An Experience-Sampling Approach to Examining Cortisol and Testosterone Profiles in Bipolar and Depressive Mood Disorders

June Gruber[1+], Keith M. Welker[2+], John R. Purcell[3], Bethany Lassetter[4], Smrithi Prasad[5], Cassandra M. Brandes[6], & Pranjal H. Mehta[7]

[1] University of Colorado Boulder, Department of Psychology and Neuroscience

[2] The Constructive Dialogue Institute

[3] Indiana University, Department of Psychological & Brain Sciences

[4] New York University, Department of Psychology

[5] Yale-NUS College, Division of Social Sciences

[6] Northwestern University, Department of Psychology

[7] University of Illinois at Chicago, Department of Psychiatry

[8] University College London, Department of Experimental Psychology

In Press, *Journal of Emotion and Psychopathology*

[+] Shared first and corresponding authorship of June Gruber and Keith Welker

Corresponding authors:

June Gruber, Ph.D.
University of Colorado Boulder
Department of Psychology and Neuroscience
345 UCB Muenzinger D321C
Boulder, CO 80309-0345
Phone: 510-410-3822
Fax: 303-492-2967
Email: june.gruber@colorado.edu

Keith M. Welker, Ph.D.
The Constructive Dialogue Institute
Email: keith@constructivedialogue.org

Abstract

Objectives: Emotion dysregulation is a core component of bipolar disorder (BD). However, there is a critical gap concerning the measurement of emotion-relevant neuroendocrine processes in this population.

Methods: The present investigation examined neuroendocrine profiles (i.e., cortisol and testosterone) alongside emotional experiences across three consecutive days in naturalistic settings among adults with bipolar I disorder (n = 28) compared to a clinical comparison group of adults with a history of major depressive disorder (MDD; n = 28) and healthy non-psychiatric controls (n = 27).

Results: The BD group exhibited decreased cortisol concentrations throughout the day, lower overall hormone output, and smaller awakening responses when compared to MDD, and flatter slopes compared to both MDD and control groups. The BD group did not differ from either group in any diurnal testosterone measures. Groups did not differ in mean positive and negative daily emotion ratings, but the BD group reported greater positive and negative variability, and the and MDD group greater NA variability, than the control group.

Conclusions: Despite similar emotion experiences during daily living, participants with BD exhibited decreased cortisol slopes relative to those with depression and controls. Cortisol levels across all measures in MDD were comparable to controls. Findings suggest the link between cortisol and mood is complex, and decreased cortisol profiles may be a marker of bipolar mood disorders.

An Experience-Sampling Approach to Examining Cortisol and Testosterone Profiles in Bipolar and Depressive Mood Disorders

Bipolar disorder (BD) is a chronic and severe psychiatric illness, ranked as one of the leading causes of disability worldwide (Ferrari et al., 2016). Despite robust support that emotion difficulties are a core component of BD (e.g., Gruber, 2011b; Johnson, 2005), less is known about the underlying psychophysiological processes of emotion difficulties in BD. Identifying specific markers of emotion-relevant disturbance may help inform explanations for both disorder specific and transdiagnostic processes. The present investigation adopted an affective neuroendocrine approach (e.g., Welker, Gruber, & Mehta, 2015) by examining two theoretically relevant hormone profiles (i.e., cortisol and testosterone) across a three-day naturalistic sampling period among individuals with BD compared to those with major depressive disorder and healthy non-psychiatric controls.

Emotion Disturbance in Bipolar Disorder: Considering a Neuroendocrine Approach

Recent psychosocial models suggest that individuals with BD experience persistent elevations in positive emotional states across different contexts (Gruber, 2011a), consistent with psychosocial models implicating heightened reward seeking and goal striving in the etiology of BD (e.g., Alloy et al., 2012, Johnson, 2005; Gruber, Johnson, Oveis, & Keltner, 2008). Specific positive emotions ratings have also been found to predict manic and depressive symptom severity in BD (Gruber et al., 2009). Abnormalities in negative emotionality in BD are mixed. On the one hand, people diagnosed with or at risk for BD do not differ from healthy controls in self-reported emotional responses to negative feedback (e.g., Ruggero & Johnson, 2006) or interpersonal criticism (Cuellar, Johnson, & Ruggero, 2009). Other studies suggest that BD is associated with increased negative emotionality in everyday life (e.g., Gruber, Kogan, Mennin,

& Murray, 2013; Myin-Germeys et al., 2003), a finding similar to people with depression (Byslma, Taylor-Clift & Rottenberg, 2011; Myin-Germeys et al., 2003).

One window into better understanding emotion disturbance in BD is through the examination of neuroendocrine function. Neuroendocrine function plays a critical role in regulating affect and behavior (e.g., Josephs, Sellers, Newman, & Mehta, 2006; Schultheiss, Wirth, Torges, Pang, Villacorta, & Welsh, 2005). Two neuroendocrine axes related to affective processes are the hypothalamic-pituitary-adrenal axis, which produces the hormone cortisol and the hypothalamic-pituitary-gonadal axis, which produces the steroid sex hormone testosterone (Welker, Gruber, & Mehta, 2015). Given the involvement of testosterone and cortisol in affective systems (e.g., Kalin et al., 1998), activity in these neuroendocrine systems may help characterize BD (e.g., Daban et al., 2005).

Research investigating cortisol function in BD has been inconsistent. For instance, some studies do not report differences between remitted BD and healthy controls (e.g., Deshauer et al., 2006; Spijker et al., 2014) whereas other studies report decreased basal cortisol in manic BD versus remitted BD samples (Cassidy, Ritchie, & Carroll, 1998) and others report increased cortisol levels in response to negative life events in BD (e.g., Staufenbiel et al., 2014). Furthermore, men with BD have been found to exhibit blunted cortisol responses in the context of mental challenge tasks compared to healthy controls that, in turn, are associated with increased frequency of psychotic symptoms within the BD group (Steen et al., 2011). One recent meta-analysis among a heterogeneous BD group (i.e. intra- and inter-episode) found increased morning levels of cortisol in BD compared to healthy controls (Girshkin et al., 2014); this meta-analysis did not differentiate between mood state, leaving it unclear whether cortisol function in BD varied depending on current mood symptom levels or whether the observed effects

generalized across distinct facets of cortisol function (e.g., morning rises in cortisol, average trait levels of cortisol, acute fluctuations, declines throughout the day, and overall cortisol production). No differences in diurnal cortisol slopes were found in the offspring of parents with BD, with increased cortisol levels only found in the afternoon (Ellenbogen, Santo, Linnen, Walker, & Hodgins, 2010). These inconsistent results suggest additional research is warranted in BD and differentiates distinct facets of cortisol functioning across the day.

With respect to testosterone, both stable profiles and dynamic changes in testosterone are thought to be associated with positive affectivity and increased appetitive behavior (e.g., Mazur & Booth, 1998; Schipper, 2012; Stanton et al., 2011). Animal research suggests that testosterone causally modulates dopaminergic activity in the nucleus accumbens and ventral tegmental area (see Welker et al., 2015 for a review)—two putative regions linked to processing reward. Additionally, exogenous testosterone administration in humans can increase sensitivity to rewards (van Honk et al., 2004) and ventral striatal activity in response to monetary reward cues (Hermans et al., 2010; Op de Macks et al., 2011). Though in its infancy, emerging work in BD reveals that elevated testosterone levels are also associated with significant increases in mania symptoms and severity in BD (e.g., Sher et al., 2012). Oral administration of testosterone has been causally linked to the onset of manic symptoms (Pope, Kouri, & Hudson, 2000). Taken together, this work suggests that both cortisol and testosterone profiles may provide important clues into understanding emotion disturbance in BD.

The Present Investigation

In the present investigation, we focus on cortisol and testosterone among inter-episode (i.e., not currently in a manic, depressed, or mixed mood in the past month) adults with BD compared to both interepisode adults with major depressive disorder (MDD) and non-psychiatric

controls. Although we primarily focus on BD, we include two comparison groups from which to compare our findings including a clinical comparison group of remitted individuals with a history of major depressive disorder (MDD) and a group of healthy controls with normative affective profiles. This work is an important contribution by examining neuroendocrine function in interepisode BD (and MDD) samples to examine neuroendocrine profiles not driven by current mood symptom severity. Drawing from the literature on hormonal mechanisms of reward dysregulation and heightened reward sensitivity and positive affectivity in BD (e.g., Gruber, 2011b; Johnson, 2005; Mehta & Prasad, 2015; Welker et al., 2015), we predicted that the BD group would be characterized by blunted cortisol and elevated testosterone profiles as compared to both MDD and healthy control groups and that this pattern would remain robust across multiple hormonal parameters sampled across several days naturalistically. We examined distinct hormone parameters to capture comprehensive features of the hormonal response in a naturalistic environment for bipolar mood disordered individuals.

Methods

Participants

Participants between the ages of 18-45 were recruited as part of a broader study on emotion and mood (Yale University IRB HIC #1309012679; University of Colorado Boulder IRB #14-0390) using posted flyers, online advertisements, and referrals from outpatient mental health centers and hospitals. Interested participants completed a brief phone screen, and those deemed potentially eligible were invited to the laboratory to determine final study eligibility. 89 were eligible for the broader protocol, and 88 completed the study procedures. Of these, 5 participants (1 BD, 2 MDD, and 2 Controls) were excluded for insufficient hormone and daily experience-sampling methodological (ESM) data, leaving the final participant size of 83.

