January 2012

Emotion-modulated startle and the course of major and minor depression

April Taylor-Clift
University of South Florida, aprilclift@gmail.com

Emotion-modulated startle and the course of major and minor depression

by

April Taylor-Clift

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
College of Arts and Sciences
University of South Florida

Major Professor: Jonathan Rottenberg, Ph.D.
Jennifer Bosson, Ph.D.
David Drobes, Ph.D.
Kristen Salomon, Ph.D.
J. Kevin Thompson, Ph.D.

Date of approval:
June 4, 2010

Keywords: psychophysiology, depression course, subthreshhold depression, affect, longitudinal

Dedication

Dedicated to my parents who taught me the importance of knowledge and wisdom and to my husband who sacrificed so much to keep me going.

Table of Contents

List of Tables

List of Figures

Abstract

Major depressive disorder (MDD) is highly recurrent. Researchers have proposed that certain traits predispose people to repeated episodes of this disorder. The current study examined the hypothesis that maladaptive emotional responding to stimuli would predict a worse depression outcome over six months. Participants were 58 individuals—18 controls, 22 individuals with MDD, and 18 individuals with minor depression (mD; subthreshold depression)—who participated in a diagnostic interview and emotion-modulated startle procedure at time one, and who returned for a second diagnostic interview six months later at time two. An identical emotion-modulated startle procedure was then repeated at time two with 33 individuals—12 controls, 14 individuals with MDD, and 7 individuals with mD. Startle probes were presented during unpleasant, neutral, and pleasant pictures, as well as during inter-trial intervals (ITI) in the absence of pictures. We used eye-blink startle responses to predict the time two level of depression severity and the likelihood of depression recurrence. Time one startle in the context of neutral pictures predicted depression outcomes at time two, such that larger time one startle responses during neutral pictures were associated with the presence of a time two depressive episode and higher time two self-report scores of depression severity (Beck Depression Scale scores). In addition, startle responses during ITIs (occurring in the absence of pictures) also predicted depression outcome, but in the opposite direction.

Specifically, larger time one startle responses during ITIs were associated with better time two depression outcomes. We discuss the implications of these results.

Introduction

Major depressive disorder (MDD) is a disorder of mood that affects one in four women and one in ten men annually and is a leading cause of disability world-wide (American Psychiatric Association, 1994; Murray & Lopez, 1997). MDD symptoms include persistent sad mood and/or loss of interest or pleasure in daily activities, as well as associated somatic and cognitive symptoms, including appetite or weight changes, sleep difficulties, psychomotor agitation or retardation, lack of energy, feelings of worthlessness or guilt, concentration difficulties, and suicidal ideation (APA, 1994). MDD is also highly recurrent. Over 70% of depressed patients have more than one episode, and most depressed patients remain symptomatic following a major depressive episode (Judd et al., 1998). Furthermore, approximately 40% of individuals with three or more episodes of depression relapse within 12 to 15 weeks of recovery (Keller et al., 1992; Mueller et al., 1996). The recurrent nature of MDD has naturally motivated a search for stable vulnerability markers that may explain why some persons are vulnerable to repeated depressive episodes.

Recently, research has focused on individual differences in emotion processing as a risk factor for depression. According to functionalist theories of emotion, emotions reflect activity in multiple response systems (including facial expressive behavior, experience, central and peripheral physiological activity, and cognitive processes) that

serve to organize an individual's behavior, prepare the individual for adaptive responses to threats and rewards, and optimize the individual's adaptation within his environment (Lazarus, 1991; Levenson, 1994; Tooby & Cosmides, 1990). Based on this definition of emotion, functionalist theories posit that emotions represent evolutionarily adaptive, specific solutions to problems. Emotions, therefore, yield important consequences for the organism, such that appropriate and flexible emotional responses will produce beneficial consequences, whereas maladaptive and inflexible emotion responses will result in negative outcomes (Keltner & Gross, 1999).

Emotional responding is associated with long-term adaptation for a number of mental disorders. For example, high levels of negative affect predict the development of alcohol use disorders (Jackson & Sher, 2003) and the development of post-disaster PTSD symptoms in young adults (Weems, Pina, & Costa, 2007). High levels of familial expressed emotion predict a more negative outcome for patients with schizophrenia, mood disorders, and eating disorders (Butzlaff & Hooley, 1998). Furthermore, a number of effective therapies, such as Mindfulness-Based Cognitive Therapy (Segal, Williams, & Teasdale, 2002) and Acceptance and Commitment Therapy (Hayes, Strosahl, & Wilson, 1999), are designed to alter the client's typical emotional responses, for example helping the client identify and accept—rather than control or suppress—emotional responses.

Although emotional processes have been demonstrated to play a significant role in the development and maintenance of several forms of psychopathology, several uncertainties remain. One complication in connecting emotion and psychopathology is that the multiple response systems of emotion do not always act in concert. For example,

Egloff and colleagues (Egloff, Wilhelm, Neubauer, Mauss, & Gross, 2002) found that implicit measures of anxiety corresponded highly to physiological responses, whereas an explicit measure of anxiety did not. A second complication is that there is no consensus as to what represents "adaptive" emotional processes (e.g., Kring & Werner, 2004; Porges, Doussard-Roosevelt, & Maita, 1994). For instance, researchers have yet to agree on the context within which certain emotion regulation strategies best serve the individual, and various therapies seek to teach either suppression of negative emotionality (e.g., Cognitive Behavioral Therapy; Beck, Rush, Shaw, & Emery, 1979) or acceptance of emotions (e.g., mindfulness based cognitive therapy; Segal et al., 2002). Finally, and most germane to the present context, prior studies aimed at explaining the role of emotion in the course of depressive disorders have not produced uniform results (reviewed below, see also: Morris, Bylsma, & Rottenberg, 2009). The goal of the current project is to utilize a physiological index of emotion reactivity to predict the course of major and minor depression.

Theories of Emotional Reactivity and the Course of Depressive Disorders

There are three main views of emotional reactivity deficits in depression. While each view has implications for the course of mood disorders, most evaluations of these views have been cross sectional (prospective tests are less common because of the high resource costs in conducting longitudinal work). The positive attenuation view hypothesizes that depressed individuals' responses to positively valenced stimuli are attenuated compared to the responses of healthy individuals (Berenbaum & Oltmanns, 1992; Rottenberg, Kasch, Gross, & Gotlib, 2002; Sloan, Strauss, Quirk, & Sajatovic,

1997; Sloan, Strauss, & Wisner, 2001). This type of affective pathology predicts a pattern of reduced responding to appetitive stimuli, such that depressed individuals will exhibit a less robust response to rewarding stimuli in the environment (Henriques & Davidson, 2000). The major implication of the positive attenuation hypothesis for depression course is that those depressed individuals who display less reactivity to positive stimuli will exhibit more severe, longer lasting, or more recurrent depressive episodes.

The negative potentiation view theorizes that the emotional responses of depressed individuals to negatively valenced stimuli are potentiated, or exaggerated, compared to the responses of healthy individuals (e.g., Beck, 1967). This theory of emotional pathology forms the basis for one of the most widely used treatments for depression, cognitive behavior therapy. In spite of its popularity as a conceptualization, cross-sectional support for the negative potentiation view remains limited (Dunn, Dalgleish, Lawrence, Cusack, & Ogilvie, 2004; Sloan, Strauss, Quirk, & Sajatovic, 1997; Sloan, Strauss, & Wisner, 2001). The main implication of the negative potentiation hypothesis for course is that those depressed individuals who display heightened reactivity to negatively valenced stimuli will exhibit a more pernicious course of the disorder (Gunthert, Cohen, Butler, & Beck, 2005).

Lastly, the emotion context insensitivity (ECI) theory hypothesizes that major depression is characterized by a general blunting of emotional responding to both positively and negatively valenced stimuli (Rottenberg, Gross, & Gotlib, 2005). This view is supported when evidence for positive attenuation is combined with evidence of

reduced reactivity to negative stimuli in MDD (Dickens, McGowan, & Dale, 2003; Gehricke & Shapiro, 2000; Peeters, Nicolson, Berkhof, Delespaul, & deVries, 2003; Rottenberg, Gross, Wilhelm, Najmi, & Gotlib, 2002; Rottenberg et al., 2002; Wexler, Levenson, Warrenburg, & Price, 1994). A meta-analytic review recently found consistent cross-sectional evidence for ECI (Bylsma, Morris, & Rottenberg, 2008). The functional implications of the ECI theory are that those depressed individuals who show blunted emotional responding to *both* negatively valenced and positively valenced stimuli will exhibit a worse course of disorder.

Emotional Reactivity and the Course of Depressive Disorders: Data

In spite of increasing interest in the functional role that emotional responding plays in the course and maintenance of psychopathology, findings regarding the role of emotion reactivity and the course of depressive disorders remain somewhat inconsistent (Morris et al., 2009). Because emotional responding involves multiple systems (e.g. cognitive, physiological, experiential, etc.), researchers have used various methodologies to measure the different components of the emotional response. These different methods often yield conflicting findings.

Several studies have supported the positive attenuation view. For instance, blunted behavioral and physiological responding to an amusing film (but not a sad or fearful film) predicted MDD non-recovery among outpatients 6 months later (Rottenberg et al., 2002). Studies relying on self-report data have found conflicting associations between positive attenuation and MDD course. For instance, higher levels of anhedonia were associated with greater chances of depression remission 7 months later in one study

(Clark, Fawcett, Salazar-Grueso, & Fawcett, 1984), but more recent studies have found higher anhedonia and lower reactivity to positive stimuli to be associated with a more pernicious course of MDD (Kasch, Rottenberg, Arnow, & Gotlib, 2002; McFarland, Shankman, Tenke, Bruder, & Klein, 2006; Moos & Cronkite, 1999; Spijker, Bijl, de Graaf, & Nolen, 2001).

Several studies support the negative potentiation view. For instance, a study involving retrospective daily self-report assessments of stressful life events found that greater reactivity to negative events predicted worse outcome to cognitive therapy (Gunthert, Cohen, Butler, & Beck, 2005). Higher self-reports on the personality trait neuroticism, indicative of greater responding toward negative stimuli, have been associated with worse depression outcome in a number of studies (Berlanga, 1999; Heinze, Torres, Apiquian, & Caballero, 1999; Enns & Cox, 2005; Katon et al., 1994; Lara, Leader, & Klein, 1997; Leskelä et al., 2006; Sakado, Sato, Uehara, Sakado, & Someya, 1999; Scott, Williams, Brittlebank, & Ferrier, 1995). Finally, higher levels of brain activation in the rostral anterior cingulate (an area associated with stress responding) predicted a worse MDD outcome (Kennedy, Koeppe, Young, & Zubieta, 2006; Koeppe, Young, & Zubieta, 2006).

Lastly, a number of studies provide support for the ECI theory, which involves both positive *and* negative attenuation. For instance, blunted reactivity to sad stimuli predicted a worse depression outcome (Rottenberg, Salomon, Gross, & Gotlib, 2005). Likewise, blunted physiologic reactivity to both positive and negative stimuli predicted worse response to antidepressant medication treatment (Fraguas et al., 2007).

Furthermore, depressed persons who exhibited blunted emotion reactivity to negative and positive daily life events before the onset of evidence-based pharmacotherapy had higher depressive symptom severity one month later and were less likely to exhibit full recovery from MDD over an 18-month follow up period (Peeters, Berkhof, Rottenberg, & Nicolson, in press).

As mentioned previously, the conflicting results may relate to the multi-system nature of emotional responding. Alternatively, the varying results of these studies may be due in part to between-study differences in depression subpopulations. For example, emotional reactivity may vary as a function of depression subtypes (e.g., melancholic versus atypical). Between-sample variations in depression severity may also be important, as depression severity has been linked to variations in emotional reactivity in cross-sectional studies (e.g., Rottenberg, Kasch, Gross, & Gotlib, 2002; Rottenberg, Chambers, Allen, & Manber, 2007). For instance, it is unclear whether subthreshold forms of depression are associated with different patterns of emotional reactivity than more severe forms of depression. Due in part to this uncertainty, some researchers argue against the use of subthreshold or dysphoric populations as an analog to depression (Sweeney, 1986; Anderson, & Bailey, 1986). Because researchers utilize either dysphoric individuals *or* major depressed individuals, we have virtually no comparative emotion data on these depressive subpopulations.

Thus far, the literature on emotional responding and the course of depressive disorders has focused almost exclusively on the role of emotion in predicting the course of *major* depression. By contrast, the role of emotion in predicting the course of less

severe mood disorders, such as minor depression, has been ignored. Minor depression (mD) is defined in the Appendix of the DSM-IV-TR as a period of two or more weeks during which at least two to four of the nine symptoms for a major depressive episode are present, and one symptom must be either depressed mood or lack of interest or pleasure in most or all daily activities (APA, 1994). Evidence of mD prevalence rates is not clear. Estimates vary depending on the degree of adherence to DSM-IV diagnostic criteria. Data from the nationally representative sample of the National Comorbidity Survey found lifetime prevalence rates for mD with no prior history of MDD of 10% (Kessler, Zhao, Blazer, & Swartz, 1997). Several studies argue for the need to consider minor depressive mood states in emotion research: mD results in substantial impairments that are similar to those of MDD, and individuals with mD may use outpatient services as frequently as individuals with MDD (González-Tejera, 2005). Individuals with mD also often experience incomplete resolution of episodes (Kessler et al., 1997), just as in MDD. In addition, minor depression results in significant functional disability and interferes with employment attendance to a similar degree as mild MDD (Cuijpers, de Graaf, & van Dorsselaer, 2004; Kessler et al., 1997). Individuals with mD are at increased risk for developing MDD compared to individuals with no depressive symptoms (Cuijpers et al., 2004). Indeed, the odds ratio for developing a first-time major depressive episode following a diagnosis of mD are as large as 5 (Fogel, Eaton, & Ford, 2006). It is perplexing that certain individuals with mD do not go on to develop MDD. Differences in emotional responding may help explain this variation in diagnostic outcomes.

Emotion-modulated Startle as a Measure of Emotional Responding

The obligatory startle response is a defensive reflex to an aversive stimulus that is characterized by a set of evolutionarily adaptive defensive behaviors, such as blinking of the eyes, forward and downward movement of the head, and drawing in of the shoulders (Landis & Hunt, 1939). In humans the startle response is typically quantified by the magnitude of the eye blink response to the aversive stimulus, called a startle probe, which is often a brief noise burst. Components of the startle response are immediate, typically occurring within 20ms of the aversive stimulus. Importantly, although the startle response can be elicited reliably, the magnitude of the response varies considerably depending on the current affective state of the organism.

Studies have continually found that emotional states elicited by affective stimuli (such as valenced pictures) modulate the amount of displayed startle magnitude (Bradley, Cuthbert, & Lang, 1990; Bradley & Lang, 2000; Vrana, Spence, & Lang, 1988). The emotional component of modulated startle appears to be mediated primarily by the amygdala, whereas the nucleus accumbens and nucleus reticularis pontis caudalis have been implicated in the role of motivation and appetitive responding (Davis, 1989, 2006; Davis, Gendelman, Tischler, & Gendelman, 1982; Lang, 1995; Lang, Bradley, & Cuthbert, 1990; Patrick, 1994; Pissiota et al., 2003). The amygdala has been implicated in emotion experience and processing in a variety of contexts (Phillips, Drevets, Rauch, & Lane, 2003; Davidson, Jackson, & Kalin, 2000). Furthermore, amygdala dysfunction is associated with a number of mood disorders, including depressive disorders (Drevets,

2003; Xiang et al., 2008). Therefore, abnormal emotion-modulated startle may provide an important index for abnormalities in amygdala dysfunction.

The emotion-modulated paradigm is ideal for measuring emotion for several reasons. Measuring emotional reactivity physiologically with the startle paradigm reduces some of the error and biases inherent in self-report measures of emotion (Lang, Bradley, & Cuthbert, 1990), and provides a relatively unobtrusive, immediate, and continuous sampling of emotional responding. The startle response is rooted in primal circuits in the brain (Lang, Davis, & Ohman, 2000), and indeed has been studied in rats (Brown, Kalish, & Farber, 1951), a variety of primates (e.g., Davis, Antoniadis, Amaral & Winslow, 2008), and even fish (e.g., Burgess, Johnson, & Granato, 2009). The paradigm allows for measurement of the motivational component of emotional reactivity, a component that is central to emotional deficits in major depressive disorder (e.g., anhedonia). In addition, there is some evidence that the emotion-modulated startle is relatively stable ($r = .89$, $p < .0001$ over four weeks; Larson, Ruffalo, Nietert, & Davidson, 2000). Lastly, the emotion-modulated startle paradigm can be, and frequently is, combined with other measures of psychophysiology to measure the multiple systems involved in emotional responding. Specifically, the emotion-modulated startle task can be accompanied by measurements of heart rate and skin conductance, as well as self-report measures of emotion. Multiple measures are essential for a complete understanding of the emotional response.

