Examining the Relationship Between Cortical Thickness and Memory Abilities in Major Depressive Disorder

Amy Ransohoff

University of Michigan

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of Bachelor of Arts

with Honors in Psychology from the

University of Michigan

2012

Advisors: Dr. Scott A. Langenecker & Dr. Sara L. Weisenbach

Abstract

The presence of memory deficits in Major Depressive Disorder (MDD) is well documented. Previous studies indicate particular deficits in the area of verbal learning as an indicator of overall impairment in various memory functions. Knowledge of the morphological changes taking place in the brain during depression is also a burgeoning area of research. While several previous studies have focused on subcortical structures, research on the relationship between memory abilities and cortical surface changes in depression has increased. This study proposed that cortical thinning in a young depressed sample would be associated with poorer performance on the California Verbal Learning Test-II (CVLT-II), a list-learning task with a short-term and long-term memory component. CVLT results and cortical thickness measurements (using FreeSurfer imaging software) were obtained (N = 21). Results demonstrated that this particular depressed group did not suffer from memory impairment compared to a normative sample. A correlation between cortical thickness and memory task performance was observed in the left inferior frontal and bilateral medial frontal lobes, but these finding were no longer significant after whole brain correction (p < .05). Future studies including a larger sample size and comparison with matched healthy controls could better assess the hypotheses examined in this study.

Key words: depression, cortical thickness, memory, structural MRI imaging

An Examination of Cortical Thickness and Memory Abilities in Major Depressive Disorder

In a world filled with stressors, it is natural to feel sad from time to time. Constant reminders of economic turmoil, rising levels of chronic health problems, and even unpredictable weather can send any person into an occasional mood slump. However, depression as a diagnosable disorder has implications beyond just sadness. Symptom severity as well as duration of sadness can impair both one's cognitive abilities and daily life functioning. While many are familiar with the concept of sadness as it relates to depression, there are other key symptoms used to define depression as a diagnosis.

The Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, see Appendix A) categorizes Major Depressive Disorder (MDD) by "episodes," or periods of time that occur for more than two weeks with depressive symptoms that are markedly different from previous mood. Each episode must be marked by the presence of depressed mood or loss of interest or pleasure, along with four or more additional symptoms including weight loss or weight gain, hypersomnia or insomnia, psychomotor agitation or retardation, fatigue, difficulty concentrating, or thoughts of death or suicide. All symptoms must occur in isolation of those from a pre-existing psychological disorder (such as bipolar disorder and schizophrenia), substance abuse, or bereavement due to the loss of a loved one. If a person experiences two or more episodes of MDD, it is classified as recurrent major depression (First, Spitzer, Gibbon, & Williams, 1994). A definition for remission of symptoms varies and is often dependent upon resolution of symptoms upon which the initial diagnosis was made. One example of an assessment instrument used to assess the severity of depressive symptoms is the Beck Depression Inventory-II (Beck, Steel, Ball & Ramieri, 1996), a 21-item questionnaire that determines the patient's emotional,

cognitive, and overall daily functioning (see Appendix A). Symptom severity measures can be self-administered by the patient or scored by a test administrator with patient input. Other examples of mood assessment measures include the Schedule for Affective Disorders and Schizophrenia (SADS) and the Hamilton Rating Scale for Depression (HDRS-17, see Appendix A). Based on MDD episode criteria, full remission from an episode is marked by 2 weeks to 6 months without symptoms, while recovery is marked by 2 to 6 months without symptoms (Keller, 2003). The great temporal variability in these criteria makes it unclear if there is a fixed distinction between sickness and wellness. Furthermore, the recovery process can be very gradual and nonlinear for the individual experiencing the episode.

MDD has the ability to deeply impact an individual's well being, but its effects can be far-reaching and long-lasting. In a nationwide, door-to-door survey of 9090 subjects, Kessler and colleagues (2003) determined lifetime rates of MDD at 16.6%, with nearly 90% of participants reporting their depression as moderate to very severe, and episodes averaging 16 weeks in length. Specific individual difficulties included high levels of comorbidity with other mental health disorders, predominantly anxiety disorders, substance abuse, and problems with impulse control.

Along with increased risk of psychiatric comorbidity, MDD can impact many life domains and lead to dysfunction in everyday tasks and responsibilities. This same study administered the Sheehan Disability Scale (see Appendix A) to its participants, in order to assess the extent to which depression interfered with ability to function in everyday domains. Nearly 30% of the depressed sample reported moderate impairment in all of the following areas of function: home, work, relationship, social, and overall (Kessler, Berglund, Demler, Jin, & Merikangas, 2003). This study demonstrates that the perception of cognitive impairment in MDD

patients is extensive, and memory impairment is a likely factor that contributes to this perception.

Historical Concepts Related to Memory

To understand memory deficits in depression, it is essential to first understand the theories and empirical support behind memory as it functions in healthy individuals. There are several important long-term memory distinctions; one of which is the distinction between explicit and implicit memory. Explicit memory consists of memory for specific past events and facts, whereas implicit memory consists of any information that is subconsciously remembered, such as learning to drive a car. Explicit memory is assessed with the use of items such as list-learning tasks during neuropsychological testing (Tulvig, 2002). However, while these memory sub-types address different types of items people remember, they do not speak to *how* information enters memory

The information processing model of memory outlined in Kellogg (1995) postulates three components of memory. First, information is perceived as environmental input. This could be the sound of a car alarm, image on a TV screen or even a very strong smell. Second, sensory registers (visual, auditory, haptic, or otherwise) transfer this input from immediate attention into short-term memory (STM) store. The STM is brief, lasting only approximately 20 seconds and limited on average to seven items. Finally, some items in STM are converted to long-term memory (LTM). Information is more likely to enter LTM if it is: 1) rehearsed in a strategic way, 2) coded by remembering visual or phonetic aspects of an item, and 3) retrieved from memory by utilizing strategic processes.

The ability to maintain items in LTM is dependent upon how many items are remembered and how many are forgotten. It is also possible that “forgotten” items were never

encoded. Memory of items may be affected by where in the learning sequence the information is presented. For example, the serial position effect refers to the observation that the probability of recalling items included in an explicit memory task, such as a list of words, resembles a U-shaped curve; items learned first (primacy effect) and last (recency effect) are most likely to be recalled when prompted. The primacy effect is thought to reflect recall for items that are likely to be consolidated into LTM, while the recency effect is mediated by working memory processes (items still in STM). The interference of new knowledge impacts what is remembered. Conversely, in proactive interference, past learning can impact new learning. In the context of a word learning task, presentation of List A, followed by List B, results in poorer recall of List B. Likewise, in retroactive interference, new learning impacts old learning, so the presentation of List A, then List B, results in poorer recall of List A. Proactive interference is thought to be a function of interference with weaker memory traces, whereas retroactive interference is mostly influenced by the similarity between List A and List B. The more similar List B is to List A, the more likely that memories from List A will interfere with those from List B (Kellogg, 1995).

There are also theories that LTM has unlimited storage capacity, as proposed in the schema theory, which serves to integrate similar ideas into a single entity (thus requiring less storage) and help procedural behavior become automated. Schemas also work to reduce working memory load by incorporating new elements into already existing schemata. The combined tasks of the schema structure allow LTM to assist in helping STM introduce new and related concepts become learned (Chi, Glaser, & Rees, 1982).

The mechanisms involved in STM can be defined on a cognitive and anatomical level. The two key distinctions between STM and LTM include the smaller size (5 to 9 items) and duration of memory stores in STM (seconds to minutes only). Baddeley's working memory

model proposes that executive functioning (responsible for allocation of attention) controls and/or overlaps with STM. Embedded within the executive functioning system are three components: 1) central executive, 2) phonological loop, and 3) visual-spatial sketchpad. The central executive modulates the function of the auditory loop and visual-spatial sketchpad and is responsible for assimilating newly attained knowledge with older knowledge. The phonological loop is responsible for language input, either by bringing words into STM, or using silent rehearsal (the "articulatory loop") to transfer words into LTM and prevent the memory from decaying. The scratchpad retains visual input other than language, including colors, shapes, and locations of objects in the STM. The visual-spatial sketchpad also brings items into LTM, which can be responsible for the memorization of tasks such as directions while driving or ability to navigate a building. The central executive's two "slave systems" allow the brain to retain environmental input to allow for encoding and later retrieval (Baddeley & Hitch, 1974).

An alternative neuroanatomical model demonstrates that STM occurs through interactions between brain structures. By way of consolidation, the hippocampus takes information received from stimulus inputs in the environment and strengthens their association by becoming activated when the environmental stimulus is salient. While the hippocampus plays a facilitating role in this process, LTM storage is subsequently rerouted to the neocortex over time. Hippocampal-neocortical interaction has been supported by positron emission tomography (PET), which demonstrates increased hippocampal activation during the encoding portion of memory tasks (Gazzaniga, Ivry, Mangun, 2008).

Neocortex function is not only important in LTM storage, but also is activated during the working-memory process and may be indicative of working-memory plasticity. Additional proof supporting the role of neocortical (specifically, fronto-parietal) activation as related to memory

processing has been found in many studies. For example, Olesen and colleagues (2004) trained two groups of subjects to improve their working memory skills. In experiment one, the group learned visuo-spatial working memory tasks (Spatial Span Board, Stroop, and Raven's Advanced Progressive Matrices; see Appendix A) over a five-week period. Each participant was assessed in the PET scanner twice before, and once after the completion of training. The working memory task completed during PET/fMRI was intentionally created to produce a low number of errors. Due to the intentional easy task design, accuracy did not improve between Time 1 to Time 3, but reaction time decreased. Additionally, frontal and parietal activation were observed in all three scan-sessions.

In experiment two, subjects completed the same task (with the exception that the same tasks completed during PET/fMRI was designed to be more difficult and involved using memorized cues to complete). Both accuracy and reaction times improved over the scanning periods. In contrast with the first experiment, increased activation after working memory training was seen in the following areas: middle frontal gyrus, superior and inferior parietal cortex, thalamic nuclei, and caudate nucleus. Increases in activity due to working memory training were interpreted to be the result of cortical plasticity. However, as working memory encompasses several processes, including encoding, control of attention, maintenance of information, and resistance to interference, cortical activation may reflect some or all of these separate abilities, and it would be difficult to attribute these activation changes solely to plasticity (Olesen, Westerberg & Klingberg, 2004).

