting the most extreme cluster mass to generate a distribution of observed cluster sizes under the null hypothesis of no group difference (Bullmore et al., 1999). For within-group analyses, the permutation was made between conditions. This null distribution was used to estimate the probability of observing the clusters actually detected when the control and MDD groups were compared. Only clusters significant at $p < 0.05$ (corrected) are reported.

Relative to the conventional practice of analyzing data from a few electrodes and time intervals in multi-factorial ANOVAs, mass univariate analysis offers important advantages. First, because it includes data from every electrode, it is an unbiased approach. By contrast, analyzing electrodes based on visual inspection or a selective reading of the prior literature

is known to drastically increase the chance of false positives in ERP studies (Luck and Gaspelin, 2017). Second, because every electrode is analyzed it is possible to detect unexpected effects that might be missed in a conventional analysis focused on a handful of sites of interest (Groppe et al., 2011a,b). Third, the mass univariate approach yields results that are more easily interpreted. Rather than dealing with unwieldy higher-order interactions from multi-factor ANOVAs—which can be difficult to unpack and are often associated with false positives (Luck and Gaspelin, 2017)—the mass univariate approach highlights clusters of electrodes that show reliable differences between conditions and/or groups, limiting the need for numerous follow-up analyses and the attendant increase in Type I error. These advantages are particularly helpful when addressing a research question without an extensive prior ERP literature, such as the study of source memory in depression.

The primary disadvantage is that mass univariate testing can be overly conservative. This may be counterintuitive given that *t*-tests are conducted at every electrode, but the rigorous methods used to correct for multiple comparisons make this approach less powerful than uncorrected *t*-tests at electrodes selected *a priori* (Groppe et al., 2011a). However, simulation studies show that the cluster-based permutation correction used here is more powerful than other methods for multiple comparison correction (e.g., Bonferroni correction, false discovery rate), particularly when effects are broadly distributed. Importantly, cluster-based correction still controls the family-wise error rate at 0.05 when the null hypothesis is true (Groppe et al., 2011b). To facilitate visual inspection and comparison with prior work, we plotted ERP waveforms from the most significant electrode in each cluster identified by the mass univariate approach.

2.6. Individual Differences

In the depressed group, we used Pearson correlations to examine relationships between source accuracy, left parietal ERPs, sleep quality, depressive severity and brooding rumination.

3. Results

3. 1. Behavior

3.1.1. Encoding—There was a main effect of *Task* for response accuracy , $F(1, 46) = 13.12$, $p < 0.001$, $\eta_p^2 = 0.22$, as the percent correct was lower for mobility vs. animacy judgments in both groups (controls: mobility = 92±7%, animacy = 96±5%; MDD: mobility = 93±5%, animacy = 96±4%). Neither the main effect of *Group* nor the *Group* x *Task* interaction was reliable, $Fs < 1$. The main effect of *Task* was also significant for correct RT , $F(1, 44) = 49.26$, $p < 0.001$, $\eta_p^2 = 0.53$, with slower responses for mobility vs. animacy judgments in both groups (controls: mobility = 1852±223 ms, animacy = 1722±232 ms; MDD: mobility = 1757±357 ms, animacy = 1615±339 ms). As with response accuracy, neither the main effect of *Group* nor the *Group* x *Task* interaction was significant, $Fs < 1.44$, $ps > 0.23$.

3.1.2. Retrieval: Odd/Even trials—There were no group differences on Odd/Even trials, which were characterized by high accuracy (controls: 98±4%; MDD: 99±1%; $t(46) = -0.82$,

$p = 0.42$), fast correct RTs (controls: 868±302 ms; MDD: 781±228 ms; $t(44) = 1.10$, $p = 0.28$), and confident responding(percentage of responses made with high confidence: controls: 99.61±0.68%; MDD: 99.87±0.35%; $t(46) = -1.68$, $p = 0.10$).

3.1.3. Retrieval: accuracy—The source accuracy resultsare depicted in Figure 2A. The left panel showsdata for words from the mobility task, while the right panel shows data for words from the animacy task. Despite the lack of a *Group* x *Cue* x *Task* interaction, $F < 1$, there was a *Group* x *Cue* interaction for words from the mobility task, $F(1, 46) = 5.92$, $p = 0.02$, $\eta_p2 = 0.11$, but not for words from the animacy task, $F < 1$. Instead, accuracy for words from the animacy task was characterized by a main effect of *Cue*, $F(1, 46) = 36.81$, $p < 0.001$, $\eta_p2 = 0.44$.

The nature of these results is highlighted in Figure 2B, which plots "Question minus Side" accuracy difference scores. For words from the mobility task (left panel), the difference score was marginally greater than zero in the MDD group, $t(23) = 1.77$, $p = 0.090$, but marginally lower than zero in the controls, $t(23) = -1.67$, $p = 0.108$, and a between-groups t-test on the difference scores was significant, $t(46) = 2.43$, $p = 0.019$, $d = 0.70$. By contrast, for words from the animacy task (right panel), the difference scores were more negative than zero in both groups, $ts < -4.29$, $ps < 0.001$, and there was no group difference, $t(46) = 0.92$, $p = 0.36$, $d = 0.27$.

3.1.5. Retrieval: confidence—Figure 3 shows the confidence data. These were characterized by a main effectof *Task*, $F(1, 46) = 7.91$, $p = 0.007$, $\eta_p2 = 0.15$, reflecting higher confidence in response to words from the mobility vs.animacy task , as well as aneffect of *Cue* , $F(1,46) = 24.85$, $p < 0.001$, $\eta_p2 = 0.35$, due to higher confidence under the Question vs. the Side cue. However, this was qualified by a *Group* x *Cue* interaction, $F(1, 46) = 4.24$, $p = 0.05$, $\eta_p2 = 0.08$. Follow-up t-tests showed lower confidence for depressedadults vs.controls under the Side cue, $t(46) = 2.25$, $p = 0.03$, $d = 0.65$, but not the Question cue, $t(46) < 1$.

3.1.6. Retrieval: correct RT—Analysis of correct RT data revealed only a main effect of *Cue*, $F(1, 44) = 267.92$, $p < 0.001$, $\eta_p2 = 0.86$, due to slower correct responses under the Question cue (1686±527 ms) vs. the Side cue (974±324 ms). The main effect of *Group* was not significant, $F < 1$, and neitherwere any interactions with *Group*, $Fs < 1.64$, $ps > 0.20$.

3.1.3. Retrieval: guessing—Thirty-oneparticipants (13 controls, 18 MDD) guessed at least once in every cell defined by the*Group* x *Cue* x *Task* ANOVA (the Odd/Even condition was omitted as it elicited only a single guess). Analysis of their data yielded a main effect of *Cue*, $F(1, 29) = 7.75$, $p = 0.009$, $\eta_p2 = 0.21$, reflecting fewer guesses under the Question (5.92±3.53) vs.the Side (7.47±4.15) cue, and an effect of *Task*, $F(1, 29) = 15.98$, $p < 0.001$, $\eta_p2 = 0.36$, reflecting fewer guesses in response to words fromthe mobility (5.82±3.90) vs. the animacy (7.56±3.75) task.Neither the main effect of *Group*nor any interactions involving *Group* were significant($ps > 0.06$).

3.1.7. Behavioral summary—At encoding, mobility judgments were made more slowly and less accuratelythan animacy judgments.T his appears to have influenced retrieval as both

groups guessed less and responded more confidentlyto words from the mobility task. Furthermore, participants responded more slowly, guessed less, and were more confident under the Question cue vs.the Side cue. The cue effects on accuracy varied depending on the encoding task. Both groups were less accurate under the Question vs.the Side cue when responding to words from the animacy task. By contrast,accuracy to words from the mobility task was characterized by a *Group* x *Cue* interaction: in depressed adults, but not controls, accuracy was better under the Question vs.Side cue. Thus , the mobility task and the Question cue led to few guessesand confident responding in all participants, and the combination of these factors boostedaccuracy in the MDDgroup. Depressed adults expressed less confidence in their memories than controls did under the Side cue, but not under the Question cue.