Of the final 83 study participants, 28 were diagnosed with BD type I in interepisode status (i.e., not currently manic, depressed, or mixed mood in the past month, *M*=24.55 months, *SD*=25.44), 28 participants with MDD currently interepisode (i.e., not currently depressed *M*=34.15 months, *SD*=34.06), and 27 healthy non-psychiatric controls (i.e., CTL) who did not meet current or past criteria for any DSM-IV-TR Axis I disorder. See **Table 1** for participant characteristics and supplementary materials for details on clinical measures, cognitive functioning, and exclusion criteria.

Saliva Collection

To assess diurnal hormone fluctuations, participants provided six saliva samples each day (i.e., at wake, wake + 30 minutes, 11am, 3pm, 8pm, and at bedtime) for three consecutive weekdays (i.e., between Monday-Friday) for a total of 18 saliva samples.[1] Prior to sample collection, participants were instructed to abstain from eating, drinking, smoking, chewing tobacco, brushing teeth, and chewing gum for at least 1 hour and were asked to not consume alcohol or take naps during this period (Schultheiss & Stanton, 2009). Participants passively drooled approximately 2 mL of saliva into polypropylene tubes. Participants were thoroughly trained to store saliva samples in home freezers or in lunch boxes with ice packs during transit and when a freezer was not immediately available (consistent with recommendations for naturalistic collection; Granger et al., 2004; Schultheiss & Stanton, 2009). Several additional actions were taken to ensure compliance in at-home saliva collection procedures following previously published guidelines (Prasad et al., 2021). Specifically, participants were provided with detailed written instructions and verbally reviewed them with the experimenter and

[1] Participants also provided a baseline saliva sample during the initial laboratory visit though the time of the collection for this sample varied considerably (i.e., 10:46-15:26 h) and so it was not included in the main analyses given the diurnal saliva samples are a more reliable index of hormonal profiles (Al-Dujaili & Sharp, 2012).

confirmed they understood, used a Dymo time stamper (Dymo Corporation, Stamford, Connecticut) to ensure compliance for saliva sample times, and received reminder text messages 5-10 minutes prior to 11am, 3pm, and 8pm saliva sample times as a reminder. Saliva samples were frozen at -80° C at Yale University[2] until they were shipped overnight on dry ice for subsequent assays where they were frozen at -35°C until assayed using commercially available enzyme-linked immunoassay kits from IBL International (cortisol) and Salimetrics (testosterone). All assays were performed in duplicate over an average of 45 days (*SD*=30) following storage at the University of Oregon. The average intra-assay coefficients of variation (CVs) were acceptable for testosterone (*M*=7.30%) and cortisol (*M*=6.16%), as were the inter-assay CVs (36.96% and 21.87%, respectively).[3]

Daily ESM Event Ratings

Participants provided five daily ESM event ratings each day for three consecutive days, totaling in 15 ratings (i.e., wake up + 30 minutes, 11:00am, 3:00pm, 8:00pm, and 10:00pm each day). Each event rating sampled four broad domains, including self-reported positive affect (PA), negative affect (NA), arousal, social context, and activity type (described below). Using these measures, we followed previously validated approaches for computing measures of intra-individual variability calculated offline by taking the standard deviation of all daily ESM emotion self-reports. This approach has been used previously in examining emotional variability and has been found to suggest that excessive variability in positive emotion is linked to poorer psychological health (Gruber et al., 2013).

[2] Samples were temporarily stored in a -40°C freezer at Yale University until being relocated (which took place approximately once a week) to a -80°C freezer at a different lab site. After being shipped to the University of Oregon for long-term storage they were stored at -80°C until being assayed.

[3] These CVs are consistent with other assays performed in our lab using the same manufacturers (see Welker et al., 2016), as well as those reported in the extant literature (e.g., Granger et al., 2004).

Specifically, participants completed two self-report items for current level of PA or NA (i.e., "how are you feeling right now?") on a 5-point Likert scale from 1 (*not positive*/*negative*) to 5 (*very positive*/*negative*). Participants reported their level of arousal in response to a 5-point Likert scale item "How active or physically aroused are you feeling right now?" from 1 (*not at all*) to 5 (*extremely*). Participants indicated details of other people in their vicinity through a multiple-choice question asking them "Who are you with?" with seven non mutually exclusive response options: "alone," "family," "friends," "peers," "co-workers," "strangers," or "other." Finally, participants also reported their current activities in response to an open-ended question asking, "What are you doing?". Example responses included "working on school project," "sleeping, showering," "reading," "Watching TV," and "stocking products/talking to customers." These activities were classified into ten different context categories by a research assistant (recreation, working, errands, eating, socializing, resting, exercising, shopping, bathing, and other; Gruber et al., 2013). To confirm reliability of the context activities, a separate researcher independently coded and categorized them, demonstrating good reliability (average κ=0.75). Descriptive information is presented in supplementary materials.

Daily ESM Beginning and End of Day Ratings

Participants provided two daily ESM reports at the beginning and end of each day. The first asked participants how many hours of sleep they had the previous night and what time they awoke. The second, at bedtime, asked if participants exercised or consumed caffeine or alcohol (yes or no) at any point in the day. If participants exercised, they listed the time, activity, duration, and described the activity on a 7-point Likert scale from 1 (not strenuous) to 7 (extremely strenuous). If participants consumed caffeine or alcohol, they indicated the time of

consumption and amount (e.g., "2 small cups of coffee"; "1 glass of wine"). Descriptive statistics of this ESM data are presented in the Supplementary Materials[4].

Procedure

The current study procedure had three stages. First, participants arrived at the laboratory, provided informed consent, and completed diagnostic and cognitive assessments (see supplementary materials). They next completed an unrelated set of laboratory tasks[5] and participated in a ~20-minute comprehensive training and acclimatization session with a senior research assistant (Day 0) including review of the Dymo (Dymo Corporation, Stamford, CT) time stamper to synchronize saliva sampling with time of day, daily ESM items, and a full practice trial including proper saliva sample provision technique (i.e., passive drooling). Participants were encouraged to contact the experimenter with questions during practice trials outside the lab on Day 0. Participants were provided with a daily ESM packet of 15 daily reports,[6] 18 vials for salivary data collection, ice packs and a lunch box, and an iPod Touch.[7] Second, participants completed three consecutive weekdays of the experience sampling method (ESM) study protocol (Days 1-3). Third, participants came back to the lab to return equipment, be reassessed for mood status during the ESM study period, and review diaries with the experimenter to confirm authentic data reporting and identifying any flaws/debris in saliva samples (e.g., discoloration, food particles). After completing an experimental task and questionnaires unrelated to the present study, participants were compensated and debriefed.

[4]We note that testosterone and cortisol were not associated with caffeine use, alcohol use, or exercise on any day of the study (*p*s from .062 to .959). As such, we did not include them as additional covariates in our main analyses.
[5] This included a baseline measure of physiological signals (unpublished), an emotional EEG task (unpublished), and a task studying cooperative behavior (e.g., Ong et al., 2017).
[6] Only the wake-up saliva samples lack corresponding ESM entries, as participants were not yet engaged in daily activity.
[7] Although unrelated to the current study, the iPod touch was used to record audio files of ambient interactions and events for an interval of 30 seconds every 12 minutes.

Results

Data Analytic Approach

Multilevel models were used to examine group differences in cortisol and testosterone concentrations throughout the day, along with overall cortisol output across the day and hormonal awakening response. Three-level multilevel models were used to examine testosterone and cortisol concentrations, with sampling time (level 1) nested in each day (level 2), which was in turn nested in the person (level 3). Analyses predicting overall output and awakening responses used two-level models, with days (level 1) nested within person (level 2). Multilevel models were conducted using the *lme4* package in R (Bates & Maecheler, 2010). To account for all possible comparisons of diurnal hormone profiles between groups, we conducted two models for each analysis, using the BD (due to our research question emphasis) and CTL groups as the reference group in each. This approach allowed us to examine interactions between group and time (mean centered), as well as interactions between group and a quadratic time variable (referred to as Time 2), which were used to reveal group differences in hormone changes across time.[8] The two-level models were analyzed with group coded as a factor, which allowed for least square means comparisons (similar to an ANOVA) between daily hormone output and awakening responses of each group. Simple slopes tests were conducted using the *reghelper* package in R (Hughes, 2017). Cortisol and testosterone outputs throughout the day were calculated using the area under the curve to ground (AUC-G) calculations provided by Pruessner and colleagues (Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003). Cortisol awakening responses (CAR) were calculated by regressing the cortisol concentrations from 30 minutes after waking on concentrations at wake and saving the unstandardized residuals for each

[8] The results converged and did not vary in statistical significance depending on whether time was a fixed or random variable.

day, similar to other studies assessing hormonal responses (e.g., Carré, Campbell, Lozoya, Goetz, & Welker, 2013)[9]

Demographic and Clinical Characteristics

As seen in **Table 1**, BD, MDD, and CTL participants did not differ significantly in age, gender, ethnicity, or mania symptoms in the last week. BD and MDD groups scored higher than CTLs on depressive symptoms in the last week, but not each other[10]. The BD group scored lower on global functioning (GAF) than MDD and CTL groups, and the MDD group scored lower than CTLs. BD and MDD groups did not differ in illness duration, though the MDD group had an earlier average age of onset compared to the BD group.

Preliminary Analyses

Preliminary analyses investigated missing data, distributions of hormonal data, and self-reported measures of PA (ICC = .42), NA (ICC = .36), and Arousal (ICC = .37). First, missing data for testosterone and cortisol was minimal (M=2.38%, SD=5.30%), and there were no differences in missing data between the BD (M=1.79%, SD=3.40%), MDD (M=1.98%, SD=3.28%), and CTL groups (M=3.40%, SD=7.98%) [$F(2,80)=.75$, $p=.478$, $\eta_p^2=.02$]. For the analyzed daily ESM data (PA, NA, and Arousal), there were no differences in missing data between the BD (M=1.03%, SD=2.38%), MDD (M=2.70%, SD=5.49%), and CTL groups (M=3.95%, SD=10.66%) [$F(2,80)=1.21$, $p=.305$, $\eta_p^2=.03$] with minimal missing data across all participants (M=2.54%, SD=7.02%).