Like the fronto-parietal cortex, the hippocampus may support several diverse memory functions. A meta-analysis of 52 PET studies investigating how activation in healthy control subjects was associated with1) process (encoding vs. retrieval), 2) type of stimulus involved

(verbal vs. figural), and 3) hemispheric laterality of activation (left vs. right). A summary of these studies indicated that the caudal region of the hippocampus is more activated during retrieval, while the rostral region is more highly activated during encoding. In addition, retrieval skills do not appear to be lateralized, though encoding, particularly for verbal information, appears to be specialized for the left hemisphere (Lepage, Habib, & Tulving, 1998). The meta-analysis concluded that the hippocampus plays an essential role in both the encoding and retrieval aspects of memory. However, many of the studies included in the meta-analysis assessed retrieval of items over minutes or hours, versus days. Thus it is difficult to determine if hippocampal activation in these studies reflected true LTM processing and retrieval as distinct from STM. In other words, inclusion of retrieval over a few minutes is unlikely to be comparable to retrieval after several weeks (since only LTM storage after minutes is assessed in the present study). In reality, though, it may be very difficult to disassociate the components of STM from LTM in otherwise healthy individuals, as these functions are integrated together within the cortex.

Specific Memory Deficits in Depression

STM impairment in depression has been widely implicated in several previous studies, using tasks of varying design and complexity. For example, Brand and colleagues (1992) used a 15-word list adapted from the Rey Auditory Verbal Learning Test (RAVLT, see Appendix A) to examine immediate and delayed recall, delayed recognition memory, recognition speed, omissions, and repetition errors. The experimenters first administered a list of 15 “meaningful monosyllabic words” to 24 MDD and matched control participants. The battery included three learning trials, a delayed recall trial, and a recognition list of 30 words (15 old, 15 new). MDD participants demonstrated significantly poorer recall, recognition, and immediate recall after one

learning trial, but only a significant difference in immediate recall after three trials. Overall, the MDD deficits indicated that recall and recognition memory tend to be impaired in depression at the beginning of a learning task, and while recognition becomes easier over time, recall remains a significant impairment across a task (Brand, Jolles, & Giespen-de Wied, 1992).

One task that has been frequently used to measure STM is the California Verbal Learning Test (Delis, Kramer, Kaplan & Ober, 2000; CVLT-II), which is structured based upon the RAVLT. During the CVLT, subjects listen to a set of 16 words (falling into one of four semantic categories) and are then asked to recite all of the words they can recall from the list (List A). This process is repeated four more times, for a total of five learning trials. The learning trials are followed by one trial in which participants must recall words from a distractor list (List B). Subjects are then asked to recall the words from List A (Short Delay Free Recall), followed by a recall trial of List A words with semantic prompts (Short Delay Cued Recall). Following a 20 minute delay, free and cued recall trials are again administered, followed by a recognition memory trial (Recognition Discrimination), in which participants must distinguish words from List A amongst distractor words, including words from List B and semantically similar words. A meta-analysis of previous studies using the CVLT indicated some trends among individuals with neurological deficits. Elwood and colleagues (1995) tested for memory deficits in various neurological and psychological disorders across several domains: 1) learning (Trials 1-5), 2) delayed recall, 3) recognition, 4) and intrusions (when words not presented on the initial list are incorrectly recalled). Supporting the findings of Brand and colleagues (1992), MDD participants showed significant deficits in primacy (but not recency) recall, suggesting particular difficulty in transferring knowledge into long-term store in MDD.

Other studies have reported even more extensive difficulties in CVLT performance in MDD samples. Otto and colleagues (1994) determined that CVLT scores in a MDD sample were one-half to one standard deviation below the mean, relative to previously recorded data on healthy control (HC) performance. While there was no significant correlation between depression severity (as measured by the HDRS-17) and CVLT performance, there was a significant correlation between CVLT scores and self-reported mood impairment on the Cognitive Failures Questionnaire (CFQ), a self-administered questionnaire used to assess perception, memory, and motor lapses in everyday life (see Appendix A). This relationship would indicate that perceived depressive behavior, rather than specific cognitive deficits resulting from MDD may be linked to poorer performance on cognitive tasks (Otto, Bruder, Fava, Delis & Outikin, 1994).

Potential Mechanisms of Memory Impairment in Depression

One consideration of memory deficits during MDD involves discerning the root of impairment, both with respect to why it occurs and what specific depression symptoms impact memory task completion. Considine and colleagues explored this question in a sample of 45 MDD participants with matched controls, in which the two groups were administered the Test of Memory Malingering (TOMM) and the CVLT-II. The TOMM is a memory task in which 50 images are shown to a participant, followed by 50 "forced choice" images in which participants must discriminate between 2 pictures to decide which was previously administered. With feedback, participants are expected to improve from Trial 1 to Trial 2 (see Appendix A). In essence, the task is designed to assess the effort of the participants in a forced choice format, rather than measure true learning and recall abilities. Results of this study demonstrated no significant difference in TOMM performance for MDD or HC participants, but significant

impairment for the MDD group on several CVLT measures, including: Trials 2-5, Short Delay Free Recall, Long Delay Cued Recall, and Long Delay Recognition. Given that performance on the effort-based task was similar, this study supports the idea that encoding and retrieval deficits in depressed individuals are not related to effort expended (Considine et al., 2011).

Length and severity of depression may have a cumulative effect on memory impairment over time. Basso and Bornstein (1999) examined how memory loss is distinguished in MDD patients suffering from a single episode of depression (MDD-SE) as opposed to MDD with recurrent episodes (MDD-RE). To test for this distinction, the study investigated group differences for participants diagnosed with MDD-SE versus MDD-RE. Neuropsychological measures for participants included: Vocabulary and Block design (subtests of the Wechsler Adult Intelligence Scale), F-A-S verbal fluency**,** and Trials A and B (see Appendix A). Results indicated that MDD-RE showed significant detriments in immediate recall, as well as short-and long-delay cued recall when compared with MDD-SE participants. CVLT scores from the RE group were not only significantly lower than scores from the SE group; they were also below the published CVLT norms (indicating mild impairment). However, scores in the SE group were well within the normal range. This study provides us with valuable information about the cumulative effects of MDD on ability to complete memory tasks. Whether the cognitive deficits indicate a risk factor for multiple depressive episodes, or occur as a result of depressive episodes, is yet to be determined.

Structural Hippocampal Changes in MDD Potentially Linked to Memory Difficulties

Beyond the memory deficits that have been documented in MDD, there is also evidence that specific brain changes occur during depressive episodes. As previously mentioned, the hippocampus plays a mediating role in transferring environmental input to the brain from STM

to LTM. In fact, there is already a firmly established, proposed mechanism by which hippocampal damage may occur. Hypothalamic-pituitary-adrenal axis (HPA) dysfunction has been implicated in acute MDD along with other psychiatric disorders. Possible mechanistic theories for this damage include unchecked stress-induced glucocorticoid (GC) emission through the body that, at high levels, could lead to neurotoxicity in the hippocampus. Furthermore, on a subtle level, if the GC levels do not result in cell death, hypercortisolemia can still impact synaptic organization and cell packing density of the hippocampus. Hippocampal neurotoxicity has been previously confirmed in stress-induced rats (Sapolsky, Krey, McEwen, 1986), which is why hippocampal volume has been a primary focus as the mechanism by which memory can be impacted in MDD.

In the first of a series of hippocampal studies conducted by Sheline and colleagues, the ratio of total hippocampal volume to total brain volume was measured in MDD and HC middle-aged women. Results demonstrated that left and right hippocampal gray matter (GM) increases were greater in MDD participants while total cerebral volumes remained the same relative to the HC group. The discovery of apparent GM increases seemed contradictory in light of previous research that had demonstrated hippocampal decreases following stress in rat studies. However, the Sheline study found a higher level of low-signal foci (LSF) in MDD compared to HC participants. LSF are areas within the hippocampus that appear darker during structural MRI scans and are often indistinguishable from surrounding cerebrospinal fluid. Though this study failed to demonstrate hippocampal volume deficits, the LSF points did indicate potential structural abnormalities in MDD participants (Sheline, Wang, Gado, Csernasky, & Vannier, 1996).

Sheline and colleagues also tested remitted depressed females to investigate how taking antidepressants impacts hippocampal volume (although not all participants underwent treatment) MRI scans demonstrated that more time spent with untreated depression was correlated with *decreased* hippocampal GM volume (Sheline, Gado, & Kraemer, 2003). However, there was no significant relationship between number of days of pharmacological treatment and hippocampal GM volume. What would warrant further investigation would be a clear-cut definition of treatment (length of treatment, medication type, if therapy were involved), in a prospective study to determine if certain treatment types resulted in less hippocampal damage than others (Sheline, Gado, & Kraemer, 2003). What the study *did* emphasize was that preventative treatment might be essential in people at risk for depression in order to curb hippocampal damage before it begins.

Another study by Bremner and colleagues (2004) investigated hippocampal functioning using a memory task during PET scanning in HC and MDD participants. All subjects (18 MDD, 9 HC participants) underwent 4 separate FDG-PET scans. Before scan 1 and scan 2, subjects were read a list of 10 paired words and subsequently recalled certain aspects of the list four times, examples of which included remembering the number of times words contained a particular letter (such as "d") during the list readings. During scans 3 and 4, participants were read a paragraph, asked to remember the story with an image in their head, and had to recall the paragraph five minutes after scans 3 and 4. There were no significant differences in paragraph recall scores or weakly encoded material for MDD and HC subjects. However, when compared with MDD, HC participants showed greater activation in the right hippocampus, along with the cerebellum, bilateral anterior cingulate, and amygdala during paragraph recall. In comparison, MDD participants showed decreased activation in the right middle/inferior frontal gyrus,

cerebellum, and left inferior parietal lobule (Bremner, Vythilingam, Vermetten, Vaccarino, & Charney, 2004). Because performance on the memory task was similar between HC and MDD groups, this study suggests that functional abnormalities in memory processing may be present, even when performance deficits are not.