In healthy subjects, the startle probe that is presented anywhere from 3 to 6 seconds after the onset of the valenced stimuli elicits a greater startle response if the

subject is viewing unpleasant pictures (e.g., snakes, injured humans) than if the subject is viewing pleasant (e.g., families, food) or neutral pictures (e.g., landscape, garden tools), and pleasant pictures elicit smaller startle responses than neutral pictures. Essentially, the startle response is primed by negative affective states (e.g., unpleasant pictures) and attenuated by positive affective states (e.g., pleasant pictures). This pattern of emotion-modulated startle does not appear to result from unequal allocation of attention to the startle probe (Vrana & Lang, 1990; Bradley et al., 1990). The emotion-modulated startle response persists even in the presence of subjects' habituation to the startle probe, and is thus thought to represent underlying evolutionarily developed, motivational processes inherent in emotional responding (Lang, 1995). Indeed, the typical explanation for this pattern of emotion reactivity is that startle responses are compatible with aversive motivation (i.e., fight or flight mechanisms) and incompatible with appetitive motivational states (i.e., approach behaviors; Lang, Bradley & Cuthbert, 1998). This linear startle response pattern can be conceptualized as normative emotional reactivity to affective stimuli.

Use of the emotion-modulated startle paradigm has provided insight into emotional responding deficits for a number of disorders. Research evidence indicates disorder-specific patterns of atypical startle responding for a variety of psychopathology. For example, a lack of startle potentiation during unpleasant picture viewing characterizes schizophrenic patients, who also display overall deficient habituation to the acoustic startle probe (Schlenker, Cohen, & Hopmann, 1995; Taiminen et al., 2000). Incarcerated, psychopathic males display an abnormal pattern of startle modulation

compared to controls, generally showing equivalent startle responses for unpleasant and pleasant stimuli and heightened responding to neutral stimuli (Patrick, Bradley, & Lang, 1993). Patients with borderline personality disorder show increased potentiation to negative stimuli compared to healthy controls (Hazlett, Dawson, Schell, & Nuechterlein, 2007). These varied results provide some indication of the utility of the emotion-modulated startle paradigm for differentiating disorders according to emotional reactivity.

Studies of the emotion-modulated startle response in clinically depressed outpatients show reduced potentiation to unpleasant stimuli and heightened responding (i.e., lack of attenuation) to pleasant stimuli when compared to responses to neutral stimuli (Dichter, Tomarken, Shelton, & Sutton, 2004; Dichter & Tomarken, 2008), a pattern consistent with the emotion context insensitivity hypothesis (ECI; Rottenberg, Gross, et al., 2005). A flattened pattern of emotion-modulated startle has been documented in samples at various levels of depression severity (e.g., severely depressed individuals, Allen, Trinder, & Brennan, 1999; Kaviani, Gray, Checkley, Wilson, & Kumari, 2004; Forbes, Miller, Cohn, Fox, & Kovacs, 2005; individuals with non-clinical depression, Mneimne, McDermut, & Powers, 2008). Furthermore, we recently found that blunted emotion-modulated startle in depressed individuals exists even when concurrent anxiety disorders are present (Taylor-Clift, Morris, Rottenberg, & Kovacs, 2011). A blunted startle pattern indicative of ECI has even been observed in non-symptomatic Bipolar I outpatients and their unaffected siblings (Giakoumaki et al., 2010). Thus, blunted startle responding may represent a marker of emotional dysfunction associated with depressive disorders and those at risk.

Very few studies have utilized the emotion-modulated startle paradigm to predict future psychiatric outcomes. For instance, emotion-modulated startle patterns predict treatment outcomes for individuals with specific phobias (de Jong, Merckelbach, & Arntz, 1991; de Jong, Visser, & Merckelbach, 1996). Exaggerated startle responding under threat predicts the development of post-traumatic stress disorder in some (Pole et al., 2009), but not all (e.g., Griffin, 2008; Guthrie & Bryant, 2005) investigations.

Fewer studies have examined the startle paradigm as a predictor of depression outcome. A handful of studies have found relationships between baseline startle responding (i.e., startle reflexes occurring between picture presentations and in the absence of any affective stimuli) and depression outcome. For instance, higher baseline startle responses among depressed patients were associated with a positive response to antidepressant therapy (Quednow et al., 2004). Consistent with these results, lower baseline startle responses in formerly depressed individuals predicted increased depressive symptoms two years later (O'Brien-Simpson, Parsia, Simmons, & Allen, 2009). Therefore, in addition to the promise of emotion modulation of startle as a predictive tool for depression outcome, baseline startle responding may also hold predictive value.

Specific Aims

The general aim of the study was to examine whether emotion reactivity would predict the course of a major and/or minor depressive episode. The specific aims for the proposed study were as follows:

Primary aims. To determine whether emotion-modulated startle responding would predict recovery from a major or minor depressive episode six months later.

Hypothesis 1a. Major and minor depressed individuals who display the normative linear pattern of emotion-modulated startle at time one would be more likely to experience recovery six months later, at time two. This hypothesis will utilize data to predict recovery from a categorical perspective (recovered versus not recovered).

Hypothesis 1b. Based on the ECI view (context inappropriate responses to both unpleasant and pleasant stimuli will be maladaptive), we predicted that increased time one startle reactivity during pleasant stimuli and decreased time one reactivity during unpleasant stimuli would be associated with a worse symptomatic outcome at time two. This hypothesis uses data to predict recovery from a dimensional perspective (recovered versus not recovered).

Hypothesis 1c. Consistent with prior studies of startle and course of depression, higher levels of baseline startle responding at time one would predict depression recovery and/or decrease in depression symptomatology at time two.

Secondary aims. To examine the correspondence between time one emotion-modulated startle and time two emotion-modulated startle. The literature provides little guidance regarding the stability of emotion-modulated startle responding over periods greater than one month. Therefore, analyses for Secondary Aims were exploratory. *Question of interest:* Are emotion modulated startle and other measures of physiology stable over the six month time period?

Self-report measures. Comparisons of self-report ratings of picture arousal and picture valence have not produced clear group differences between depressed and nondepressed persons, even in cases where there are group differences in startle modulation (e.g. Forbes et al., 2005; Dichter et al., 2004). However, studies have not examined whether self-report picture ratings have utility for predicting depression course or severity. Given the weakness of the literature, we made no predictions regarding self-report ratings of arousal or valence.

Heart rate and skin conductance hypotheses. In addition to collecting emotion-modulated startle, we collected measures of heart rate (HR) and skin conductance (SC) during picture viewing. These additional measures served a twofold purpose. First, skin conductance levels vary as a function of arousal, such that levels increase with the presentation of arousing stimuli and decrease in response to neutral stimuli (Dawson, Schell, & Filion, 2007). Therefore, SC serves as a test of equivalent levels of arousal for pleasant and unpleasant picture conditions. Specifically, we expect that SC values will be significantly higher during affective pictures as compared to neutral pictures. Because of the lack of research findings regarding the predictive utility of SC, we made no specific predictions about the relation between SC at time one and depression outcome at time two.

Second, HR responses over the presentation of an affective stimulus generally follow a complex waveform (Berntson, Quigley, & Lozano, 2007). Specifically, HR initially decelerates (D1) following the onset of a stimulus, indicating orientation of attention to that stimulus (Graham and Clifton, 1966). HR subsequently accelerates (A1),

which is thought to be indicative of an increased defensive response (Sokolov, 1963). Lastly, HR decelerates a second time (D2). Less is known about this secondary deceleration and it has received little attention in the research literature. This deceleration could represent a return of HR to pre-stimulus onset levels. Alternatively, D2 may be indicative of an anticipatory response to an expected stimulus (Berg & Donohue, 1992). Specifically, D2 response is enhanced prior to a cued presentation of an aversive stimulus (Somsen, 1983).

If D1 is indicative of orienting to a stimulus, we would not expect differences in this variable between controls, depressed individuals, or those individuals who recovered from depressive symptoms. Therefore, we made no specific hypotheses for D1. Because A1 is thought to represent defensive preparation towards a stimulus, we predicted those individuals who were still depressed at time two would show lower time one A1 responses than those who recovered and controls. Lastly, we did not make directional predictions regarding D2 due to the questionable purpose of the D2 HR response.

Method

Time One Procedures

At time one, participants were 88 individuals diagnosed with the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition with Psychotic Screen (SCID-I/P W/ PSY SCREEN; First, Spitzer, Gibbon, & Williams, 2002), as either currently experiencing a major depressive episode (n = 33), currently experiencing a minor depressive episode (n = 25), or as healthy controls with no history of psychopathology (n = 30). Within two weeks of diagnosis, participants completed a 36-picture emotion-modulated startle session (see below). Participants were then told that they would be contacted in 6 months. Sample sizes at time one were collected in anticipation of a 20% attrition rate. However, attrition rates were closer to 35%.

Time Two Procedures

Participants were re-contacted approximately 6 months following their initial SCID interview. Individuals were invited to return to the lab to participate in the follow-up interview. Participants were re-diagnosed to assess depression recovery status at time two using a modified version of the SCID. Depression recovery status criteria were employed following guidelines set forth by the NIMH Collaborative Program on the Psychobiology of Depression (Keller et al., 1992). These guidelines require that in order for an individual to be considered fully recovered, no more than

two depression symptoms may be present and only to a mild degree. To be considered mild, symptoms may not interfere significantly with the individual's daily life. Previous data suggest that up to 50% of depressed participants can be expected to recover over a six-month follow-up period (Coryell et al., 1994). Utilizing these criteria allowed us to determine any remission during the 6-month period, as well as current remission at the time of time two startle testing. Individuals were considered to be remitted if they were without significant depression symptoms for four or more consecutive weeks.

By means of systematic questioning, the interviewer probed for information on a week-by-week basis for each symptom to establish the approximate dates and time periods of critical transition points (e.g., offset of a previous episode, onset of a new episode, etc.; see Appendix for examples). Despite the arduousness of this type of specific memory recall, past studies have shown acceptable reliability (Keller, Lavori, & Friedman, 1987; Winokur, Coryell, & Keller, 1993) and feasibility (Rottenberg, Wilhelm, Gross, & Gotlib, 2002; Rottenberg, Salomon, et al., 2005) for this kind of questioning at one- and six-month intervals. This interview procedure provides both categorical (depressed, recovered) and continuous variables (number and severity of symptoms) for use as dependent measures to test the hypotheses. The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) was also administered to obtain an additional depression severity score.

To examine changes in other important domains of functioning, several questionnaires from the time one battery were re-administered at time two. The Beck

Anxiety Inventory (BAI) was administered to assess anxiety symptom severity (Beck, Epstein, Brown, & Steer, 1988). A Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) was also administered. The PANAS is a 20 item self-report scale measuring dispositional forms of positive and negative affect.

Physiological Recording Session

Following administration of the modified SCID and self-report questionnaires, participants completed an emotion-modulated startle session procedure identical to that administered at time one. The only difference was that picture order was varied systematically, such that individuals who received order one at time one received order two at time two, and individuals who received order two at time one received order three at time two, etc. Emotion-modulated startle procedures were similar to those conducted with prior studies of MDD populations (e.g., Allen et al., 1999; Dichter et al., 2004). An acoustic startle-probe was administered to participants following the onset of affective pictures displayed on a monitor. Participants also rated the subjective valence and arousal of each picture after each picture was presented. In addition, as mentioned previously, HR and SC were recorded throughout the startle modulation session to afford secondary measures of physiological arousal.

The participant sat in a comfortable chair for the startle procedure. Recording procedures followed general recommendations for startle research (Blumenthal et al., 2005). Electrodes used to measure the startle eyeblink response were applied as well as skin conductance and heart rate sensors. First, three “large” (8 mm) Beckman-type electrodes were placed between the participant’s wrist and elbow to measure cardiac

activity. Next, two large electrodes were applied to the palm of the participant's non-dominant hand to measure of skin conductance response. Lastly, two "small" (4 mm) Beckman-type electrodes were placed just beneath the lower eyelid of the left eye to record the contraction of the orbicularis oculi muscle (the eyeblink component of the startle response). Impedance levels were kept below 20 K Ohms. Once the electrodes were attached, headphones were placed over the participant's ears. Recording set-up took on average ten to fifteen minutes.

Affective picture presentation was modeled after well established and widely used emotion-modulated startle methodology (Bradley, Lang, & Cuthbert, 1993; Lang, Bradley, & Cuthbert, 1998; Bradley, Codispoti, Cuthbert, & Lang, 2001). Specifically, following a five-minute period of acclimation, two neutral pictures were presented to familiarize the participant with the picture-viewing sequence, to allow the participant to practice the rating procedure, and to control for habituation effects. The participant then viewed one of four randomized sequences of 36 digitized pictures, including twelve each of neutral (e.g., mushrooms; 7224, 2570, 7187, 7175, 7491, 7217, 7030, 7150, 2580, 7006, 7025, 7009 for males; 5530, 7035, 7031, 7004, 5740, 2840, 7040, 9360, 7010, 7491, 7185, 6150 for females) pleasant (e.g., puppies; slides 8180, 8260, 7501, 8300, 8501, 8470, 1650, 8080, 4689, 4320, 8380, 4653 for males; 8180, 8400, 8030, 8185, 8080, 5910, 5460, 8210, 8200, 4660, 8034, 7502 for females), and unpleasant (e.g., burn victim; slides 6260, 3060, 3071, 3170, 3530, 3080, 6313, 3015, 6570, 9570, 3053, 9410 for males; 3053, 6312, 9921, 3060, 3120, 3015, 9571, 2730, 9433, 9050, 3100, 3010 for females) from the International Affective Picture System (IAPS; CSEA-NIMH, 1999).

These pictures were chosen based on standardized normative ratings of picture valence, and positive and negative pictures were matched on mean normative ratings of picture arousal (Lang, Bradley, & Cuthbert, 2005). Because these ratings varied according to gender, different picture sets were chosen for males and females.

The normed, mean level of arousal for both sets of valenced pictures individually was 6.02 out of a possible rating of 9—although, out of the entire 700 IAPS slides, the slide with the highest rated arousal only has a rating of just 7.35 (Lang, Bradley & Cuthbert, 2005). Participants rated the arousal and valence of each individual picture with the self-assessment manikin (SAM; Lang, 1980; Bradley & Lang, 1994). The SAM is a picture-based, interactive, computer program in which valence and arousal are rated on a twenty-one-point scale (Cook, Atkinson & Lang, 1987). Subjects made use of pictorial SAM rating scales ranging from a smiling figure to a frowning figure (for valence) and ranging from an excited figure to a sleepy figure (for arousal). The SAM is reliable with split-half correlations of $r = .94$ for valence and $r = .93$ for arousal (Lang, 1985). In addition, the valence and arousal scores from the SAM correlate highly with the widely used, longer Semantic Differential Scale (Bradley & Lang, 1994).

Pictures were displayed on a large (20-inch) computer monitor placed on a table directly in front of the participant. The presentation sequence was as follows: (1) 2-second baseline; (2) 6-second picture viewing; (3) 20 seconds to rate valence and arousal using the SAM; (4) variable (15-second average) inter-trial intervals prior to presentation of the next picture. During nine of the inter-trial intervals (three times per block), a startle probe was presented an average of 7.11 seconds following picture offset. This

post-offset probe provided an assessment of the baseline startle response (startle in the absence of picture viewing). In addition, the random offset probes served to make it less predictable to participants when the noise bursts would be delivered. The startle response was elicited by a binaural acoustic stimulus (50 milliseconds of white noise at 100db with an instantaneous rise time) during nine of the twelve images in each category, and during nine of the inter-trial intervals. Probe times occurred randomly between 3500ms and 5500ms after picture onset. These "late onset" startle probes provide the optimal assessment of the affective influence of valenced pictures (Larson, Ruffalo, Nietert, & Davidson, 2000; Bradley & Lang, 2000; Bradley, Cuthbert, & Lang, 1990). The varying temporal presentation of the startle probe makes it more difficult for the participant to anticipate the onset of the probe during picture presentation. Heart rate and skin conductance were measured continuously throughout each picture-viewing interval. Picture viewing took approximately 40 minutes.

Data Reduction

Signal processing of startle data followed guidelines and recommendations set forth in Blumenthal and colleagues' committee report (2005). First, EMG data was amplified using a high-resolution (A/D) converter. Next EMG signal was filtered to remove background noise and therefore maximize the signal to noise ratio. The EMG signal was then rectified to prevent negative and positive values from canceling the other out. Lastly, the signal was digitally smoothed. EMG signals for each individual were analyzed for onset latency and, most importantly for the purposes of the current study, for mean peak amplitude by condition. As is common convention, values were transformed

into standardized *T* scores to account for the large between-subjects differences in baseline startle values (Blumenthal et al., 2005). Each group's mean blink amplitude for each startle condition served as the between-subjects factor of unpleasant, neutral, and pleasant slide valence.

Hypothesis Testing

Hypothesis 1a. To test the hypothesis that depressed individuals who display the normative linear pattern of emotion-modulated startle at time one will be more likely to experience recovery at time two, a 3 (time one Picture valence: pleasant, neutral, unpleasant) by 3 (time two Group Status: recovered, still depressed, control) repeated-measures Analysis of Variance (ANOVA) was conducted to test for a group by linear or group by quadratic trend. Either of these polynomial trends would indicate a difference in the expected linear pattern of startle modulation. Following significant results of this test, within group contrasts were conducted to determine the nature of within group linearity. This analysis allowed for comparison of differences in the linear pattern of startle at time one across valence conditions. Specifically, we expected that the initial ANOVA would show group polynomial differences. We further expected that follow-up analyses would show a normative linear pattern for controls and recovered individuals, but a lack of linearity (e.g., a blunting of startle responding across pictures) for still depressed individuals. A sample of predicted results are presented below, in Figure 1.