[9] For the AUC-G and CAR, missing saliva samples and concentrations prevented calculation of these metrics on some days, although this was relatively minimal for testosterone AUC-G (M_{days}=.42, SD=.78), cortisol AUC-G (M_{days}=.37, SD=.79), and the CAR (M_{days}=.05, SD=.22). Moreover, most participants had CAR estimates available from all days (95.18%), but no participants had missing data for more than one day. Three participants could not have cortisol AUC-G estimated from any days, seven had two days with missing AUC-G, and eight had one day of missing cortisol AUC-G. Testosterone AUC-G estimates were missing on all days for three participants, missing on two days for six participants, and missing one day for fourteen participants.

[10] All between-group models for our main outcomes were run with YMRS or IDS-C scores as covariates and no theoretically meaningful pattern of changes resulted.

Second, we visually inspected histograms of testosterone and cortisol for skewness. Both were substantially positively skewed and were thus transformed using a natural log transformation (e.g., Denson et al., 2013), and a constant of + 1 to cortisol concentrations prior to the log transformation due to their low absolute values. At each time of day, we examined the ICCs of all testosterone and cortisol concentrations at the same time points across the three days (e.g., Ross, Murphy, Adam, Chen, & Miller, 2014). Consistency among testosterone concentrations (average ICC=0.78, *SD*=0.03) and cortisol concentrations (average ICC=0.44, *SD*=0.09) were generally good.

Third, consistent with previous experience-sampling work investigating positive and negative emotionality in daily life among adults with BD and MDD (Gruber et al., 2013), we examined potential group differences in our single item assessments of PA, NA, and Arousal. These group differences consisted of averages across all time-points across the study and variability across these time-points (SD).

For mean PA and NA, we note that the Group main effect for mean NA was significant [$F(2, 80)=5.79$, $p=.004$, $\eta_p^2=.06$], with follow-up pairwise comparisons indicating that MDD (*M*=1.97, *SD*=0.59) and BD (*M*=1.80, *SD*=0.60) groups reported increased NA relative to CTLs (*M*=1.47, *SD*=.47) (*p*s=.001 and .030, respectively). However, BD and MDD groups did not differ in NA (*p*=.256). For mean PA, there was a similar trend toward a significant overall Group main effect [$F(2,83)=2.49$, $p=.089$, $\eta_p^2=.06$], whereby the MDD group reported lower PA (*M*=3.26, *SD*=.58) compared to CTLs (*M*=3.65, *SD*=.68; *p*=.029), but no other group differences were significant (*p*s≥.196).

For PA and NA variability, we note that PA variability (i.e., intra-individual standard deviation over the three-day period), there was an overall Group main effect [$F(2,83)=5.05$,

p=.009, η_p^2=.11] characterized by elevated PA variability in the BD group (M=.85, SD=.32) compared to CTLs (M=.62, SD=.24; p=.002). The MDD group (M=.72, SD=.24) did not differ in PA variability compared to the BD (p=.087) or the CTL (p=.150) groups. Groups also differed in NA variability [F(2,83)=3.24, p=.044, η_p^2=.07], with CTLs (M=.53, SD=.36) reporting significantly less NA variability than BD (M=.75, SD=.40, p=.029) and MDD (M=.75, SD=.31; p=.031) groups. BD and MDD groups did not differ in NA variability (p=.978). We noted no group differences in mean Arousal (F(2,83)=.190, p=.827, η_p^2=.00) or Arousal variability [F(2,83)=0.38, p=.685, η_p^2=.01].

Group Differences in Diurnal Cortisol and Testosterone Profiles

We first examined differences in diurnal profiles of testosterone and cortisol between groups (see **Table 2)**. To illustrate differences in diurnal profiles of hormones within groups, we conducted quadratic growth models within each group (**Figure 1**). The full results of these within group models are presented in supplemental analyses, along with quadratic and linear plots of the hormonal profiles across groups. For main effects of clinical groups, we compared hormone concentrations of least squares means. For cortisol, our analyses indicated a significant contrast between BD and MDD groups, reflecting elevated mean cortisol in MDD (M=.50, SE=.02) compared to BD (M=.42, SE=.02; B=-.11, t(80)=-3.21, p=.002). There was also a significant linear Time x MDD vs. BD interaction (B=-.03, t(1206)=-4.38, p<.001), and a linear Time x CTL vs. BD interaction (B=.03, t(1206)=4.26, p<.001), indicating the linear cortisol slope of the BD group significantly differed from MDD and CTL groups. Simple slopes analysis revealed the BD group experienced a less pronounced decline in cortisol (B=-.09, t(1206)=-17.18, p<.001) compared to CTL (B=-.12, t(1206)=-22.75, p<.001) and MDD groups (B=-.12, t(1206)=-23.36, p <.001). However, CTL vs. BD between-group comparison did not reveal

differences in overall cortisol concentrations (*B*=.05, *t*(80)=1.54, *p*=.127). Results also revealed a significant quadratic Time x MDD vs. BD interaction (*B*=-.01, *t*(1206)=-1.98, *p*=.048). Simple slopes indicated the presence of a quadratic effect that followed a negative decelerating pattern in the BD (*B*=.01, *t*(1206)=2.36 *p*=.018), but not MDD group (*B*=.00, *t*(1206)=-.44, *p*=.659). The quadratic Time x BD vs. CTL interaction was nonsignificant (*p*=.708). The CTL vs. MDD group comparison was not significant (*p*=.108), nor did it interact with linear or quadratic Time (*p*s≥.115).

Our analyses with testosterone concentrations indicated elevated testosterone in the MDD (*M*=4.59, *SE*=.07) relative to CTL group (*M*=4.42, *SE*=.07; *B*=.22, *t*(79)=2.02, *p*=.046). Although Group did not interact with linear time, there was a significant quadratic Time x MDD vs. CTL interaction (*B*=.02, *t*(1200)=2.46, *p*=.014), indicating a difference in the quadratic testosterone declines (the diurnal slope curvature). Specifically, simple slopes analysis revealed there was a pronounced decelerating quadratic decline in the CTL (*B*=.02, *t*(1200)= 4.31, *p* <.001) but not MDD group (*B*=.00, *t*(1200)=.88, *p*=.377).

Group Differences in Cortisol Awakening Responses

Next, we examined group differences in the cortisol awakening response. Two-level multilevel models revealed an overall main effect for group (*F*(1,80)=6.34, *p*=.003; **Figure 2**). Contrast comparisons revealed significantly lower cortisol awakening responses for the BD (*M*=-.18, *SE*=.08) compared to MDD group (*M*=.16, *SE*=.07; *t*(79)=3.18, *p*=.006). CTLs (*M*=.03, *SE*=.08) did not significantly differ from BD or MDD groups (*t*(79)=1.94, *p*=.135 and *t*(79)=-1.24, *p*=.434, respectively).

Group Differences in Overall Cortisol and Testosterone Production

Hormonal output throughout the day (AUC-G) between groups were examined. For cortisol, there was an overall effect of group ($F(2,77)=4.49$, $p=.014$, $\eta_p^{2=}.99$ (**Figure 2**). Contrasts revealed significantly more cortisol output in the MDD ($M=35430.85$, $SE=2109.08$) compared to BD group ($M=26585.66$, $SE=2072.563$, $t(77)=2.99$, $p=.010$).[11] CTLs ($M=30475.64$, $SE=2191.35$) did not significantly differ from MDD and BD groups ($t(77)=-1.63$, $p=.240$ and $t(77)=1.29$, $p=.405$, respectively). There were no significant group differences in testosterone overall output ($F(2,76)=2.27$, $p=.110$, see **Figure 2**).[12]

Discussion

The present investigation examined cortisol and testosterone profiles across three consecutive days in everyday life among adults with BD, MDD, and no psychiatric history. Partially consistent with our prediction, adults with BD exhibited flatter cortisol slopes compared to MDD and CTL groups. However, BD subjects only differed from our MDD group across cortisol concentrations throughout the day, overall cortisol hormone output, and cortisol awakening response, but did not differ from the CTL group. Overall, these findings are consistent with previous work documenting dampened cortisol diurnal slopes in men and women with BD (Havermans, Nicolson, Berkhof, & deVries, 2011). However, the design of the current study afforded the ability to assess diurnal rhythms in cortisol, shedding light on how cortisol function over extended periods is characterized in BD. Yet, these findings were not consistent with a recent meta-analysis suggesting that individuals with BD have elevated, rather than

[11] After controlling for SES and hormonal medication usage, the difference in testosterone AUC between the BD and MDD groups became nonsignificant ($p=.065$).

[12] Because our affect and arousal experience-sampling reports invite examining the associations between hormones and daily affectivity, we used multilevel modeling and regression analyses to explore these associations, which are presented in the **Supplemental Materials**. Of note, the supplemental analyses found that moment-to-moment cortisol concentrations were negatively associated with positive affect across all participants. Additionally, several of our participants (23.6%) were using medications known to affect hormones (e.g., thyroid medication, birth control). Therefore, we conducted follow-up analyses in all models with hormones controlling for SES and hormone-relevant medication usage. Unless otherwise specified above, doing so did not change the significance of any results.

flattened cortisol awakening responses (Girshkin et al., 2014). As negative life events have been linked to mood episode onset and recovery in BD (Urošević, Abramson, Harmon-Jones, & Alloy, 2008), future work should test whether cortisol may covary such events to examine whether dampened cortisol levels sustain during the presence of discrete negative stimuli.

