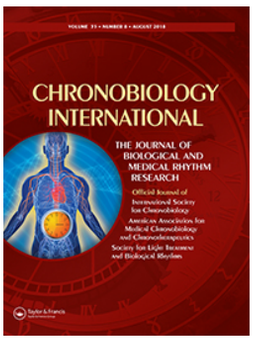

Taylor & Francis
Taylor & Francis Group

Chronobiology International

The Journal of Biological and Medical Rhythm Research

ISSN: 0742-0528 (Print) 1525-6073 (Online) Journal homepage: http://www.tandfonline.com/loi/icbi20

Evidence of a diurnal rhythm in implicit reward learning

Alexis E. Whitton, Malavika Mehta, Manon L. Ironside, Greg Murray & Diego A. Pizzagalli

To cite this article: Alexis E. Whitton, Malavika Mehta, Manon L. Ironside, Greg Murray & Diego A. Pizzagalli (2018) Evidence of a diurnal rhythm in implicit reward learning, Chronobiology International, 35:8, 1104-1114, DOI: 10.1080/07420528.2018.1459662

To link to this article: https://doi.org/10.1080/07420528.2018.1459662

Published online: 24 Apr 2018.

Submit your article to this journal

Article views: 72

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=icbi20

CHRONOBIOLOGY INTERNATIONAL
2018, VOL. 35, NO. 8, 1104–1114
https://doi.org/10.1080/07420528.2018.1459662

Taylor & Francis
Taylor & Francis Group

Check for updates

Evidence of a diurnal rhythm in implicit reward learning

Alexis E. Whitton[a], Malavika Mehta[a], Manon L. Ironside[a], Greg Murray[b], and Diego A. Pizzagalli[a]

[a]McLean Hospital & Harvard Medical School, Belmont, MA, USA; [b]Department of Psychological Sciences, Swinburne University, Hawthorn, Victoria, Australia

ABSTRACT

Many aspects of hedonic behavior, including self-administration of natural and drug rewards, as well as human positive affect, follow a diurnal cycle that peaks during the species-specific active period. This variation has been linked to circadian modulation of the mesolimbic dopamine system, and is hypothesized to serve an adaptive function by driving an organism to engage with the environment during times where the opportunity for obtaining rewards is high. However, relatively little is known about whether more complex facets of hedonic behavior – in particular, reward learning – follow the same diurnal cycle. The current study aimed to address this gap by examining evidence for diurnal variation in reward learning on a well-validated probabilistic reward learning task (PRT). PRT data from a large normative sample (N = 516) of non-clinical individuals, recruited across eight studies, were examined for the current study. The PRT uses an asymmetrical reinforcement ratio to induce a behavioral response bias, and reward learning was operationalized as the strength of this response bias across blocks of the task. Results revealed significant diurnal variation in reward learning, however in contrast to patterns previously observed in other aspects of hedonic behavior, reward learning was lowest in the middle of the day. Although a diurnal pattern was also observed on a measure of more general task performance (discriminability), this did not account for the variation observed in reward learning. Taken together, these findings point to a distinct diurnal pattern in reward learning that differs from that observed in other aspects of hedonic behavior. The results of this study have important implications for our understanding of clinical disorders characterized by both circadian and reward learning disturbances, and future research is needed to confirm whether this diurnal variation has a truly circadian origin.

ARTICLE HISTORY
Received 30 January 2018
Revised 26 March 2018
Accepted 27 March 2018

KEYWORDS
Reward learning; anhedonia; circadian rhythms; dopamine

Introduction

Across species, hedonic behaviors such as feeding, mating and socializing, are known to follow a diurnal cycle. Evolutionary theories posit that this daily rhythm in response to rewarding stimuli (also known as "reward potential"; Sleipness et al., 2005) serves an adaptive function, as it drives an organism to engage with the environment at times when the opportunity of obtaining rewards is high. As such, hedonic behaviors typically peak during the species-specific active period and wane nearing periods of rest (for a review, see Webb et al., 2009). This diurnal fluctuation in reward-based behavior is hypothesized to arise due to circadian modulation of the mesolimbic dopamine (DA) system by the master circadian oscillator – the hypothalamic suprachiasmatic nucleus (SCN). This endogenous circadian clock is synchronized to the environment via photic stimulation that varies in accordance with the light/dark cycle, and entrains multiple aspects of physiology and behavior to this cycle through internal messengers (e.g. melatonin; Cajochen et al., 2003).