3.2. ERPs

3.2.1. Conceptual and perceptual retrieval, collapsed over encoding task—To test our *a priori*hypothesis, we conducted between -group tests on "Question minus Odd/ Even" and "Side minus Odd/Even" difference waves, collapsed across the encoding tasks as in prior studies. Contrary to our expectations, neither contrast revealed significant group differences (smallest cluster $p = 0.29$). Thus, we present the resultscollapsed across groups in Figures 4 and 5.

Figure 4 depictsthe "Question minus Odd/Even" contrast and Table 2 lists electrodes where condition effects were observed. From 400–800 ms (top panel) there were two significant clusters. As expected, Question hits elicited more positive ERPs than Odd/Even hits over left parietal electrodes, consistent with many prior studies of recollection(Rugg and Curran, 2007) . In addition, there was a relative negativity for Question hits vs. Odd/Even hits over right frontal electrodes in this time window. As shown in the middle and bottom panels, later intervals(800 – 1400 ms,1400–2000 ms) were characterized by a sustained LPN that was larger (more negative) for Question hits vs.Odd/Even hits and centered over left frontal and right occipital sites.

For the "Side minus Odd/Even" contrast(Figure 5, Table 3), no reliable differences between conditions were seen from 400–800 ms, but strong effects were observed from 800–1400 ms and 1400–2000 ms. In these windows, a stronger LPN in response to Side vs.Odd/Even hits was seen over the posterior midline, extending from anterior parietal to occipital sites.

3.2.2. *Question* minus *Side*, animacy task—Our second ERP analysis was designed to more closely trackthe source memory results. Accuracy for words from the animacy task was worse under the Question vs.Side cue in both groups (Figures 2A and 2B, right panels). To probe the neural correlates of thiseffect, we computed "Question minus Side" difference waves for words from the animacy task. A between-groups test revealed no reliable differences (smallest cluster $p = 0.20$), thus we present data collapsed across groups in Figure 6 (also see Table 4). There wasa broadly distributed negativity that was focused over left fronto-central scalp from 400–800 ms, dispersed over bilateral fronto-central scalp from 800–1400 ms, and separated into left fronto-central and right centro-parietal clusters from 1400–2000 ms. Inspection of waveforms revealed a stable pattern: relative to Side hits,

Question hits elicited more negative potentials, with below-baseline activity especially evident from 1400–2000 ms.

3.2.3. *Question* minus *Side*, mobility task—Finally, we computed "Question minus Side" difference scores for words for the mobility task and compared responses across the two groups. As shown in Figure 7 (see also Table 5), this contrast was associated with group differences over left centro-parietal scalp between 400–800 ms (top panel) and 800–1400 ms (middle panel). In these intervals, the depressed and healthy groups generated similar responsesfor Question hits , but the depressed group showed a weaker response for Side hits. Indeed, follow-up *Group* x *Cue* ANOVAs on mean amplitudes averaged over all electrodes in these clusters (bottompanel) revealed *Group* x *Cue* interactions in both intervals(Fs > 14, ps < 0.001). Follow-up t-tests confirmed thatleft centro -parietal ERP amplitudes did not differ between the groups for Question hits (ts < 0.7, ps > 0.52, ds <0.19)in either interval , but controls generated larger ERPs than depressed participants for Side hits in both intervals (ts > 2.2, ps < 0.025, ds > 0.67). Moreover, the MDD group generated higher amplitude ERPs for Question vs. Side hits in both intervals(ts > 3.1, ps < 0.006, ds > 0.54), but controls showed the opposite pattern—larger ERPs for Side vs.Question hits (ts > 2.0, ps< 0.056, ds >0.20).

3.3 Individual Differences

Analysis of individual differences within the MDD group focused on the "Question minus Side" comparison for words from the mobility task, as this was where group differences in accuracy and left centro-parietal ERP amplitudes emerged. We first computed Pearson correlations to examine associations between source memory accuracy and depressive severity (BDI-II total), brooding rumination (RRS-Brooding), and sleep disruption (PSQI total). To correct for multiple comparisons, alpha was set to 0.0167 (0.05/3). This analysis did not yield any significant results (ps > 0.11).

Next we examined relationships with "Question minus Side" ERP difference waves, computed for words from the mobility task. To maximize sensitivity, we did not restrict the analysis to sites that showed group differences (Figure 7). Instead, we extracted data from electrodes that showed condition effects in the MDD group considered alone (Figure 8A). These electrodes were predominantly located over left parieto-occipital scalp, and significant clusters were found in all three time windows. We found a marginal relationship between accuracy and ERP amplitude from 1400–2000 ms, $r = 0.39$, $p < 0.06$; no other relationships with accuracy or confidence emerged (all other $|r|$s < 0.26, ps > 0.22). However, there were negative correlations between PSQI scores and ERP amplitudes from 400–800 ms (Figure 8B; r= −0.48, $p = 0.0191$) and 800–1400 ms (Figure 8C; r= −0.50, $p = 0.0156$). Note that only the relationship from 800–1400 ms meets the corrected threshold of $p < 0.0167$.

To confirm that these results did not simply reflect depressive severity, we computed hierarchical regressions with ERP amplitude as the criterion, entering BDI–II and PSQI scores in steps 1 and 2. PSQI predicted ERP amplitude after accounting for BDI-II in both intervals (400–800 ms; β= -0.45, p <0.0–5; 800–1400 ms; β= −0.49, $p = 0.03$), and adding PSQI improved both models (R^2s > 0.16, Fs > 4.5, ps < 0.05). Thus, "Question minus

Side"ERP amplitude over left parieto-occipitalscalp in response to words from the mobility taskwas lowest in depressed adults who reported chronic sleep disruption, and this wasnot due to greater depressive severity.

4. Discussion

This study yielded three key findings. First, source memory accuracy was characterized by a *Group* x *Cue* interaction for words from the mobility task: depressed adults showed nominally better memory under the Question vs. Side cue but controls showed the opposite, leading to a significant group difference for "Question minus Side" accuracy scores. Second, a similar pattern was observed for left centro-parietal ERP amplitudes in response to words from the mobility task. From 400–800 ms and 800–1400 ms, *Group*x *Cue* interactions emerged because controls showed larger amplitude ERPs than depressed adults in response to the Side cue, but there was no group difference under the Question cue. Third, in the MDD group the amplitude of left parietal "Question minus Side" difference scores (in response to words from the mobility task) was negatively correlated with sleep quality over the last month. These results do not confirm our prediction that MDD would selectively affect conceptual source memory, but they are consistent with predictions from the cognitive initiative framework (Hertel, 1997) and the ERP data provide useful insight into underlying neural mechanisms.

The rationale for this interpretation is as follows. First, the cognitive initiative framework predicts that memory in MDD will be most accurate when the encoding task demands sustained attention, as this should keep participants engaged and prevent them from ruminating. In the current study, the mobility task was more demanding than the animacy task, and it appears to have promoted deeper encoding (Craik & Tulving, 1975). This claim is based on the fact that encoding RT was slower on mobility (vs. animacy) trials, and because both tasks required analysis of the semantic properties of the words, we regard longer RTs as evidence of more sustained analysis—deeper processing—in the mobility task (for a similar argument applied to pleasantness vs. animacy judgments, see Dobbins and Wagner, 2005). If this reasoning is correct and the mobility task truly elicited deeper encoding than the animacy task, then one would expect words from the mobility task to elicit fewer guesses and more high confidence responses at retrieval that words from the animacy task. Indeed, this is what was observed.