Contrary to our hypotheses, we did not find testosterone concentrations differences between the BD and other groups. This is somewhat surprising given heightened testosterone levels have been associated with increased mania symptoms (Sher et al., 2012) and studies causally linking oral testosterone administration to the onset of mild to moderate hypomanic symptoms (Pope, Kouri, & Hudson, 2000). Additionally, lack of elevated testosterone is inconsistent with previous theorizing linking elevated testosterone with appetitive and reward seeking behavior (Welker, Gruber, & Mehta, 2015) that is common to models of BD (Alloy & Abramson, 2010; Johnson, 2005). It is possible that elevated testosterone may be closely tied with manic symptom severity and hence not detectable in the current study which ensured interepisode mood status. Unexpectedly, the MDD group showed elevated testosterone compared to the healthy controls. This finding is somewhat unexpected given the rising prevalence of testosterone replacement therapy as a solution to MDD (Zarrouf, Artz, Griffith, et al., 2009), suggesting that further research may be needed on depression and diurnal testosterone function.

Although not part of our central aims, BD participants did not report increased levels of positive emotionality compared to controls. This is consistent with recent experience-sampling finds of increased positive affectivity in remitted BD compared to remitted major depressive disorder, but no differences from controls (Gruber et al., 2013). However, the present study utilized a statistically weak single-item assessment of positive affect whereas previous research has used multi-item composites of positive affect and/or looked at more specific or discrete types

of positive affectivity (Gruber & Johnson, 2009; Gruber et al., 2008). Thus, it remains a viable possibility that people with BD may have a propensity toward heightened positive emotionality compared with healthy adults, all else being equal (Gruber, 2011a). Future research could usefully explore the hypothesis that laboratory studies control for stressful stimuli, thereby masking a tendency toward negative affect in BD and exposing elevated propensity for positive affect relative to the general population.

Limitations and Future Directions

Our findings should be interpreted within the confines of several limitations. First, the present study did not explicitly assess the role of contextual factors and as such future studies are thus warranted to investigate these critical interactions between context and neuroendocrine-related functioning including contexts known to alter hormonal reactivity of both testosterone and cortisol hormone profiles, such as stressors, competition, and aggressive provocation (Carré, et al., 2013; Carré, Iselin, Welker, Hariri, & Dodge, 2014).

Second, although the present sample sizes are commendable given the severe nature of the psychiatric groups recruited and intensive within-subjects experience sampling study design, there is low statistical power to assess smaller effect sizes that may have captured more nuanced associations between self-reported affect and hormone profiles. Future studies replicating these results in larger samples are warranted. Moreover, the fact that the present study included a third clinical comparison group to examine emotion difficulties within the mood disorder family represents a strength and an important first step in this trans-diagnostic mission to identify shared and dimensional features across individuals (Insel et al., 2010).

Third, we note that the self-report ESM data was not time-stamped in parallel manner to the saliva data. As such, it is possible that the accuracy of the self-report emotion data could be

raised as compared to more electronic methods of collecting this information. We note that future studies would be well-advised to use innovative developments in ESM data acquisition.

Fourth, given the challenges of accessing an unmedicated clinical sample, we were unable to investigate the influence of medication effects on results. Indeed, future studies with larger sample sizes, assessment of blood serum levels for psychotropic medications, and random assignment of individuals on different medication classes are warranted. Given that our sample was predominantly female (66.27%), using more accurate hormone assessment methods than enzyme-linked immunoassays such as mass spectrometry may have yielded more valid testosterone measurements of female's testosterone (Welker et al., 2016), and consequently, more conclusive differences in neuroendocrine profiles. Despite these limitations, the present research underscored the importance of adopting a neuroendocrinology approach and framework to provide insights into mood disturbances in everyday life.

Author Disclosure

Role of funding source: Research reported in this publication was supported by a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (JG), a CTSA Grant Number UL1 RR024139 from the National Center for Research Resources (JG), National Institute of Mental Health under award number T32MH103213 (JRP), Indiana CTSI Predoctoral Award Grant UL1 TR001108, (JRP), F31 MH122122 (JPR), and the National Science Foundation under grant numbers 1451848 and 1303743 (PHM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH or any other entity.

Conflict of interest: The authors declare no financial disclosures or conflicts of interest with respect to the authorship or the publication of this article.

Acknowledgements: The authors would like to acknowledge Nicholas A. Snyder for his contribution of performing hormone assays, Alta du Pont and Natasha Hansen for assisting with clinical inter-reliability ratings, and Benjamin Billingsley, Minhae Kim, Corinne Sejourne, Celinda Smith, and Jordan Thomas for data collection/cleaning.

Contributors: J.G. and P.H.M. developed the study concept and design. Testing and data collection were performed by J.R.P. and overseen by J.G. Data analysis was performed by K.M.W., B.L., S.P., and C.M.B. Manuscript was drafted by J.G., K.M.W., J. R.P., and P.H.M. All authors provided manuscript revisions and approved the final product.

References

Al-Dujaili, E. A. S., & Sharp, M. A. (2012). Female salivary testosterone: Measurement, challenges and applications. *Steroids-From Physiology to Clinical Medicine*, 129-167.

Alloy, L. B., Bender, R. E., Whitehouse, W. G., Wagner, C. A., Liu, R. T., Grant, D. A., Jager-Hyman, S., Molz, A., Choi, J. Y., Harmon-Jones, E., & Abramson, L. Y. (2012). High behavioral approach system (BAS) sensitivity, reward responsiveness, and goal-striving predict first onset of bipolar spectrum disorders: A prospective behavioral high-risk design. *Journal of Abnormal Psychology, 121,* 339-351. doi: 10.1037/a0025877

Bates D., Maecheler M. (2010). The lme4 Package. Available online at: http://r-forge.rproject.org/projects/lme4/

Carré, J. M., Iselin, A. M. R., Welker, K. M., Hariri, A. R., & Dodge, K. A. (2014). Testosterone reactivity to provocation mediates the effect of early intervention on aggressive behavior. *Psychological science, 25*(5), 1140-1146. doi: 10.1177/0956797614525642

Carré, J. M., Campbell, J. A., Lozoya, E., Goetz, S. M. M., & Welker, K. M. (2013). Changes in testosterone mediate the effect of winning on aggressive behavior. *Psychoneuroendocrinology, 38*(10), 2034-2041. doi: 016/j.psyneuen.2013.03.008

Cassidy, F., Ritchie, J. C., & Carroll, B. J. (1998). Plasma dexamethasone concentration and cortisol response during manic episodes. *Biological Psychiatry*, *43*(10), 747-754. doi: 10.1016/S0006-3223(97)00274-6

Cuellar, A. K., Johnson, S. L., & Ruggero, C. J. (2009). Affective reactivity in response to criticism in remitted bipolar disorder: a laboratory analog of expressed emotion. *Journal of Clinical Psychology, 65*(9), 925-941. doi: 10.1002/jclp.20596

Daban, C., Vieta, E., Mackin, P., & Young, A. H. (2005). Hypothalamic-pituitary-adrenal axis and bipolar disorder. *Psychiatric Clinics of North America, 28*(2), 469-480. doi:

Denson, T. F., Mehta, P. H., & Tan, D. H. (2013). Endogenous testosterone and cortisol jointly influence reactive aggression in women. *Psychoneuroendocrinology, 38*(3), 416-424.

Deshauer, D., Duffy, A., Meaney, M., Sharma, S., & Grof, P. (2006). Salivary cortisol secretion in remitted bipolar patients and offspring of bipolar parents. *Bipolar Disorders*, *8*(4), 345-349. doi: 10.1111/j.1399-5618.2006.00338.x

Ellenbogen, M. A., Santo, J. B., Linnen, A. M., Walker, C. D., & Hodgins, S. (2010). High cortisol levels in the offspring of parents with bipolar disorder during two weeks of daily sampling. *Bipolar Disorders, 12*(1), 77-86.

Ferrari, A. J., Stockings, E., Khoo, J. P., Erskine, H. E., Degenhardt, L., Vos, T., & Whiteford, H. A. (2016). The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. *Bipolar Disorders*, *18*(5), 440-450.

Girshkin, L., Matheson, S. L., Shepherd, A. M., & Green, M. J. (2014). Morning cortisol levels in schizophrenia and bipolar disorder: a meta-analysis. *Psychoneuroendocrinology*, *49*, 187-206. doi: 10.1016/j.psyneuen.2014.07.013

Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T., & Schwartz, E. B. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology*, *29*(10), 1229-1240.

Gruber, J. (2011a). A review and synthesis of positive emotion and reward disturbance in bipolar disorder. *Clinical Psychology & Psychotherapy, 18*(5), 356-365. doi: 10.1002/cpp.776

Gruber, J. (2011b). Can feeling too good be bad? Positive emotion persistence (PEP) in bipolar disorder. *Current Directions in Psychological Science*, *20*(4), 217–221. doi: 10.1177/0963721411414632

Gruber, J., Culver, J. L., Johnson, S. L., Nam, J. Y., Keller, K. L., & Ketter, T. A. (2009). Do positive emotions predict symptomatic change in bipolar disorder?. *Bipolar Disorders*, *11*(3), 330-336.

Gruber, J., Gilbert, K. E., Youngstrom, E. A., Kogos Youngstrom, J., Feeny, N. C., & Findling, R. L. (2013). Reward dysregulation and mood symptoms in an adolescent outpatient sample. *Journal of Abnormal Child Psychology, 41*(7), 1053-1065. doi: 10.1007/s10802-013-9746-8.