Arguably the most comprehensive article highlighting brain regions associated with depressive effects also comes from a meta-analysis compiled by Sheline and colleagues (2003) highlighting brain changes in early-onset and late life depression (examining samples of both young and older adults). Results of the meta-analysis concluded that participants have demonstrated volumetric losses in the frontal cortex (up to 7%), hippocampus (8-19%), loss of normal symmetry in the amygdala, and basal ganglia losses, particularly in late-life depressive samples (Sheline, 2003). Mechanism of cortical loss may be related to the HPA-axis, with areas that contain high amounts of GC receptors, including the hippocampus, amygdala, and prefrontal cortex, most strongly implicated. Finally, the article addresses the question of whether brain changes lead to depression, or vice versa, suggesting that damage to regions important to emotion (by GC increases or otherwise) cause emotional changes over time that may lead to depression.

Cortical Volume Changes in MDD Potentially Linked to Memory Difficulties

With some lingering questions as to how hippocampal dysfunction relates to and might negatively impact MDD, we turn to other brain structures as potential key players in the role of STM dysfunction in MDD. This shifts our focus to cortical GM and white matter (WM), a topic that has been less studied as related to memory, but still has important implications for learning. Because GM and WM surround the brain in its entirety (including poor demarcations of different

gyri and Brodmann areas), specific parcellation or tracing techniques must be used to quantify them.

One of the more common ways to quantify GM/WM involves calculating density via voxel-based morphometry (VBM), developed by Ashburner and Friston (1999). This method has become one of the leading methods to measure volumetric concentration of GM/WM, and it has been used previously to study other brain-altering disorders, including schizophrenia and autism. MRI image analysis must go through a series of steps before volumetric concentration can be adequately measured. First, the image must be spatially normalized, in which all of the brain images taken during the experiment are overlaid with a general brain image template in order to make sure measurements are not skewed by individual anatomical differences. Next, image partitioning examines voxel intensities in order to differentiate GM from WM and cerebrospinal fluid (CSF). Next, the GM segments are preprocessed on a voxel-by-voxel basis. Finally, the statistical analysis uses a General Linear Model (GLM) to determine GM density in relation to variables being assessed.

A specific program developed to quantify cortical thickness and distinguish GM/WM boundaries is FreeSurfer imaging software (developed by Fischl & Dale, 2000). In the past, cortical thickness measurements were manually conducted, a labor-intensive process that requires a well-trained individual to determine reliable measurements of the cortex. For this reason, Fischl and colleagues developed an automated program that allowed for efficiently measuring cerebral thickness and statistically comparing thickness across several subjects in a shorter amount of time. The FreeSurfer program was developed by measuring cortical thickness in HC participants and validating the measurements by comparing them with post-mortem findings in matched subjects.

FreeSurfer works by quantifying pial surface thickness through an estimation of GM voxels per total MRI volume. Once these boundaries are determined, the computer program measures thickness for individual subjects and then averages all subjects by aligning each image. Cortical GM thickness was calculated in the left hemisphere for 30 HCs subjects (ages ranging from 20-37). Overall, areas of greater thickness were found in the gyral regions (M = 2.7 mm) and lesser thickness in the sulcal regions (M = 2.2 mm). There *was* natural variance in several areas of the brain among subjects, predominantly in the anterior, ventral, temporal, prefrontal cortices. To investigate whether this was true inter-subject variability or variance due to measurement noise, the researchers re-tested the same subjects, and compared the results from these findings in a recent study using manually calculated postmortem sulcal and posterior bank measurements, as measured by a trained anatomist. Findings demonstrated a high level of agreement between current procedures used and postmortem findings. Some technical considerations to be mindful of when analyzing cortical thickness in FreeSurfer include adjusting contrast-to-noise ratio, as well as being aware of the normal variability in degrees of myelination that occur across the cortical surface (Fischl & Dale, 2000).

Several studies have already used the VBM technique to examine GM and WM deficits in depression. Shah and colleagues used VBM to measure GM density in current, treatment-resistant patients (2 years + MDD), recovered MDD patients, and HCs. In addition to MRI acquisition, all participants underwent a full neuropsychological battery including the RAVLT (see Appendix A). Results demonstrated that subjects with chronic MDD had lower scores on the RAVLT and related measures when compared with the recovered and HC groups. VBM analysis determined several regions of GM reductions for non-remitted MDD participants, including in the left temporal neocortex and left anterior hippocampus. However, there were also areas of

higher density in the cuneus/precuneus grey regions in MDD participants. Remitted MDD and HC subjects showed no significant differences in overall GM density. Some potential explanations suggested for the brain atrophy reported in this study include external factors related to MDD (reduced/increased food intake, weight loss/gain), and dysregulation in the HPA-axis as a factor in cell death. Future research might determine the cognitive implications of cortical atrophy in MDD. This study suggests that cortical abnormalities observed in MDD may be a result (rather than a preceding factor) of the illness, indicating the need to treat MDD as soon as possible to lessen the chance of permanent brain changes during the disease (Shah, Ebmeier, Glabus, & Goodwin, 1998).

Salvadore and colleagues (2011) addressed inconsistencies in previous VBM analyses of cortical deterioration and/or diminished size by using a larger sample to investigate GM and WM density *and* volumetric abnormalities in unmedicated remitted and unremitted and MDD participants (rMDD and dMDD, respectively). GM density reduced in dMDD when compared with HCs and rMDD. The rMDD participants had certain areas of *higher* GM density when compared with HC. Decreased WM volume was noted in the middle frontal gyrus for the dMDD group. Finally, unmedicated participants demonstrated specific GM abnormalities in the dorso-lateral prefrontal cortex (DFPLC) and in the medial prefrontal cortical regions within the visceromotor network, noted by elevated glutamatergic levels monitored via PET scan (Salvadore et al., 2011). This study leaves us with some lingering questions about cortical changes in MDD. For example, we must try and surmise if GM reductions are reversed in remission, or if repeated depressive episode experiences serve as an indicator of hereditary or acquired cortical deficits.

To build upon knowledge of cortical abnormalities in MDD, Peterson and colleagues assessed whether individuals with *familial risk* for depression had significant levels of thinning in the limbic system and frontal cortices. This study examined three generations of (related) individuals with histories of extensive MDD (high-risk) or no history of MDD (low-risk). Participants also completed neuropsychological testing assessing inattentiveness, hyperactivity, and impulsivity using the DuPaul-Barkley Attention Deficit Hyperactivity disorder assessment (see Appendix A). MRI imaging revealed significant cortical thinning in the lateral right hemisphere for the high-risk group, and observed that thinning was associated with severity of depression but not with length of depressive episode. Furthermore, right hemisphere thinning was related to decreased performance in the attention and visual memory tasks. Some potential explanations for the cortical deficits in this study include exaggerated arousal and vigilance responses that may also impact processing and recall abilities. There is also a possibility that having familial depression may limit development of cortical regions that support these cognitive skills. These findings suggest two important concepts: that cortical thinning may precede MDD symptoms and may indicate a risk factor for depression, and that cortical thinning may be indicative of other cognitive deficits, such as difficulties with attention and impulsivity (Peterson et al., 2005).

Anatomical Deficits Correlated with Memory Changes

Research over the past several years has expanded to focus on how measures of cortical thickness and volume relate to cognitive abilities. One study focused on memory retention skills as they relate to cerebral and hippocampal volume in HC participants aged 20-88 (Walhovd et al., 2006). A total sample of 71 participants were administered the standard CVLT-II with the additional component of an approximate 83-day delay administered (without prior warning) by

phone. Cerebral volume was measured using general linear models (GLMs), calculating the effects of thickness at each vertex on different memory variables, and a segmentation process used to differentiate the hippocampus from other brain structures. Results demonstrated that areas of cortical thickness in the left hemisphere were associated with higher recall of the words after several months, but not after the 5- or 30-minute delay trials. The sample was divided into a "low" and "high" memory group based on number of items remembered after the 83-day (average) delay. In this sample, the "high" group showed greater thickness in the hippocampal and parahippocampal gyrus, gyrus rectus, middle frontal gyrus, parieto-occipital sulcus and the lingual gyrus of both hemispheres. Two potential explanations for these distinctions include a predisposed, anatomical advantage for participants with increased cortical volume. It is possible, if less likely, that these anatomical changes arose as a result of the testing and reflect changes associated with memory processes, including learning and rehearsal.

Another study expanded upon the research examining cortical thickness in memory by including participants with a history of mood disorder. The study utilized VBM to detect GM and cognitive performance deficits for MDD participants between ages 30 and 45. A group of 15 MDD and 14 HC matched participants were administered a neuropsychological battery that included: tonic and phasic alertness (tAL/pAL), divided attention test (DA), verbal and spatial span, and Wisconsin Card Sort Test (WCST, see Appendix A). Additionally participants completed tests to assess emotional functioning, including the Montgomery-Asberg Depression Rating Scale (MADRS; Appendix A), a series of ten statements used to assess emotional, cognitive, and general lifestyle functioning. Behavioral results indicated that MDD patients performed worse on the DA, verbal working memory, and spatial working memory tasks, with no significant differences for tonic or phasic alertness.

Based upon the sample, structural abnormalities appeared between three to seven years of the participant's first depressive episode, with MDD participants showing reduced GM concentrations in the bilateral inferior frontal gyrus, left inferior temporal gyrus, and right orbitiofrontal cortex, medial frontal gyrus, and transverse temporal gyrus. The MDD group also showed less GM volume in the bilateral thalamus and left hippocampal and cingulate gyri. There was also a significant correlation between higher MADRS scores and reduced GM volume in the right orbitofrontal cortex, but no significant correlation between structural abnormalities and alertness, working memory, or inhibition (Vasic, Walter, Hose, & Wolf, 2009). Although this study did not find a significant correlation between GM abnormalities and cognitive deficits, the sample size was small ($N = 15$). Future research could use a larger sample to detect a potential correlation.