Next, the cognitive initiative framework predicts that depressed adults should also benefit from sustained engagement during retrieval (Hertel, 1997). In the current study, participants were cued to retrieve the encoding judgment made for each word (Question cue), as well as the side of the screen on which each word had been presented (Side cue). Correct RT was markedly slower in response to the Question vs. Side cue, and the Question cue elicited fewer guesses and more high confidence responses than the Side cue. Furthermore, depressed adults were less confident than controls when responding to the Side cue, but not the Question cue. Collectively, these data strongly suggest that the Question cue elicited more sustained cognitive processing relative to the Side cue, with positive consequences for retrieval in general and for depressed adults in particular.

What is the nature of the sustained cognitive processes triggered by the Question cue? A series of studies on "depth of retrieval" conducted by Jacoby and colleagues (Halamish et al., 2011; Jacoby et al., 2005; Shimizu and Jacoby, 2005) and replicated by others (Marsh et al., 2009) provides valuable insight. Briefly, these studies indicate that when participants attempt to remember words associated with deep vs. shallow encoding, they re-instantiate the encoding processes to constrain the retrieval search so that it is most likely to recover relevant information. To demonstrate this effect, participants performed deep or shallow encoding for neutral words (phase one), completed source or "old/new" recognition memory tests for the deeply and shallowly encoded words in separate blocks of trials (phase two), and then completed a surprise recognition memory test for the "new" words used as lures in the prior memory tests (phase three) (Jacoby et al., 2005). The key finding was that phase three memory was better for lures that had been presented with deeply vs. shallowly encoded old words, which implies that the lures underwent deep vs. shallow processing themselves during phase two retrieval.

The interpretation of these findings is that by re-engaging encoding processes at retrieval, participants exercise "front-end" control over memory: they increase the likelihood that diagnostically relevant information about the encoding episode will come to mind for old words, while the absence of such information for new words should make it easier to correctly reject them as lures (Halamish et al., 2012; Marsh et al., 2009). This reduces the need for "back-end" control, in which post-retrieval monitoring is used to discard recovered material that is deemed inaccurate or irrelevant. In the current study, the Question cue explicitly directed participants to focus on the cognitive processes engaged during encoding, while the Side cue did not. Consequently, the Question cue—but not the Side cue—should reliably drive deep retrieval.

The ERP data corroborate this statement. Although we did not detect reliable group differences, the "Question minus Odd/Even" and "Side minus Odd/Even" contrasts yielded a spatially distinct pattern of LPN effects consistent with recapitulation of encoding processes in the former contrast, but not the latter. Specifically, both subtractions revealed LPNs over posterior sites, but only the Question vs. Odd/Even contrast showed an LPN over left frontal scalp, thus replicating work by Bergström and colleagues (2013). Although ERPs were not recorded during the encoding phase, semantic decisions such as those made during the animacy and mobility tasks are well-known to elicit sustained negative potentials over left frontal scalp during encoding (Friedman and Johnson, 2000). Thus, while the midline posterior LPNs common to both contrasts may reflect activation of retrosplenial cortex and precuneus in the service of processes (e.g., visual imagery) that support source memory in general (Ritchey et al., 2015), it appears likely that only the Question cue triggered reinstatement of encoding activity.

Furthermore, the possibility that deep retrieval reflects enhanced front-end rather than back-end control over retrieval (Halamish et al., 2012; Jacoby et al., 2005; Marsh et al., 2009) is supported by the time course of the ERP group differences. As reviewed in Rugg and Curran (2007), prior studies have consistently linked recollection to larger left-parietal ERPs from about 400–800 ms. Specifically, this ERP effect emerges when conditions associated with successful recollection of old items ("Remember" responses, hits for stimuli that underwent

deep encoding) are contrasted with correct rejections of new items, but it does not emerge when conditions associated primarily with familiarity ("Know" responses, hits for stimuli that underwent shallow encoding) are contrasted with correct rejections. By contrast, post-retrieval monitoring processes emerge over right frontal scalp at approximately 800 ms post-stimulus and can last for over a second (Hayama et al., 2008; Wilding and Rugg, 1996). In the current study, the group difference in "Question minus Side" accuracy for words from the mobility task was paralleled by a group difference in left centro-parietal ERPs that began at 400–800 ms post-stimulus; there were no group differences were over right frontal scalp during later epochs, although long-lasting positive ERPs were evident over these electrodes from about 800 ms onward (Supplemental Figure S2). Consequently, the ERP results are consistent with presence of enhanced front-end control over retrieval on *Question*/mobility trialsin depressed adults. This enhancement began relatively early and may have contributed to improved source accuracy in the MDD group.

The amplitude of the left parietal "Question minus Side"/mobility ERP effect was negatively correlated with sleep disruption over the last month in depressed adults, and this remained significant after controlling for depressive severity. We are unaware of prior ERP or imaging data on the relationship between sleep and episodic retrieval in MDD, but work in healthy adults has shown that sleep deprivation leads to decreased activation of left parietal cortex in the context of memory tasks (Chee et al., 2006); our data appear broadly consistent with these findings. We did not find a relationship between sleep and source accuracy or confidence, but we believe this topic merits continued investigation. Sleep disruption is common in depression (Tsuno et al., 2005), it has a strong negative effect on executive function (Durmer and Dinges, 2005), and impaired executive function likely contributes to retrieval deficits in depression (Dalgleish et al., 2007). Thus, further work on the relationship between disrupted sleep and episodic retrieval in depression is clearly warranted.

Finally, the "Question minus Side" contrast for words from the animacy task elicited broadly distributed, sustained negative-going potentials with a fronto-central focus, with a shift towards the left hemisphere in the 400–800 ms and 1400–2000 ms intervals. To our knowledge, this ERP effect has not been reported in prior studies of episodic retrieval and its functional correlates are unclear. However, mediofrontal negativities are elicited by outcomes that are worse than expected, and they may reflect increased cognitive control in an effort to improve performance (Potts et al., 2006). Therefore, because accuracy was relatively poor for animacy words presented under the Question cue vs. the Side cue, we speculate that these ERP effects may also reflect increased cognitive control in an attempt to improve performance.

Some limitations deserve comment. First, given the challenges associated with timely recruitment of an unmedicated MDD sample, we invited participants from a wide age range (18–62 years). Although there was no group difference in age, source memory typically declines in older adults (McIntyre and Craik, 1987) and thus it may be preferable to study the effects of MDD exclusively in younger adults. Second, this design does not involve the presentation of new items at retrieval. This confers one advantage: because all the words are "old", there is no reason to suspect that source accuracy is confounded with gross differences in recognition memory, which can be the case in designs that require old/new

judgments before source decisions are rendered (Murnane and Bayen, 1996). However, the lack of new items makes it difficult to tease apart condition effects on accuracy vs. response bias, and we cannot determine whether familiarity made any contribution to performance. We were aware of this limitation during the design of the study but decided that it was not critical because MDD affects recollection more than familiarity (Hertel and Milan, 1994; MacQueen et al., 2002); nevertheless, it would be preferable to include new items in future work. Third, the use of semantic retrieval as a control condition (as in prior studies: Bergström et al., 2013; Simons et al., 2005a,b) may be suboptimal. In particular, accuracy was near ceiling and correct RT quite fast on Odd/Even vs. Question or Side trials, raising the possibility that the LPN findings in Figures 4 and 5 may be influenced by differences in time on task. We think it is unlikely that RT can entirely explain the distinct LPN distributions seen for the "Side minus Odd/Even" and "Question minus Odd/Even", but we acknowledge that it would be preferable to have a control condition that yielded accuracy and RT results more similar to what is observed on Question and Side trials. Fourth, with 24 individuals per group, the power of this study is relatively low. The sample sizes reflect the challenges associated with recruiting unmedicated depressed participants without substantial comorbidity, and they are in line with prior studies. However, we agree with the stated need for better-powered studies in psychology and neuroscience (Button et al., 2013), and it is possible that we might have detected stronger negative effects of MDD with more participants. Other methodological changes that would likely facilitate the detection of negative effects of MDD include the use of less structured encoding and retrieval tasks, as well as the inclusion of emotionally positive stimuli, which would be expected to enhance memory in the controls more than in the depressed adults (Dillon et al., 2014). Finally, it is important to emphasize again that our *a priori* hypotheses were not supported and thus our second ERP analysis was a post-hoc analysis, which raises the likelihood of Type I errors.