Circadian pacemakers in the SCN govern several lower order processes, such as sleep, body temperature, hormone levels and digestion (Moore and Eichler, 1972; Stephan and Zucker, 1972). In the context of reward, the SCN is hypothesized to entrain basic appetitive drives, including reward seeking and consumption, which tend to follow an inverted U-shaped rhythm that peaks during the active period. For example, studies in rats have shown that fluctuations in the consumption of both natural (e.g. palatable food) and drug rewards (e.g. cocaine) exhibit clear circadian variation, with intake peaking during the night when the species is most active (for a review, see McClung, 2007). At

CONTACT Alexis E. Whitton and Diego A. Pizzagalli awhitton@mclean.harvard.edu; dap@mclean.harvard.edu Laboratory for Affective and Translational Neuroscience, McLean Hospital & Harvard Medical School, Belmont, MA 02478, USA

© 2018 Taylor & Francis Group, LLC

the molecular level, several aspects of DA transmission have been found to display diurnal rhythms that may underpin this variation in reward-based behavior, including the expression of DA receptors and the density and function of the DA transporter (McClung, 2007).

Similar diurnal cycles in reward-based behavior have been observed in humans. Most notably, research in psychiatrically healthy adults has shown that positive affect – an experiential manifestation of reward potential – also exhibits a quadratic pattern across the day, with more positive states between noon and the early evening compared to the morning or later in the evening (Clark et al., 1989; Murray et al., 2002, 2009; Watson et al., 1999). This rhythm appears to be specific to positive affect, as similar variations have not been observed in negative affect (Murray et al., 2002; Watson et al., 1999). Studies that have constrained social interaction or those that have used forced desynchrony protocols to separate endogenous from activity-related influences on mood (e.g. Murray et al., 2009) suggest that this variation in positive affect is a truly biologically driven rhythm that may reflect circadian modulation of the mesolimbic DA system.

Although several studies converge on an inverted U-shaped diurnal pattern in reward-based behavior, growing evidence suggests that reward processing is multifaceted and extends beyond these basic appetitive drives, and it is unknown whether other facets of hedonic behavior follow the same diurnal cycles. A critical aspect of an organism's hedonic repertoire is reward learning, which refers to the ability to adaptively modulate behavior based on reinforcement. This process is critically reliant on the phasic firing of DA neurons in the ventral tegmental area, which increase firing in response to rewarding stimuli or reward-predicting cues (Cohen et al., 2012; Eshel et al., 2015; Schultz et al., 2015). In particular, during learning, the sign and magnitude of this DA signal is dependent on the degree to which the reward is anticipated or deviates from prior reward expectancies, with unexpected rewards eliciting a strong phasic DA signal known as a reward prediction error (Schultz et al., 2015). Once learning has occurred, this DA signal is elicited by reward-predictive cues and diminishes in response to the reward itself (Schultz et al., 1997). As such, current theories posit that phasic DA firing acts as a teaching signal that promotes learning about antecedent cues (Starkweather et al., 2017; Steinberg et al., 2013). Although these neural systems overlap closely with those implicated in basic motivational drives, very little is known about whether reward learning is modulated by the circadian processes described previously. This is an important area of research as disruption in both circadian systems and reward learning are two major factors that have been implicated in mood disorders, particularly as a mechanism underpinning depression and anhedonia (Alloy et al., 2015; Whitton et al., 2015). Accordingly, gaining a better understanding of circadian influences on reward learning may facilitate the dissection of molecular links between circadian oscillators and mood-controlling systems, and thereby improve our understanding of mood pathology.

To date, only one study (using the Iowa Gambling task) has examined diurnal variation in reward-based decision making, and in contrast to other aspects of reward processing, performance did not exhibit diurnal variation (Byrne and Murray, 2017). Furthermore, two human fMRI studies examining circadian modulation in neural reward response in DA-rich regions of the striatum have generated conflicting findings. In a small pilot study (n = 11), Hasler and colleagues (2014) found ventral striatal responses to receipt of monetary reward in a non-learning task to be higher during an afternoon scan (~5 pm) relative to a morning scan (~10 am). This finding contrasts with findings from a second study by Byrne, Hughes and colleagues (2017) who found evidence of decreased activation in the putamen – a region critically implicated in reward learning – upon receipt of a monetary reward when participants were scanned in the afternoon (2 pm) relative to the morning (10 am) and evening (7 pm). These mixed findings indicate that further research is needed to better understand how the circadian system may influence reward-based learning.