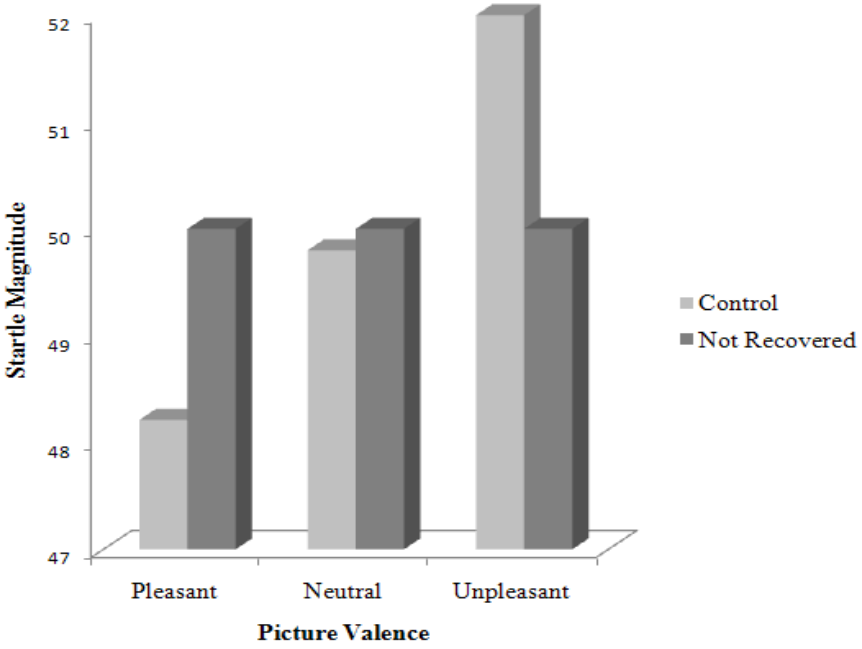

Figure 1. An example of hypothesized results for Hypothesis 1a.

Hypothesis 1b. We predicted that increased time one startle reactivity to pleasant stimuli and decreased time one reactivity to unpleasant stimuli would be associated with a worse symptomatic outcome at time two (see Figure 2 below for hypothesized results). To test these dimensional hypotheses, a series of regression analyses were conducted to determine the unique contribution of time one startle reactivity during pleasant, neutral, and unpleasant stimuli to time two symptomatology and severity.

Hypothesis 1c. A series of regression analyses were conducted to determine whether, as predicted, lower time one baseline startle responses were associated with an increase in depressive symptoms and severity at time two.

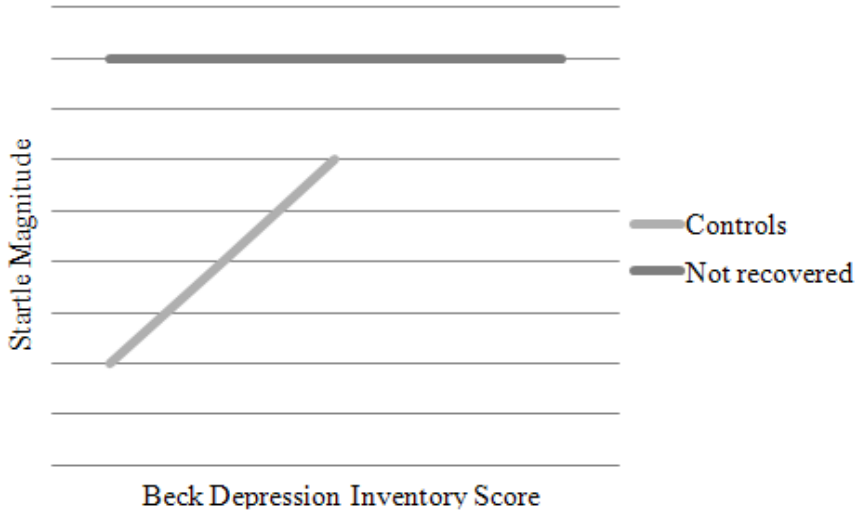

Figure 2. Example of predicted results for Hypothesis 1b.

Secondary Aims. In order to examine the relationship between time one and time two startle modulation, correlation and regression analyses tested for between- and within- group stability of startle responses. We expected to find stronger correspondence between time one and time two measures of startle responding for controls compared to depressed or formally depressed individuals.

Heart Rate and Skin Conductance. Because skin conductance levels vary as a function of arousal, we expected that skin conductance values would be significantly higher during affective pictures as compared to neutral pictures. To test this supposition and to determine whether time one skin conductance would predict time two diagnostic status, we conducted a 3 (Group) by 3 (valence condition) repeated measures ANOVA. Planned contrasts tested the hypothesis that skin conductance peaks during affective picture presentation would be greater than to peak skin conductance values during neutral

picture presentation. A significant group by valence interaction was followed by one-way ANOVAs to determine the nature of the difference.

As described previously, we did not expect to find time two diagnostic group differences in the time one HR variable D1 or D2. Therefore, we made no specific hypotheses for these variables and planned exploratory analyses of possible group differences using repeated measures ANOVAs for each HR variable. Significant group by valence interactions would be followed by one-way ANOVAs.

Because A1 is thought to represent defensive preparation towards a stimulus, we predicted that time two never-recovered, depressed individuals would have higher time one levels of A1 during pleasant pictures and lower time one levels of A1 during unpleasant pictures (consistent with the ECI theory of emotional reactivity). To test these hypotheses, we conducted a 3 (time one valence: negative, neutral, positive) by 3 (time two group) repeated measures ANOVA. Significant results were investigated through follow-up tests.

Results

To set a context for our main analyses, we first provide descriptive statistics on our time one and time two samples, and we describe attrition analyses comparing participants who did and did not participate in the time two procedures. We then describe the clinical outcome of our sample in terms of rates of depression recovery and other measures of symptom severity. Prior to testing our hypotheses, we report the results of a manipulation check (i.e., picture ratings differed in the expected directions; pleasant pictures < neutral pictures < unpleasant pictures for self-reported valence and pleasant pictures = unpleasant pictures > neutral pictures for arousal). We then describe the results of Hypothesis 1a testing, which predicted that individuals who displayed the normative pattern of emotion-modulates startle at time 1 would be more likely to experience recovery by time 2, followed by tests of Hypothesis 1b, which predicted that decreased time one attenuation (i.e., increased reactivity) during pleasant pictures and decreased time one potentiation (i.e., decreased reactivity) during unpleasant pictures would be associated with higher levels of depression severity at time two. Next, we list results from tests of Hypothesis 1c that lower time one baseline startle responses would be associated with increased time two depression severity.

We then describe analyses and results of skin conductance and heart rate testing. Although we made no specific predictions regarding the utility of time one skin

conductance for predicting time two depression status, we expected that SC values would be significantly higher during affective pictures as compared to neutral pictures. We also tested the hypothesis that only the acceleration variable of HR would predict time two depression status. Because of the novelty of this type of longitudinal design with these variables, we tested all time one skin conductance variables and HR measures (initial acceleration, A1; deceleration, D1; and secondary acceleration, A2) for predicting time two depression status. Finally, we describe the results of secondary correlation analyses between time one and time two measures of physiology, which sought to determine test-retest reliability of these physiological variables.

Sample Characteristics

Participants were recruited from fliers and online postings in and around the Tampa Bay community (Table 1). Of the approximately 450 potential participants who were screened by telephone, 171 were invited into the lab to complete the Structured Clinical Interview based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV-TR) Axis I Disorders, Research Version, Patient Edition with Psychotic Screen (SCID-I/P W/ PSY SCREEN; First et al., 2002). Fifty-six persons that completed the SCID did not meet study inclusion or exclusion criteria. Participants were excluded for history of a major head injury, hearing impairment, diagnosis of bipolar disorder, substance abuse occurring within 6 months prior to entry into the study, or any history of primary psychotic symptoms.

Provisional DSM-IV-TR criteria for an mD diagnosis recommend an absence of past episodes of MDD. To improve study feasibility, we loosened this criterion and 36%

of mD participants had experienced at least one major depressive episode (MDE). These subjects were only included if there was a period of at least eight weeks with no residual depressive symptoms between the major depressive episode and the minor depressive episode. In all cases of mD with a past MDE, MDEs occurred at least one year prior to the current mD episode.

Table 1. Time one sample demographic characteristics.

	Group		
Variable	MDD (*n* = 33)	mD (*n* = 25)	Controls (*n* = 30)
Age, M (SD)	29.88 (11.25)	25.85 (6.88)	26.72 (7.79)
% Caucasian	46.9%	73.1%	61.8%
% Female	88.2%	65.4%	75%
Education	5.29 (1.85)[a]	5.23 (1.75)[a]	5.63 (2.14)[a]
Income	5.31 (3.44)[b]	4.40 (3.07)[b]	5.93 (3.89)[b]
% Married	17.6%	19.2%	28.1%
% Antidepressants	14.7%	5.3%	3.7%[c]
% Psychotherapy	11.8%	5.3%	0%

[a]Education was assessed on an 8-point scale with higher numbers representing more education—a score of 5.63 reflects graduation from a 2-year or a technical college.
[b]Income was assessed on a 12-point scale—a score of 5.93 represents an income of between $25,000 and $34,999.
[c]One control participant was taking an antidepressant to (successfully) control migraine headaches.

Of those 115 individuals who completed the first SCID interview at time one, 88 (77%) participants—including 30 controls, 33 MDD individuals, and 25 mD individuals—completed the time one startle procedure. Of those not completing the startle procedure, some individuals declined to participate (n = 5), but the large majority failed to attend the session and were unable to be rescheduled (n = 22). Of those individuals who completed the time one startle procedure, 52 individuals—17 controls,

18 MDD individuals, and 17 individuals with mD—returned for the second SCID interview (time two). Following the second interview, a total of 31 individuals—12 controls, 13 MDD individuals, and 6 mD individuals—chose to participate in a second startle session. Of those who did not participate in the second startle session, four were unable to be scheduled, while the remaining 19 individuals did not want to participate. The original sample was collected in anticipation of a 20% attrition rate, based on previous studies in this laboratory. The actual attrition rate for individuals not completing the time two SCID was 35%.

Table 2. Time two sample demographic characteristics.

	Group		
Variable	MDD at T1 (*n* = 18)	mD at T1 (*n* = 17)	Controls (*n* = 17)
% of T1 group	65%	56%	53%
Age, M (SD)	30.18 (10.61)	26.44 (8.16)	29.22 (9.06)
% Caucasian	63.6%	77.8%	33.3%*
% Female	95.5%	66.7%‡	77.8%
Education	5.27 (1.67)[a]	5.67 (1.88)[a]	6.00 (2.33)[a]
Income	5.37 (3.86)[b]	5.29 (2.87)[b]	7.36 (3.48)[b]
% Married	22.7%	27.8%	27.8%
% Antidepressants	22.7%	11.1%	5.6%[c]
% Psychotherapy	3.4%	0%	0%

[a]Education was assessed on an 8-point scale with higher numbers representing more education—a score of 6.00 reflects graduation from a junior college.
[b]Income was assessed on a 12-point scale—a score of 7.36 represents an income of between $35,000 and $44,999.
[c]One control participant was still taking an antidepressant to (successfully) control migraine headaches.
*$p < .01$
‡$p = .06$

The final sample that participated in the time two follow-up interview (Table 2) was again primarily Caucasian (57%) females (80%). At time two there was a significant difference between controls (33%) and the clinical groups (64% for MDD and 78% for mD) in terms of percentage of Caucasian participants (p <.01). There was a near significant difference between MDD individuals (95%) and the other groups (67% of mD individuals and 78% of controls) in terms of percentage of females (p = .06). Therefore, sex and ethnicity were entered as covariates in all subsequent analyses. No other significant demographic differences existed between groups (all ps>.17).

Missing Data

Due to time constraints, time one BAI data were missing for one individual. Time one startle magnitude data were missing for two individuals because they were categorized as nonresponders (i.e., the startle probe failed to elicit a startle response during at least two startle probes during each valenced picture category). Due to equipment malfunction, heart rate data were missing for four individuals. Three individuals had skin conditions that eliminated the possibility of measuring skin conductance (e.g., blisters, hyperhidrosis). Through an unfortunate photocopying error, time two PA and NA scores were missing for 19 individuals. Three individuals chose not to complete time two PANAS questionnaires, and time two BAI data were missing for an additional 4 individuals. All missing data was excluded casewise.

Attrition Analyses

Before testing primary hypotheses, attrition analyses were conducted to determine if those who did and did not participate in the time two SCID differed on descriptive

variables. A series of 2 (Group: participated, did not participate) by 3 (valence condition) repeated measures ANCOVAs were conducted with gender and ethnicity as covariates. Individuals who participated in the time two interview (Table 3) did not differ from those who did not participate in terms of BDI score (p>.34), BAI score (p>.75), NA (p> .87), PA (p>.66), age (p>.88), educational level (p>.25), sex (χ^2 = 1.33, p =.25), marital status (χ^2 = 4.15, p =.39), occupation (χ^2 = 11.10, p =.20), income (χ^2 = 9.30, p =.60), or treatment at time one (χ^2 = 4.53, p =.21). Lastly, the three diagnostic groups did not differ in terms of retention at time two (χ^2 = 1.03, p =.60). Thus, those who were lost to follow up did not differ in any way from those who were retained in the sample.

Table 3. Demographic characteristics of time two sample by time two diagnostic grouping.

	T2 Group		
Variable	Ever Remitted (n = 17)	Never Remitted (n = 24)	Controls (n = 17)
Age, M (SD)	26 (7.15)	30.29 (10.70)	29.24 (9.34)
% Caucasian	82.4%	62.5%	29.4%*
% Female	82.4%	83.3%	13 (76.5%)
Education	4.88 (1.80)[a]	5.79 (1.64)[a]	6.12 (2.34)[a]
Income	5.00 (3.90)[b]	5.58 (3.11)[b]	7.36 (3.48)[b]
% Married	11.8%	33.3%	29.4%
% Antidepressants	23.5%	16.7%	0%
% Psychotherapy	5.9%	4.2%	0%

[a]Education was assessed on an 8-point scale with higher numbers representing more education—a score of 6.00 reflects graduation from a junior college.
[b]Income was assessed on a 12-point scale—a score of 7.36 represents an income of between $35,000 and $44,999.
*p = .06

In addition, we also examined whether those who did and those who did not participate in the time two protocol differed on any predictor variable. These included tests of differences for startle magnitude during pleasant, unpleasant, and neutral pictures, as well as baseline startle, heart rate and skin conductance measures, and self-report ratings of picture valence and arousal. All were nonsignificant (all *p*s > .17), with one exception. Those who did and did not complete the follow-up interview differed in time one raw startle magnitude during neutral pictures [$F(1, 83) = 5.02$, $p < .05$], such that those who participated had higher raw startle magnitude during neutral pictures (M=200.38) compared to those who did not (M=134.07).

Clinical Outcome at Follow-up

Groups differed significantly at follow up on several clinical characteristics (Table 4). As expected, controls were less likely to experience a major depressive episode during the 6-month follow-up period than the mood-disordered participants. In fact, only one control participant experienced a depressive episode during the 6-month follow-up period. Individuals diagnosed as MDD or mD at time one did not differ significantly in the likelihood of a MDE any time in the 6-month follow-up period (p = .21) or at time of testing at time two (p = .78). Persons with a MDD diagnosis at time one met MDD criteria for significantly more weeks over the follow up period than did persons with a mD diagnosis at time one ($p < .05$) who in turn met MDD criteria for significantly more weeks than did time one controls ($p <.001$). Additionally, time one MDD individuals experienced more total MDD week-by-week symptoms over the follow-up period than time one mD individuals ($p<.001$), and time one mD individuals

had a greater number of symptoms than controls (p<.01). Time one MDD and mD individuals did not differ significantly in likelihood to remit at any point over the 6-month period.

We conducted a one-way analysis of variance in order to test for group differences in self-report measures of symptoms. Time one BDI [F (2, 49) = 63.47, p < .01], BAI [F (2, 48) = 4.05, p < .05], PA [F (2, 49) = 10.97, p < .01], and NA [F (2, 49) = 9.08, p < .01] scores differed significantly between groups (Table 4). BDI scores differed significantly for all three groups in the expected direction (with the MDD group having the highest score, followed by the mD group, and controls; ps < .01). For BAI scores, controls differed only from mD individuals (p < .05), who reported higher anxiety scores than controls. PA scores differed significantly and in the expected direction between controls and MDD individuals (p < .01), and approached significance between mD individuals and MDD individuals (p = .08), as well as between controls and mD individuals (p = .07), with controls having the highest levels of positive affect, followed by mD and then MDD individuals. NA of MDD and mD individuals was significantly higher than that of controls (p < .01), but mD and MDD did not differ from one another.