Present Study Aims and Hypotheses

While a great deal of research has examined cortical and subcortical deficits as they relate to depression, few have focused on overall neocortical thickness as it relates to specific performance variables on memory tasks. To address this challenge, the present study will examine the differences in cortical thickness measurements and memory task abilities (CVLT-II) in an actively depressed group ($N = 21$). Based upon the previous research indicating both cortical thinning and memory deficits in MDD, it is hypothesized that decreased memory scores will be correlated with decreased cortical thickness. Specific regions for which we anticipate this correlation include: the hippocampal and parahippocampal gyrus, middle frontal gyrus, lingual gyrus, and the cingulate gyrus. Specifically, we expected to see cortical decreases in the tasks assessing recency phase of learning, STM, and LTM.

Method

Participants

This study consisted of 21 participants, (16 female, 5 male) with MDD. All subjects were previously recruited through the University of Michigan Depression Center either for cross-sectional or treatment studies of MDD. Each of these studies focused on emotional processing and executive functioning deficits in depression. Visit one included the Structured Clinical Interview for DSM-IV (SCID-I; First et al., 1994) administered for diagnostic purposes. Depression severity was assessed using the 17-item HDRS (see Appendix A). Eligible participants went on to complete a targeted neuropsychological battery, including the CVLT-II and related measures. Participants also underwent functional MRI with a GE Signa 3 T scanner, which included the structural scans being used for the present study, typically within one or two weeks of completion of the neuropsychological battery. Participants were compensated between $90 and $270 for these components of the study.

Those included were young adult to middle age (M = 32.24, SD = 13.35), with a mean education level of 16.05 years (SD = 1.95). All participants with a SCID-I diagnosis of MDD and a cut-off score ≥ 15 on the HDRS-17 qualified for inclusion, making them eligible also for this study (M *HDRS* =17.36, SD = 4.05). Exclusion criteria for this group included a history of serious medical illness bipolar disorder, history of drug, alcohol, or tobacco dependence within the past five years, comorbid non-mood psychiatric disorder with the exception of an anxiety disorder, or head injury with loss of consciousness > 5 minutes. Because the participants in this sample had been previously selected for studies with differing recruitment criteria and was cross-sectional in design, some (n = 7) were receiving psychotropic medication at the time of the study, which included a combination of SSRIs, benzodiazepines, and sedative hypnotics.

Measures

California Verbal Learning Test-II (CVLT-II; Delis et al., 2000). All subjects completed the CVLT, along with other neuropsychological measures. During the CVLT, subjects listened to a set of 16 words and were asked to recite all the words they could recall from the list. After repeating this process four times, subjects were asked to recall the words after a distractor list was read, and finally, to recall the list of words with semantic category cueing After a 20-minute delay, the subjects were asked to recall the list again, using free recall, cued recall, and recognition discrimination, without having it re-read. Specific items examined included standardized variables for number of items remembered after five trials (Trial 5 z-score, learning score), cumulative number of items remembered in the first five trials (Trials 1-5 T-score, STM score), and long term recognition (Long Delay Free Recall, LTM score), and number of items recognized from a list of correct items plus distractors, after a 20-minute lapse.

MRI Procedure

During the MRI session, subjects underwent MRI scanning appropriate for volumetric measures. A GE Signa 3T scanner was used to obtain images, with the vast majority acquired sagitally with between 114-124 high-resolution Fast SPGR IR anatomic images. The image matrix was 256 x 256 over a 24 cm field of view with fixed voxel sizes of 1.00-1.02 x 1.00-1.02 x 1.2 mm voxel. Typically, subject scans took place within one month of administration of the neuropsychological battery, which included CVLT administration.

FreeSurfer Protocol. FreeSurfer was used to quantify whole-brain cortical thickness. Freesurfer's main task is to quantify the cortical surface thickness through an estimation of WM voxels per total MRI volume. Processing for each image is completed through a series of steps: the Talairach transform, followed by WM and pial edits, and finally skull stripping (Fischl & Dale, 2000). During Talairach transform, WM/GM boundaries are determined, and each subject

must be adjusted to fit a computer-generated template. WM edits are manually made corrections to select missing or eliminate non-WM that appears on the MRI image. Pial edits are done to ensure that pial matter is not being calculated as GM. Finally, skull-stripping is performed in cases where skull is being picked up on the images within the WM/GM surfaces.

Manual edits on each participant averaged around 6 hours per subject, while FreeSurfer processing lasted for approximately 36 hours per subject, totaling to approximately 42 hours per subject. Inter-rater reliability for manual edits was established with Tricia Merkley, M.S. of the VA Ann Arbor Healthcare System, who is experienced in use of FreeSurfer. Inter-rater reliability was performed by calculating left and right hemisphere thickness in five randomly selected individual subjects (including images edited by each of the different raters). Cortical thickness measures were automatically quantified in 34 regions, with special attention paid to several regions: middle temporal, parahippocampal, superior frontal, rostral middle frontal, lateral orbitofrontal, isthmus cingulate, entorhinal, caudal middle frontal, and caudal anterior cingulate regions. The average thickness of each hemisphere was compared in SPSS 19 utilizing the Analyze-Scale-Reliability measure. Cronbach's Alpha scores ranged from .890 to .989 between all five subjects (see Table 3), indicating a high level of inter-rater reliability (George & Mallery, 2003). Once reliability was established, FreeSurfer edits were conducted by this writer and three research associates following a 3-hour tutorial and training session reviewing FreeSurfer guidelines and editing techniques. A "tracking" spreadsheet was utilized to ensure that each subject underwent identical processing (initial reconstruction, Talairach transform, and three final reconstructions with WM, skull and pial edits).

Statistical Analysis

Once subject processing was completed through all of the first-level FreeSurfer protocol, second-level processing was completed within Query, Design, Estimate, Contrast (QDEC). QDEC is a FreeSurfer application to quantify cortical thickness for each subject, and then to correlate these values with CVLT task performance. Variables used in this specific model were the following CVLT sub-scores for each participant: Trial 5 z-score, Trial 1-5 cumulative score, and Long Delay Free Recall. Covariates for the model included age, education, and gender. The QDEC program then generated three-dimensional right and left hemisphere models indicating areas of correlation between cortical thickness and CVLT score (see Figures 1-3). Statistical thresholds were uncorrected at $p < .05$ due to the small sample size. Finally, cluster-based analysis was used to label coordinates within cortical regions that were significantly correlated with the CVLT measures.

Results

Behavioral Demographic Results

A total of 21 MDD participants were included in this study. Three CVLT measures were focused on in this study, including the Trial 5 z-score, Trial 1-5 total T-score, and Long Delay Free Recall z-score. Mean and standard deviation scores for each measure are listed in Table 1. All three standardized scores were within normal range, indicating no significant task impairment in this particular sample. Correlations were also conducted between HDRS-17 scores and CVLT measures (see Table 2). Correlations for the CVLT 5 z-score and Long Delay Free Recall were negative, indicating that an increase in depressive severity was related to a decrease in these CVLT sub-scores.

Structural MRI Results

The QDEC program was utilized to evaluate areas of correlation between CVLT sub-scores and cortical thickness (see Tables 4-6). For the Trial 5 z-score, areas of positive

correlation included the frontal caudal middle region in the right hemisphere, and the rostral middle and supramarginal region in the left hemisphere. Areas of negative correlation included the postcentral occipital region in the left hemisphere and caudal middle and superior frontal regions in the right hemisphere.

For the Trial 1-5 T-score, there was a positive correlation for the superior frontal and lateral occipital cortices in the left hemisphere, and the insula and frontal superior region in the right hemisphere. Areas of negative correlation included the cuneus and postcentral occipital region in the right hemisphere, and superior temporal and pars opercularis in both hemispheres.

Finally, areas of positive correlation for the Long Delay Free Recall z-score were found in the pars opercularis, fusiform gyrus, and posterior temporal region in the left hemisphere, and frontal rostral middle and superior region in the right hemisphere. Areas of negative correlation for this measure included the posterior cingulate, fusiform gyrus, frontal superior regions in the left hemisphere and precentral frontal region in the right hemisphere. When testing for multiple comparisons using a Monte Carlo simulation at the $p < .05$ significance level, these correlations were no longer significant.

Discussion

This study sought to determine the relationship between scales of auditory memory learning and consolidation abilities and cortical thickness in a MDD sample. Each subscale included in the study assessed a different skill related to memory. As previously mentioned, the Trial 5 z-score assessed learning ability, the Trial 1-5 cumulative T-score assessed learning and cumulative STM storage, and the Long Delay Free Recall Score assessed LTM storage. For each subscale score, there were several correlated regions, both positive and negative. There were several regions of positively correlated cortical thickness including: bilateral frontal caudal,

rostral, and superior frontal gyrus and the posterior temporal and lateral occipital regions in the left hemisphere. Regions of negatively correlated cortical thickness included the caudal middle, superior, and precentral frontal regions in the right hemisphere, superior temporal gyrus in the left hemisphere, and pars opercularis region bilaterally.

Behavioral Results Summary

Memory Differences in Depression based on Age. An important consideration to make in this study are the results of the CVLT testing for this sample, and how that may have impacted the potential for cortical thickness correlations. The CVLT Trial 5 z-score and CVLT Long Delay Free Recall z-score were both positive (0.39 and 0.41, respectively), indicating that they fell above the standardized mean for those particular sub-scores based upon educational attainment and gender. These results beg the question, *who* is experiencing memory deficits in a depressed sample? There are several ways of conceptualizing this problem. One is to consider that depression, compounded with age, is more likely to demonstrate the memory deficits that may not yet be visible in a younger population. The average sample age (late 20s) may have been too young to demonstrate cumulative memory difficulties. Salloway and colleagues (1996) explored the question of age by conducting both structural MRI and neuropsychological testing (including the CVLT) on an elderly early onset and late onset depressed group. The late onset group performed worse when compared to early onset on several of the CVLT sub-scores including two examined in this study, CVLT 1-5 T-score and Long Delay Free Recall z-score. There were also decrements in other tasks, including COWAT, a verbal fluency task (see Appendix A). The results of this study indicate that depression may pose as a risk factor for memory impairment, particularly in specific depressed populations, such as late onset depression.