Gruber, J., Harvey, A. G., & Purcell, A. L. (2011). What goes up can come down? A preliminary investigation of emotion reactivity and emotion recovery in bipolar disorder. *Journal of Affective Disorders* 133:457-466. doi: 10.1016/j.jad.2011.05.009

Gruber, J., Johnson, S. L., Oveis, C, & Keltner, D. (2008). Risk for mania and positive emotional responding: Too much of a good thing? *Emotion, 8(1)*, 23-33. doi: 10.1037/1528-3542.8.1.23

Gruber, J., Kogan, A., Mennin, D., & Murray, G. (2013). Real-world emotion? An experience-sampling approach to emotion disturbance and regulation in bipolar disorder. *Journal of Abnormal Psychology, 122(4)*, 971-983. doi: 10.1037/a0034425

Gruber, J., Oveis, C., Keltner, D., & Johnson, S. L. (2011). A discrete emotions approach to positive emotion disturbance in depression. *Cognition and Emotion, 25*(1), 40-52. doi: 10.1080/02699931003615984

Havermans, R., Nicolson, N. A., Berkhof, J., & deVries, M. W. (2011). Patterns of salivary cortisol secretion and responses to daily events in patients with remitted bipolar disorder. *Psychoneuroendocrinology*, *36(2),* 258-265. doi: 10.1016/j.psyneuen.2010.07.016

Hermans, E. J., Bos, P. A. , Ossewaarde, L., Ramsey, N. F., Fernández, G., van Honk, J. (2010). Effects of exogenous testosterone on the ventral striatal BOLD response during reward anticipation in healthy women. *NeuroImage, 52*, 277-283. doi: 10.1016/j.neuroimage.2010.04.019

Hughes, J. (2017). Reghelper: helper functions for regression analysis. R package, version 0.3.3. https://CRAN.R-project.org/package=reghelper

Insel, T. R., & Wang, P. S. (2010). Rethinking mental illness. *The Journal of American Medicine Association*, *303*(19), 1970-1971. doi: 10.1001/jama.2010.555

Johnson, S. L. (2005). Mania and dysregulation in goal pursuit: a review. *Clinical Psychology Review*, *25*(2), 241–262. doi: 10.1016/j.cpr.2004.11.002

Josephs, R. A., Sellers, J. G., Newman, M. L., & Mehta, P. H. (2006). The mismatch effect: when testosterone and status are at odds. *Journal of personality and social psychology, 90*(6), 999-1013. doi: 10.1037/0022-3514.90.6.999

Kalin, N. H., Shelton, S. E., Rickman, M., & Davidson, R. J. (1998). Individual differences in freezing and cortisol in infant and mother rhesus monkeys. *Behavioral Neuroscience*, *112*(1), 251-254. doi: 10.1037/0735-7044.112.1.251

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences, 21*(3), 353-397. doi: 10.1017/s0140525x98001228

Mehta, P.H., Prasad, S. (2015). The dual-hormone hypothesis: A brief review and future research agenda. *Current Opinion in Behavioral Sciences, 3,* 163-168. doi: 10.1016/j.cobeha.2015.04.008

Ong, D. C., Zaki, J., & Gruber, J. (2017). Increased cooperative behavior across remitted bipolar I disorder and major depression: Insights utilizing a behavioral economic trust game. *Journal of Abnormal Psychology*, *126*(1), 1.

Op de Macks, Z. A., Moor, B. G., Overgaauw, S., Güroğlu, B., Dahl, R. E., & Crone, E. A. (2011). Testosterone levels correspond with increased ventral striatum activation in response to monetary rewards in adolescents. *Developmental Cognitive Neuroscience*, *1*(4), 506-516. doi:10.1016/j.dcn.2011.06.003

Pope, H. G., Kouri, E. M., & Hudson, J. I. (2000). Effects of supraphysiologic doses of testosterone on mood and aggression in normal men: A randomized controlled trial. *Archives of General Psychiatry, 57(2),* 133-140. doi: 10.1001/archpsyc.57.2.133

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, *28*(7), 916-931. doi: 10.1016/S0306-4530(02)00108-7

Ross, K. M., Murphy, M. L., Adam, E. K., Chen, E., & Miller, G. E. (2014). How stable are diurnal cortisol activity indices in healthy individuals? Evidence from three multi-wave studies. *Psychoneuroendocrinology, 39*, 184-193. doi: 10.1016/j.psyneuen.2013.09.016

Ruggero, C. J., & Johnson, S. L. (2006). Reactivity to a laboratory stressor among individuals with bipolar I disorder in full or partial remission. *Journal of Abnormal Psychology, 115*(3), 539. doi: 10.1037/0021-843X.115.3.539

Schipper, B. (2012). *Sex hormones and choice under risk* (No. 12, 7). Working Papers, University of California, Department of Economics.

Schultheiss, O. C., & Stanton, S. J. (2009). Assessment of salivary hormones. *Methods in Social Neuroscience*, 17-44.

Schultheiss, O. C., Wirth, M. M., Torges, C. M., Pang, J. S., Villacorta, M. A., & Welsh, K. M. (2005). Effects of implicit power motivation on men's and women's implicit learning and testosterone changes after social victory or defeat. *Journal of personality and social psychology, 88*(1), 174-188. Doi: 10.1037/0022-3514.88.1.174

Sher, L., Grunebaum, M. F., Sullivan, G. M., Burke, A. K., Cooper, T. B., Mann, J. J., & Oquendo, M. A. (2012). Testosterone levels in suicide attempters with bipolar disorder. *Journal of Psychiatric Research, 46(10),* 1267-1271. doi: 10.1016/j.jpsychires.2012.06.016

Spijker, A. T., Manenschijn, L., Staufenbiel, S., Koenders, M. A., Hoencamp, E., Giltay, E., & Van Rossum, E. F. C. (2014). EPA-1041–Hair cortisol analysis in patients with bipolar disorder: towards a new endophenotype?. *European Psychiatry*, *29*, 1. doi: 10.1016/s0924-9338(14)78329-4

Stanton, S. J., O'Dhaniel, A., McLaurin, R. E., Kuhn, C. M., LaBar, K. S., Platt, M. L., & Huettel, S. A. (2011). Low-and high-testosterone individuals exhibit decreased aversion to economic risk. *Psychological science*, 22, 447-453. doi: 10.1177/0956797611401752

Staufenbiel, S. M., Koenders, M. A., Giltay, E. J., Elzinga, B. M., Manenschijn, L., Hoencamp, E., ... & Spijker, A. T. (2014). Recent negative life events increase hair cortisol concentrations in patients with bipolar disorder. *Stress*, 17(6), 451-459. doi: 10.3109/10253890.2014.968549

Steen, N. E., Lorentzen, S., Barrett, E. A., Lagerberg, T. V., Hope, S., Larsson, S, Berg, O. S., Agartz, I., Melle, I., Berg, J. P., Andreassen, O. A. (2011). Sex-specific cortisol levels in

bipolar disorder and schizophrenia during mental challenge--relationship to clinical characteristics and medication. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35(4),* 1100-1107. doi: 10.1016/j.pnpbp.2011.03.008

Urošević, S., Abramson, L. Y., Harmon-Jones, E., & Alloy, L. B. (2008). Dysregulation of the behavioral approach system (BAS) in bipolar spectrum disorders: Review of theory and evidence. *Clinical Psychology Review, 28*(7), 1188-1205.

van Honk, J., Schutter, D. J., Hermans, E. J., Putman, P., Tuiten, A., & Koppeschaar, H. (2004). Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. *Psychoneuroendocrinology*, *29*(7), 937-943. doi:10.1016/j.psyneuen.2003.08.007

Welker, K. M., Lassetter, B., Brandes, C., Prasad, S., Koop, D., & Mehta, P. H. (2016). A comparison of salivary testosterone measurement using immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology, 71,* 108-188. doi: 10.1016/j.psyneuen.2016.05.022

Welker, K. M., Gruber, J., & Mehta, P. H. (2015). A Positive Affective Neuroendocrinology (PANE) approach to reward and behavioral dysregulation. *Frontiers in Psychiatry, 6(93)*, 1-13. doi: 10.3389/fpsyt.2015.00093

Zarrouf, F. A., Artz, S., Griffith, J., Sirbu, C., & Kommor, M. (2009). Testosterone and depression: Systematic review and meta-analysis. *Journal of Psychiatric Practice, 15*(4), 289-305. doi: 10.1097/01.pra.0000358315.88931.fc

Table 1. Demographic and Clinical Characteristics by Diagnostic Group

	BD	MDD	CTL	Statistic
Demographic				
Age (Yrs)	29.93 (6.81)	27.82 (6.48)	28.37 (6.54)	F=0.76
Female (%)	64.29%	75.0%	59.26%	χ^2=1.60
Caucasian (%)	75.0%	75.0%	77.78%	χ^2=0.08
Education (Yrs)	14.23 (2.02)	15.91 (1.85)	15.96 (2.20)	F=6.55**[a,b]
Employed (%)	53.6%	71.4%	70.4%	χ^2=2.46
Partnered (%)	42.9%	60.7%	63.0%	χ^2=2.73
Number Children	0.46 (0.69)	0.11 (0.42)	0.33 (0.73)	F=2.31
Annual Income				F=3.51*[a,b]
<$10K	17.9%	7.1%	25.9%	
$10K-$25K	46.4%	32.1%	3.7%	
$26K-$50K	28.6%	25.0%	33.3%	
$51K-$75K	0.0%	7.1%	11.1%	
$76K-$100K	3.6%	21.4%	0.0%	
>$100K	3.6%	7.1%	25.9%	
Cognitive and Clinical				
MMSE	28.96 (1.67)	28.32 (1.76)	29.15 (1.37)	F=2.00
YMRS	1.32 (1.49)	1.00 (1.09)	0.56 (0.97)	F=2.78
IDS-C	3.43 (2.33)	4.61 (3.07)	2.41 (2.06)	F=5.21**[c]
GAF	78.11 (8.84)	79.11 (5.39)	86.07 (6.08)	F=10.69***[a,c]
Age at Onset (Yrs)	15.17 (3.36)	18.52 (6.94)	--	F=5.29[b]
Illness Duration (Yrs)	11.41 (5.77)	12.65 (5.72)	--	F=0.66
# Comorbid Disorders	0.14 (0.36)	0.39 (.63)	0.00 (0.00)	F=6.19**[b,c]
# Psychotropic Medications	1.57 (1.35)	0.75 (1.29)	0.00 (0.00)	F=14.45***[a,b,c]
# Depressive Episodes	15.10 (17.90)	15.39 (23.63)	--	F=0.00
# Manic and Hypomanic Episodes	13.63 (21.53)		--	--