Time two BDI [F (2, 49) = 12.68, p < .01], BAI [F (2, 29) = 6.19, p < .01] and PA [F (2, 29) = 5.44, p < .05] scores differed significantly between time one groups. At time two, time one MDD and mD individuals had significantly higher time two BDI scores compared to controls (p's < 01), but MDD and mD individuals did not differ. Controls had significantly lower BAI scores compared to time one MDD individuals (p < .05). Controls had significantly higher time two PA scores than individuals who were MDD at

time one ($p < .05$), but no other groups differed on time two PA. Lastly, groups did not differ significantly on time two NA ($ps > .21$; this analysis was compromised by low power due to missing data).

Because there is currently no widely accepted gold standard for defining depression recovery, we examined the grouping variable of interest, time two depression status, in two different ways. In the first, participants who had a mood disorder and were remitted at the time of the second interview were contrasted with those not in remission at time two. This first grouping variable presumes that those individuals who are able to remit and stay remitted for an extended period of time will differ in a significant way from those individuals who either do not remit at all or remit for a shorter period of time. Because controls were not expected to develop a depressive episode, they were placed in a third group. However, one control, who did develop a depressive episode, was included in analyses for remitted individuals. Only twelve individuals from the clinical groups and one individual from the control group were in remission at the time of the second diagnostic interview.

Second, we created an additional grouping that included participants as remitted if they achieved a full remission for one month at any point during the 6-month follow up. The second grouping variable presumes that those individuals who are able to remit for even a short amount of time are significantly different from those depressed individuals who do not remit at all over a 6 month period. The ability to recover, even for a short period of time, may be indicative of more extensive affective resources as compared to those who stay more chronically depressed. This second grouping variable made fuller

use of the detailed weekly codings in the longitudinal data set. This grouping identified four additional participants who experienced a one-month period of full remission at some point during the 6-month period, for a total of 16 individuals from the clinical group and one individual from the control group.

To identify whether ever remitted and no current MDE individuals differed from non-remitted and time two MDE individuals, respectively, on any variables that were not of primary interest, we conducted a series of analyses of variance (ANOVAs) with diagnostic group as the between-subjects variable and education level, income level and age as the within-subjects variable or Cramer's V tests, depending on whether the variable was continuous or categorical. Groups did not differ significantly on gender, education, income, or marital status (ps > .15). Furthermore, diagnostic groups did not differ on time one or time two antidepressant use.

Table 4. Clinical characteristics by time one diagnostic group.

	Group					
	MDD at T1		mD at T1		Controls	
Variable	(*n* = 18)		(*n* = 17)		(*n* = 17)	
MDE at T2	7 (32%)		5 (28%)		0	
Any MDE	15 (68%)		9 (50%)		1 (4%)	
Weeks MDD	9.95 (7.30)		4.11 (4.70)		0.33 (1.03)	
Total Symptoms*	94.18 (47.53)		47.44 (29.10)		3.06 (9.17)	
Full Remission‡	8 (36.4%)		8 (44.4%)		1 (4%)	
	T1	T2 (n = 16)	T1	T2 (n = 4)	T1	T2 (n = 13)
BDI	32.05 (9.72)	20.82 (13.85)	21.88 (8.31)	16.65 (10.68)	3.33 (4.98)	3.89 (5.37)
BAI	9.38 (9.81)	16.00 (12.40)	15.47 (9.18)	11.00 (6.36)	7.22 (7.31)	3.50 (3.90)
PA	20.60 (6.19)	26.00 (11.10)	25.71 (7.34)	25.25 (9.38)	32.12 (8.82)	27.46 (7.89)
NA	20.90 (8.93)	16.64 (9.68)	18.82 (6.02)	18.50 (10.77)	11.94 (2.63)	11.92 (1.93)

* Total Symptoms equals the total number of symptoms that an individual experienced over the 6-month follow-up period.

‡ Denotes full remission at any point over the follow-up period. In order to achieve full remission, an individual needed to have less than two symptoms, and neither could be depressed mood or anhedonia, for at least 4 weeks. In addition, symptoms had to be experienced to no more than a mild degree to meet criteria for full remission.

Major depressive episode (MDD); Minor depression (mD) Major depressive episode (MDE); Time one (T1); Time two (T2); Beck Depression Inventory (BDI); Beck Anxiety Inventory (BAI); Positive Affect (PA); Negative Affect (NA)

We then examined whether groups differed in terms of clinical characteristics, specifically PA, NA, BDI scores, BAI scores, or 6-month anxiety disorder diagnoses (Tables 5 and 6). When diagnostic group was defined as ever remitted, never remitted, or controls (Table 5), significant group differences were found for BDI [F (2, 49) = 58.78, p < .01], PA [F (2, 49) = 8.02, p < .01], and NA [F (2, 49) = 9.96, p < .01]. BAI approached significance [F (2, 48) = 2.92, p = .06], and along with BDI, PA, and NA, was examined further. Individuals who did not experience remission before time two had higher BDI, BAI, and NA scores, as well as lower PA scores compared to controls (all ps < .05). BAI scores showed a trend towards larger scores in those who never remitted compared to those that did (p = .07). No other time one variables differed between those who never remitted and those who did remit (ps > .37). Those who remitted before time two also had higher time one BDI and NA scores than controls (ps < .01).

Groups differed significantly on time two BDI [F (2, 49) = 27.20, p < .01], time two BAI [F (2, 29) = 6.75, p < .01], time two PA [F (2, 29) = 6.87, p < .01], and time two NA [F (2, 29) = 6.70, p < .01]. Time two BDI scores differed in the expected direction for all three groups, such that those who never remitted before time two had the highest scores, followed by those who did experience a remission during the follow-up period, and then controls (ps < .01). Controls had significantly lower time two BAI scores compared to those who never remitted (p < .01). Time two PA was lower in those who did not experience remission before time two compared to controls (p < .01). NA was highest in those who never remitted compared to controls (p = .05) and those who did experience a remission (p = .05).

When diagnostic group was defined as presence or absence of MDE at time two (Table 6), groups differed significantly on time one BDI [F (2, 49) = 49.04, $p < .01$], time one PA [F (2, 49) = 9.01, $p < .01$], and time one NA [F (2, 49) = 9.11, $p < .01$]. Again, BAI approached significance [F (2, 48) = 2.92, $p = .06$]. Specifically, controls reported significantly lower BDI, BAI, and NA scores and higher PA scores compared to those with and those without a time two MDE (p's < .05). Groups also differed significantly on time two measures of BDI [F (2, 49) = 23.54, $p < .01$], time two BAI [F (2, 29) = 6.60, $p < .01$], time two PA [F (2, 29) = 9.11, $p < .01$], and time two NA [F (2, 29) = 7.24, $p < .01$]. Time two BDI was highest in those with a time two MDE, followed by formerly depressed individuals without a time two MDE and then controls (p's < .01).

Table 5. Clinical characteristics according to time two MDE group.

	Group					
	No MDE at Time 2		MDE at Time 2		Controls	
Variable	(*n* = 25)		(*n* = 11)		(*n* = 16)	
Panic Disorder	3 (10.7%)		1 (8.3%)		1 (5.9%)	
Social Phobia	4 (14.8%)		3 (25%)		0	
Specific Phobia	2 (7.4%)		2 (16.7%)		1 (5.9%)	
OCD	0		1 (8.3%)		0	
PTSD	1 (3.7%)		1 (8.3%)		0	
GAD	6 (22.2%)		4 (33.3%)		0	
	T1	T2 (n = 13)	T1	T2 (n = 6)	T1	T2 (n = 12)
BDI	26.82 (9.61)	15.29 (11.68)	28.92 (12.03)	27.67 (9.76)	2.29 (2.39)	3.00 (3.94)
BAI	12.30 (10.19)	12.54 (11.85)	12.83 (10.03)	18.50 (8.02)	6.12 (5.79)	3.50 (3.90)
PA	24.77 (7.81)	25.92 (7.40)	19.92 (4.56)	17.33 (5.68)	32.00 (9.10)	33.17 (8.35)
NA	19.62 (7.15)	14.00 (4.79)	20.50 (8.82)	24.67 (14.40)	11.56 (2.19)	11.92 (1.93)

Major depressive episode (MDE); Obsessive Compulsive Disorder (OCD); Post-traumatic Stress Disorder (PTSD); Generalized Anxiety Disorder (GAD); Beck Depression Inventory (BDI); Time one (T1); Time two (T2); Beck Anxiety Inventory (BAI); Positive Affect (PA); Negative Affect (NA)

Controls had significantly lower time two BAI scores compared to those with a time two MDE ($p < .01$) and nearly significantly lower scores than those without a time two MDE ($p = .05$). Time two PA was lowest in those with a time two MDE compared to those without a time two MDE and controls (ps < .05).

Hypothesis Testing

Subjective ratings of picture valence and arousal. Next, using data from only those individuals who had diagnostic data at time two, a series of 3 (time one group) by 3 (valence) repeated measures ANCOVAs were conducted to determine whether time one subjective ratings differed between time one groups. For subjective ratings of valence, the data was spherical. Therefore, a Huynh-Feldt adjustment was made (Huynh & Feldt, 1976). The results indicated a near-significant group by valence interaction [F (2, 47) = 2.48, $p = .08$, $\eta_p^2 = .09$] for subjective ratings of valence. Although this analysis did not reach conventional levels of significance, we conducted follow-up tests because of the small sample size and potential importance for interpreting physiological results. Follow-up analyses indicated that time one groups differed on their valence ratings of unpleasant pictures, but not pleasant or neutral pictures. Specifically, individuals with time one MDD reported unpleasant pictures to be more unpleasant than did individuals with mD ($p < .05$). The group by arousal interaction for subjective ratings of arousal was nonsignificant ($p > .10$).

We also performed a manipulation check on subjective ratings of valence and arousal. The ANCOVAs yielded the expected main effects (ps < .01). For valence, unpleasant pictures were rated as least pleasant while pleasant picture were rated

Table 6. Clinical characteristics according to time two remission group.

	Group					
	Ever Remitted‡		Never Remitted		Controls	
Variable	(*n* = 15)		(*n* = 21)		(*n* = 16)	
Panic Dx	1 (5.9%)		3 (12.5%)		1 (5.9%)	
Social Phobia	2 (12.5%)		5 (20.8%)		0	
Specific Phobia	1 (5.9%)		3 (12.5%)		1 (5.9%)	
OCD	0		1 (4.2%)		0	
PTSD	2 (12.5)		0		0	
GAD	5 (29.4%)		5 (20.8%)		0	
	T1	T2 (n = 16)	T1	T2 (n = 16)	T1	T2
BDI	23.29 (9.77)	11.59 (5.66)	30.52 (9.74)	24.48 (13.29)	2.29 (2.39)	3.00 (3.94)
BAI	12.18 (10.68)	11.00 (14.40)	12.68 (9.72)	16.91 (7.33)	6.12 (5.79)	3.50 (3.90)
PA	25.19 (8.30)	26.00 (7.84)	21.82 (6.20)	20.70 (7.42)	32.00 (9.10)	33.17 (8.35)
NA	18.44 (6.02)	12.00 (2.12)	20.95 (8.56)	21.45 (11.55)	11.56 (2.19)	11.92 (1.93)

‡Denotes full remission at any point over the follow-up period. In order to achieve full remission, an individual needed to have less than two symptoms, and neither could be depressed mood or anhedonia, for at least 4 weeks. In addition, symptoms had to be experienced to no more than a mild degree to meet criteria for full remission.
Major depressive episode (MDE); Obsessive Compulsive Disorder (OCD); Post-traumatic Stress Disorder (PTSD); Generalized Anxiety Disorder (GAD); Beck Depression Inventory (BDI); Time one (T1); Time two (T2); Beck Anxiety Inventory (BAI); Positive Affect (PA); Negative Affect (NA)

as most pleasant (*p*s < .001). For arousal ratings, pleasant and unpleasant pictures were rated as more arousing than neutral pictures ($p < .001$), but pleasant and unpleasant did not differ significantly from each other on rated arousal ($p > .29$).

There was a nonsignificant group by valence interaction for ratings of picture valence for the grouping variable of any 6-month remission [F (2, 48) = 0.31, p = .79, η_p^2 = .01], as well as for presence of a depressive episode at time two [F (2, 48) = 0.22, p = .86, η_p^2 = .01]. The interaction was also nonsignificant for the interaction of time one ratings of picture arousal by time two presence of depression [F (2, 48) = 0.98, p = .42, η_p^2 = .04]. However, for ratings of picture arousal, there was a near significant interaction of time one ratings by time two remitted group [F (2, 48) = 2.23, p = .07, η_p^2 = .09]. Although the results did not reach conventional levels of significance, follow-up comparisons indicated that individuals who did not remit at any time during the 6-month follow-up period rated unpleasant pictures as more arousing at time one (M = 14.35) than controls (M = 10.48; $p < .05$) and those who did remit at some point (M = 11.21; $p = .08$). Groups did not differ on ratings of arousal (Figure 1) for pleasant or neutral pictures.

Startle analyses. To test Hypothesis 1a, that blunted time one emotion-modulated startle magnitude would predict a worse course of mD and MDD, repeated-measures ANCOVAs were conducted with time one startle valence as the predictor and time two diagnostic status as the outcome variable. A 3 (time one valence: negative, neutral, positive) by 3 (Group: remitted for at least 1 month, never remitted, controls) repeated measures ANCOVA yielded a nonsignificant valence by group interaction [F (4, 44) =

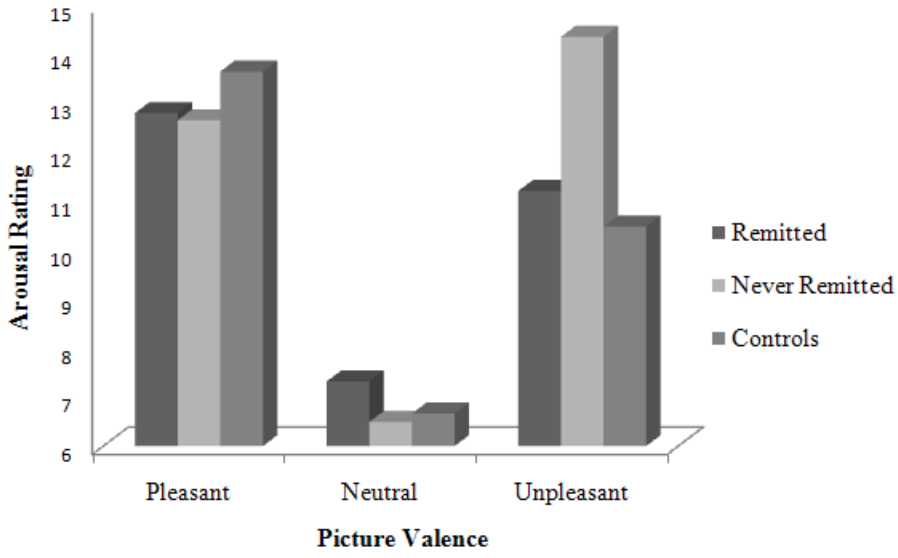

Figure 3. Time one subjective ratings of arousal for remitted and never remitted individuals and controls.

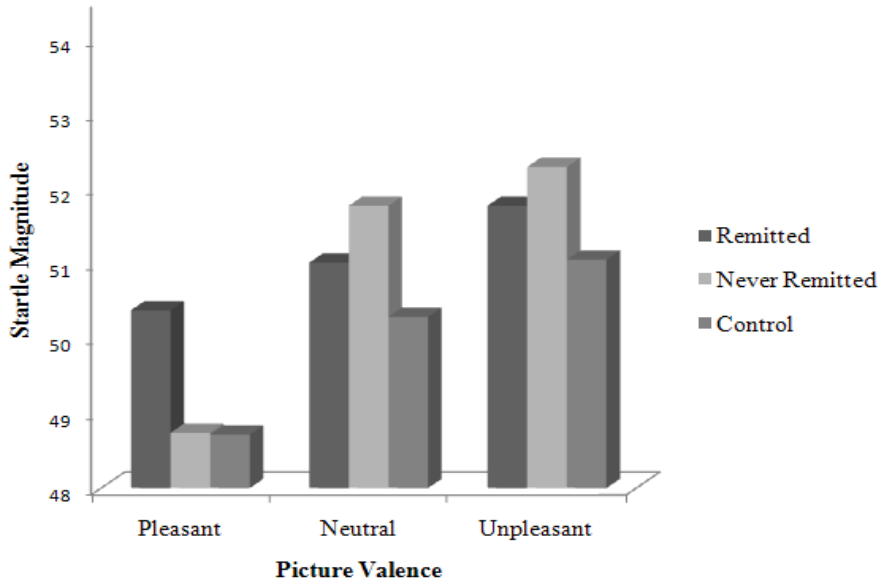

Figure 4. Time one startle responses by picture valence for remitted and never remitted depressed individuals and controls.

1.50, p =.21, η_p^2 = .06]. Additionally, within-subjects tests did not yield a significant group by linear trend [F (4, 44) = 2.12, p = .13, η_p^2 = .09; Figure 2].