In conclusion, this study advances our understanding of source memory in MDD. As predicted by the cognitive initiative framework, depressed adults performed best when deep encoding was followed by deep retrieval focused on the encoding judgments. Prior behavioral work has suggested that this combination of factors should elicit reinstatement of encoding processes and thus enhance front-end control of retrieval, so that maximally relevant information is recovered from memory. Two aspects of the ERP data support that hypothesis. First, the LPN extended over left frontal scalp for Question but not Side hits, suggesting reactivation of regions that support the semantic decisions made at encoding. Second, the group difference in "Question minus Side" accuracy for words from the mobility task was paralleled by a similar effect on left centro-parietal ERPs consistently linked to recollection and the early stages of retrieval. Finally, the amplitude of the left parietal ERP "Question minus Side"/mobility effect was negatively correlated with sleep disruption in the MDD group. Collectively, these data confirm that source memory in MDD is sensitive to factors that influence sustained attention at encoding and retrieval, and they highlight the ability of ERPs to identify specific processes that support memory and that are affected by depression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Airaksinen E, Larsson M, Lundberg I, Forsell Y. Cognitive functions in depressive disorders: Evidence from a population-based study. Psychol Med. 2004; 34:83–91. [PubMed: 14971629]

Baddeley A. The episodic buffer: a new component of working memory? Trends Cogn. Sci. 2000; 4:417–423.

Beck, AT., Steer, RA., Brown, GK. Manual for the Beck depression inventory-II. The Psychological Corporation; San Antonio, TX: 1996.

Bergström ZM, Henson RN, Taylor JR, Simons JS. Multimodal imaging reveals the spatiotemporal dynamics of recollection. Neuroimage. 2013; 68:141–153. [PubMed: 23201363]

Bremner JD, Vythilingam M, Vermetten E, Vaccarino V, Charney DS. Deficits in hippocampal and anterior cingulate functioning during verbal declarative memory encoding in midlife major depression. Am J Psychiatry. 2004; 161:637–645. [PubMed: 15056509]

Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ. 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. 1999; 18:32–42. [PubMed: 10193695]

Burt DB, Zembar MJ, Niederehe G. Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. Psychol Bull. 1995; 117:285–305. [PubMed: 7724692]

Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013; 14:365–376. [PubMed: 23571845]

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ III, CFR, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 1989; 28:193–213. [PubMed: 2748771]

Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. Parietal cortex and episodic memory: an attentional account. Nat Rev Neurosci. 2008; 9:613–625. [PubMed: 18641668]

Chee MWL, Chuah LYM, Venkatraman V, Chan WY, Philip P, Dinges DF. Functional imaging of working memory following normal sleep and after 24 and 35 h of sleep deprivation: Correlations of fronto-parietal activation with performance. Neuroimage. 2006; 31:419–428. [PubMed: 16427321]

Coltheart M. The MRC psycholinguistic database. Q J Exp Psychol Sect A. 1981; 33:497–505.

Cooney RE, Joormann J, Eugène F, Dennis EL, Gotlib IH. Neural correlates of rumination in depression. Cogn Affect Behav Neurosci. 2010; 10:470–478. [PubMed: 21098808]

Craik FIM, Tulving E. Depth of processing and the retention of words in episodic memory. J Exp Psychol Gen. 1975; 104:268–294.

Cycowicz YM, Friedman D, Snodgrass JG. Remembering the color of objects: an ERP investigation of source memory. Cereb Cortex. 2001; 11:322–334. [PubMed: 11278195]

Dalgleish T, Williams JMG, Golden AJ, Perkins N, Barrett LF, Barnard PJ, Yeung CA, Murphy V, Elward R, Tchanturia K, Watkins E. Reduced specificity of autobiographical memory and depression: the role of executive control. J Exp Psychol Gen. 2007; 136:23–42. [PubMed: 17324083]

Davidson RJ, Pizzagalli D, Nitschke JB, Putnam K. Depression: perspectives from affective neuroscience. Annu Rev Psychol. 2002; 53:545–574. [PubMed: 11752496]

Degl'Innocenti A, Bäckman L. Source memory in major depression. J Affect Disord. 1999; 54:205–209. [PubMed: 10403166]

Deldin PJ, Phillips LK, Thomas RJ. A preliminary study of sleep-disordered breathing in major depressive disorder. Sleep Med. 2006; 7:131–139. [PubMed: 16260180]

Delorme A, Makeig S. EEGLAB:An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004; 134:9–21. [PubMed: 15102499]

Dillon DG, Dobbins IG, Pizzagalli DA. Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. Soc Cogn Affect Neurosci. 2014; 9:1576–1583. [PubMed: 24078019]

Dillon DG, Pizzagalli DA. Evidence of successful modulation of brain activation and subjective experience during reappraisal of negative emotion in unmedicated depression. Psychiatry Res - Neuroimaging. 2013; 212:99–107. [PubMed: 23570916]

Dobbins IG, Wagner AD. Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. Cereb Cortex. 2005; 15:1768–1778. [PubMed: 15728740]

Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Semin Neurol. 2005; 25:117–129. [PubMed: 15798944]

Eichenbaum H, Yonelinas AP, Ranganath C. The medial temporal lobe and recognition memory. Annu Rev Neurosci. 2007; 30:123–152. [PubMed: 17417939]

Farah MJ. The neurological basis of mental imagery: a componential analysis. Cognition. 1984; 18:245–272. [PubMed: 6396031]

Farah MJ. The neural basis of mental imagery. Trends Neurosci. 1989; 12:395–399. [PubMed: 2479137]

Friedman D, Johnson R Jr. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microsc Res Tech. 2000; 51:6–28. [PubMed: 11002349]

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

Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. Psychophysiology. 2011a; 48:1711–1725. [PubMed: 21895683]

Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. Psychophysiology. 2011b; 48:1726–1737. [PubMed: 21895684]

Halamish VH, Goldsmith M, Jacoby LL. Source-constrained recall: front-end and back-end control of retrieval quality. J Exp Psychol Learn Mem Cogn. 2012; 38:1–15. [PubMed: 21859234]

Hamilton JP, Gotlib IH. Neural substrates of increased memory sensitivity for negative stimuli in major depression. Biol Psychiatry. 2008; 63:1155–1162. [PubMed: 18281017]

Han S, O'Connor AR, Eslick AN, Dobbins IG. The role of left ventrolateral prefrontal cortex during episodic decisions: semantic elaboration or resolution of episodic interference? J Cogn Neurosci. 2012; 24:223–234. [PubMed: 21916561]

Hayama HR, Johnson JD, Rugg MD. The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: specific or non-specific? Neuropsychologia. 2008; 46:1211–1223. [PubMed: 18234241]

Hertel PT. On the contributions of deficient cognitive control to memory impairments in depression. Cogn Emot. 1997; 11:569–583.

Hertel PT, Hardin TS. Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. J Exp Psychol Gen. 1990; 119:45–59. [PubMed: 2141063]

Hertel PT, Milan S. Depressive deficits in recognition: Dissociation of recollection and familiarity. J Abnorm Psychol. 1994; 103:736–742. [PubMed: 7822575]

Hertel PT, Rude SS. Depressive deficits in memory: focusing attention improves subsequent recall. J Exp Psychol Gen. 1991; 120:301–309. [PubMed: 1836493]

Holdnack, HA. Wechsler Test of Adult Reading: WTAR. The Psychological Corporation; San Antonio, TX: 2001.