Note: BD=Bipolar I disorder group (currently inter-episode); MDD=Major depressive disorder group (currently inter-episode); CTL=Healthy control group; Employed=Employed full-time or part-time; Partnered=Married or in a relationship; MMSE=Mini Mental State Examination; YMRS=Young Mania Rating Scale; IDS-C=Inventory to Diagnose Depression; GAF=Global Assessment of Functioning; Age at Onset=Age of first depressive or manic episode; # Comorbid Disorders=the number of current DSM-IV-TR Axis I comorbidities. # Medications=the number of psychotropic medications currently taken (including anticonvulsants, lithium, neuroleptics,

anxiolytics, stimulants, antidepressants, and sedative-hypnotics); Mean values are displayed with standard deviations in parentheses where applicable. [a]$p<0.05$ for BD and CTL, [b]$p<0.05$ for BD and MDD, [c]$p<0.05$ for MDD and CTL; *$p<.05$, **$p<.01$, ***$p<.001$. See supplementary materials for details on clinical and cognitive functioning measures and scoring.

Table 2. Multilevel Growth Curve Models Predicting Cortisol and Testosterone profiles between Groups.

Models Predicting Cortisol									
BD group as the reference group					CTL group as the reference group				
Effect	*B*	*df*	*t*	*p*	Effect	*B*	*df*	*t*	*p*
Intercept	.39	1206	16.49	<.001	Intercept	.45	1206	18.24	<.001
Time	-.09	1206	-17.18	<.001	Time	-.12	1206	-22.75	<.001
$Time^2$	.01	1206	2.37	.018	$Time^2$	.01	1206	1.77	.076
MDD vs. BD	.11	80	3.21	.002	BD vs. CTL	-.05	80	-1.54	.127
CTL vs. BD	.05	80	1.54	.127	MDD vs. CTL	.06	80	1.63	.108
Time X (MDD vs. BD)	-.03	1206	-4.38	<.001	Time X (BD vs. CTL)	.03	1206	4.26	< .001
Time X (CTL vs. BD)	-.03	1206	-4.26	<.001	Time X (MDD vs. CTL)	-.00	1206	-.06	.950
$Time^2$X (MDD vs. BD)	-.01	1206	-1.98	.048	$Time^2$X (BD vs. CTL)	.00	1206	.37	.708
$Time^2$X (CTL vs. BD)	-.00	1206	-.37	.708	$Time^2$X (MDD vs. CTL)	-.01	1206	-1.58	.115
Models Predicting Testosterone									
BD group as the reference group					CTL group as the reference group				
Effect	*B*	*df*	*t*	*p*	Effect	*B*	*df*	*t*	*p*
Intercept	4.19	1200	51.53	<.001	Intercept	4.14	1200	49.00	<.001
Time	-.12	1200	-17.86	<.001	Time	-.12	1200	-18.57	<.001
$Time^2$	.01	1200	2.41	.016	$Time^2$	.02	1200	4.31	<.001
Gender	.66	79	7.27	<.001	Gender	.66	79	7.27	<.001
MDD vs. BD	.17	79	1.61	.111	BD vs. CTL	.05	79	.44	.663
CTL vs. BD	-.05	79	-.44	.663	MDD vs. CTL	.22	79	2.02	.046
Time X (MDD vs. BD)	-.00	1200	-.25	.806	Time X (BD vs. CTL)	.01	1200	.79	.431
Time X (CTL vs. BD)	-.01	1200	-.79	.431	Time X (MDD vs. CTL)	.01	1200	.54	.588
$Time^2$X (MDD vs. BD)	-.01	1200	-1.07	.286	$Time^2$X (BD vs. CTL)	-.01	1200	-1.42	.155
$Time^2$X (CTL vs. BD)	.01	1200	1.42	.155	$Time^2$X (MDD vs. CTL)	-.02	1200	-2.46	.014

Note: BD=Bipolar I disorder group (currently inter-episode); MDD=Major depressive disorder group (currently inter-episode);

CTL=Healthy control group.

Figure 1. Quadratic Diurnal Changes in Cortisol and Testosterone by Diagnostic Group. Log-transformed hormone concentrations are plotted as quadratic slopes across time. MDD=Major depressive disorder group. BD=Bipolar disorder group, Control=Healthy control group. The bottom panel reports results for Testosterone controlling for gender and error bars represent standard errors around the points of prediction.

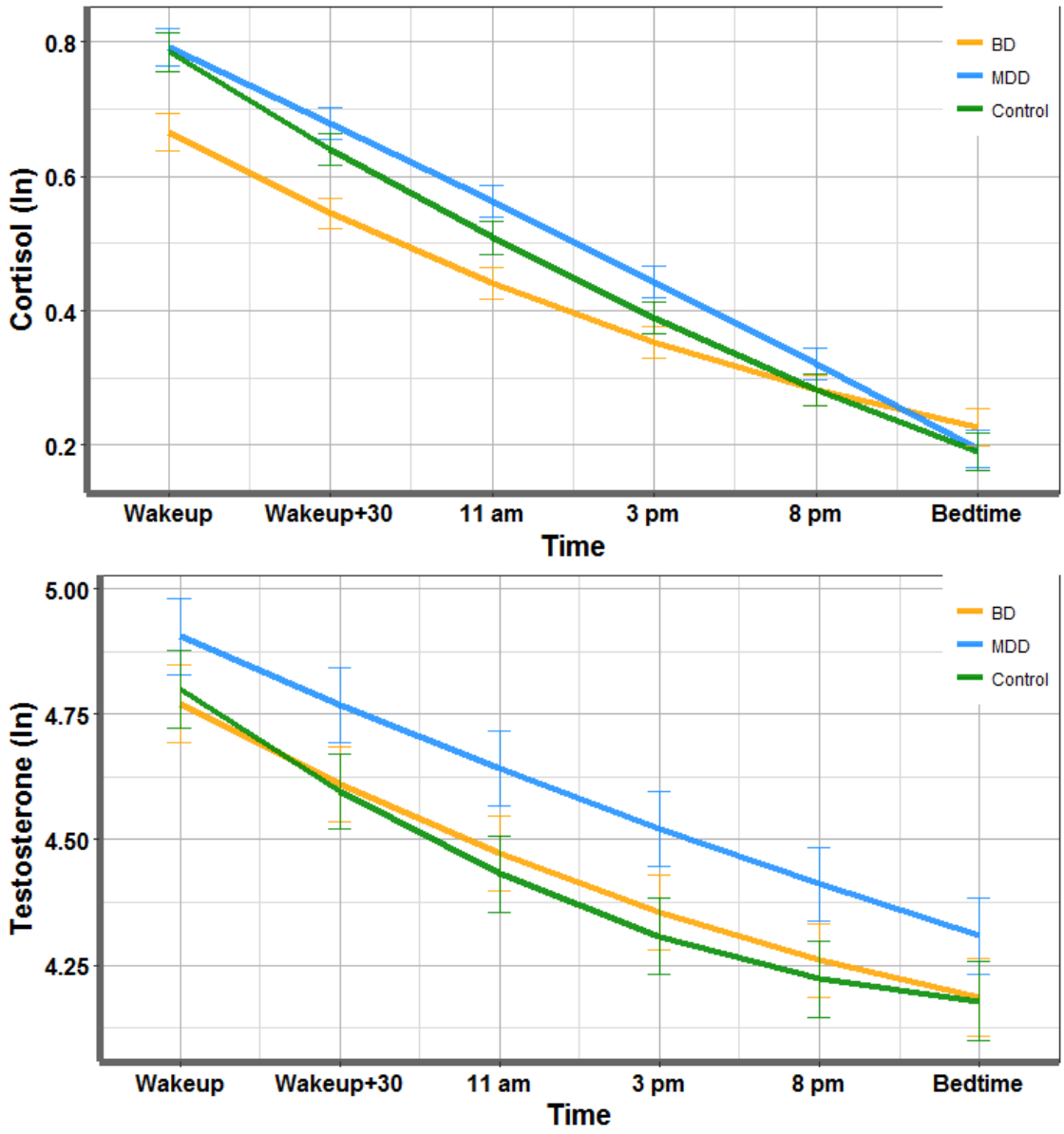
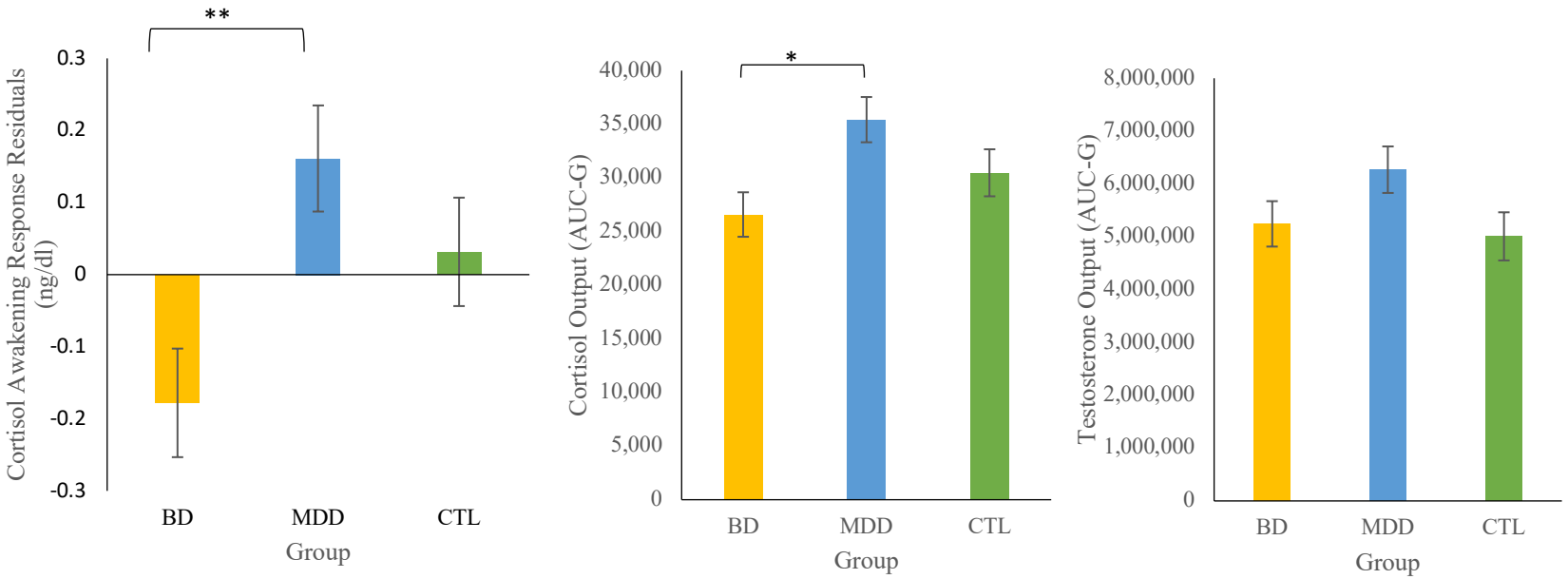