When current MDE at time two was the grouping variable, the interaction of valence by group verged on significance [F (4, 47) = 2.27, p =.07, η_p^2 = .09]. Although there was no group by linear trend effect [F (4, 46) = 0.10, p =.88, η_p^2 = .006], there was a significant group by quadratic trend [F (4, 46) = 4.16, p < .05, η_p^2 = .15]. Figure 3 depicts that while individuals with a time two current MDE display a quadratic pattern of startle reactivity across picture conditions, controls and individuals without a time two

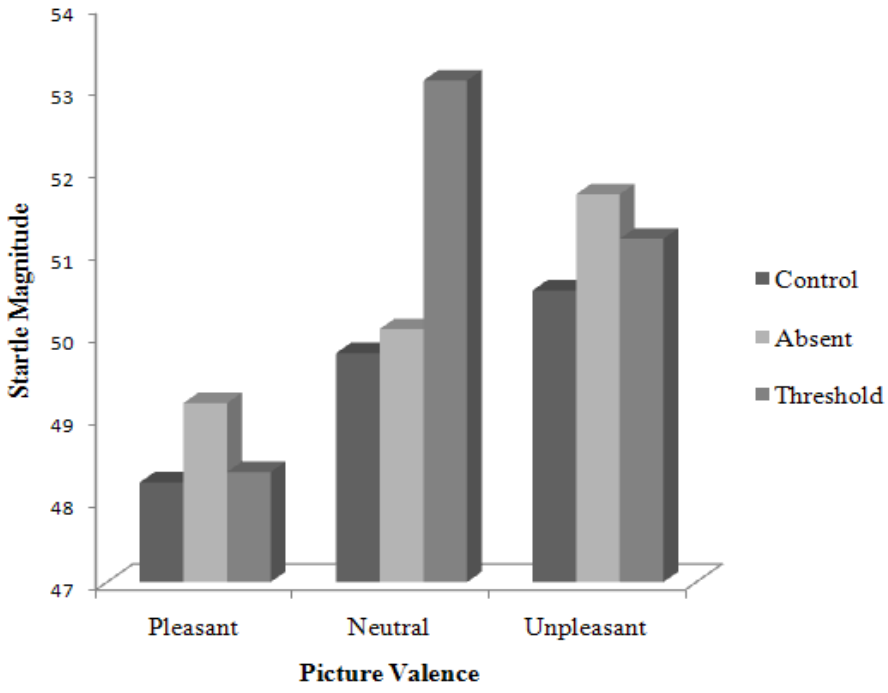

Figure 5. Time one startle responses by picture valence for depressed individuals, previously depressed individuals and controls.

MDE do not display this pattern. Inspection of the results indicates that depressed individuals show an unusual elevation of startle responses during neutral stimuli in comparison to other valence conditions. Follow-up contrasts confirmed the between groups difference for startle magnitude during neutral pictures ($p < .01$), such that individuals with a current MDE at time two displayed higher startle reactivity at time one to neutral pictures (M = 53.42) compared to controls (M = 49.72) and remitted individuals (M = 49.99). There were no significant between-group differences for pleasant or unpleasant pictures.

In line with Hypothesis 1b, that blunted startle responding to pleasant and unpleasant pictures would be associated with higher time two depression severity (measured continuously), we examined time one startle magnitude from all conditions, separately, as predictors of time two BDI scores. Prior to testing, assumptions of regression were examined. Examination of the data revealed that BDI scores were not normally distributed. This is not surprising, given that controls with low BDI scores were combined with depression groups. To remedy this non-normality, BDI scores were transformed into z-scores with a normal distribution. Next, the variables already established to be associated with time two status (gender, ethnicity and time one diagnostic category) were entered into a model to predict time two BDI z-scores. The model (Model 1; Table 7) was significant [F (3, 47) = 3.17, $p < .05$], accounting for 17% of the variance (12% adjusted).

Two startle conditions predicted time two BDI z-scores: startle magnitude during neutral pictures and baseline startle magnitude. Each of these was entered separately into

a model with gender, ethnicity and time one diagnostic status. When startle magnitude during neutral pictures was added to the model (Model 2), the model was significant [F (3, 47) = 2.99, $p < .05$], accounting for 16% of the variance (11% adjusted). Standardized β weights indicate that startle magnitude during neutral pictures predicted time two depression severity above and beyond time one diagnostic status, with higher time one startle during neutral pictures predicting higher time two depression severity. When baseline startle was added to the model (with gender, ethnicity and time one diagnostic status), the model remained significant [F (3, 47) = 8.00, $p < .01$] and predicted 34% (28% adjusted) of the total variance. Furthermore, standardized β weights (Table 7) indicate that time one baseline startle provided a statistically significant contribution to the model predicting time two BDI scores, with lower time one baseline startle predicting with higher time two depression severity.

Because startle magnitudes for valenced conditions can be correlated with baseline startle magnitude, we tested the strength of the correlation between baseline startle and startle during neutral pictures; baseline and neutral startle were moderately negatively correlated (r = -.32, p <.05). In order to assess the unique contribution of both baseline startle and neutral startle, both were entered into a model together to predict time two BDI z-scores. The model remained significant [F (3, 47) = 5.51, $p < .01$] and accounted for 39% (31% adjusted) of the total variance. Importantly, in this analysis, baseline startle remained an independent predictor of depression severity, while neutral startle became marginally significant.

Table 7. Regression models with time one neutral startle predicting time two BDI.

Model	Variable	R^2	*t*-value	*p*-value	β-weight	Correlations		
						Zero-order	Partial	Part
1	T1 Group	.17	2.90	.006	.38	.39	.39	.37
	Gender		.56	.58	.07	.05	.08	.07
	Ethnicity		1.99	.05	.26	.29	.28	.26
2	T1 Group	.16	2.50	.02	.32	.39	.35	.31
	Gender		.33	.75	.04	.05	.05	.04
	Ethnicity		2.27	.03	.28	.29	.32	.28
	T1 Neutral Startle		2.36	.02	.30	.34	.33	.29
3	T1 Group	.34	1.47	.15	.20	.39	.21	.18
	Gender		.87	.39	.11	.05	.13	.10
	Ethnicity		1.63	.11	.20	.29	.23	.20
	T1 Baseline Startle		-2.83	.002	-.39	-.51	-.39	-.34
4	T1 Group	.39	1.20	.24	.19	.39	.18	.14
	Gender		.56	.58	.07	.05	.08	.06
	Ethnicity		1.93	.06	.24	.30	.28	.23
	T1 Baseline Startle		-2.31	.03	-.33	-.51	-.33	-.27
	T1 Neutral Startle		1.76	.09	.23	.35	.26	.21

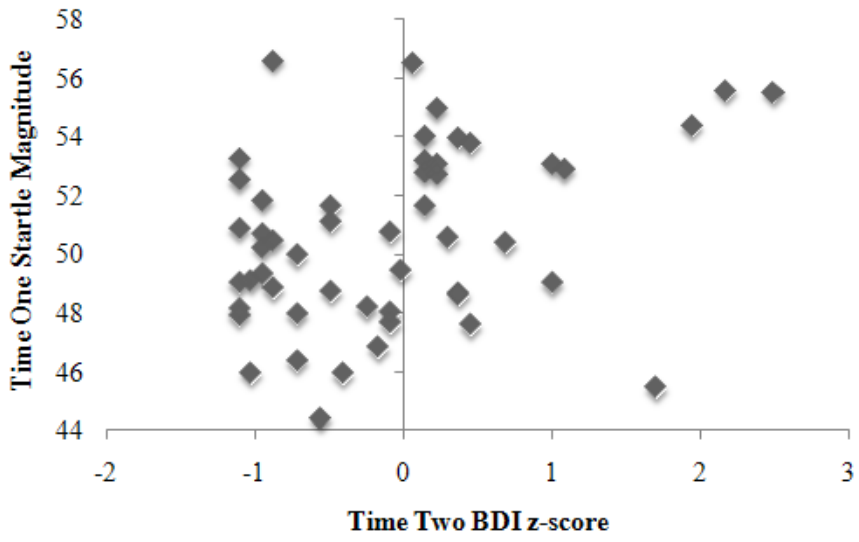

Figure 6. Time one neutral startle magnitude and time two Beck Depression Inventory score.

Taken together, results from analyses of startle responding suggest that abnormal responding during neutral pictures predicts time two depression status. This pattern was found in both categorical and continuous analyses. Higher startle responses during neutral pictures were indicative of risk for a depressive episode after 6 months. Furthermore, higher time one startle responses during neutral pictures were associated with higher time two depression severity (as indexed by the BDI-II). In addition, consistent with Hypothesis 1c, lower time one baseline startle responses (i.e., startle responses that occur between pictures in the absence of affective stimuli) predicted higher time two depression severity.

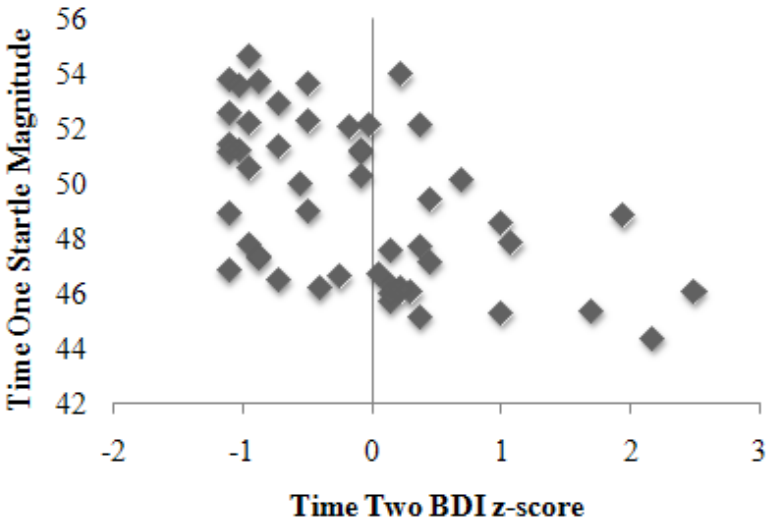

Figure 7. Time one startle during baseline and Beck Depression Inventory score.

Heart rate analyses. For heart rate (HR) analyses, there were several variables of interest. As described previously, the pattern of HR over the course of each picture viewing trial assumes a waveform with an initial deceleration (D1), a subsequent acceleration (A), followed by a second deceleration (D2). In addition to these variables, a baseline HR was collected in the absence of picture stimuli during a resting period before the start of the protocol.

A one-way ANCOVA found no difference between time two groups (ever remitted, never remitted, controls) in time one baseline HR (p = .85). Next, a 3 (time two group: ever remitted, never remitted, controls) by 3 (time one HR change across the picture viewing trial: D1, A, D2) by 3 (time one valence: pleasant, neutral, unpleasant) repeated measures ANCOVA was conducted to examine whether differences in the pattern of time one HR predicted time two diagnosis (Figures 6-9).

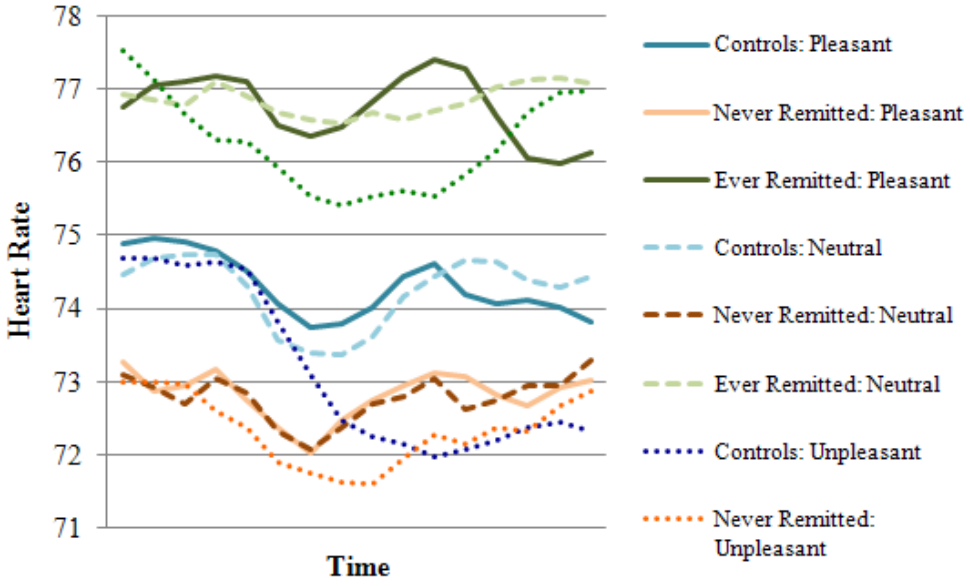

Figure 8. Heart rate wave forms for all conditions and groups.

There was no main effect for HR (p = .38) or valence (p = .85), nor was there an interaction for time one HR by time two diagnosis (p = .71), valence by group (p = .19), or time one HR by valence (p = .80). There was, however, a near significant three-way interaction of time two remitted group by time one HR by time one valence [F (8, 44) = 2.03, p = .07, η_p^2 = .09] but not for time two presence of a depressive episode [F (8, 44) = 1.76, p = .10, η_p^2 = .15].

Although the near significant interaction did not reach conventional levels of significant, we completed follow-up analyses. A follow-up ANOVA of HR variables for controls and never remitted individuals indicated significant differences in A ($p < .05$) and D2 ($p < .01$) for unpleasant pictures (Figure 9). Specifically, controls displayed less acceleration and secondary deceleration in response to unpleasant pictures as compared to those individuals who never remitted.

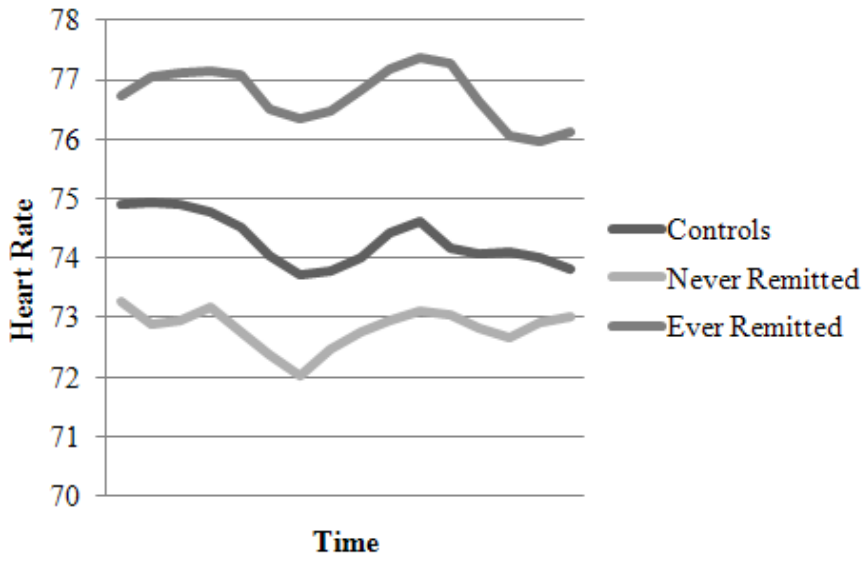

Figure 9. Heart rate waveforms during pleasant pictures.

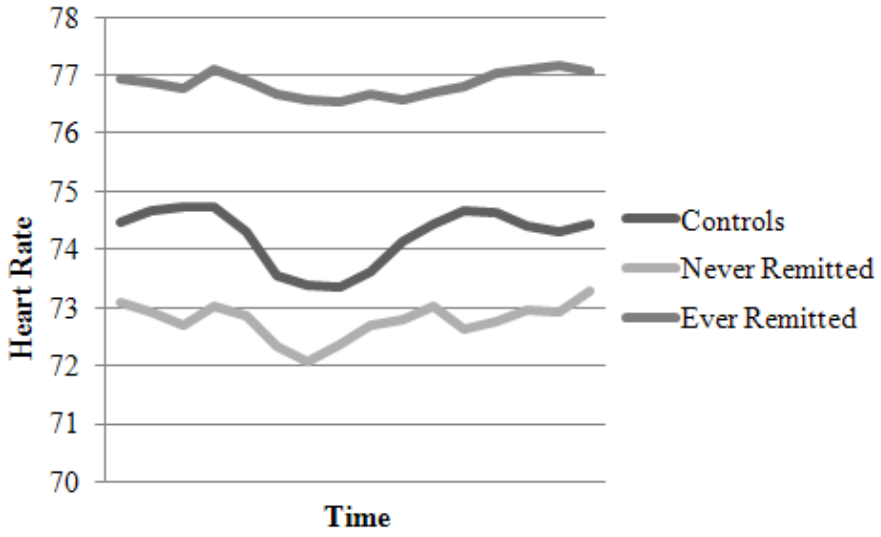

Figure 10. Heart rate waveforms during neutral pictures.