Memory Differences in Depression Based on Depressive Severity. Another important variable to consider is the impact of single episode (MDD-SE) versus recurrent episode (MDD-RE) on memory abilities, which was not controlled for in this study. To reiterate the importance of a study cited earlier, Basso and Bornstein's research (1999) on MDD-SE versus MDD-RE impairment CVLT testing revealed subtle and specific differences in performance: MDD-RE performance fell in the mildly impaired range based on published norms, while MDD-SE performance was within normal range (0.50 deviation below the mean). Additional differences in CVLT performance included poorer short-term recall, cued recall, and recognition for the MDD-RE group. Nonetheless, CVLT performance was similar on other measures, including semantic and serial cluster ratios and proactive and retroactive inhibition. The subtle memory differences that appear as a result of multiple episodes of depression may not have been adequately accounted for in the present study.

Similar discrepancies in MDD participants hold true for differences in performance in relationship to other variables, such as depressive severity at the time of testing. For example, neuropsychological testing completed by Beblo and colleagues determined improved performance on several measures for depressed participants who received treatment during the testing process. Specific enhanced performance was found in the subtests in the WAIS-R and COWAT (both listed in Appendix A), (Beblo, Baumann, Bogerst, Wallesch, & Herman, 1999). While learning and memory were not specifically assessed in this study, it suggests that several deficits may vary based upon depression severity. Had the present study separated the sample by number of previous depressive episodes, we may have observed slight differences in performance. Furthermore, memory deficits in depression can be highly selective, and even if not revealed by the specific measures in this study, they still may have been present in the sample.

Areas of Positive Correlation of Cortical Thickness with Memory Performance

A notable distinction in the positively correlated regions of cortical thickness was the localization of clusters in the frontal lobe and cingulate gyrus, whereas negatively correlated regions were spread out more arbitrarily throughout the cortex. The positively correlated regions found in this study were consistent with those discussed in previous literature, specifically in the frontal cortex in both hemispheres as documented by several previous studies (Bremner et al., 2004, Sheline, 2003, Walhovd et al., 2006). A similar study specifically focused on CVLT performance in patients with frontal lobe lesions. Deficits in this group included impaired learning and recall among patients with lesions in left posterior dorsolateral frontal and posterior medial frontal regions (Alexander, Stuss, & Fansabedian, 2003). The left posterior lesion group is particularly notable, as the strongest positive correlation in this study's sample was found in the left inferior frontal gyrus.

How exactly does thinning in the frontal cortex impacts related memory structures? One study reported losses in specific frontal cortical regions (rostral and caudal orbitofrontal cortex), including glial and neuronal cell decreases among patients with depression. The proposed mechanism by which these changes may occur is the Limbic-Cortical-Striatal-Palladial-Thalamic Tract (LCSPT) as an area of focus for several mental disorders. Specific structures of interest within the tract include the amygdala, hippocampus, mediodorsal nucleus in the thalamus, and medial and ventrolateral sections of the prefrontal cortex, caudate, putamen, and global pallidus. The article proposes that underactive dopamine in the forebrain by way of the LCSPT leads to disinhibition of the thalamus, prefrontal cortex and amygdala, potentially resulting in guilty or ruminative thoughts and motoric slowing (Sheline, 2003). These dysfunctions, particularly ruminative thoughts, could impair memory task completion. Even though we did not see as many

significant correlations in regions outside of the frontal cortex, it may be that disinhibition-related memory is interfering with memory in some patients, clouding the relationship between thickness, volume, and performance.

Potential Mechanisms of Cortical Thinning. Our present findings lead us to ask whether depressive symptoms precede or are a *result* of cortical thinning. A great deal of literature purports that cortical thinning may indicate a risk factor for developing depressive symptoms, which may explain why results were not significant in this particular instance. Peterson and colleagues (2009) proposed a model in which genetic increases in cortical thinning lead to cognitive problems including visual memory deficits and inattention, which in turn increases the risk of developing depression. This proposed mechanism of cortical thinning would make sense in light of the fact that memory deficits were not evident in this group, which may have had a lower load of familial risk for illness. Another proposed mechanism for cortical changes indicates that morphological changes may not be evident until multiple episodes have occurred. In an article previously cited, Shah and colleagues (1998) examined cortical density in treatment-resistant participants. Results indicated that left temporal cortical density was lower in treatment-resistant depression when compared with HC and remitted depressed subjects. These findings would suggest that the mechanism of cognitive changes in depression is temporal and related to length of depressive episodes (perhaps by way of HPA dysfunction), rather than genetic (or other) pre-disposing factors. Adequate understanding of cortical thinning and memory difficulties in depression, and the differences in how they function, will be essential to suitable treatment models in the future. If the proposed Peterson model holds true, then more can be done to target individuals at genetic risk for depression before onset of the illness. On the

other hand, if cortical thinning follows a depressive episode (and worsens over time), then treatment should focus on reducing depressive symptoms once they have already began.

Limitations

Although we did find areas of increased cortical thickness in consistent regions that correlated with CVLT sub-scores, these findings were not significant after correcting for multiple comparisons. It is likely that the findings were not significant due to the small sample size and inclusion of only MDD participants. Accurate and reliable FreeSurfer processing required both a great attention to detail and a large amount of time invested per subject. As emphasized earlier, an important element of FreeSurfer edits involves detailed correction to brain images in order to ensure that GM and WM has been segmented and quantified correctly. These edits often took several days for each participant, and editing participants simultaneously often caused the server to slow down or crash. These limitations prevented us from comparing depressed subjects with controls, as well as analyzing a larger depressed subject group.

Another limitation of our research was the problem of inter-rater reliability. We countered this by comparing cortical thickness measurements across five subjects, with each rater responsible for processing one subject. Despite these efforts, there are many ways in which the FreeSurfer editing process can be somewhat subjective. Edits to brain images are conducted manually and require close attention to detail, familiarity with three-dimensional brain anatomy, and consistency in editing practices between raters. The potentially subjective nature of this process, alongside the fact that edits made directly on brain images are permanent and cannot be reverted, made it difficult to assume total reliability for all edited subjects.

Future Directions

There are several ways in which the present research project could be expanded upon. A larger sample size would increase power to detect a significant finding and help raters to solidify their comfort-level and reliability with FreeSurfer processing. Along the same vein, the addition of matched controls would provide an understanding of memory task performance and cortical thickness in healthy brains, and ways in which they might deviate in those with MDD.

Finally, the addition of more memory and clinical variables (with increased statistical power from a larger sample size) would advance our understanding of how other aspects of memory and cognitive functioning impact cortical thickness. For example, the CVLT assesses several other aspects of memory functioning, such as semantic clustering (or the ability to remember groups of words in categories). Measuring these types of variables, alongside short and long delay recall measurements, would give us a more intricate understanding of how different aspects of memory are impacted in depression. In addition, many have speculated that the CVLT does not work as well at segregating functions of STM and LTM. On the whole, this research gives a solid foundation for techniques to study how depression impacts both morphological changes in the brain, along with overall cognitive functioning. The continuation of this research is essential because it demonstrates the importance of treating depression early to curb the risk of these cortical and functional changes, and there are a myriad of exciting ways in which the present study can be built upon in the future.

References

Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry: The Methods. *NeuroImage*, *11*, 805-821.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, 8*, 47–89. New York: Academic Press.

Basso, M. R., & Bornstein, R. A. (1999). Relative Memory Deficits in Recurrent Versus First Episode Major Depression on a Word-List Learning Task. *Neuropsychology*, *13*(4), 557-563.

Beblo, T., Baumann, B., Bogerts, B., Wallesch C. W., & Herman, M. (1999). Neuropsychological correlates of major depression: A short-term follow-up. *Cognitive Neuropsychiatry, 4*, 333-341.

Brand, A. N., Jolles, J., & Giespen-de Wied, C. (1992). Recall and recognition memory deficits in depression. *Journal of Affective Disorders*, *25*, 77-86.

Bremner, J. D., Vythilingam, M., Vermetten, E., Vaccarino, V., & Charney, D. (2004). Deficits in Hippocampal and Anterior Cingulate Functioning During Verbal Declarative Memory Encoding in Midlife Major Depression. *Am J Psychiatry*, *161*, 637-645.

Chi, M., Glaser, R. & Rees, E. (1982) Expertise in problem solving. In: R. Sternberg (ed), *Advances in the Psychology of Human Intelligence*, pp. 7–75. Hillsdale, NJ: Erlbaum.

Considine, C. M., Weisenbach, S. L., Walker, S. J., McFadden, E. M., Franti, L. M., Biellauskus, L. A., & Maixner, D. F. (2011). Auditory Memory Decrements, Without Dissimulation, among Patients with Major Depressive Disorder. *Archives of Clinical Neuropsychology*, *26*, 445-453.

Elwood, R. W. (1995). The California Verbal Learning Test: Psychometric Characteristics and Clinical Application. *Neuropsychology Review*, *5*(3), 173-200.

Fischl, B., & Dale, A. M. (2000, September 26). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *PNAS*, *97*(20), 11050-11055.

Gazzaniga, M., Ivry, R. B., & Mangun, G. R. (2008). *Cognitive Neuroscience: The Biology of the Mind* (3rd ed., pp. 15-37).

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K., Rush, A. J., & Walters, E. (2003) The Epidemiology of Major Depressive Disorder: results From the National Comorbidity Survey Replication. *Journal of American Medicine Association*, *298*(23), 3095-3105.

Keller, M. B. Past, present, and future directions for defining optimal treatment outcome in depression: Remission and beyond. *JAMA*, *289*, 3152-3160.

Kellogg, R. T. (1995). *Cognitive Psychology* (2nd ed., pp. 25-50). N.p.: Thousand Oaks.

Kizilbash, A. H., Vanderploeg, R. D., & Curtiss, G. (2002). The effects of depression and anxiety on memory performance. *Archives of Clinical Neuropsychology, 17*(1), 57-67.