The current study aimed to address this gap in the literature by examining evidence for diurnal variation in learning on a widely used and well-validated probabilistic reward learning task (PRT; Pizzagalli et al., 2005, 2008). The current study represents a secondary analysis of data collected from a large normative sample (N = 516) of non-

clinical individuals. Rooted in signal detection theory (McCarthy and Davison, 1979; Tripp and Alsop, 1999), the PRT uses an asymmetrical reinforcement schedule to induce a behavioral response bias toward a more frequently rewarded ("rich") stimulus. Prior studies have shown that conditions that reduce reward-related phasic DA signaling (e.g. low-dose pramipexole; Der-Avakian et al., 2013; Pizzagalli et al., 2008) inhibit the acquisition of this response bias. Conversely, stimulants thought to increase phasic DA signaling (e.g. nicotine, amphetamine) have been found to potentiate response bias (Barr et al., 2008; Der-Avakian et al., 2013). Along similar lines, blunted response bias has been observed in psychiatric conditions hypothesized to involve disruptions in striatal DA reward signaling, such as depression and anhedonia (Liu et al., 2011; Morris et al., 2015; Pizzagalli et al., 2008; Vrieze et al., 2013b). Furthermore, multi-modal neuroimaging studies have shown that individual differences in the strength of this response bias are linked to (1) DA release (Vrieze et al., 2013a), (2) DA transporter binding (Kaiser et al., 2017) and (3) resting state functional connectivity within nodes of the reward system (Kaiser et al., 2017).

The primary aim of the current study was to evaluate whether reward learning on the PRT showed evidence of a diurnal pattern paralleling fluctuations observed in preclinical studies of reward self-administration, as well as human studies of positive affect, which have been shown to have a peak in the middle of the species-specific active period. In line with this, we predicted that response bias would show a quadratic pattern, with the strongest bias evident in the mid-to-late afternoon, mirroring the peak observed in positive affect. Furthermore, since prior evidence has identified a circadian rhythm in general alertness (Van Dongen and Dinges, 2005), we also evaluated stimulus discriminability on the PRT (a measure of general task performance that is independent of reward responsivity) as well as reaction time and accuracy, as proxy measures of alertness and vigilance, in order to determine whether fluctuations in reward learning might be accounted for by fluctuations in more general aspects of cognitive functioning. Finally, given evidence showing that positive affect varies in accordance with changes in levels of daylight (a factor putatively underlying seasonal variation in mood; Golder and Macy, 2011), we also examined whether reward learning varied as a function of daylight availability, defined as the number of hours from sunrise to sunset.

Materials and methods

Sample

Subjects were 516 individuals from eight studies (Unpublished study, 2007, 2008; Bogdan et al., 2010; Bogdan and Pizzagalli 2009; Goetz et al., 2013; Nikolova et al., 2012; Pizzagalli et al., 2008; Vrieze et al., 2013b), recruited from schools, universities and the community (see Table 1 for a summary of the study-specific details). The sample included 316 females (61.2%),and the age range was 17-70 years old ($M = 26.7$, $SD = 11.8$). Of those who had ethnicity data available ($N = 300$), the majority were Caucasian ($N = 219$, 73.0%). All participants were free from the use of mood altering medications (except caffeine and alcohol), major neurological conditions, medical conditions and psychiatric disorders, and had normal or corrected-to-normal vision.

Participant screening

Screening procedures differed slightly across each of the studies. In four of the eight studies, healthy control status was confirmed using the Structured Clinical Interview for DSM-IV (First et al., 1994). For the other four studies, healthy control status was defined as an absence of self-reported current or past psychiatric disorder. Five of the eight studies recruited controls through the community using flyers and advertisements, two recruited from college campuses and one recruited from a sample of graduating seniors at a high school in Bulgaria as part of a larger genetics study. In seven of the eight studies, the Mood and Anxiety Symptom Questionnaire (MASQ; Watson et al., 1995) was also administered.