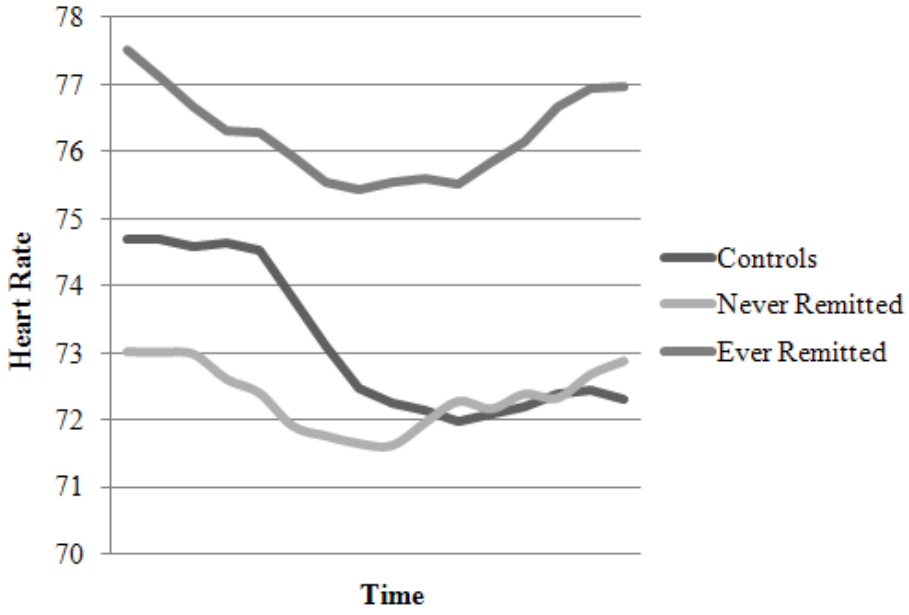

Figure 11. Heart rate waveforms during unpleasant pictures.

Skin conductance analyses. Exploratory analyses considered the ability of time one skin conductance to predict time two depression status, with a series of repeated-measures ANOVAs with time one SC as the predictor and time two diagnostic status as the outcome variable. A 3 (time one SC by valence) by 3 (Group: remitted for at least 2 months, never remitted, controls) repeated measures ANOVA failed to reveal a significant main effect of valence [F (2, 45) = .58, p =.51, η_p^2 = .01] or valence by group interaction [F (4, 45) = .64, p =.59, η_p^2 = .03]. Additionally, when group was defined as current MDE at time two, there was no significant main effect of valence [F (2, 45) = .26, p =.71, η_p^2 = .00] nor was there a group by valence interaction [F (2, 45) = 1.64, p =.19, η_p^2 = .07].

Table 8. Correlations between time one and time two subjective ratings.

Subjective Rating	T2
Time One Valence	
Pleasant	.62
Neutral	.72
Unpleasant	.85
Time One Arousal	
Pleasant	.50
Neutral	.73
Unpleasant	.56

All *p*s < .05

Time one and time two physiological correspondence. Test-retest reliability was high for ratings of picture valence and arousal (Table 8). In general, raw startle magnitudes showed greater correspondence over time one and time two as compared to startle magnitude. Consistent with analyses by Larson and colleagues (2000) we analyzed raw startle magnitudes, and all test-retest values were significant (Table 9). None of the test-retest values were significant for startle magnitude (all *p*s > .60). However, it is not surprising that startle magnitudes were not correlated between time one and time two because startle magnitude is obtained by transforming raw startle scores into within subject, standardized *T* scores. Computation of *T* scores removes between-subject variability, therefore decreasing the likelihood of detecting significant test-retest correlations.

Table 9. Correlations between time one and time two raw startle magnitudes.

Magnitude	T2 Pleasant	T2 Neutral	T2 Unpleasant	T2 Baseline
T1 Pleasant	.61			
T1 Neutral	.66	.75		
T1 Unpleasant	.52	.57	.57	
T1 Baseline	.53	.60	.55	.60

All *p*s < .01

HR variables showed varying levels of correspondence over time (Table 10). Specifically, D1 during neutral pictures (r = .53, p< .01), A1 during neutral pictures (r = .44, p < .05), and D2 during neutral (r = .44, p < .05) and unpleasant pictures (r = .48, p < .05) all showed moderate and significant levels of test-retest reliability. Lastly, SC showed good test-retest reliability for pleasant (r = .47, p < .05) and neutral (r = .50, p < .01) pictures only.

Table 10. Correlations between time one and time two heart rate measures.

	T2
T1 Baseline	.34
T1 D1	
Pleasant	-.07
Neutral	.57*
Unpleasant	-.03
T1 A1	
Pleasant	.14
Neutral	.44*
Unpleasant	.22
T1 D2	
Pleasant	.26
Neutral	.44*
Unpleasant	.48*

*p < .05

Discussion

MDD is a highly recurrent and impairing disorder. Functionalist theories of emotion predict that maladaptive emotional responses will be associated with a worse course of MDD. Studies of the emotion-modulated startle response in clinically depressed individuals have attempted to identify maladaptive emotional reactivity, and typically have found a reduction of the normal modulation of the startle response by emotional valence (Dichter et al., 2004; Dichter & Tomarken, 2008). Such a pattern is consistent with the idea that MDD is characterized by emotion context insensitivity (ECI; Rottenberg, Kasch, Gross & Gotlib, 2002) or context-inappropriate emotional responses. A flattened pattern of emotion-modulated startle has been documented in samples at various levels of depression severity (e.g., severely depressed individuals, Allen et al., 1999; Kaviani, Gray, Checkley, Wilson, & Kumari, 2004; Forbes et al., 2005; and in individuals with non-clinical depression, Mneimne, McDermut, & Powers, 2008), although it should be noted that the cross-sectional analyses of our mood disordered sample at time one did not find blunted emotion-modulated startle (Taylor-Clift, 2008).

The primary aim of the current study was to examine how emotional responding during a depressive episode predicted recovery 6 months later. To accomplish this aim, we gathered diagnostic and physiological data on individuals with Major Depressive Disorder (MDD), Minor Depressive Disorder (mD), and healthy controls at time one and

again, 6 months later, at time two. Specifically, we used the emotion-modulated startle paradigm, a well-studied physiological method for examining the motivational aspect of emotional reactivity to valenced stimuli (i.e., normed pictures with emotional content). We also collected subjective ratings of the valenced stimuli, as well as heart rate and skin conductance data during presentation of valenced stimuli.

Prior to testing our main hypotheses, we examined clinical outcomes at time two. We also conducted attrition analyses to ensure that our time two participants did not differ significantly from those individuals who were lost to attrition. Within the context of this information, we then tested our hypotheses. Based on prior studies of emotional functioning in depression (Morris, Bylsma & Rottenberg, 2009), we hypothesized that depressed individuals who displayed blunted emotional reactivity at time one, in the form of a flattened pattern of emotion-modulated startle, would be more likely to exhibit a worse course of depression, whereas those with normative emotion-modulated startle at time one would be more likely to demonstrate recovery from a depressive episode during the 6 months following the time one assessment. Specifically, we predicted that major and minor depressed individuals who showed the normative, linear pattern at time one would be more likely to experience recovery at time two (Hypothesis 1a). We also predicted that higher time one startle during pleasant pictures and lower time one startle during unpleasant pictures (opposite the normative linear pattern of startle) would predict higher time two measures of depression severity (Hypothesis 1b). Our last hypothesis for startle data predicted that higher time one levels of baseline startle would predict lower time two symptom severity (Hypothesis 1c) in line with prior results (O'Brien-Simpson

et al., 2009). Lastly, we collected startle response data at time one and time two in order to examine the stability of responses over this period of time.

The collection of self-report ratings of valenced stimuli, heart rate, and skin conductance was performed to augment our understanding of the startle analyses results. As such, we did not have hypotheses regarding the predictive utility of these measures. We did not expect time one self-report ratings to predict time two depression status, and we therefore made no predictions regarding this association. We also made no predictions regarding the ability of skin conductance values at time one to predict depression status at time two. In terms of heart rate data, we predicted only that the time one acceleration (A1) component of the heart rate waveform would predict time two depression status. Because A1 is thought to represent defensive orienting towards an affectively threatening stimulus, we predicted that lower defensive orienting to unpleasant pictures and higher defensive orienting to pleasant pictures (i.e., context inappropriate responses) would predict worse time two depression status.

Clinical Outcome of Sample at Follow-up

Groups differed at time two in ways that might be expected. In general, individuals who met MDD criteria at time one spent more time with syndromal MDD symptoms and experienced a greater number of total symptoms week by week over the 6-month follow-up. Despite the fact that the MDD group was the most burdened by symptoms, MDD and mD individuals did not differ in their likelihood of achieving a month long period of full recovery. Likewise, as expected individuals with MDD reported levels of depression symptom severity (on the BDI) compared to those with a

mD episode, who in turn reported high levels of symptom severity than controls. However, individuals with MDD did not report more severe Beck Anxiety Inventory (BAI) scores than the mD or control groups (mD individuals report higher levels of anxiety than controls).

Subjective Picture Ratings

Manipulation check analyses confirmed that participants overall rated pictures in the expected ways (e.g., negative pictures rated more negatively on valence than neutral pictures). Given the inconsistency of findings with self-report ratings in similar paradigms (Forbes et al., 2005; Dichter et al., 2004), we did not expect to find differences in time one self-report ratings for time two diagnostic groups and subjective ratings of valence did not, in fact, predict clinical outcome at time two. However, individuals who rated unpleasant pictures as *more* arousing were *less* likely to recover during the 6-month follow-up and had higher time two depression severity.

Although these results seem to suggest that individuals who were more activated by unpleasant stimuli were more likely to experience a worse depression course, these data were not supported by higher startle reactivity in the context of unpleasant pictures. As mentioned previously, this likely represents a frequently observed disconnect between subjective reporting and other measures of emotion (e.g., Sloan, Strauss, Quirk & Sajatovic, 1997; Russell, Bachorowski, & Fernandez-Dols, 2003). The disconnect between self-reported arousal and various other forms of arousal (i.e., fractionation) has been used as an argument for the complex nature of arousal as a construct (e.g., Davidson, 1978). Indeed, physiological markers of arousal are often employed because

arousal that exists after the presentation of an arousing stimulus is not always reliably accessible through self-reports (Cacioppo, Tassinary Stonebraker, & Petty, 1987). Furthermore, physiological measures have been found to predict outcomes better than self-report ratings (e.g., Fox, Cahill, & Zougkou, 2010; Rottenberg, Kasch, Gross & Gotlib, 2002). Although time two depressed individuals in the present study reported being more aroused by unpleasant pictures at time one, physiological measures did not support group differences in level of arousal (e.g., heart rate or skin conductance measures of arousal). Therefore, it was the self-report or belief (Robinson & Clore, 2002), and not necessarily the physiological level of arousal, that predicted time two depressive status. It may be the case that those depressed individuals who display a larger fractionation between physiological indicators of arousal and self-reports of arousal are at a higher risk for chronic depression.

Startle Responding

We predicted that a non-linear pattern of startle responding at time one would be associated with a worse course of depression over the 6-month follow-up. Specifically, we expected controls and those individuals who recovered from a depressive episode to show the normative linear pattern of startle modulation and for individuals who did not recover form a depressive episode to show a blunted pattern of modulated startle. Although we did not detect a group by linear trend, we did find a significant group by quadratic trend. Specifically, MDD and mD individuals who were experiencing at major depressive episode (MDE) at time two showed a quadratic pattern of startle responding, whereas those who did not have a time two depressive episode and healthy controls did

not display a quadratic pattern. The quadratic pattern of startle responding for depressed individuals with a MDE at time two was one in which startle responses during pleasant and unpleasant pictures were attenuated compared to neutral pictures. Stated in other terms, individuals who displayed exaggerated startle responding during neutral pictures and a lack of startle potentiation during unpleasant pictures were more likely to be experiencing a MDE six months later. Consistent with these results, higher startle responding during neutral pictures also predicted depression symptom severity at time two.

A handful of studies have found relationships between baseline startle responding (i.e., startle reflexes occurring in between affective pictures and in the absence of any affective stimuli) and depression outcome. For instance, higher baseline startle responses among recovered depressed patients were associated with a positive response to antidepressant therapy (Quednow et al., 2004). Consistent with these results, in formerly depressed individuals, lower baseline startle responses predicted an increase in depressive symptoms two years later (O'Brien-Simpson et al., 2009). This study extended these findings by including a sample with subthreshold depression, as well as healthy controls, and by examining the association between startle responses during valenced conditions. We found that lower startle reflexes occurring in the absence of emotional stimuli (baseline startle) predicted worse depression severity at time two.

Baseline startle responding is generally viewed as an evolutionarily adaptive measure of the defensive readiness of an organism (Landis & Hunt, 1939). Whereas higher levels of baseline startle responding appear to be indicative of hyperactive

defensive responding in individuals with anxiety disorders (e.g., Larsen, Norton, & Walker, 2002; Elsesser, Sartory, & Tackenberg, 2004), lower baseline startle in more chronically depressed patients may be indicative of disengagement from the environment in the absence of stimuli. O'Brien-Simpson and colleagues (2009) found lower levels of baseline startle in formerly depressed individuals at greatest risk for relapse, suggesting that hypoactive baseline startle persists even in the absence of depressive symptoms. This disengagement seems to put depressed individuals at risk for further depressive episodes and may be a stable marker for depression chronicity. Furthermore, baseline startle responding appears to be more heritable than emotion-modulation of startle (Anokhin, Golosheykin, & Heath, 2007), suggesting a strong genetic influence and/or a possible disruption in the neural circuitry associated with the basic startle response.

Indeed, attenuated startle responding has been associated with a number of neurological abnormalities that may provide a clue as to what brain structures are impaired in those more prone to recurrent depression. For instance, MDD has long been linked to increased levels of cortisol (hypercortisolemia) (Gibbons, 1966; Stokes, Pick, Stoll & Noll, 1975; Sachar, 1976; Carroll, Curtis, & Mendels, 1976). Cortisol is a hormone released by the adrenal gland during times of stress to prepare the body for the "fight or flight" defensive response. A dexamethasone suppression test (DST) results in decreased production of cortisol in healthy controls but has no effect on the levels of cortisol in depressed patients (Carroll, 1982). This nonsuppression after the DST indicates hyperactivity of the hypothalamic-pituitary-adrenal axis (HPA; see Gillespie & Nemeroff, 2005). This DST nonsuppression is predictive of a worse course of depression

(Arana, Baldessarini, & Ornsteen, 1985) and is a predictor of relapse (Ribeiro, Tandon, Grunhaus, & Greden, 1993). Diurnal variation in baseline startle responding has been linked to HPA axis activity (Miller & Gronfier, 2006). Specifically, Miller and Gronifer found that baseline startle responses in healthy individuals were highest in the evening—when cortisol levels were at their lowest—and lowest in the morning—when cortisol levels were highest. If MDD individuals are consistently stuck in a defensive state because of hypercortisolemia, there may be little room for an increase in defensiveness in response to a startle probe. Thus, deficient baseline startle responding may be a marker of hypercortisolemia, which is in turn, a marker of more severe, chronic depressive illness.

The amygdala has also been implicated in the impairment of the startle reflex. The amygdala appears to be associated with emotion experience and processing in a variety of contexts (Phillips, Drevets, Rauch, & Lane, 2003; Davidson, Jackson, & Kalin, 2000; Davidson, Jackson & Kalin, 2000). Researchers also argue that the amygdala is involved in the vigilance of an organism and the interpretation of potentially threatening or ambiguous stimuli in the environment (Davis & Whalen, 2001; Whalen, 1998). Furthermore, amygdala dysfunction is associated with a number of mood disorders, including depressive disorders (Drevets, 2003; Xiang et al., 2008). Amygdalar responses to emotional stimuli are blunted in depressed individuals (Thomas et al., 2001). Furthermore, amygdala volume is reduced in severely depressed, unmedicated MDD patients, and this decreased volume is directly proportional to the number of prior depressive episodes (Kronenberg et al., 2009). Therefore, underactive amygdala

functioning and reduced amygdala volume may be risk factors or consequences of chronicity in MDD. Electrical stimulation of the amygdala results in enhancement of the startle response (Rosen & Davis, 1988), and lesions to the amygdala are associated with attenuated startle responding (Angrilli et al., 1996; Funayama, Grillon, Davis, & Phelps, 2001). Taken together, this research suggests that our finding of decreased baseline startle responding in the more chronically depressed may be associated with impaired amygdala functioning.

What might seem most perplexing about our startle results is that both higher startle in the presence of neutral stimuli and lower baseline startle responding (in the ITI) were associated with worse depression outcomes. At first glance, it might seem natural to consider baseline startle responding as simply another form of "neutral" startle responding. However, it is important to understand the distinction between baseline startle and startle during neutral pictures when interpreting these results. Baseline startle is not a measure of the startle responses occurring at the beginning of the session. The first several startle probes presented to the participant are meant to habituate the participant to the noise. However, throughout a startle session, startle responses continue to decline as the participant continues to habituate to the startle probes. Baseline startle responses are collected within 9 seconds of picture offset (an average of approximately 7 seconds) in the absence of any stimuli. Neutral stimuli in the present study were generally household items (e.g., a spoon, a cup), buildings, or abstract images. Within the typical emotion-modulated startle paradigm, as was the case here, neutral stimuli serve as a reference for comparison of pleasantly and unpleasantly valenced stimuli.

Unlike baseline startle, during neutral pictures participants are asked to attend to a stimulus when the startle probe is presented. Therefore, while baseline startle represents a basic measure of the startle reflex, neutral startle represents startle during attending to non-valenced stimuli. Indeed the correlation between baseline startle responding and startle responding during neutral stimuli was only modest, at -.32, suggesting that these variables are only loosely related.