Jacoby LL, Shimizu Y, Daniels KA, Rhodes MG. Modes of cognitive control in recognition and source memory: Depth of retrieval. Psychon Bull Rev. 2005; 12:852–857. [PubMed: 16524001]

Johansson M, Mecklinger A. The late posterior negativity in ERP studies of episodic memory: action monitoring and retrieval of attribute conjunctions. Biol Psychol. 2003; 64:91–117. [PubMed: 14602357]

Johnson MK, Hashtroudi S, Lindsay DS. Source monitoring. Psychol Bull. 1993; 114:3–28. [PubMed: 8346328]

Koenigs M, Grafman J. The functional neuroanatomy of depression: distinct roles for ventromedial and dorsolateral prefrontal cortex. Behav Brain Res. 2009; 201:239–243. [PubMed: 19428640]

Lopez-Calderon J, Luck SJ. ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front Hum Neurosci. 2014; 8:213. [PubMed: 24782741]

Luck, SJ. An introduction to the event-related potential technique. 2. MIT Press; Cambridge, MA: 2014.

Luck SJ, Gaspelin NS. How to get statistically significant effects in ERP experiment (and why you shouldn't). Psychophysiology. 2017; 54:146–157. [PubMed: 28000253]

MacQueen G, Frodl T. The hippocampus in major depression: evidence for the convergence of the bench and bedside in psychiatric research? Mol. Psychiatry. 2014; 16:252–264.

MacQueen GM, Galway TM, Hay J, Young LT, Joffe RT. Recollection memory deficits in patients with major depressive disorder predicted by past depressions but not current mood state or treatment status. Psychol Med. 2002; 32:251–258. [PubMed: 11866320]

Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods. 2007; 164:177–190. [PubMed: 17517438]

Marsh RL, Meeks JT, Cook GI, Clark-Foos A, Hicks JL, Brewer GA. Retrieval constraints on the front end create differences in recollection on a subsequent test. J Mem Lang. 2009; 61:470–479.

McEwen BS. Sleep deprivation as a neurobiologic and physiologic stressor: allostasis and allostatic load. Metabolism. 2006; 55:S20–S23.

McIntyre JS, Craik FIM. Age differences in memory for item and source information. Can J Psychol. 1987:175–192. [PubMed: 3502895]

Mecklinger A, Johansson M, Parra M, Hanslmayr S. Source-retrieval requirements influence late ERP and EEG memory effects. Brain Res. 2007; 1172:110–123. [PubMed: 17822684]

Murnane K, Bayen UJ. An evaluation of empirical measures of source identification. Mem Cognit. 1996; 24:417–428.

Peirce JW. Generating stimuli for neuroscience using PsychoPy. Front Neuroinform. 2009; 2:10. [PubMed: 19198666]

Pernet CR, Latinus M, Nichols TE, Rousselet GA. Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. J Neurosci Methods. 2015; 250:85–93. [PubMed: 25128255]

Potts GF, Martin LE, Burton P, Montague PR. When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. J Cogn Neurosci. 2006; 18:1112–1119. [PubMed: 16839285]

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2017.

Reitman JS, Higman B, Lifson A, Rosenblum J. Without surreptitious rehearsal, information in short-term memory decays I. J Verbal Learning Verbal Behav. 1974; 13:365–377.

Ritchey M, Libby LA, Ranganath C. Cortico-hippocampal systems involved in memory and cognition: The PMAT framework. Prog Brain Res. 2015; 219:45–64. [PubMed: 26072233]

Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. Psychol Med. 2014; 44:2029–2040. [PubMed: 24168753]

Rugg MD, Curran T. Event-related potentials and recognition memory. Trends Cogn Sci. 2007; 11:251–257. [PubMed: 17481940]

Rugg MD, Vilberg KL. Brain networks underlying episodic memory retrieval. Curr Opin Neurobiol. 2013; 23:255–260. [PubMed: 23206590]

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. 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. 1998; 59(Suppl 20):22–33.

Shimizu Y, Jacoby LL. Similarity-guided depth of retrieval: constraining at the front end. Can J Exp Psychol. 2005; 59:17–21. [PubMed: 15832628]

Simons JS, Gilbert SJ, Owen AM, Fletcher PC, Burgess PW. Distinct roles for lateral and medial anterior prefrontal cortex in contextual recollection. J Neurophysiol. 2005a; 94:813–820. [PubMed: 15728761]

Simons JS, Owen AM, Fletcher PC, Burgess PW. Anterior prefrontal cortex and the recollection of contextual information. Neuropsychologia. 2005b; 43:1774–1783. [PubMed: 16154453]

Simons JS, Peers PV, Mazuz YS, Berryhill ME, Olson IR. Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. Cereb Cortex. 2010:479–485. [PubMed: 19542474]

Singmann, H., Bolker, B., Westfall, J., Aust, F. afex: Analysis of Factorial Experiments. R package version 0.16-1. 2016. https://CRAN.R-project.org/package=afex

Song J, Davey C, Poulsen C, Luu P, Turovets S, Anderson E, Li K, Tucker D. EEG source localization: Sensor density and head surface coverage. J Neurosci Methods. 2015; 256:9–21. [PubMed: 26300183]

Starns JJ, Hicks JL. Source dimensions are retrieved independently in multidimensional monitoring tasks. J Exp Psychol Learn Mem Cogn. 2005; 31:1213–1220. [PubMed: 16393041]

Steffens DC, Otey E, Alexopoulos GS, Butters MA, Cuthbert B, Ganguli M, Geda YE, Hendrie HC, Krishnan RR, Kumar A, Lopez OL, Lyketsos CG, Mast BT, Morris JC, Norton MC, Peavy GM, Petersen RC, Reynolds CF, Salloway S, Welsh-Bohmer KA, Yesavage J. Perspectives on depression, mild cognitive impairment, and cognitive decline. Arch Gen Psychiatry. 2006; 63:130–138. [PubMed: 16461855]

Thakral PP, Wang TH, Rugg MD. Cortical reinstatement and the confidence and accuracy of source memory. Neuroimage. 2015; 109:118–129. [PubMed: 25583615]

Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: a psychometric analysis. Cognit Ther Res. 2003; 27:247–259.

Tsuno N, Besset A, Ritchie K. Sleep and depression. J Clin Psychiatry. 2005; 66:1254–1269. [PubMed: 16259539]

Watkins PC, Mathews A, Williamson DA, Fuller RD. Mood-congruent memory in depression: emotional priming or elaboration? J Abnorm Psychol. 1992; 101:581–586. [PubMed: 1500617]

Watson D, Weber K, Assenheimer JS, Clark LA, Strauss ME, McCormick RA. Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. J Abnorm Psychol. 1995; 104:3–14. [PubMed: 7897050]

Wilding EL, Rugg MD. An event-related potential study of recognition memory with and without retrieval of source. Brain. 1996; 119:889–905. [PubMed: 8673500]

Wulff K, Gatti S, Wettstein JG, Foster RG. Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. Nat Rev Neurosci. 2010; 11:589–599. [PubMed: 20631712]

Zakzanis KK, Leach L, Kaplan E. On the nature and pattern of neurocognitive function in major depressive disorder. Neuropsychiatry Neuropsychol Behav Neurol. 1998; 11:111–119. [PubMed: 9742509]

Highlights

- Depressed adults benefit from deep encoding followed by deep retrieval.

- The influence of depression on source retrieval is reflected in left parietal ERPs.