Langenecker, S. A., Weisenbach, S. L., Giordani, B., Briceno, E. M., Guidotti Breting, L. M., Schallmo, M., & Starkman, M. (2011). Impact of chronic hypercortisolemia on affective processing. *Neuropharmacology, 62*(1), 217-255.

Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET Activations of Memory Encoding and Retrieval: The HIPER Model. *Hippocampus*, *8*, 313-322.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2003). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7*(1), 75-79.

Otto, M. W., Bruder, G. E., Fava, M., Delis, D. C., & Quitkin, F. M. Norms for Depressed Patients for the California Verbal Learning Test: Associations with Depression Severity and Self-Report of Cognitive Difficulties. *Clinical Neuropsychology*, *9*, 81-88.

Peterson, B. S., Warner, V., Bansal, R., Zhu, H., Hao, X., Liu, J., & Durkin, K. (2005). Cortical thinning in persons at increased familial risk for major depression, *PNAS*, *106*(15), 6273-6278.

Salvadore, G., Nugent, A. C., Lemaitre, H., Luckenbaugh, D. A., Tinsley, R., Cannon, D. M., & Neumeister, A. (2011). Prefrontal cortical abnormalities in currently depressed versus currently remitted patients with major depressive disorder. *NeuroImage*, *54*(4), 2643-2651.

Salloway, S., Malloy, P., Kohn, R., Gillard, E., Duffy, J., Rogg, J., & Tung, G. (1996). MRI and neuropsychological differences in early- and late-life - onset geriatric depression. *American Academy of Neurology*, *46*(6), 1567-1574.

Sapolsky, R. M., Krey, L. C., & McEwen, B. S. (1984). Glucocorticoid-sensitive hippocampal neurons are involved in terminating the adrenocortical stress response. *Proc. Natl. Acad. Sci*, *81*(19), 6174-6177.

Shah, P. J., Ebmeier, K. P., Glabus, M. F., & Goodwin, G. M. (1998). Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression. Controlled magnetic resonance imaging study. *The British Journal of Psychiatry*, *172*, 527-532.

Sheline, Y. I. (2003). Neuroimaging Studies of Mood Disorder Effects on the Brain. *Society of Biological Psychiatry*, *54*, 338-352.

Sheline, Y. I., Gado, M. H., & Kraemer, H. C. (2003). Untreated depression and hippocampal volume loss. *Am J Psychiatry*, *8*, 1516-1518.

Sheline, Y. I., Wang, P. W., Gado, M. H., Csernasky, J. G., & Vannier, M. W. (1996). Hippocampal atrophy in recurrent major depression. *Proc. Natl. Acad. Sci*, *93*, 3908-3913.

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*, 1-25.

Vasic, N., Walter, H., Hose, A., & Wolf, R. C. (2009). Gray matter reduction associated with psychopathology and cognitive dysfunction. *Journal of Affective Disorders*, *31*, 107-116.

Walhovd, K. B., Fjell, A. M., Dale, A. M., Fischl, B., Quinn, B. T., Makris, N., & Reinvang, I. (2006). Regional cortical thickness matters in recall after months more than minutes. *Neuroimage*, *31*(3).

Author's Note

Amy Ransohoff , Department of Psychology, University of Michigan, Ann Arbor.

A big thanks is in order to two terrific mentors, Drs. Langenecker and Weisenbach, who without, this study would have been entirely impossible. Their willingness to answer my trivial questions, read draft upon draft of this paper, and continually offer words of encouragement made an initially daunting task move with relative ease. It is my hope that every student who hopes to write a thesis is lucky enough to have mentors who are so attentive, accessible, and genuinely interested in supporting the work of students.

Another essential aspect of this study was technical support, which subsisted a huge amount of the work that was completed. This help was provided by (but certainly not limited to), Tricia Merkley, Chelsea Cummiford, Annie Weldon, Kortni Meyers, and Laura Gabriel. All of these individuals played an integral role in helping to de-mystify FreeSurfer processing and statistical data analysis. Their involvement in the day-in and day-out process, though often tedious, helped make this study come to life. Thank you for your patience and taking time out of your busy schedules to work with me.

Finally, a big thanks to my family, friends, and terrific boyfriend. While their most important role was positive reinforcement during late night trips to the lab and endless hours of proofreading, I would not have been as productive and successful without their encouragement. Their continuous ability to support me, even in the midst of my failures, is what enabled me to create this finished product of which I am incredibly proud. To all of you, thanks for making what seemed an impossible dream become a 52-page reality.

Additional correspondence should be sent to Amy Ransohoff, 19850 Marchmont Road, Shaker Heights, OH, 44122, amyjrans@umich.edu.

Table 1

Demographic and Behavioral Variables

	MDD (N = 21)	
Measure	Mean	SD
Age	29.13	8.62
Education	16.05	1.90
HDRS-17	17.36	4.05
CVLT 5 z-score	0.39	0.72
CVLT 1-5 Total T-score	57.88	8.03
CVLT Long Delay Free Recall z-score	0.41	1.38

Note: HDRS-17 = Hamilton Depression Rating Scale, a 17-item scale of depression severity. CVLT 5-Z = Trial 5 z-score. CVLT 1-5 Total T represents the total number of items recalled across the five learning trials (standardized in T-scores). CVLT Long Delay Free Recall z-score = the number of items (standardized in z-scores) remembered after a 20-minute delay without cues.

Table 2

Correlation Between CVLT Scores and HDRS-17 Scores

	HDRS-17
1. CVLT 5-Z	-0.32
2. CVLT 1-5 Total T	0.01
3. CVLT Long Delay Free Recall Z	-0.06

Note: CVLT = the California Verbal Learning Test-II, HDRS-17 = Hamilton Depression Rating Scale, 17-item. None of these correlations were significant at the $p < .05$ level

Table 3

Mean Cortical Thickness in Left and Right Hemispheres

	Mean Thickness (LH)	Alpha (LH)	Mean Thickness (RH)	Alpha (RH)
Subject 1	4.99	0.96	4.96	0.96
Subject 2	5.03	0.99	4.97	0.97
Subject 3	4.97	0.98	4.93	0.95
Subject 4	5.05	0.98	4.97	0.95
Subject 5	5.00	0.96	4.94	0.89
Average	5.01	0.97	4.95	0.95

Note: Mean thickness calculated as average across 24 cortical regions, including middle temporal, parahippocampal, superior frontal, rostral middle frontal, lateral orbitofrontal, isthmus cingulate, entorhinal, caudal middle frontal, and caudal anterior cingulate regions. Means listed are total averages generated by two separate raters on the same subjects. Chronbach's Alpha value listed. An alpha value of $\alpha \geq .9$ is said to reflect "excellent" internal consistency. All alpha values for these calculations fell between the "good" to "excellent" category.

Table 4

Cluster Analysis of Associations of Thickness with CVLT 5 Z-Score

Lobe/Region	Foci	x	y	z	mm^3	Z
Left Hemisphere						
Positive Correlation						
Frontal	Rostral Middle	-41	37	22	230	4.1
	Rostral Middle	-22	57	17	59	3.1
	Rostral Middle	-20	49	28	20	2.4
	Caudal Anterior Cingulate	-11	20	29	1	2.0
Parietal	Supramarginal	-41	-37	38	96	4.1
Negative Correlation						
Occipital	Postcentral	-50	-18	15	17	-2.3
Temporal	Inferior	-45	-18	-32	85	-3.4
	Transverse	-52	-17	51	32	-2.1
Right Hemisphere						
Positive Correlation						
Frontal	Caudal Anterior Cingulate	8	3	31	12	2.3
	Caudal Middle	40	10	48	71	3.0
Negative Correlation						
Frontal	Caudal Middle	26	4	47	30	-3.4
	Superior	12	-1	46	50	-2.6
	Pars Orbitalis	31	48	-11	63	-2.1
Occipital	Precuneus	5	-57	18	40	-2.4
	Precuneus	21	-57	16	20	-2.3
	Precuneus	10	56	57	3	-2.0.
Temporal	Superior	62	-7	-1	4	-2.0

Note: n = 18 areas of interest.

Table 5

Cluster Analysis of Associations of Thickness with CVLT 1-5 T-Score

Lobe/Region	Foci	x	y	z	mm^3	Z
Left Hemisphere						
Positive Correlation						
Frontal	Superior	-134	21	33	46	3.5
	Superior	-11	49	11	78	2.6
	Lateral Orbital	-37	29	-14	6	-2.0
	Medial Orbital	-13	45	-6	42	2.2
	Rostral Middle	-40	27	24	19	2.4
	Precentral	-35	-16	40	1	-2.0
Occipital	Lateral	27	-89	0	28	2.8
Temporal	Insula	-35	-30	30	11	2.3
	Lingual	-13	-59	0	3	2.0
Negative Correlation						
Frontal	Pars Opercularis	-48	15	9	11	-2.1
	Pars Opercularis	-54	21	18	32	-3.1
	Lateral Orbital	-37	29	-14	6	-2.0
Occipital	Cuneus	-3	-30	11	8	-2.1
	Postcentral	-55	-14	15	11	-2.1
Temporal	Superior	-50	-12	-4	113	-3.2
	Superior	-49	4	-17	39	-2.2
	Pole	54	5	-31	8	-2.2
Right Hemisphere						
Positive Correlation						
Frontal	Superior	-14	21	33	46	3.5
	Superior	-11	49	11	78	2.6
	Medial Orbital	-13	45	-6	42	2.2
	Rostral Middle	-40	27	24	19	2.4
Occipital	Lateral	27	-89	0	28	-2.0
Temporal	Insula	-35	-30	30	11	2.8
	Lingual	-13	-59	-0.4	3	2.3
Negative Correlation						
Frontal	Pars Opercularis	-48	15	9	11	-2.1
	Pars Opercularis	-54	21	18	32	-3.1
	Lateral Orbital	-37	29	-14	6	-2
Occipital	Cuneus	-3	-30	11	8	-2.1
	Postcentral	-55	-14	15	11	-2.1
Temporal	Superior	-50	-12	-4	113	-3.2

	Superior	-49	4	-17	39	-2.2
	Pole	54	5	-31	8	-2.2

Note: n = 33 areas of interest.