Figure 2. Cortisol Awakening Responses, Cortisol Output, and Testosterone Output by Diagnostic Group. Plotted bars represent the least squares means of daily output of hormones from multilevel models, and error bars represent standard errors of least squares means. *p<.05, **p<.01, BD=Bipolar disorder type I group (currently inter-episode), MDD=Major depressive disorder group (currently inter-episode), CTL=Healthy Control Group, AUC-G=Area under the curve with respect to ground. Analyses with testosterone control for gender.



Online Supplemental Materials for:

An Experience-Sampling Approach to Examining Cortisol and Testosterone Profiles in Bipolar and Depressive Mood Disorders

June Gruber[1+], Keith M. Welker[2+], John R. Purcell[3], Bethany Lassetter[4], Smrithi Prasad[5], Cassandra M. Brandes[6], & Pranjal H. Mehta[7]

[1] University of Colorado Boulder, Department of Psychology and Neuroscience

[2] University of Massachusetts Boston, Department of Psychology

[3] Indiana University, Department of Psychological & Brain Sciences

[4] University of Toronto, Department of Psychology

[5] University of Southern California, Marshall School of Business

[6] Northwestern University, Department of Psychology

[7] University College London, Department of Experimental Psychology

Measures of Clinical and Cognitive Functioning

Diagnostic Evaluation. Diagnoses were confirmed using the Structured Clinical Interview for DSM-IV (SCID-IV; First, Spitzer, Gibbon, & Williams, 2007). Interviews were administered by a trained post-baccalaureate research assistant (Author JP). Following accepted practices in inter-rater reliability (e.g., Gruber & Weinstock, 2018), a subset of participants from the broader study protocol were rated by a second reviewer to establish inter-rater reliability. The clinical characteristics reported reflect the interviewer's scores after consensus meetings in which raters discussed discrepancies, corrected errors that arose, and solidified training during informal consensus meetings. (See Ong, Zaki, & Gruber, 2017 for inter-rater reliability values and details from the broader study protocol). During the SCID-IV, illness duration and lifetime number of mood episodes were also collected (See **Table 1**).

Current Axis I comorbidities for the BD group included social phobia (n=1), specific phobia (n=1), obsessive-compulsive disorder (n=1), and generalized anxiety disorder (n=1) and for the MDD group included dysthymia (n=1), panic disorder (n=1), agoraphobia (n=1), social phobia (n=1), specific phobia (n=1), obsessive-compulsive disorder (n=1), and generalized anxiety disorder (n=1).

Mood Symptoms. Current symptoms of mania were measured using the Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978). Current symptoms of depression were measured using the Inventory of Depressive Symptomatology (IDS-C; Trivedi et al., 2004). The YMRS is an 11-item, clinician-rated measure of current manic symptoms with scores ranging from 0 to 60, whereas the IDS-C is a 30-item, clinician-rated measure of current depressive symptoms with scores ranging from 0 to 84. Current inter-episode mood status was

verified according to both current SCID-IV criteria for the past month and cutoff scores on the YMRS (≤ 7), and IDS-C (≤ 11) for the past week.

Global Functioning. The Global Assessment of Functioning (GAF; *DSM-IV Axis V*) Scale was used to assess general functioning in the past week. The GAF assesses overall psychological, social, and occupational functioning on a scale from 1 (lowest level of functioning) to 100 (highest level of functioning).

Cognitive Functioning. Cognitive functioning was assessed using the Mini Mental Status Examination, a brief objective measure of cognitive status and impairment (MMSE; Folstein, Folstein, & McHugh, 1975) with scores calculated as the total number of trials correct. All participants exceeded the eligibility cutoff score (≥ 24; Folstein et al., 1975).

Exclusion Criteria

Exclusion criteria for all groups included a self-reported history of severe head trauma, stroke, neurological disease, brain tumors or surgery, severe medical illness (e.g., autoimmune disorder, blindness, cardiovascular disease, HIV/AIDS), untreated endocrine-related disorders (e.g., Cushing's disease, Hypothyroidism, Thyroid Cancer, Hoshimoto's Disease), currently breastfeeding or pregnant, or current alcohol or substance abuse or dependence in the past six months. BD and MDD participants were not excluded based on current comorbid Axis I disorders in the past month (aside from current substance or alcohol use disorders) to ensure ecological validity given that mood disorders are commonly comorbid with other disorders (e.g., Kessler, Chiu, Demler, & Walters, 2005).

Exploratory Associations Between Daily Hormones and Affectivity

Our first set of supplemental analyses examined the effects of positive affectivity, negative affectivity, and arousal on cortisol and testosterone function. To explore these effects, we used a series of three-level multilevel models regressing cortisol and testosterone on positive affect, negative affect, and arousal, as well as their interactions with linear and quadratic time. These analyses tested whether self-reported affectivity in the daily ESM portion of the study was associated with overall hormones as well as differences in diurnal hormone function. These models were tested across all participants and within each group. When looking at cortisol function (See **Table S1**), the only substantial effect was positive emotionality interacted with time (linear) to predict cortisol (B = 0.02, t(936) = 3.18, p = .002). Simple slopes analyses indicated that the linear diurnal decline of cortisol was greater when positive emotionality was low (-1 SD; B = -0.19, t(936) = -28.55, p < .001) rather than high (+1 SD; B = -0.16, t(936) = -20.83, p < .001). This interactive pattern held in the BD group (B = .02, t(313) =0 2.69, p = .008) and MDD group (B = 0.03, t(309) = 3.24, p = .001), but not the healthy controls (B = -0.01, t(292) = -.61, p = .539).

For our analyses predicting testosterone (See **Table S2),** there was a significant positive effect of arousal across all groups (B = 0.03, t(929) = 2.81, p = .005), suggesting greater T was linked to elevated emotional arousal. Analyses within the control group indicated a positive association between negative emotionality and testosterone (B = 0.08, t(289) = 2.55, p = .011). There was also a linear time x negative emotionality interaction (B = 0.04, t(289) = 2.36, p = .019), whereby the testosterone declines across the day were more pronounced when negative emotionality was low (B = -0.18, t(189) = -10.57, p < .001), compared to high (B = -0.12, t(289) = -7.25, p < .001). In the MDD group, there was a significant effect of arousal similar to the effect found across all groups (B = 0.04, t(305) = 2.25, p = .025). There was also a quadratic time

x arousal interaction (B = -0.01, t(305) = -2.04, p = .042). Simple slopes analysis indicated that there was a quadratic pattern of testosterone when arousal was low (-1 SD, B = 0.03, t(305) = 3.65, p < .001), but not high (+1 SD, B = 0.00, t(305) = .33, p = .740). There were no significant emotionality effects or interaction effects with time within the BD group.