Probabilistic reward task (PRT)

The PRT consists of trials in which cartoon faces are presented in the center of the monitor. Trialsbegin with a fixation cross (500 ms), followed by a face with no mouth. After a 500 ms delay,

Table 1. Study-specific PRT parameters and assessment of healthy control status.

Study	*N*	# blocks	Trials per block	Short mouth length	Long mouth length	Reinforcer amount	Sample type	Assessment of healthy control status
1. Unpublished study (2007)	24	3	100	11.5 mm	13.0 mm	5 cents	College undergraduate	Self-reported absence of current or past psychiatric, neurological or learning disorder
2. Unpublished study (2008)	31	3	100	11.5 mm	13.0 mm	5 cents	Community	SCID-IV non-patient edition
3. Pizzagalli et al. (2008)	26	3	100	11.5 mm	13.0 mm	5 cents	Community	SCID-IV non-patient edition
4. Bogdan and Pizzagalli (2009)	32	3	80	10.0 mm	11.0 mm	5 cents	Community	Self-reported absence of current or past psychiatric, neurological or learning disorder
5. Bogdan et al. (2010)	158	3	100	11.5 mm	13.0 mm	5 cents	College and community	SCID-IV non-patient edition
6. Nikolova et al. (2012)	96	3	80	10.0 mm	11.0 mm	0.15 lev	Graduating high school seniors	Self-reported absence of current or past depression or bipolar disorder, absence of drug use in the past month
7. Goetz et al. (2013)	84	3	100	11.5 mm	13.0 mm	5 cents	College undergraduate	Self-reported absence of psychological distress
8. Vrieze et al. (2013b)	65	3	100	11.5 mm	13.0 mm	5 eurocents	College and community	SCID-IV non-patient edition

SCID-IV = Structured Clinical Interview for DSM-IV Axis I disorders, non-patient edition.

either a short mouth or a long mouth is presented for 100 ms. Participants are instructed to press a key to indicate whether the short or the long mouth was presented. For each block of trials, 40% of correct trials are followed by a monetary reward. Participants are instructed to try to win as much money as possible. Long and short trials are presented at equal frequency, however, unbeknownst to participants, correct identification of one mouth length (the "rich" stimulus) is rewarded three times more frequently than correct identification of the other mouth length (the "lean" stimulus). Prior studies using this task have shown that this asymmetrical reinforcement schedule elicits a behavioral response bias toward the rich stimulus across the course of the task (Pizzagalli et al., 2005, 2008). Prior to analysis, data were evaluated against several quality control criteria. Specifically, participants were only included in analyses if they demonstrated above-chance accuracy (i.e. >55%) to ensure that they were exposed to the intended asymmetrical (3:1) reinforcement schedule. Participants were also excluded if they had greater than 10% of reaction time outlier trials (defined as trials on which responses were <150 ms or >2500 ms).

Across the eight studies, the parameters of the PRT differed slightly depending on the amount of time allotted for task administration and the overall remuneration (study-specific task parameters are outlined in Table 1). Briefly, for six of the studies the short and long mouth lengths were 11.5 mm and 13 mm, respectively, and for the other two studies the short and long mouth lengths were 10 mm and 11 mm, respectively. For six studies, monetary reward was 5 cents, for one study the reward was 5 eurocents (~$0.06 USD) and for one study the reward was 0.15 Bulgarian lev (~$0.09 USD). Finally, six of the studies used three blocks of 100 trials and two used three blocks of 80 trials. An overview of the sequence and timing of events leading up to PRT administration, across the eight studies, is shown in Figure 1.

Mood and Anxiety Symptom Questionnaire – short form

The MASQ is a 62-item self-report scale that was used to assess variation in mood and anxiety symptoms (Watson et al., 1995). It was developed to evaluate predictions of the tripartite model of anxiety and depression. The MASQ has four subscales: (1) General Distress Anxiety (11 items); (2) General Distress Depression (12 items); (3) Anxious Arousal (17 items) and (4) Anhedonic Depression (22 items). Participants are asked to rate the extent to which they experienced each symptom during the past 2 weeks