Table 6

Cluster Analysis of Associations of Thickness with Long Delay Free Recall Score

Lobe/Region	Foci	x	y	z	mm^3	Z
Left Hemisphere						
Positive Correlation		-22	56	20	3	2.0
Frontal	Rostral Middle	-22	56	20	3	2.0
Parietal	Supramarginal	-56	-21	28	55	2.7
Negative Correlation						
Frontal	Lateral Orbital	-35	28	-17	65	-2.6
	Caudal Middle	-25	11	48	66	-2.6
		-38	3	40	55	-2.5
	Pars opercularis	-52	21	18	33	-2.9
	Superior	-9	49	39	16	-2.4
	Paracentral	-14	-41	71	16	-2.4
	Lateral	-24	83	12	75	-2.6
Occipital	Postcentral	-38	-30	57	17	-2.2
	Inferior	-39	-77	12	87	-2.5
Parietal	Inferior	36	62	40	1	-2.0
	Precuneus	-8	-57	10	7	-2.2
	Postcentral	-38	-30	57	17	-2.2
	Supramarginal	-53	-46	38	0.4	-2.0
	Inferior	-48	-44	-17	69	-2.5
Temporal	Insula	-35	2	-18	16	-2.6
	Pars triangularis	-42	33	-4	0.5	-2.0
	Fusiform Gyrus	-36	-6	-40	278	-3.6
	Posterior	-4	-5	36	278	-3.5
Right Hemisphere						
Positive Correlation						
Frontal	Rostral Middle	36	36	26	81	2.9
	Superior	10	61	4	60	2.7
	Superior	17	48	-3	1	2.0
	Medial Orbital	5	44	-19	5	2.2
	Caudal Anterior Cingulate	4	3	29	2	2.1
Parietal	Posterior Cingulate	4	24	33	71	-2.5
Negative Correlation						
Frontal	Precentral	36	-13	52	70	-2.4

Occipital	Postcentral	42	-20	41	3	-2.0

Note: n = 28 areas of interest.

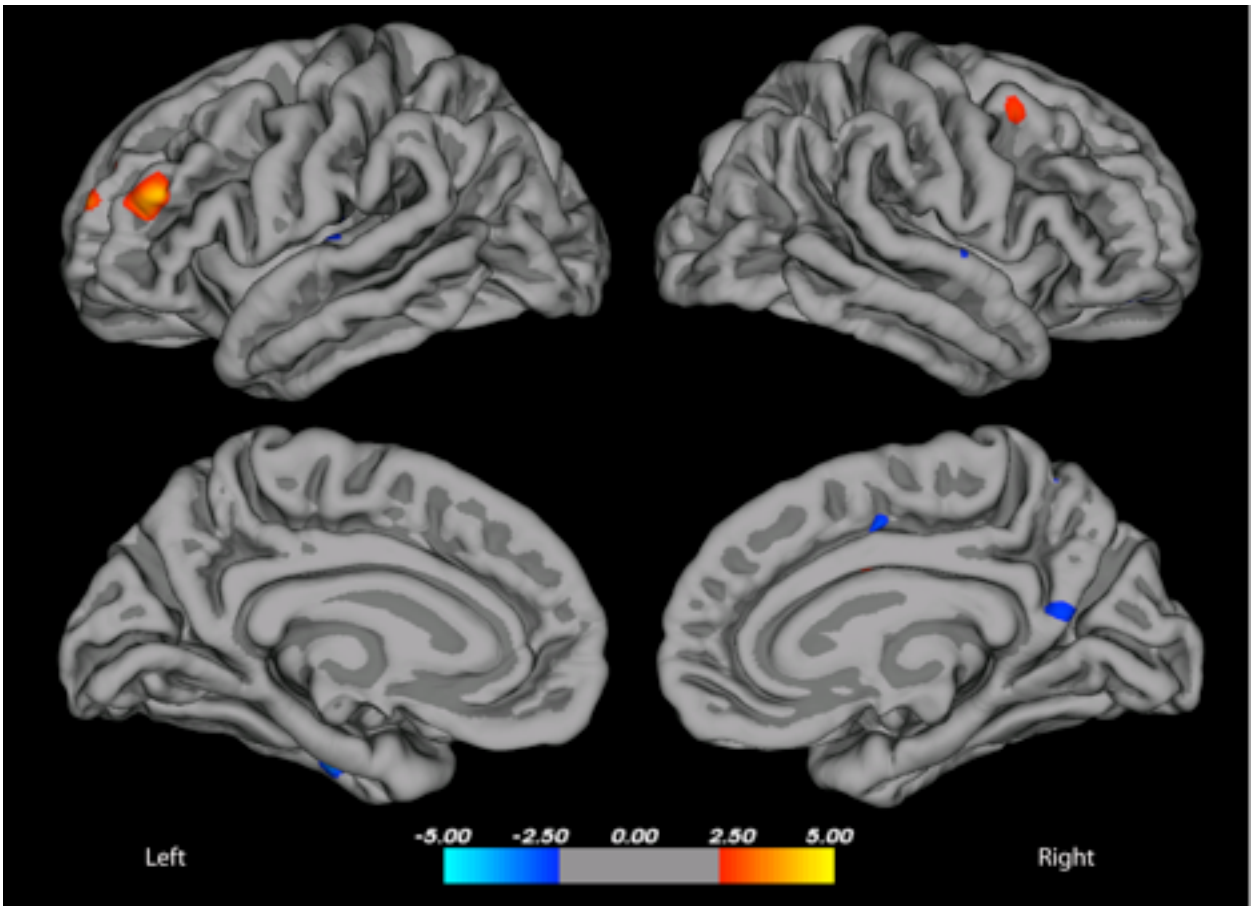

Figure 1: Correlation of cortical thickness and CVLT 5 z-score. Lateral and medial view of correlation of cortical thickness and CVLT 5 z-score. Significance threshold measured as -log10(*p*), where p is the significance. No correlations were significant at the corrected $p < .05$ level.

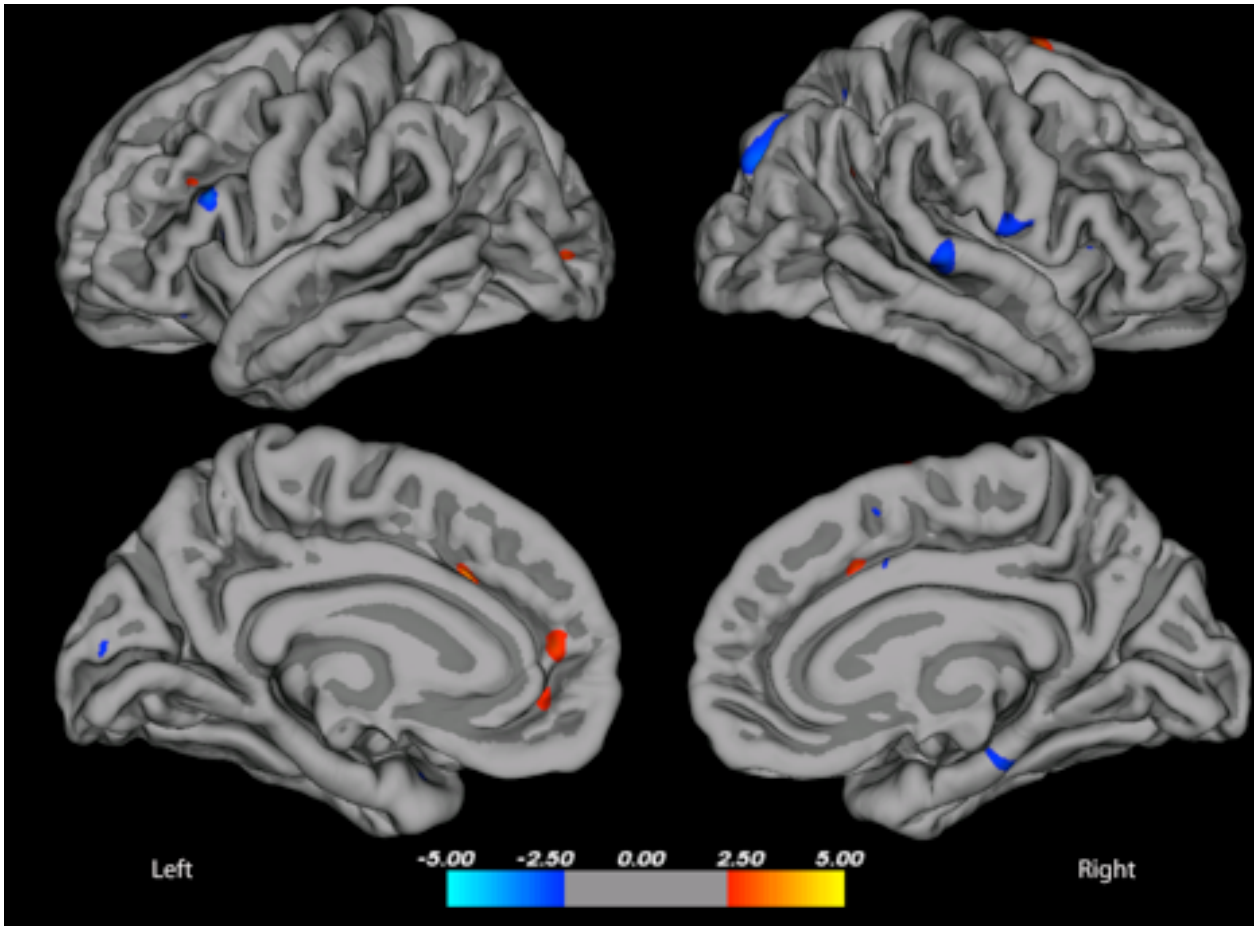

Figure 2: Correlation of cortical thickness and CVLT 1-5 Total Recall. Lateral and medial view of correlation of cortical thickness with CVLT 1-5 T-score. Significance threshold measured as -log10(*p*), where p is the significance. No correlations were significant at the corrected $p < .05$ level.