- Chronic sleep disruption may contribute to poor episodic memory in depression.
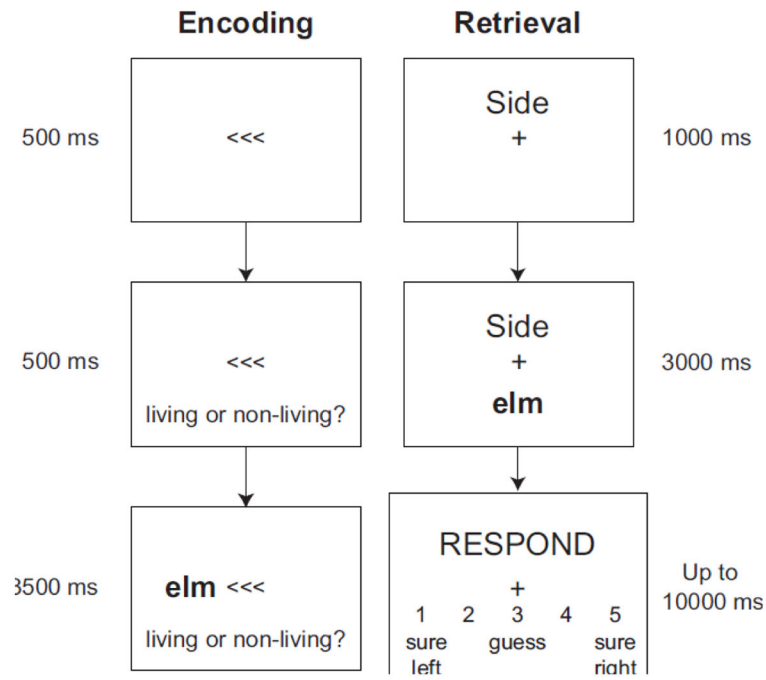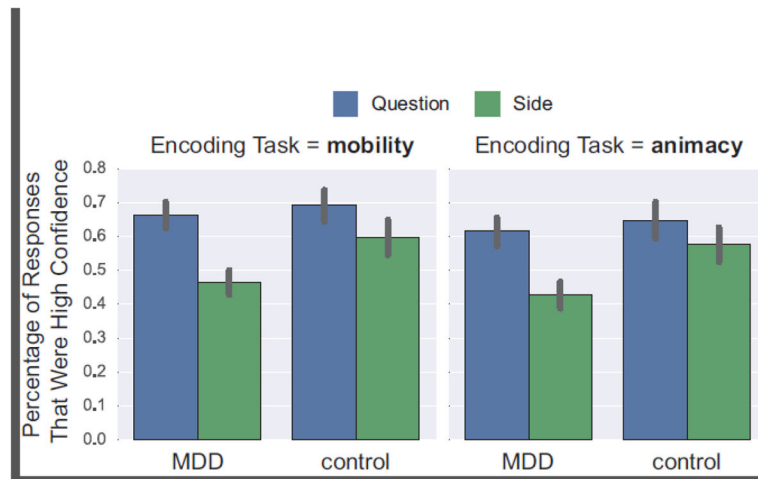
Figure 1.
Encoding (left) and retrieval (right) trial structures. Encoding trials included centrally presented arrows pointing to the side on which the word would appear, presentation of the encoding task, either "living or non-living?" (animacy judgment) or "mobile or immobile?" (mobility judgment), and then presentation of a word directly above the encoding question. Retrieval trials began with presentation of one of three cues (Side, Question, or Odd/Even), followed by presentation of a word from the preceding encoding block on Side and Question trials, or a numeral (e.g., "seventy-seven") on Odd/Even trials. Finally, a response screen was presented until the participant responded or 10 seconds had elapsed. The response options for a Side trial are shown. On Question trials, "left" and "right" were replaced with "living/non-living" and "mobile/immobile"; on Odd/Even trials they were replaced with "odd" and "even".
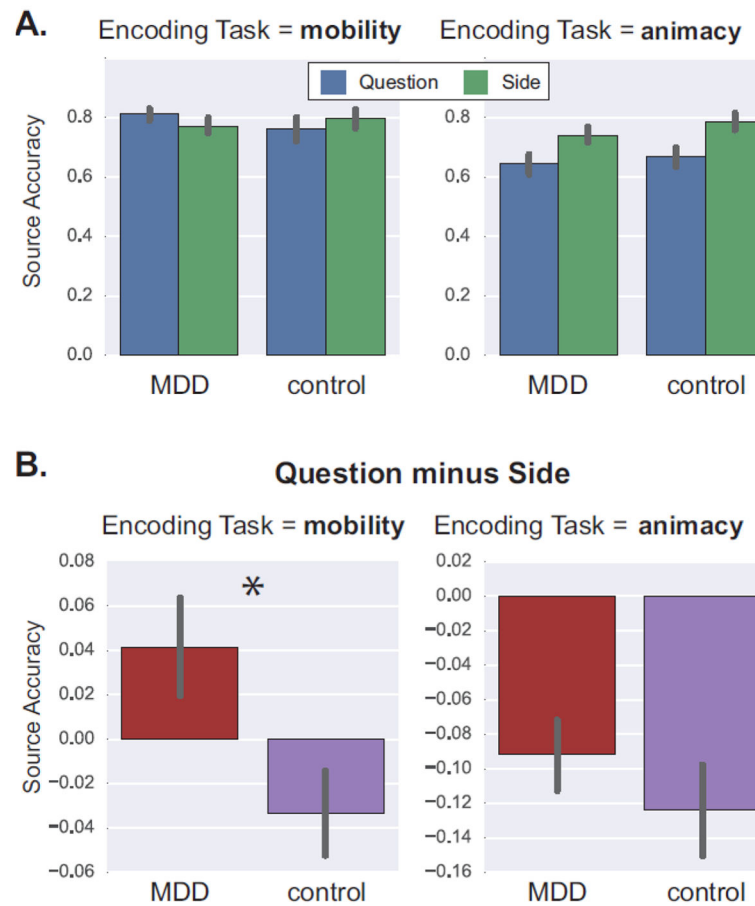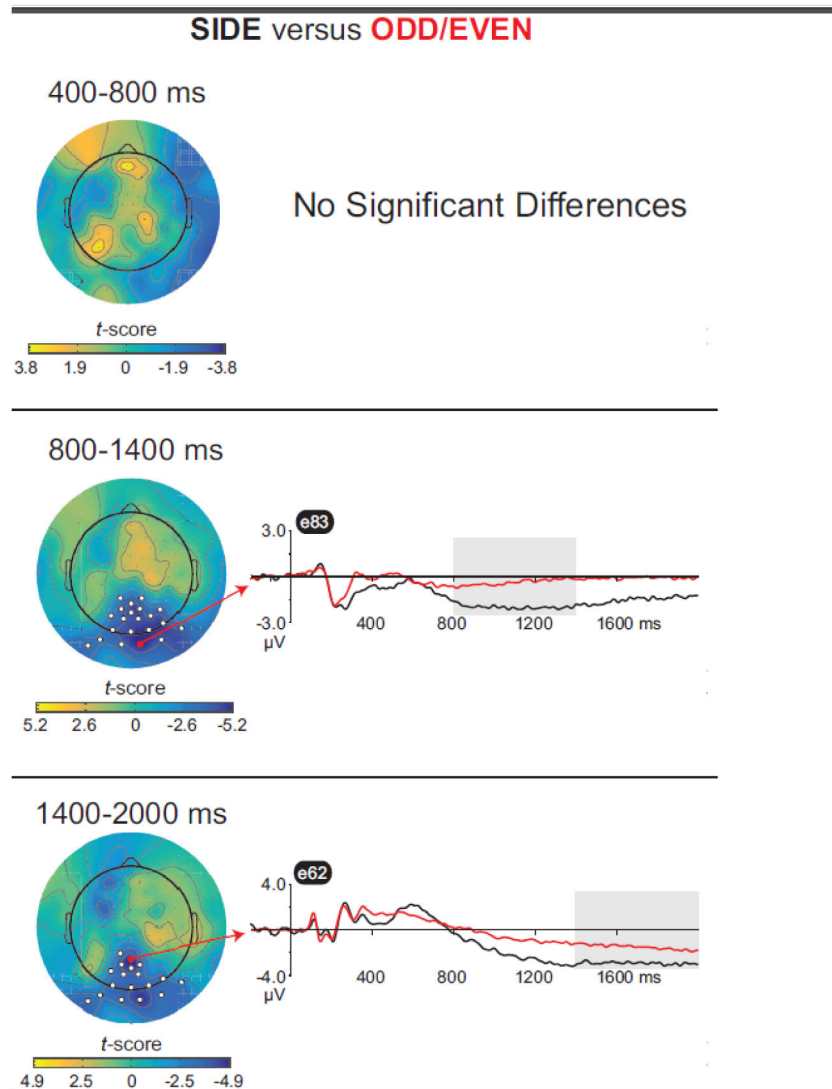
Figure 2.