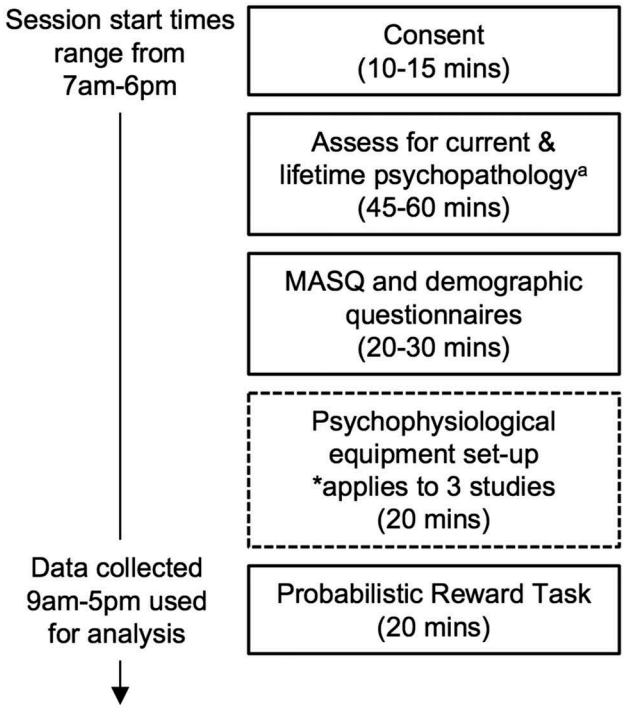

Figure 1. Broad schematic overview of the timing and sequence of events leading up to PRT administration, across the 8 studies. Note that although the starting times of experimental sessions ranged from 7am to 6pm, only PRT data obtained between the hours of 9am and 5pm were analyzed, in order to ensure that a sufficient number of cases were included in each hourly time bin. aFor unpublished study (2008) and Bogdan et al. (2010), consent and clinical assessment were conducted at a separate session that occurred approximately one week prior to PRT administration.

on a Likert scale from 1 (*not at all*) to 5 (*extremely*). The four MASQ subscales have been found to possess good psychometric properties in student, adult volunteer and clinical samples (Watson et al., 1995).

PRT data reduction

Signal detection analysis (Macmillan and Creelman, 1991) was used to calculate response bias (the tendency to bias responding to the rich stimulus) and discriminability (the ability to accurately distinguish between the two mouth sizes) for each block of the PRT using the following formulae:

$$\text{Response bias: } \log b = \frac{1}{2}\log\left(\frac{Rich_{correct} * Lean_{incorrect}}{Rich_{incorrect} * Lean_{correct}}\right)$$

$$\text{Discriminability: } \log d = \frac{1}{2}\log\left(\frac{Rich_{correct} * Lean_{correct}}{Rich_{incorrect} * Lean_{incorrect}}\right)$$

To compute response bias and discriminability for cases that contained a zero in the formula, 0.5 was added to each cell in the matrix (Hautus, 1995). Mean accuracy and reaction time were also computed.

Statistical analyses

Data were first divided into hourly bins based on the time of day at which the PRT was completed. The primary analyses focused on hourly bins containing at least 20 cases, which resulted in the inclusion of cases sampled exclusively between 9 am and 5 pm. Separate mixed repeated measures analyses of variance (ANOVA) were used to examine diurnal patterns in response bias, discriminability, accuracy and reaction time over the course of the day, with *Time* (9 am, 10 am, …, 4 pm, 5 pm) as a between-subjects factor and *Block* (1, 2, 3) as the repeated factor. Individual one-way ANOVAs were also conducted on each of the four subscales of the MASQ to determine whether there were any comparable diurnal patterns in self-reported depressive, anxious or anhedonic symptoms. Follow-up analyses including "study" as a covariate were also conducted. To examine putative seasonal variations in reward learning, Pearson correlation was used to assess the association between day length (at the location and date of testing), and overall response bias (averaged across all three blocks). Day length was computed using the "geosphere" package (Hijmans, 2017) in r (version 3.4.1), which calculates the number of hours between sunrise and sunset, as a function of latitude and day of the year (range 1–365). Data were collected across five geographic locations, therefore the following latitudes (in conjunction with the assessment date) were used to compute day length at the time of testing: Durham, NC, USA: 35.994; Twinsburg, OH, USA: 41.3126; Cambridge, MA, USA: 42.3736; Yambol, Bulgaria: 42.3736; and Leuven, Belgium: 50.8798.

Results

Sample

Demographic characteristics of the entire sample are shown in Table 2. Of this sample, 485 subjects completed the PRT between 9 am and 5 pm and were included in primary analyses.

Table 2. Sample characteristics.