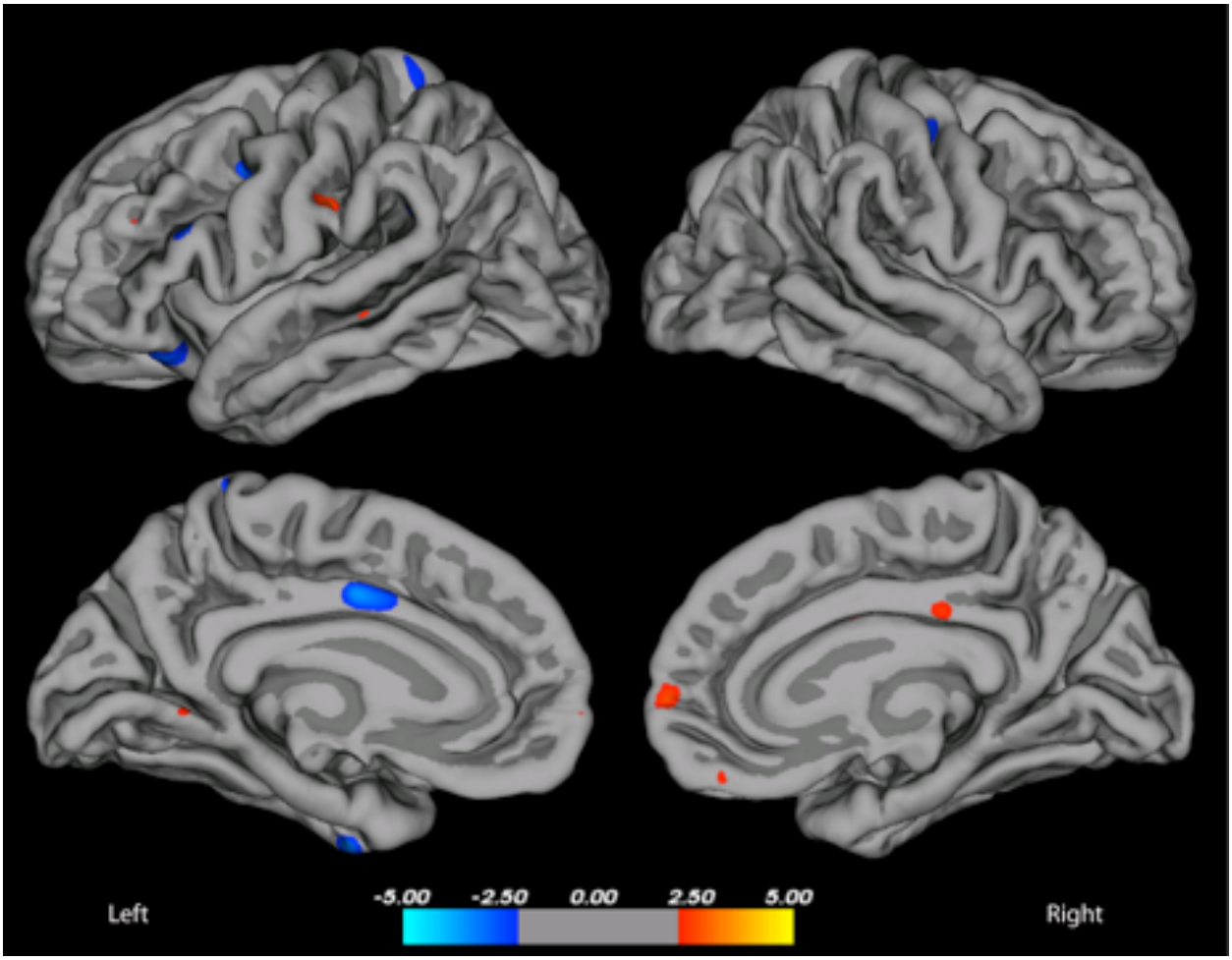

Figure 3: Correlation of cortical thickness and CVLT Long Delay Free Recall. Lateral and medial view of correlation of cortical thickness with CVLT Long Delay Free Recall z-score. Significance threshold measured as -log10(p), where p is the significance. No correlations were significant at the corrected $p < .05$ level.

Appendix A

1. Beck Depression Inventory (BDI-II): A self-administered inventory meant to assess symptoms related to depression, including changes in emotions, cognition, and ability to complete tasks. The inventory has 21 questions total and each item can be scored between 0-3 points. (Beck, Steer, Ball & Raneri, 1996).
2. California Verbal Learning Test-II (CVLT): Examinees listen to a set of 16 words and must recite all the words they recall from the list. This process is repeated four times, at which point examinees must recall the words after an intrusive list (List B) is read, and finally, recall the words in groups or categories. After a 20-minute lapse, the subjects are again asked to recall the words in various ways without having the list re-read. (Delis, et al., 2000).
3. Cognitive Failures Questionnaire (CFQ): Self-administered questionnaire used to assess perception, memory, and motor lapses in everyday life; shown to correlate with symptoms of stress (Broadbent, Cooper, FitzGerald & Parkes, 1982).
4. Divided attention test (DA): Tests one's ability to focus on simultaneously presented stimuli. Participants must focus both on a 4 x 4 visual dot formation and alternating high and low auditory pitches. When either the pitches or dot formation "match," the participant must respond by pressing a button as quickly as possible (Zimmermann & Fimm, 1994).
5. DuPaul-Barkeley Attention Deficit Hyperactivity Disorder Assessment: Neuropsychological test that assesses attentiveness, hyperactivity, and impulsivity (Barkley, DuPaul & McMurray, 1990).
6. F-A-S verbal fluency: Phonemic verbal fluency test during which the participant is given a letter of the alphabet and must recite all words beginning with that letter that come to mind within one minute. A decrease in production of words has previously been implicated in

several neuropsychological disorders, and as a natural part of the aging process (Thurstone, 1938).

7. Hamilton Rating Scale for Depression (HDRS-17): Semi-structured clinician-administered interview designed to assess depression severity. Contains 17 variables measured on either five-point or three-point scales.
8. Montgomery-Asberg Depression Rating Scale (MADRS): Depression inventory consisting of ten statements graded on a 1-5 point scale. Questions address issues related to emotional, cognitive, and general lifestyle functioning (Montgomery and Asberg, 1979).
9. Phasic alertness: Tests subject's ability to respond quickly to an auditory stimulus. An example of this involves presenting a cross on a screen where the subject responds by pressing a button as quickly as possible (Zimmermann & Fimm, 1994).
10. Rey Auditory Verbal Learning Test (RAVLT): Adaption of the original RAVLT (Binder et al., 1993) using a 15-word list-learning task to examine STM, delayed recall and delayed recognition (LTM), recognition speed, omissions, and repetition errors. In the development of the instrument, the experimenters first administered a list of 15 "meaningful monosyllabic words" to 24 MDD participants and an equal number of controls, includes 5 learning trials, a delayed recall trial, and a recognition list of 30 words (Lezak et al., 1983).
11. Raven's Advanced Progressive Matrices: Test involving 12 items meant to assess general intelligence, with each sequential item becoming progressively more difficult. Focuses on skills such as, vocabulary and spatial perception (Raven, 1935).
12. Schedule for Affective Disorders and Schizophrenia (SADS): An inventory assessing symptoms for several psychopathological illnesses. Assess both severity and duration of

mental health problems with written descriptions and assigned numeric value (Endicott & Spitzer, 1978).

13. Sheehan Disability Scale: Used to assess level of functioning with disability, including mental illness or physical impairment. Scoring system measures work/school, social life, and family and daily responsibilities each on a 10-point scale. Can be self-scored or administered by an examiner (Sheehan, 1983).
14. Spatial Span (Wechsler Memory Scale-III subtest): Requires subjects to remember the order in which a series of blocks are "tapped" by the test administrator, and then repeat this pattern back accurately directly after watching (Wechsler, 1997).
15. Stroop Test: Three-level test used to assess how interference impacts speed and learning of verbal recitation. Levels assess: 1) speed in reciting words, 2) speed in reciting colors and 3) speed in reciting color when word and color do not match, respectively (Stroop, 1935).
16. Structured Clinical Interview for the DSM-IV (SCID-I): Diagnostic tool used to determine DSM-IV mental disorders and personality disorders. Assesses past psychiatric history and description of past and present symptoms. Takes 1 to 2 hours for patients and ½ to 1 hour for non-psychiatric patients (First et al., 1994).
17. Test of Memory Malingering (TOMM): Test used to assess effort during task completion. Fifty images are shown to a participant, followed by fifty "forced choice" images in which participants must discriminate between 2 pictures to decide which one they were previously administered. With feedback, participants are expected to improve from Trial 1 to Trial 2 (Tombaugh, 1996).

18. Tonic Alertness (tAL): Tests the ability to respond quickly to visually presented stimulus. An example of this involves presenting a cross on a screen where the subject responds by pressing a button as quickly as possible (Zimmermann & Fimm, 1994).
19. Vocabulary (WAIS-III Subtest): Asks examinees to define words, scored on a 0-2 point scale (0 being definition totally incorrect, 2 meaning definition is completely correct); assesses one's ability to retain and utilized previously learned words (Wechsler, 1997).
20. Wechsler Block Design Test (WAIS-III Subtest): Sub-test of the WAIS-III in which participants are shown a geometric design and asked to replicate it with a set of pre-selected blocks. Responses are scored for accuracy and time (Wechsler, 1997)
21. Wechsler Adults Intelligence Scale-R (WAIS-R: Set of seven subtests designed to test a person's ability to learn and adapt to novel situations, ability on certain areas may be related to cognitive or psychological deficits. Full battery when administered to healthy controls ranges from 60-90 minutes, but sub-tests can be administered individually (Wechsler, 1981).
22. Wisconsin Card Sorting Test: Tests subject's "set-shifting" ability or adaption to pattern changes with instructor reinforcement. Instructor presents four "key cards" to which the participant can match the presented cards with a set of 128 cards in front of them. The participant must match the deck with the four key cars based on three patterns: color, word, and number (although this information is not presented to the subject). Once participant correctly matches by one pattern type for a set of 10 cards, the set "shifts" and the participant must now correctly match the new category. This continues until the participant has successfully matched the cards 6 times (Grant & Berg, 1948).