Table S1. Prediction of Cortisol from Affective ESM Reports

	All Groups			BD Group		
	B	*t*	*p*	*B*	*t*	*p*
Intercept	0.40	29.84	< .001	0.35	16.91	< .001
Time	-0.17	-37.00	< .001	-0.13	-16.29	< .001
Time^2	0.05	15.03	< .001	0.04	7.70	< .001
Positivity	-0.01	-1.01	0.312	-0.03	-1.96	0.051
Negativity	0.01	0.65	0.513	-0.03	-1.54	0.124
Arousal	0.00	0.08	0.934	0.00	0.03	0.976
Time x Positivity	0.02	3.18	0.002	0.02	2.69	0.008
Time x Negativity	0.00	0.69	0.488	0.01	1.07	0.286
Time x Arousal	0.01	1.54	0.124	0.00	0.43	0.669
Time^2 x Positivity	0.00	1.00	0.317	0.00	-0.61	0.542
Time^2 x Negativity	0.00	0.24	0.809	0.00	0.00	0.999
Time^2 x Arousal	0.00	-0.34	0.734	0.00	0.90	0.371
	MDD Group			CTL Group		
	B	*t*	*p*	*B*	*t*	*p*
Intercept	0.44	17.39	< .001	0.40	19.78	< .001
Time	-0.20	-26.68	< .001	-0.19	-23.80	< .001
Time^2	0.05	9.87	< .001	0.05	9.41	< .001
Positivity	0.01	0.49	0.622	0.01	0.31	0.754
Negativity	0.03	1.73	0.084	0.04	1.83	0.068
Arousal	0.01	0.47	0.642	-0.01	-0.42	0.677
Time x Positivity	0.03	3.24	0.001	-0.01	-0.61	0.539
Time x Negativity	0.02	1.82	0.07	-0.02	-1.74	0.083
Time x Arousal	0.00	0.54	0.589	0.01	1.03	0.304
Time^2 x Positivity	-0.01	-1.80	0.072	0.00	0.62	0.537
Time^2 x Negativity	0.00	-0.39	0.694	0.00	0.41	0.683
Time^2 x Arousal	0.00	-0.78	0.436	0.00	-0.42	0.675

Table S2. Prediction of Testosterone from Time and Affective ESM Reports

	All Groups			BD Group		
	B	*t*	*p*	*B*	*t*	*p*
Intercept	4.22	80.49	< .001	4.20	45.90	< .001
Gender	0.62	6.96	< .001	0.56	3.72	< .001
Time	-0.15	-21.48	< .001	-0.14	-10.52	< .001
Time^2	0.03	5.99	< .001	0.03	3.42	< .001
Positivity	0.00	-0.03	0.974	-0.01	-0.20	0.846
Negativity	0.03	1.71	0.088	0.01	0.23	0.820
Arousal	0.03	2.81	0.005	0.03	1.51	0.131
Time x Positivity	0.01	1.24	0.215	0.00	-0.10	0.922
Time x Negativity	0.01	1.48	0.140	0.01	0.63	0.528
Time x Arousal	0.00	-0.59	0.559	0.00	-0.08	0.938
Time^2 x Positivity	0.00	-0.46	0.644	0.00	0.41	0.680
Time^2 x Negativity	0.01	-1.40	0.163	0.00	-0.36	0.720
Time^2 x Arousal	0.00	-1.05	0.296	0.00	0.25	0.805
	MDD Group			CTL Group		
	B	*t*	*p*	*B*	*t*	*p*
Intercept	4.32	43.94	< .001	4.11	51.26	< .001
Gender	0.66	3.38	0.002	0.69	5.56	< .001
Time	-0.14	-14.24	< .001	-0.15	-12.83	< .001
Time^2	0.02	2.85	0.005	0.03	3.95	< .001
Positivity	-0.03	-1.08	0.282	0.03	0.92	0.361
Negativity	0.00	0.12	0.902	0.08	2.55	0.011
Arousal	0.04	2.25	0.025	0.02	1.00	0.319
Time x Positivity	0.02	1.39	0.167	0.02	1.49	0.139
Time x Negativity	0.00	0.06	0.955	0.04	2.36	0.019
Time x Arousal	0.01	0.75	0.455	-0.02	-1.81	0.071
Time^2 x Positivity	-0.01	-1.08	0.281	-0.01	-0.56	0.573
Time^2 x Negativity	0.00	-0.12	0.904	-0.02	-1.77	0.078
Time^2 x Arousal	-0.01	-2.04	0.042	0.00	-0.15	0.883

Affective Associations with Cortisol Awakening Response and Overall Hormonal Output

To examine how cortisol awakening responses and overall cortisol and testosterone output were associated with PA, NA, and Arousal, we conducted a series of multilevel models regressing each of these three daily affective reports on gender, time, a quadratic time term, the cortisol awakening response, cortisol AUC, and testosterone AUC. These hormonal parameters were used as predictors rather than outcomes because multilevel modeling does not allow for outcomes that are not within Level 1. These models were conducted across all groups and within specific clinical groups. Overall, the results did not reveal a consistent pattern. The results are presented in **Table S3**. Of note, Testosterone AUC was positively associated with negative affect within the entire sample (B= 1.00E-07, t(128) = 2.09, p = .039), whereas the CAR was negatively associated with negative affect in the MDD group (B= -0.52, t(40) = -2.84, p = .007) and the AUC for cortisol was positively associated with negative affect in the CTL group (B = 2.49E-05, t(41) = 2.93, p = .006). In the CTL group, the cortisol awakening response was positively associated with arousal (B = 0.45, t(41) = 3.05, p= .004).

Table S3. Positive Affect, Negative Affect, and Arousal ESM Reports by Day-level Hormone Parameters.

	All Groups			BD Group			MDD Group			CTL Group		
Prediction of PA	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*
Intercept	3.59	20.96	< .001	3.75	12.80	< .001	3.17	12.02	< .001	4.02	11.13	< .001
Gender	0.29	1.60	0.114	0.80	2.63	0.014	-0.15	-0.54	0.597	-0.02	-0.06	0.955
Time	0.14	6.95	< .001	0.14	3.29	0.001	0.15	4.27	< .001	0.15	4.82	< .001
Time^2	-0.06	-4.12	< .001	-0.06	-2.22	0.027	-0.05	-2.25	0.025	-0.06	-2.82	0.005
CAR	-0.03	-0.38	0.702	-0.12	-0.84	0.404	0.06	-0.31	0.761	0.14	0.93	0.358
AUC Cortisol	0.00	-0.08	0.937	0.00	-0.68	0.501	0.00	0.74	0.464	0.00	-0.87	0.392
AUC Testosterone	0.00	-1.18	0.240	0.00	-1.36	0.180	0.00	-0.19	0.851	0.00	-0.11	0.916
	All Groups			BD Group			MDD Group			CTL Group		
Prediction of NA	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*
Intercept	1.54	9.77	< .001	3.75	12.80	< .001	1.97	7.45	< .001	0.56	2.19	0.029
Gender	-0.44	-2.65	0.010	0.80	2.63	0.014	-0.02	-0.06	0.950	-0.11	-0.40	0.693
Time	-0.04	-2.06	0.040	0.14	3.29	0.001	-0.03	-0.80	0.425	-0.05	-1.55	0.122
Time^2	0.01	0.42	0.975	-0.06	-2.22	0.027	-0.01	-0.28	0.778	0.00	0.23	0.815
CAR	-0.12	-1.51	0.133	-0.12	-0.84	0.404	-0.52	-2.84	0.007	-0.22	-1.99	0.054
AUC Cortisol	0.00	0.62	0.536	0.00	-0.68	0.501	0.00	0.34	0.733	2.49E-05	2.93	0.006
AUC Testosterone	1.00E-07	2.09	0.039	0.00	-1.36	0.180	0.00	0.29	0.774	0.00	0.75	0.455
	All Groups			BD Group			MDD Group			CTL Group		
Prediction of Arousal	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*	*B*	*t*	*p*
Intercept	2.44	12.87	< .001	2.22	6.83	< .001	2.55	8.01	< .001	2.64	7.37	< .001
Gender	-0.07	-0.32	0.750	-0.35	-1.00	0.328	-0.18	-0.48	0.632	-0.01	-0.03	0.978
Time	0.12	5.04	< .001	0.00	0.00	0.997	0.21	5.24	< .001	0.15	3.77	< .001
Time^2	-0.19	-12.25	< .001	-0.13	-4.87	< .001	-0.24	-9.27	< .001	-0.19	-7.42	< .001
CAR	0.01	0.11	0.909	-0.11	-0.76	0.452	-0.29	-1.34	0.189	0.46	3.05	0.004
AUC Cortisol	0.00	0.41	0.683	0.00	1.33	0.191	0.00	0.88	0.383	-2.17E-05	-1.90	0.065
AUC Testosterone	0.00	0.94	0.349	0.00	0.75	0.456	0.00	-0.23	0.821	1.00E-07	1.86	0.070

References

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition* (pp. 94-1). New York, NY, USA: SCID-I/P.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198. doi:10.1016/0022-3956(75)90026-6

Gruber, J., & Weinstock, L. M. (2018). Interrater reliability in bipolar disorder research: current practices and suggestions for enhancing best practices. *International Journal of Bipolar Disorders*, *6*(1), 1-3.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627. doi:10.1001/archpsyc.62.6.617

Ong, D. C., Zaki, J., & Gruber, J. (2017). Increased cooperative behavior across remitted bipolar I disorder and major depression: Insights utilizing a behavioral economic trust game. *Journal of Abnormal Psychology*, *126*(1), 1.

Trivedi, M. H., Rush, A. J., Crismon, M. L., Kashner, T. M., Toprac, M. G., Carmody, T. J., Key, T., Biggs, M. M., Shores-Wilson, K., Witte, B., Suppes, T., Miller, A. L., Altshuler, K.Z., Shon, S. P. & Suppes, T. (2004). Clinical results for patients with major depressive disorder in the Texas Medication Algorithm Project. *Archives of General Psychiatry, 61*(7), 669-680. doi: 10.1001/archpsyc.61.7.669

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *The British Journal of Psychiatry*, *133*(5), 429–435. doi: 10.1192/bjp.133.5.429