Demographics (N = 516)		
Female, N (%)	316	(61.2)
Age, M (SD)	26.7	(11.8)
Caucasian, N (%)	219[a]	(73.0)
MASQ scores, M (SD) (N = 448)		
Depressive distress	20.6	(8.4)
Anxious distress	18.2	(6.2)
Anhedonia	53.4	(13.7)
Anxious arousal	22.8	(7.7)

[a]Percentage computed based on the portion of the sample that had ethnicity data available (n = 300).

Diurnal variation in response bias

A significant main effect of *Time* emerged for response bias, $F(7, 477) = 2.40$, $p = 0.02$, $\eta_p^2 = 0.03$. *Post hoc* tests showed that overall response bias was significantly lower at 12 pm relative to 10 am ($p = 0.049$), 11 am ($p = 0.01$), 1 pm ($p = 0.002$) and 2 pm ($p = 0.02$). After applying a Bonferroni correction to control for 24 comparisons, response bias at 12 pm remained significantly lower than response bias at 1 pm (Bonferroni-adjusted $p = 0.048$; Figure 2A). Follow-up analyses confirmed that the main effect of *Time* remained significant when *Study* was entered as a covariate ($p = 0.04$). A finer-grained examination of response bias from 11 am to 2 pm in 20-min increments revealed that compared to response bias at 11:20–11:40 am, response bias was significantly lower at 12:00–12:20 pm (adjusted $p = 0.003$) and 12:20–12:40 pm (adjusted $p = 0.005$), and steadily increased thereafter (Figure 2B).

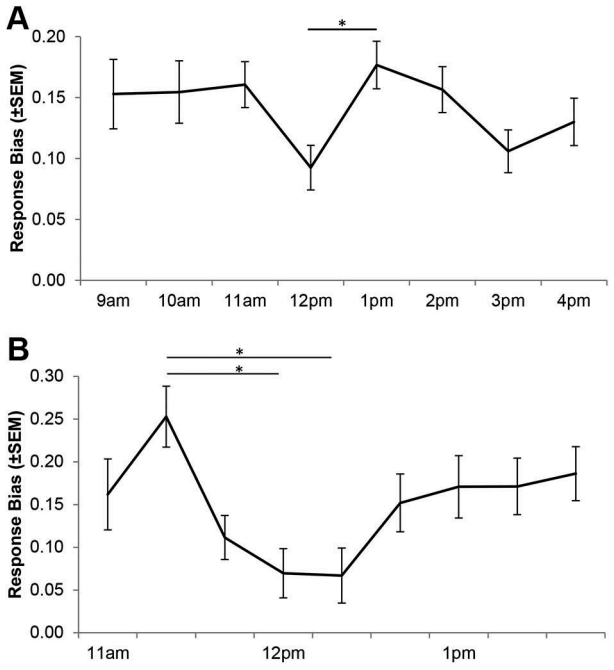

Figure 2. Mean response bias (±SEM) per hour from 9am to 5pm is shown in panel A. A finer-grained analysis of the decrease in response bias during the middle of the day is shown in 20-minute increments from 11am to 2pm in panel B. Asterisks (*) indicate pairwise comparisons that were significant at p < 0.05 after correcting for multiple comparisons.

Diurnal variation in discriminability

A main effect of *Time* also emerged for discriminability, $F(7, 477) = 2.94$, $p = 0.005$, $\eta_p^2 = 0.04$. Bonferroni-corrected *post hoc* tests showed evidence of a steady decrease in discriminability as a function of time, with discriminability being significantly lower at 3 pm relative to 9 am (adjusted $p = 0.01$) and 12 pm (adjusted $p = 0.02$). Follow-up analyses confirmed that when controlling for overall discriminability, the main effect of *Time* for response bias remained significant, $F(7, 476) = 2.48$, $p = 0.02$, $\eta_p^2 = 0.04$, indicating that fluctuations in response bias across the day were not simply attributable to fluctuations in discriminability (i.e. they were not driven by diurnal variation in more general task performance). Providing further evidence for the independence of these diurnal rhythms, we found that the reverse was also true, in that the main effect of *Time* for discriminability remained significant when overall response bias was controlled for, $F(7, 476) = 3.06$, $p = 0.004$, $\eta_p^2 = 0.04$.