While we cannot claim that startle responses were blunted in those who later failed to remit from their depressive episode, we also cannot state that these individuals displayed normative, linear startle responding. On the contrary, never remitted, depressed individuals seemed to display a startle response indicative of reacting to neutral stimuli as if they were aversive. One possibility is that those who experienced a time two MDE interpreted neutral pictures as more threatening. It is especially interesting, then, that groups did not differ on their subjective ratings of picture valence for neutral stimuli. However, as previously mentioned, subjective ratings and physiological measures often do not correspond.

These results must be interpreted with caution. This is the first study to examine the predictive utility of startle responses to each valenced condition separately. However, other research paradigms indicate that depressed individuals may interpret ambiguous information more negatively than non-depressed individuals (Lawson & MacLeod, 1999), and that such a bias exists in individuals who are at-risk for developing depression (Dearing & Gotlib, 2009). Although this bias is often argued to be a cognitive one (requiring prolonged attending to a stimulus), the misinterpretation of ambiguous stimuli

has even been replicated in a startle reflex paradigm (Leppänen, Milders, Bell, Terriere, & Hietanen, 2004), indicating that the phenomenon may be pre-attentive. Therefore, depressed individuals' higher startle responding in the context of neutral stimuli could indicate that they interpreted neutral stimuli as negative stimuli, which in turn activated a defensive response. This possibility corresponds to our hypothesis of reduced amygdala functioning in more chronically depressed individuals, as the amygdala is associated with the resolution of ambiguous stimuli (reviewed above).

Skin Conductance

Skin conductance did not predict time two diagnostic status. Time one skin conductance levels also did not differ according to valence of pictures in the time two sample. According to the emotion-modulated startle paradigm, skin conductance levels in the context of pleasant and unpleasant stimuli should be equivalent, indicating similar levels of arousal in response to affective stimuli. In addition, arousal ratings of pleasant and unpleasant pictures should be significantly greater than arousal ratings of neutral pictures. SC data analyses did not match this expected outcome. However, our other manipulation checks (e.g., subjective ratings, HR), as well as time one manipulation checks (Taylor-Clift, 2008) confirm the valence and arousal categorizations of our stimuli.

Heart Rate

Heart rate (HR) data were collected as a secondary measure of emotion processing. In healthy controls, the phasic heart rate response involves an initial deceleration (D1), a subsequent acceleration (A1), and a secondary deceleration in HR

following the onset of a stimulus (D2). The current study found differences only in the A and D2 components of the phasic HR response. Specifically, individuals who never remitted from their time one depressive episode showed greater initial defensive and late anticipatory responding in the context of unpleasant pictures. Somewhat similar results were found in a study of individuals with and without Generalized Anxiety Disorder (GAD), in which those individuals with GAD showed a distinct acceleration in response to threatening words, while controls showed a deceleration during the same time period (Thayer, Friedman, Borkovec, Johnsen & Molina, 2000). During the presentation of a stimulus that served as a cue for upcoming unpleasant stimuli, those with GAD showed a larger D2 response, indicating a greater amount of anticipatory preparation. Therefore, it's possible that higher D2 responding in chronically depressed individuals is indicative of anticipatory preparation in response to the presentation of unpleasant stimuli.

Test-retest Reliability of Psychophysiological Measures

To date, there are very few studies on the stability of the startle response across two or more time points (Larson, Ruffalo, Nietert & Davidson, 2000) and no results regarding stability in individuals with MDD. Therefore, the current study included exploratory correlational analyses of time one and time two physiology. Correspondence differed according to the measure. Consistent with reliability results conducted by the authors of the picture stimuli (Lang, Bradley, & Cuthbert, 2005), correspondence was high for subjective ratings of both arousal and valence, and all correlations were significant. Although startle magnitude (*T* scores) at time one and time two were not correlated, startle magnitude (raw values) showed strong correlations across time points.

These results are remarkably consistent with results of Larson and colleagues who examined test-retest stability of startle responses over the course of one month. They found test-retest correlations of .55, .55, and .49 for unpleasant, neutral and pleasant pictures, respectively, compared to our results of .57, .75, and .61 for unpleasant, neutral, and pleasant pictures respectively. This adds to the evidence of raw startle magnitude as a stable marker of defensive responding.

Measures of heart rate (HR; including baseline HR) were correlated only during viewing of neutral pictures, and for the secondary deceleration during unpleasant pictures. Like startle studies, there are very few longitudinal studies on the stability of the heart rate waveform, or heart period (HP). There are several longitudinal studies of HR, however, which indicate that cardiovascular measures, such as HP and HR, are generally stable across time in developing infants (Fracasso, Porges, Lamb & Rosenberg, 1994), during physiologically stressful tasks across ten years (cold pressor; Sherwood et al., 1997), and across various settings (Gerin et al., 1998). In this respect, we are surprised by the lack of test-retest reliability between time one and time two baseline HR measures.

Strengths and Limitations

The current study is one of the few studies of startle modulation to have examined the course of depression in a longitudinal study for the purpose of predicting recovery. The current study utilized multiple sources of emotion data to predict several outcomes of depression. Furthermore, the inclusion of a subthreshold depression group allowed for more variability in depression measures and emotional reactivity than in previous studies.

Another strength of the study involves the method of diagnostic assessment at the follow-up interview. We developed an extensive interview based on the Structured Clinical Interview for the DSM-IV (SCID; First, Spitzer, Gibbon & Williams, 2002), which included diagnostic assessment for the entire six-month follow-up period. The interview collected diagnostic information for mood symptoms week-by-week. The resulting data provided a more complete context for the course of the depressive episode and included information that is not captured by the traditional SCID.

There were also a number of potential limitations with the current study. Attrition was substantial, and above the expected 20% of our original sample for which we had planned, despite all best efforts. We conducted attrition analyses to ensure that those who participated at time two did not differ from those who did not participate at time two. There were no significant differences in any predictor variable, with the exception of startle responding in the context of neutral pictures. Individuals who participated in the time two follow-up had significantly higher startle responses to neutral pictures than did those who did not participate at time two. Interpretation of startle responses during neutral stimuli should be mindful of possible attrition effects; it is conceivable that inclusion of attritted individuals with lower startle responses in the context of neutral stimuli would have eliminated the significant group differences in startle responding during neutral stimuli. Our greater than anticipated attrition rates also limited our ability to conduct certain categorical and covariation analyses. Categorical analyses were limited by small sample sizes. In fact, observed power for repeated measures ANOVAs of group by startle interactions was only .06.

Summary

These results emphasize the importance of multi-method studies of emotion when predicting depression course. While self-report ratings of unpleasant pictures predicted depression course, startle responding during neutral pictures predicted time two diagnostic status. Higher time one startle responding during neutral pictures predicted a worse time two outcome. In addition, baseline startle responding also predicted time two depression severity, such that lower time one startle responding in the absence of picture stimuli predicted a worse outcome at time two. Perhaps most compelling, this association between time one baseline startle and time two depression severity corroborated prior results (O'Brien-Simpson, Parsia, Simmons & Allen, 2009) and may be indicative of an overall neurologically based deficit in defensive responding, associated with dysfunctional HPA axis activity and/or amygdala dysfunction. Overall, raw startle magnitude showed high test-retest reliability, indicating that startle may serve as an important marker of emotional responding. Although the current results did not provide much support for the Emotion Context Insensitivity (ECI) theory of blunted emotional responding being associated with worse depression course, neither do our data support the theories of positive attenuation or negative potentiation. In fact, the interpretation of neutral and ambiguous stimuli as threatening supports the theory of negative interpretive bias (e.g., Lawson & MacLeod, 1999). Given the novelty of the current design, these results should be interpreted with caution. However, results do provide more evidence that low baseline startle responding may serve as a marker of more chronic forms of depression.

References

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective startle modulation in clinical depression: Preliminary findings. *Biological Psychiatry, 46,* 542-550.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (Fourth Edition--Text Revision ed.). Washington: American Psychiatric Association.

Angrilli, A., Mauri, A., Palomba, D., Flor, H., Birbaumer, N., Sartori, G., & di Paola, F. (1996). Startle reflex and emotion modulation impairment after a right amygdala lesion. *Brain, 119,* 1991–2000.

Anokhin, A. P., Golosheykin, S., & Heath, A. C. (2007). Genetic and environmental influences on emotion-modulated startle reflex: A twin study. *Psychophysiology, 44,* 106-112.

Arana, G. W., Baldessarini, J. R., & Ornsteen, M. (1985). The dexamethasone suppression test for diagnosis and prognosis in psychiatry. *Archives of General Psychiatry, 42,* 1193–204.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Hoeber.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal for Consulting and Clinical Psychology, 56,* 893-897.

Beck, A. T., Rush, A. J., Shaw, B. R., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-2nd Edition (BDI-II) Manual*. San Antonio: The Psychological Corporation.

Berenbaum, H., & Oltmanns, T. F. (1992). Emotional experience and expression in schizophrenia and depression. *Journal of Abnormal Psychology, 101,* 37-44.

Berg, W. K., & Donohue, R. L. (1992). Anticipatory processes in infants: Cardiac components. In B. A. Campbell, H. Hayne & R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 61-80). Hillsdale, England: Lawrence Erlbaum Associates.

Berlanga, C., Heinze, G., Torres, M., Apiquian, C., & Caballero, A. (1999). Personality and clinical predictors of recurrence of depression. *Psychiatric Services, 50,* 376-380.

Berntson, G. G., Quigley, K. S., & Lozano, D. (2007). Cardiovascular Psychophysiology. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 182-210). New York, NY: Cambridge University Press.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42*, 1-15.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*, 276-298.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1990). Startle reflex modification: Emotion or attention? *Psychophysiology, 27*, 513-522.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 49-59.

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology, 37*, 204-215.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: Habituation in humans. *Behavioral Neuroscience, 107*, 970-980.

Brown, J. S., Kalish, H. I., & Farber, I. E. (1951). Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *Journal of Experimental Psychology, 41*, 317-328.

Burgess, H. A., Johnson, S. L., & Granato, M. (2009). Unidirectional startle responses and disrupted left-right co-ordination of motor behaviors in robo3 mutant zebrafish. *Genes, Brain & Behavior, 8*, 500-511.

Butzlaff, R. L., & Hooley, J. M. (1998). Expressed emotion and psychiatric relapse: A meta-analysis. *Archives of General Psychiatry, 55*, 547-552.

Bylsma, L. M., Morris, B., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in Major Depressive Disorder. *Clinical Psychology Review, 28*, 676-691.

Cacioppo, J. T. Tassinary, L. G., Stonebraker, T. B., & Petty, R. E. (1987). Self-report and cardiovascular measures of arousal: Fractionation during residual arousal. *Biological Psychology, 25,* 135-151.

Carroll, B. J., Curtis, G. C., Mendels, J. (1976). Neuroendocrine regulation in depression: II. Discrimination of depressed from nondepressed patients. *Archives of General Psychiatry, 33,* 1051-1058.

Center for the Study of Emotion and Attention [CSEA-NIMH] (1999). The international affective picture system: Digitized photographs. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Clark, D. C., Fawcett, J., Salazar-Grueso, E., & Fawcett, E. (1984). Seven-month clinical outcome of anhedonic and normally hedonic depressed inpatients. *American Journal of Psychiatry, 141*, 1216-1220.

Coryell, W., Akiskal, H. S., Leon, A. C., Winokur, G., Maser, J. D., Mueller, T. L., et al. (1994). The time course of nonchronic major depressive disorder: Uniformity across episodes and samples. National Institute of Mental Health Collaborative Program on the Psychobiology of Depression—Clinical Studies. *Archives of General Psychiatry, 51*, 405-410.

Cuijpers, P., de Graaf, R., & van Dorsselaer, S. (2004). Minor depression: Risk profiles, functional disability, health care use and risk of developing major depression. *Journal of Affective Disorders, 79*, 71-79.

Davidson, R. J. (1978). Specificity and patterning in biobehavioral systems: Implications for behavior change. *American Psychologist, 33*, 430-436.

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, and regulation: Perspectives from affective neuroscience. *Psychological Bulletin, 126*, 890-909.

Davis, M. (1989). Neural systems involved in fear-potentiated startle. *Annals of the New York Academy of Sciences, 563*, 165-183.

Davis, M. (2006). Neural systems involved in fear and anxiety measured with fear-potentiated startle. *American Psychologist, 61*, 741-756.

Davis, M., Antoniadis, E. A., Amaral, D. G., & Winslow, J. T. (2008). Acoustic startle reflex in rhesus monkeys: A review. *Reviews in the Neurosciences, 19*, 171-185.

Davis, M., Gendelman, D. S., Tischler, M. D., & Gendelman, P. M. (1982). A primary acoustic startle circuit: Lesion and stimulation studies. *Journal of Neuroscience, 6*, 791– 805.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159-181). New York, NY: Cambridge University Press.

de Jong, P. J., Merckelbach, H., & Arntz, A. (1991). Eyeblink startle response in spider phobics before and after treatment: A pilot study. *Journal of Psychopathology & Behavioral Assessment, 13*, 213-223.

de Jong, P. J., Visser, S., & Merckelbach, H. (1996). Startle and spider phobia: Unilateral probes and the prediction of treatment effects. *Journal of Psychophysiology, 10*, 150-160.

Davis, M. & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6*, 13-34.

Dearing, K. F., & Gotlib, I. H. (2009). Interpretation of ambiguous information in girls at risk for depression. *Journal of Abnormal Child Psychology: An official publication of the International Society for Research in Child and Adolescent Psychopathology, 37*, 79-91.

Dichter, G. S., & Tomarken, A. J. (2008). The chronometry of affective startle modulation in unipolar depression. *Journal of Abnormal Psychology, 117*, 1-15.

Dichter, G. S., Tomarken, A. J., Shelton, R. C., & Sutton, S. K. (2004). Early- and late-onset startle modulation in unipolar depression. *Psychophysiology, 41*, 433-440.

Dickens, C., McGowan, L., & Dale, S. (2003). Impact of depression on experimental pain perception: A systematic review of the literature with meta-analysis. *Psychosomatic Medicine, 65*, 369-375.

Drevets, W. C. (2003). Neuroimaging abnormalities in the amygdala in mood disorders. *Annals of the New York Academy of Sciences, 985*, 420-444.

Dunn, B. D., Dalgleish, T., Lawrence, A. D., Cusack, R., & Ogilvie, A. D. (2004). Categorical and dimensional reports of experienced affect to emotion-inducing pictures in depression. *Journal of Abnormal Psychology, 113,* 654-660.

Egloff, B., Wilhelm, F. H., Neubauer, D. H., Mauss, I. B., & Gross, J. J. (2002). Implicit anxiety measure predicts cardiovascular reactivity to an evaluated speaking task. *Emotion, 2,* 3-11.

Enns, M. W., & Cox, B. J. (2005). Psychosocial and clinical predictors of symptom persistence vs remission in major depressive disorder. *Canadian Journal of Psychiatry, 50,* 769-777.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR Axis I disorders, research version, patient edition with psychotic screen (SCID-I/P W/ PSY SCREEN).* New York, NY: Biometrics Research, New York State Psychiatric Institute.

Fogel, J., Eaton, W. W., & Ford, D. E. (2006). Minor depression as a predictor of the first onset of major depressive disorder over a fifteen-year follow-up. *Acta Psychiatrica Scandinavica, 113,* 36-43.

Forbes, E. E., Miller, A., Cohn, J. F., Fox, N. A., & Kovacs, M. (2005). Affect-modulated startle in adults with childhood-onset depression: Relations to bipolar course and number of lifetime depressive episodes. *Psychiatric Research, 134,* 11-25.

Fox, E., Cahill, S., & Zougkou, K. (2010). Preconscious processing biases predict emotional reactivity to stress. *Biological Psychiatry, 67,* 371-377.

Fracasso, M. P., Porges, S. W., Lamb, M. E., & Rosenberg, A. A. (1994). Cardiac activity in infancy: Reliability and stability of individual differences. *Infant Behavior & Development, 17*, 277-284.

Fraguas, R., Jr., Marci, C., Fava, M., Iosifescu, D. V., Bankier, B., Loh, R., et al. (2007). Autonomic reactivity to induced emotion as potential predictor of response to antidepressant treatment. *Psychiatry Research, 151*, 169-172.

Funayama, E. S., Grillon, C., Davis, M., & Phelps, E. A. (2001). A double dissociation in the affective modulation of startle in humans: Effects of unilateral temporal lobectomy. *Journal of Cognitive Neuroscience, 13*, 721–729.

Gehricke, J. G., & Shapiro, D. (2000). Reduced facial expression and social context in major depression: Discrepancies between facial muscle activity and self-reported emotion. *Psychiatry Research, 95*, 157-167.

Gerin, W., Christenheld, N., Pieper, C., DeRafael, D. A., Su, O., Stroessner, S. J., et al. (1998). The generalizability of cardiovascular responses across settings. *Journal of Psychosomatic Research, 44*, 209-218.

Giakoumaki, S. G., Bitsios, P., Frangou, S., Roussos, P., Aasen, I., Galea, A., et al. (in press). Low baseline startle and deficient affective startle modulation in remitted bipolar disorder patients and their unaffected siblings. *Psychophysiology*.

Gibbons, J. L. (1966). The secretion rate of corticosterone in depressive illness. *Journal of Psychosomatic Research, 10*, 263-266.