Source memory accuracy (A) under Question (blue) and Side (green) cues, and for (B) "Question minus Side" difference scores. In both panels, the left column shows data for words from the mobility task, and the right column shows data for words from the animacy task. Bar heights correspond to mean accuracy, error bars = SEM, *$p = 0.019$.
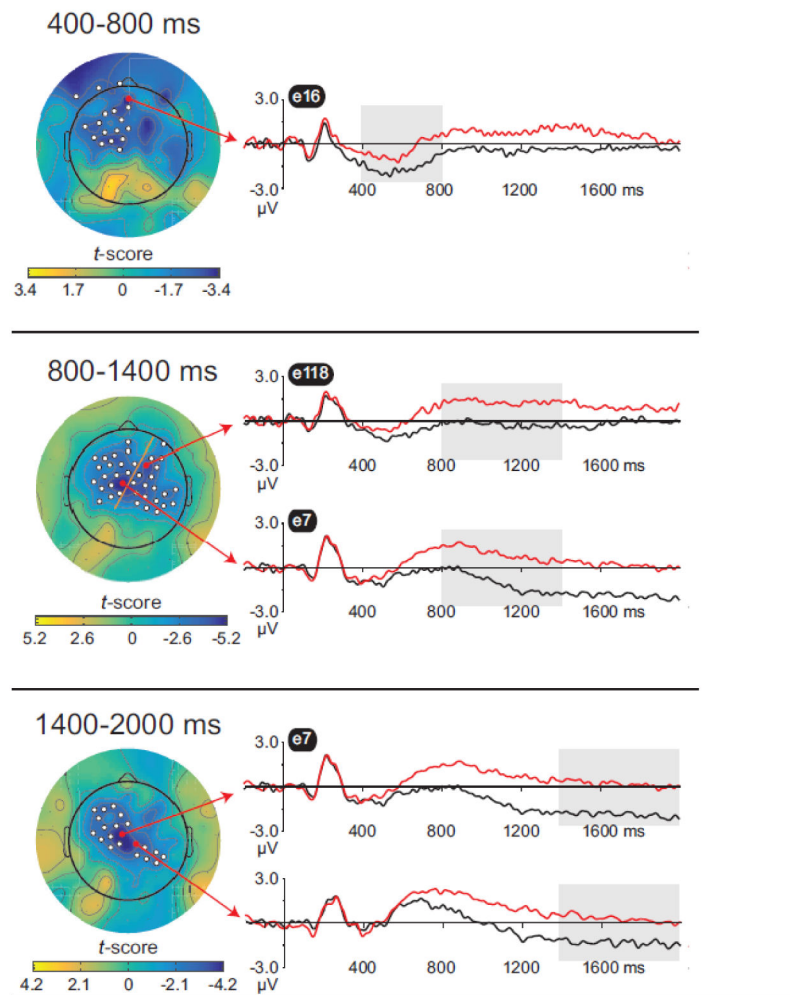
Figure 3.
Percentage of high confidence responses. Responses to words from the mobility and animacy tasks are in the left and right columns, respectively. Responses to the Question cue are in blue, responses to the Side cue are in green. Error bars = SEM, *$ps < 0.001$.

Figure 4.
Mass univariate analysis of *Question* minus *Odd/Even* difference waves, from 400–800 ms (top), 800–1400 ms (middle), and 1400–2000 ms (bottom). Data are collapsed across groups as there were no between-group differences. In Figures 4–7, electrodes in significant clusters are marked with white circles. Waveforms from the electrode that showed the strongest condition effect per cluster are plotted separately for each condition, with the time window shaded in gray. Electrode numbers (e.g., "e109") correspond to the EGI net—see Figure S1 for a complete map.
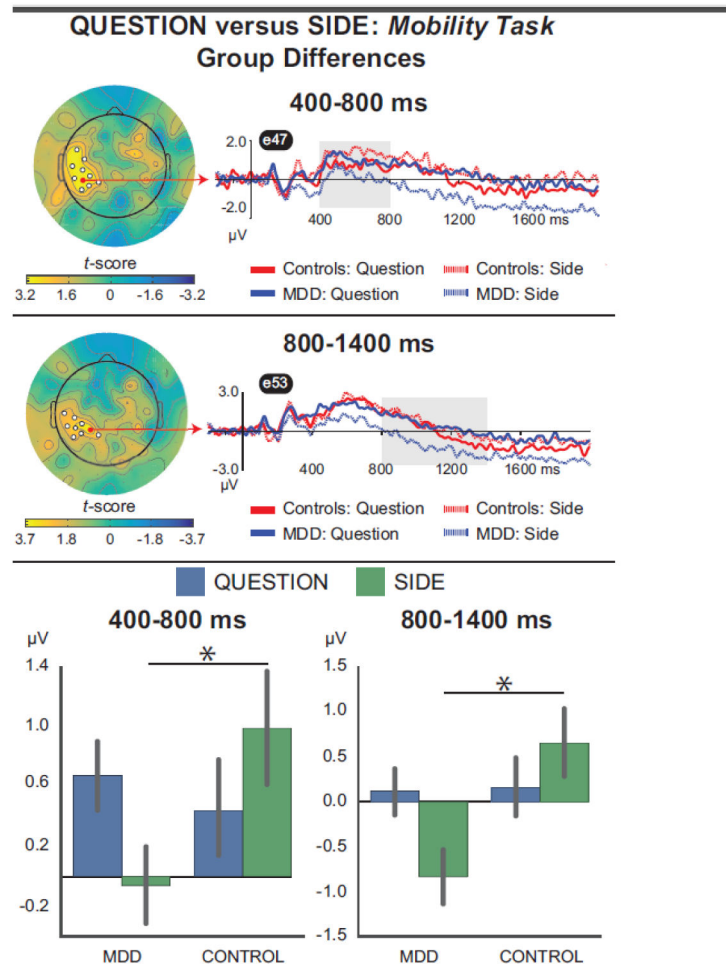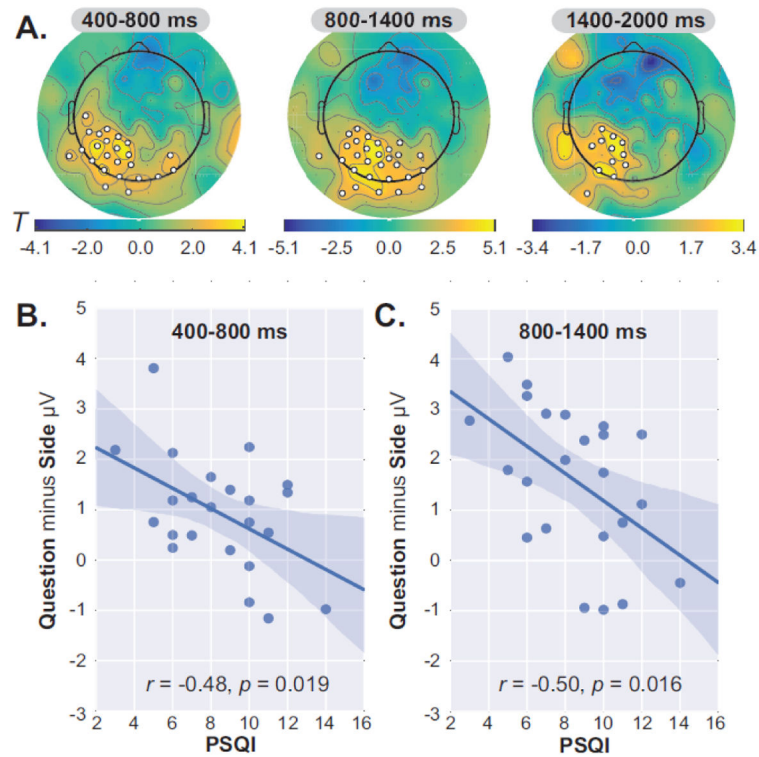
Figure 5.
Mass univariate analysis of *Side* minus *Odd/Even* difference waves, from 400–800 ms (top),
800–1400 ms (middle), and 1400–2000 ms (bottom).