Diurnal variation in accuracy and reaction time

A *Time* × *Block* ANOVA on mean accuracy (averaged across rich and lean accuracy) revealed a main effect of *Time*, $F(7, 477) = 2.86$, $p = 0.006$, $\eta_p^2 = 0.04$. Bonferroni-corrected *post hoc* tests revealed that accuracy followed a similar temporal pattern to discriminability, being significantly lower later in the day. Specifically, accuracy at 3 pm was found to be significantly worse than at 9 am (adjusted $p = 0.04$) and 12 pm (adjusted $p = 0.007$). As with discriminability, we found that the effect of *Time* for response bias remained significant when controlling for overall accuracy, $F(7, 476) = 2.58$, $p = 0.01$, $\eta_p^2 = 0.04$. We also confirmed that the reverse was true, where the effect of *Time* on accuracy remained when controlling for overall response bias, $F(7, 476) = 3.08$, $p = 0.004$, $\eta_p^2 = 0.04$. No effects or interactions

involving *Time* emerged for reaction time (all *p*s > 0.05).

Diurnal variations in self-reported depressive, anxious and anhedonic symptoms

Prior to analyses, MASQ subscales were log-transformed to ensure normality. Results showed no significant main effects of *Time* for any of the MASQ subscales (all *p*s > 0.05), indicating that changes in PRT response bias over the course of the day were likely independent of variation in state levels of depression, anxiety or anhedonia.

Assessment of seasonal variation in response bias, discriminability and anhedonia

Day length did not correlate with overall response bias, or any of the MASQ subscales (all *r*s < 0.1 ; *p*s > 0.05), however, day length was correlated with overall PRT discriminability ($r = 0.17$, $p < 0.001$), due to better discriminability at times and locations that had a greater daylight availability.

Discussion

Prior evidence points to a clear diurnal rhythm in the self-administration of natural and drug rewards (Webb et al., 2009) as well as in human positive affect (Clark et al., 1989; Murray et al., 2009; Watson et al., 1999), but little is known about potential biological rhythmicity in reward learning. Therefore, the aim of the current study was to evaluate whether a similar diurnal pattern exists for reward learning, which was assessed using a behavioral reinforcement learning paradigm that has been shown to be sensitive to pharmacological challenges known to affect reward-related phasic DA signaling (Barr et al., 2008; Der-Avakian et al., 2013; Pizzagalli et al., 2008; Santesso et al., 2009), DA release (Vrieze et al., 2013a), DA transporter function (Kaiser et al., 2017), and resting state functional connectivity within the reward circuit (Kaiser et al., 2017). We also sought to examine the degree to which any diurnal variation in reward learning was attributable to, or independent from, any concurrent diurnal pattern in more general task performance, vigilance and alertness (indexed by discriminability, accuracy and reaction time). Finally, given prior evidence linking seasonal variation in day length to variation in positive affect (Golder and Macy, 2011), we examined whether differences in day length at the time and location of testing was associated with variation in reward learning, discriminability and mood and anxiety symptoms on the MASQ. As hypothesized, we observed a diurnal pattern in PRT response bias, however contrary to predictions, this diurnal pattern did not follow the same timing as that observed in prior studies examining other aspects of reward-based behavior and positive affect. Specifically, a significant decrease in response bias was seen at 12 pm, which then increased steadily from 1 pm onwards. Although a significant diurnal pattern was also observed for discriminability and accuracy, these followed a different time course and did not appear to account for the decrease in response bias. Similarly, time-dependent variation in discriminability and accuracy did not appear to be attributable to variation in response bias, further supporting the independence of these diurnal patterns. Furthermore, there was no evidence of diurnal variation in reaction time or any of the MASQ subscales (which probe depressive and anxiety symptoms over the past 2 weeks). Finally, although we observed evidence for a link between greater daylight availability and improved PRT discriminability, the same link was not observed for response bias. Taken together, these findings reveal evidence of a diurnal variation in behavioral reward learning that is distinct from concurrent variation in aspects of more general task performance or self-reported symptomatology. Importantly, the timing of this variation differs from that observed in other aspects of reward-based behavior, raising the possibility that reward learning may follow a different rhythm.