González-Tejera, G., Canino, G., Ramírez, R., Chávez, L., Shrout, P., Bird, H., Bravo, M., Martínez-Taboas, A., Ribera, J., & Bauermeister, J. (2005). Examining minor

and major depression in adolescents. *Journal of Child Psychology & Psychiatry, 46*, 888-899.

Griffin, M. G. (2008). A prospective assessment of auditory startle alterations in rape and physical assault survivors. *Journal of Traumatic Stress, 21*, 91-99.

Gunthert, K. C., Cohen, L. H., Butler, A. C., & Beck, J. S. (2005). Predictive role of daily coping and affective reactivity in cognitive therapy outcome: Application of a daily process design to psychotherapy research. *Behavior Therapy, 36*, 77-88.

Gunthert, K. C., Cohen, L. H., Butler, A. C., & Beck, J. S. (2005). Predictive role of daily coping and affective reactivity in cognitive therapy outcome: Application of a daily process design to psychotherapy research. *Behavior Therapy, 36*, 77-88.

Guthrie, R. M., & Bryant, R. A. (2005). Auditory startle response in firefighters before and after trauma exposure. *American Journal of Psychiatry, 162*, 283-290.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford Press.

Hazlett, E. A., Dawson, M. E., Schell, A. M., & Nuechterlein, K. H. (2007). Probing attentional dysfunctions in schizophrenia: Startle modification during a continuous performance test. *Psychophysiology, 45*, 632-642.

Henriques, J. B., & Davidson, R. J. (2000). Decreased responsiveness to reward in depression. *Cognition and Emotion, 14*, 711-724.

Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1*, 69-82.

Jackson, K. M., & Sher, K. J. (2003). Alcohol use disorders and psychological distress: A prospective state-trait analysis. *Journal of Abnormal Psychology, 112*, 599-613.

Judd, L. L., Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., et al. (1998). Special Issue: George Winokur. *Journal of Affective Disorders, 50*, 97-108.

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology, 111*, 589-597.

Katon, W., Lin, E., Von Korff, M., Bush, T., Walker, E., Simon, G., et al. (1994). The predictors of persistence of depression in primary care. *Journal of Affective Disorders, 31*, 81-90.

Kaviani, H., Gray, J. A., Checkley, S. A., Wilson, G. D., & Kumari, V. (2004). Affective modulation of the startle response in depression, anhedonia, and anxiety. *Journal of Affective Disorders, 83*, 21-31.

Keller, M. B., Lavori, P. W., & Friedman, B. (1987). The longitudinal interval follow-up evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry, 44*, 540-548.

Keller, M. B., Lavori, P. W., Mueller, T. I., Endicott, J., Coryell, W., Hirschfeld, R., et al. (1992). Time to recovery, chronicity, and levels of psychopathology in major depression: A 5-year prospective follow-up of 431 subjects. *Archives of General Psychiatry, 49*, 809-816.

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition and Emotion, 13,* 467-480.

Kennedy, S. E., Koeppe, R. A., Young, E. A., & Zubieta, J. (2006). Dysregulation of endogenous opioid emotion regulation circuitry in major depression in women. *Archives of General Psychiatry, 63,* 1199-1208.

Kessler, R. C., Zhao, S., Blazer, D. G., & Swartz, M. (1997). Prevalence, correlates, and course of minor depression and major depression in the national comorbidity survey. *Journal of Affective Disorders, 45,* 19-30.

Kring, A. M., & Werner, K. H. (2004). Emotion regulation in psychopathology. In P. Philippot & R. S. Feldman (Eds.), *The Regulation of Emotion* (pp. 359-385). New York, NY: LEA.

Kronenberg, G., van Elst, L. T., Regen, F., Deuschle, M., Heuser, I., & Colla, M. (2009). Reduced amygdala volume in newly admitted psychiatric in-patients with unipolar major depression. *Journal of Psychiatric Research, 43,* 1112-1117.

Landis, C., & Hunt, W. A. (1939). *The startle pattern*. New York, NY: Farrar and Company, Inc.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119-137). Norwood, NJ: Ablex.

Lang, P. J. (1985). *The psychophysiology of emotion: Anxiety and the anxiety disorders*. Hillsdale, NJ: Lawrence Erlbaum.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50,* 372-385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377-395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*, 1248-1263.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual.* (No. Technical Report A-6). Gainsville, FL: University of Florida.

Lang, P. J., Davis, M., & Ohman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders, 61,* 137-159.

Lara, M. E., Leader, J., & Klein, D. (1997). The association between social support and course of depression: Is it confounded with personality? *Journal of Abnormal Psychology, 106,* 478-482.

Larson, C. L., Ruffalo, D., Nietert, J. Y., & Davidson, R. J. (2000). Temporal stability of the emotion-modulated startle response. *Psychophysiology, 37,* 92-101.

Lawson, C., & MacLeod, C. (1999). Depression and the interpretation of ambiguity. *Behaviour Research and Therapy, 37,* 463-474.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Leppänen, J. M., Milders, M., Bell, J. S., Terriere, E., & Hietanen, J. K. (2004). Depression biases the recognition of emotionally neutral faces. *Psychiatry Research, 128,* 123-133.

Leskelä, U., Rytsälä, H., Komulainen, E., Melartin, T., Sokero, P., Lestelä-Mielonen, P., et al. (2006). The influence of adversity and perceived social support on the outcome of major depressive disorder in subjects with different levels of depressive symptoms. *Psychological Medicine, 36,* 779-788.

Levenson, R. W. (1994). Human emotions: A functional view. In P. Ekman & R. J. Davidson (Ed.), *The nature of emotion: Fundamental questions* (pp. 123-126). New York, NY: Oxford University Press.

McFarland, B. R., Shankman, S. A., Tenke, C. E., Bruder, G. E., & Klein, D. N. (2006). Behavioral activation system deficits predict the six-month course of depression. *Journal of Affective Disorders, 91,* 229-234.

Miller, M. W., & Gronfier, C. (2006). Diurnal variation of the startle reflex in relation to HPA-axis activity in humans. *Psychophysiology, 43,* 297-301.

Mneimne, M., McDermut, W., & A.S., P. (2008). Affective ratings and startle modulation in people with nonclinical depression. *Emotion, 8,* 552-559.

Moos, R. H., & Cronkite, R. C. (1999). Symptom-based predictors of a 10-year chronic course of treated depression. *Journal of Nervous Mental Disorders, 187,* 360-368.

Morris, B. H., Bylsma, L. M., & Rottenberg, J. (2009). Does emotion predict the course of Major Depressive Disorder? A review of prospective studies. *British Journal of Clinical Psychology, 48,* 255–273.

Mueller, T. I., Keller, M. B., Leon, A. C., Solomon, D. A., Shea, M. T., Coryell, W., et al. (1996). Recovery after 5 years of unremitting major depressive disorder. *Archives of General Psychiatry, 53,* 794-799.

Murray, C. J., & Lopez, A. D. (1997). Global mortality, disability, and the contribution of risk factors: Global burden of disease study. *Lancet, 349*, 1436-1442.

O'Brien-Simpson, L., Parsia, P. D., Simmons, J. G., & Allen, N. B. (2009). Recurrence of major depressive disorder is predicted by inhibited startle magnitude while recovered *Journal of Affective Disorders, 112*, 243-249.

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. *Journal of Abnormal Psychology*, *102*, 82-92.

Peeters, F., Berkhof, J., Rottenberg, J., & Nicolson, N. A. (in press). Ambulatory emotional reactivity to daily life events predicts the eighteen-month course of major depressive disorder. *Behavioral Research and Therapy.*

Peeters, F., Nicolson, N. A., Berkhof, J., Delespaul, P., & deVries, M. (2003). Effects of daily events on mood states in major depressive disorder. *Journal of Abnormal Psychology, 112*, 203-211.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry, 54*, 504-514.

Pissiota, A., Frans, Ö., Michelgård, Å., Appel, L., Långström, B., Flaten, M. A., et al. (2003). Amygdala and anterior cingulate cortex activation during affective startle modulation: A PET study of fear. *European Journal of Neuroscience, 18*, 1325-1331.

Pole, N., Neylan, T. C., Otte, C., Henn-Hasse, C., Metzler, T. J., & Marmar, C. R. (2009). Prospective prediction of posttraumatic stress disorder symptoms using fear potentiated auditory startle responses. *Biological Psychiatry, 65*, 235-240.

Porges, S. W., Doussard-Roosevelt, J. A., & Maita, A. K. (1994). Vagal tone and the physiological regulation of emotion. *Monographs of the Society for Research in Child Development, 59*, 167-186, 250-283.

Quednow, B. B., Kuhn, K., Stelzenmueller, R., Hoenig, K., Maier, W., & Wagner, M. (2004). Effects of serotonergic and noradrenergic antidepressants on auditory startle response in patients with major depression. *Psychopharmacology, 175*, 399-406.

Ribeiro, S. C., Tandon, R., Grunhaus, L., & Greden, J. F. (1993). The DST as a predictor of outcome in depression: a meta-analysis. *American Journal of Psychiatry, 150*, 1618-1629.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an acceptability model of emotional self-report. *Psychological Bulletin, 128*, 934-960.

Rosen, J. B. & Davis, M. (1988). Enhancement of acoustic startle by electrical stimulation of the amygdala. *Behavioral Neuroscience, 102*, 195-202.

Rottenberg, J., Chambers, A. S., Allen, J. J. B., & Manberc, R. (2007). Cardiac vagal control in the severity and course of depression: The importance of symptomatic heterogeneity *Journal of Affective Disorders, 103*, 173-179.

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depression disorder. *Journal of Abnormal Psychology, 114*, 627-639.

Rottenberg, J., Gross, J. J., Wilhelm, F. H., Najmi, S., & Gotlib, I. H. (2002). Crying threshold and intensity in major depressive disorder. *Journal of Abnormal Psychology, 111,* 302-312.

Rottenberg, J., Kasch, K., Gross, J., & Gotlib, I. (2002). Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder. *Emotion, 2*, 135-146.

Rottenberg, J., Kasch, K. L., Gross, J. J., & Gotlib, I. H. (2002). Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder. *Emotion, 2,* 135-146.

Rottenberg, J., Salomon, K., Gross, J. J., & Gotlib, I. H. (2005). Vagal withdrawal to a sad film predicts subsequent recovery from depression. *Psychophysiology, 42*, 277-281.

Rottenberg, J., Wilhelm, F. H., Gross, J. J., & Gotlib, I. H. (2002). Respiratory sinus arrhythmia as a predictor of outcome in major depressive disorder. *Journal of Affective Disorders, 71*, 265-272.

Russell, J. A., Bachorowski, J.-A., & Fernandez-Dols, J.-M. (2003). Facial and vocal expressions of emotions. *Annual Review of Psychology, 54,* 359-349.

Sachar, E. J. (1976). Neuroendocrine dysfunction in depressive illness. *Annual Review of Medicine, 27,* 389-396.

Sakado, K., Sato, T., Uehara, T., Sakado, M., & Someya, T. (1999). Perceived parenting pattern and response to antidepressants in patients with major depression. *Journal of Affective Disorders, 52*, 59-66.

Schlenker, R., Cohen, R., & Hopmann, G. (1995). Affective modulation of the startle reflex in schizophrenic patients. *European Archives of Psychiatry & Clinical Neuroscience, 245*, 309-318.

Scott, J., Williams, J. M. G., Brittlebank, A., & Ferrier, N. (1995). The relationship between premorbid neuroticism, cognitive dysfunction and persistence of depression: A 1 year follow up. *Journal of Affective Disorders, 33*, 167-172.

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse*. New York, NY: Guilford Press.

Sherwood, A., Girdler, S. S., Bragdon, E. E., West, S. G., Brownley, K. A., Hinderliter, A. L., et al. (1997). Ten-year stability of cadiovascular responses to laboratory stressors. *Psychophysiology, 34*, 185-191.

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders, 46*, 135-141.

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders, 46*, 135-141.

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished response to pleasant stimuli by depressed women. *Journal of Abnormal Psychology, 110*, 488-493.

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished response to pleasant stimuli by depressed women. *Journal of Abnormal Psychology, 110*, 488-493.

Sokolov, Y. N. (1963). *Perception and the conditioned reflex* (S. J. Waydenfeld, Trans.). New York, NY: Macmillan.

Somsen, R. J., Van der Molen, M. W. & Orlebeke (1983). Phasic heart rate changes in reaction time, shock avoidance and unavoidable shock tasks: Are hypothetical generalizations about different S1-S2 tasks justified? *Psychophysiology, 20*, 88-94.

Spijker, J., Bijl, R. V., de Graaf, R., & Nolen, W. A. (2001). Determinants of poor 1-year outcome of DSM-III-R major depression in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica, 103*, 122-130.

Stokes, P. E., Pick, G. R., Stoll, P. M., Nunn, W. D. (1975). Pituitary-adrenal function in depressed patients: Resistance to dexamethosone supression. *Journal of Psychiatric Research, 12*, 271-281.

Sweeney, P. D., Anderson K. & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology, 50*, 974-991.

Taiminen, T., Jaakelainen, S., Ilonen, T., Meyer, H., Karlsson, H., & Lauerma, H., et al. (2000). Habituation of the blink reflex in first episode schizophrenia, psychotic depression, and non-psychotic depression. *Schizophrenia Research, 44*, 69-79.

Taylor-Clift, A. (2008). *Emotion-modulated startle in major and minor depression: The role of mood severity in emotion reactivity* (Master Thesis). Available from USF Electronic Theses and Dissertations. (Digital ID No. SFE0002419).

Taylor-Clift, A., Morris, B., Rottenberg, J., & Kovacs, M. (in press). Emotion-modulated startle in anxiety disorders is blunted by comorbid depressive episodes. *Psychological Medicine*.

Thayer, J. F., Friedman, B. H., Borkovec, T. D., Johnsen, B. H., & Molina, S. (2000). Phasic heart period reactions to cued threat and nonthreat stimuli in generalized anxiety disorder. *Psychophysiology, 37*, 361-368.

Thomas, K. M., Drevets, W. C., Dahl, R. E., Ryan, N. D., Birmaher, B., Eccard, C. H., Axelson, D., Whalen, P. J., & Casey, B. J. (2001). Amygdala response to fearful faces in anxious and depressed children. *Archives of General Psychiatry, 58*, 1057-1063.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology & Sociobiology, 11*, 375-424.

Vrana, S. R., & Lang, P. J. (1990). Fear imagery and the startle-probe reflex. *Journal of Abnormal Psychology, 99*, 189-197.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology, 97*, 487-491.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 47*, 1063-1070.

Weems, C. F., Pina, A. A., & Costa, N. M. (2007). Predisaster trait anxiety and negative affect predict posttraumatic stress in youths after hurricane Katrina. *Journal of Consulting and Clinical Psychology, 75*, 154-159.

Wexler, B. E., Levenson, L., Warrenburg, S., & Price, L. H. (1994). Decreased perceptual sensitivity to emotion-evoking stimuli in depression. *Psychiatry Research, 51*, 127-138.

Whalen, P. J. (1998). Fear, vigilance and ambiguity: Initial neuroimagin studies of the human amygdala. *Current Direction in Psychological Science, 7*, 177-188.

Winokur, G., Coryell, W., & Keller, M. (1993). A prospective follow-up of patients with bipolar and primary unipolar affective disorder. *Archives of General Psychiatry, 50*, 457-465.

Appendix: Additional Data Analyses

Time Two Startle and Time Two Diagnostic Status

Although the small sample size precluded testing categorical analyses of time two physiological variables and time two diagnoses, correlation analyses were conducted between time two measures of severity and subjective picture ratings, startle magnitude, raw startle magnitude, heart rate variables, and skin conductance for all pictures. Correlations are listed in Table A.11.

Table A.11. Correlations between time two physiological measures and time two Beck Depression Inventory scores.

	Rating		*Startle*	*SC*	*Heart Rate*		
Condition	Arousal	Valence			D1	A1	D2
Pleasant	-.31	-.13	-.15	-.09	-.17	-.21	-.35§
Neutral	-.44*	-.11	.18	-.26	-.32	-.19	-.19
Unpleasant	-.38*	.13	.17	-.03	.19	.23	.33

* $p < .05$, § $p = .06$, all others ns
Startle values are for startle magnitudes; SC = Skin Conductance; D1 = First deceleration in heart rate; A1 = Acceleration in heart rate; D2 = Second deceleration.

For time two subjective ratings of picture valence and time two BDI, there were no significant correlations (all *p*s > .46). For ratings of picture arousal and time two BDI, correlations for all picture conditions were significant or near significant, such that lower time two BDI scores were associated with rating pictures as more highly arousing at time

two. There were no significant correlations between time two BDI and time two startle magnitude or time two raw magnitude (all *p*s > .17).

There were several near significant associations between time two HR and time two BDI scores. Specifically, higher time two BDI scores predicted a smaller D1 during neutral pictures (r = -.32, *p* = .09), a smaller D2 during pleasant pictures (*r* = -.35, *p* = .06), and a larger D2 during unpleasant pictures (*r* = .33, *p* = .08; all other *p*s > .23). Lastly, there were no significant correlations between time two SC and time two measures of severity (all *p*s > .52).