Figure 6.
Mass univariate analysis of *Question* minus *Side* difference waves, for words from the animacy task, from 400–800 ms (top), 800–1400 ms (middle), and 1400–2000 ms (bottom). The orange line (middle panel) separates two clusters identified in the 800–1400 ms interval.

Figure 7.
Topographies show group differences in the mass univariate analysis of *Question* minus *Side* difference waves, for words from the mobility task, from 400–800 ms (top) and 800–1400 ms (middle); bottom panel shows the data averaged over all significant electrodes in each cluster, in each interval. The data were characterized by *Group* x *Cue* interactions: depressed and healthy adults generated similar responses to the Question cue, but depressed adults generated significantly weaker responses to the Side cue (*$p < 0.025$).

Figure 8.

Relationship between left parietal ERP amplitudes and sleep disruption in MDD. (A)
Topographic maps of Question minus Side differences for words from the mobility task in
the MDD group, with electrodes that showed significant condition effects highlighted in
white. (B) Negative relationships between PSQI scores and mean Question minus Side ERP
amplitudes from the significant electrodes shown in panel A from (B) 400–800 ms and (C)
800–1400 ms.

Table 1

Demographics and Mean (SD) Self-Report Data

Variable	Controls $n = 24$	Depressed $n = 24$	P	Effect size
Gender	13 f, 11 m	15 f, 9 m	0.89	0.085
Age	30.58 (11.09)	29.79 (10.62)	0.81	0.071
Education (years)	16.92 (1.98)	16.29 (2.44)	0.35	0.275
BDI-II	1.29 (2.22)	25.38 (8.69)	< .001	3.797
MASQ-GDA	13.04 (2.10)	21.38 (7.04)	< .001	1.589
MASQ-AA	17.65 (0.98)	24.00 (8.24)	.001	1.070
MASQ-GDD	13.65 (2.08)	38.46 (10.0)	< .001	3.402
MASQ-AD	45.61 (12.29)	86.54 (8.74)	< .001	3.852
RRS-Dep	17.96 (4.73)	32.96 (4.51)	< .001	3.247
RRS-Brood	7.75 (2.38)	12.54 (2.99)	< .001	1.772
RRS-Reflect	9.04 (3.80)	12.25 (2.97)	0.002	0.940
PSQI*	3.00 (2.00)	8.48 (2.73)	< .001	2.298
WTAR	116.73 (11.58)	117.09 (7.84)	0.90	0.072

Note. f = female, m = male; BDI-II = Beck Depression Inventory II; MASQ = Mood and Anxiety Symptoms Questionnaire (GDD = General Distress: Depressive symptoms, AD = Anhedonic Depression, GDA = General Distress: Anxious symptoms, AA = Anxious Arousal); RRS = Ruminative Response Scale (Dep = depression subscale, Brood = brooding subscale, Reflect = reflection subscale); PSQI = Pittsburgh Sleep Quality Index; WTAR = Wechsler Test of Adult Reading. Statistics reflect between-group *t*-tests except for gender (chi-square).

*
PSQI scores <= 5 indicate good sleep quality, scores > 5 indicate poor sleep quality. Of 24 depressed participants: 2 met criteria for generalized anxiety (past 6 months); 2 reported agoraphobia (past month); 2 reported social anxiety (past month); 2 reported panic attacks (last month); and 7 reported at least one panic attack (lifetime). Effect size: Cramer's V for Gender, otherwise Cohen's *d*.

Table 2

Results of the "Question minus Odd/Even" contrast

Time Window (ms)	Cluster location	Electrode Numbers	Cluster *p*-value (corrected)
400–800	Left parietal	51, 52, 53, 58, 59, 60, 61, 62, 64, 65, **66**, 67, 70, 71, 72, 76, 77, 78, 84, 85	0.005
	Right frontal	1, 2, **109**, 111, 114, 115, 116, 117, 118, 120, 121, 123, 124	0.021
800–1400	Left frontal	19, 20, **23**, 24, 27, 28, 29, 30, 31, 34, 35	0.026
	Right occipital	75, 77, 78, 79, **83**, 84, 89, 90, 95	0.048
1400–2000	Left frontal	6, 7, 12, 13, 18, 19, 20, 21, 22, **23**, 24, 27, 28, 29, 30, 31, 34, 35, 37, 106	0.001
	Right occipital	75, 76, 77, 78, 79, 80, **83**, 84, 88, 89, 90, 94, 95	0.018

Note. Results are collapsed across groups, as there were no significant differences between healthy and depressed adults. *Cluster location* gives the approximate position of the most significant electrode in a cluster. Waveforms in Figure 4 are from **bold** electrodes, which showed the strongest difference (smallest *p*-value) between conditions. A map of the EGI 128 channel net can be found in the Supplement (Figure S1).

Table 3

Results of the "Side minus Odd/Even" contrast

Time Window (ms)	Cluster location	Electrode Numbers	Cluster *p*-value
800–1400	Posterior midline	61, 62, 66, 67, 68, 69, 70, 71, 72, 74, 75, 76, 77, 78, **82**, 83, 84, 89, 90, 91, 95	<0.001
1400–2000	Posterior midline	61, **62**, 66, 67, 68, 69, 70, 71, 72, 74, 75, 76, 77, 82, 83, 89, 90, 95	0.001

Note. Results are collapsed across groups, as there were no significant differences between healthy and depressed adults. No significant effects were observed from 400–800 ms. *Cluster location* gives the approximate position of the most significant electrode per cluster. Waveforms in Figure 5 are from **bold** electrodes, which showed the strongest difference (smallest *p*-value) between conditions.

Table 4

Results of the "Question minus Side" Analysis for Words from the **Animacy** Task

Time Window (ms)	Cluster location	Electrode Numbers	Cluster p-value (corrected)
400–800	Left fronto-central	6, 7, 11, 12, 13, **16**, 19, 21, 24, 25, 28, 29, 30, 31, 32, 34, 35, 36	0.006
800–1400	Left fronto-central	2, 5, 6, **7**, 11, 12, 13, 16, 19, 20, 24, 27, 28, 29, 30, 31, 35, 36, 37, 42	0.008
	Right fronto-central	79, 80, 86, 87, 92, 93, 98, 102, 103, 104, 105, 106, 109, 111, 112, 117, **118**, 123, 124	0.002
1400–2000	Left fronto-central	6, **7**, 12, 13, 19, 24, 27, 28, 29, 30, 31, 35, 36, 37	0.002
	Right centro-parietal	79, **80**, 86, 87, 92, 93, 98	0.044

Note. Because no significant between-group differences were observed, this analysis was collapsed across healthy and depressed adults. *Cluster location* gives the approximate position of the most significant electrode per cluster. Waveforms in Figure 6 are from **bold** electrodes, which showed the strongest difference (smallest *p*-value) between conditions

Table 5

Group Differences in "Question minus Side" Analysis for Words from the **Mobility** Task

Time Window (ms)	Cluster location	Electrode Numbers	Cluster p-value (corrected)
400–800	Left centro-parietal	34, 35, 40, 41, 42, 46, **47**, 51, 52, 53	0.025
800–1400	Left centro-parietal	40, 41, 42, 46, 47, 51, 52, **53**, 61	0.037

Note. Cluster location gives the approximate position of the most significant electrode per cluster. Waveforms in Figure 7 are from **bold** electrodes, which showed the strongest difference (smallest *p*-value) between conditions. No significant group differences were observed between 1400–2000 ms.