Current theories posit that the increase in motivated behavior that occurs during the species-specific active period is biologically advantageous as it drives an organism to engage with the environment at times when opportunities for obtaining rewards is maximal (Sleipness et al., 2005). Our data are suggestive of a contrasting diurnal pattern in reward learning, with a nadir occurring in the middle of the day. Therefore, a critical question is whether this pattern arises because the same circadian system is driving reward learning in an

opposing manner, or whether this diurnal variation is entrained by a different system. Support for the former comes from a recent study by Byrne, Hughes and colleagues (2017) who found evidence of relatively *reduced* reward-related activation in the dorsal striatum on a (non-learning) monetary choice task when participants were scanned at 2 pm relative to 10 am or 7 pm. This finding was conceptualized as a form of reward prediction error, wherein the magnitude of neural reward responses was modulated by circadian variation in reward priming, and therefore smallest at times of day when rewards were most expected (for humans, in the afternoon). A similar finding has been observed in patterns of cue-induced drug craving in abstinent heroin users, which was lowest at noon relative to the morning and evening (Ren et al., 2009). The relationship between fMRI blood oxygen level-dependent fluctuations in the striatum and DA activity is complex (Knutson and Gibbs, 2007). However, at the molecular level, it is possible that these midday dips in the magnitude of striatal reward responses or cue-induced craving – both of which rely on phasic DA signals – may arise because the phasic DA signal is diminished against a background of elevated basal DA firing, which has been found to peak around the middle of the day (Webb et al., 2009). Extending this conceptualization to our data, it is possible that reward learning is reduced in the middle of the day because the phasic DA teaching signal is diminished via tonically increased DA (Bilder et al., 2004), which may slow the acquisition of reward contingencies. These speculations await empirical evaluation.

An alternative explanation is that the diurnal variation in reward learning observed in the current study is driven by a separate system. Recent work has pointed to the existence of a dopaminergic ultradian (~4 h) oscillator that operates independently from the SCN, and is hypothesized to govern daily rhythms in arousal and activity (Blum et al., 2014). Emphasizing the dopaminergic origin of these rhythms, perturbation of the DA system via administration of amphetamines has been found to lengthen these ultradian cycles, whereas reducing DA activity via the D2 antagonist haloperidol has been found to significantly shorten these cycles. Although the dopaminergic ultradian oscillator is thought to act in concert with the circadian system to regulate daily activity patterns, it has been suggested that elevations in DA tone may cause it to become desynchronized with the circadian system. Blum and colleagues (2014) propose that disruption in this dopaminergic oscillating system, rather than the circadian system *per se*, underpins the aberrant sleep-wake cycles and arousal patterns observed in individuals with psychiatric disorders. Since several lines of evidence now point to significant disruptions in reward learning in individuals with mood disorders (for a review, see Whitton et al., 2015), future studies that examine variation in reward learning across the full 24-h period are needed to determine whether reward learning may fluctuate according to this more ultradian, as opposed to circadian, time course.

Several factors must be kept in mind when interpreting the findings of the current study. First, although our data are suggestive of a distinct daily rhythm in reward learning, our study design does not allow us to conclude that this rhythm has a circadian origin. The next step is to test the endogeneity of this rhythm using rigorous chronobiological protocols that control for exogenous influences on reward learning, such as activity levels and social interaction, and that use within-subjects designs. Second, as this was a secondary analysis of existing data, we did not restrict participant's daily activity and it is likely that our sample contains individuals with varying sleep-wake cycles, which may impact diurnal variation in reward learning. Third, although all participants were free from the use of psychotropic medication and illicit substances, and were asked to refrain from alcohol 24 h before testing, we did not limit or control intake of other substances that may impact circadian variation in reward processes, such as caffeine, or nicotine, and future research controlling for these critical factors is needed to rule out their potential confounding effects. Finally, although we did not observe any significant diurnal variation in self-reported anxious, depressive and anhedonic symptoms on the MASQ subscales, this may be partly attributable to (1) a floor effect, since the sample was comprised of healthy controls and (2) the fact that

these subscales assess self-reported mood over the past 2 weeks. An important next step is to examine these findings in the context of clinical samples with mood pathology and using temporally more granular assessment methods (e.g. ecological momentary assessment).

To conclude, findings from the current study point to a distinct diurnal pattern in reward learning, the timing of which differs from that observed in basic appetitive drives, such as reward seeking and consumption. Furthermore, we found that this diurnal rhythm is likely not driven by concurrent fluctuations in general aspects of task performance that may be influenced by diurnal variation in general alertness or vigilance. These findings have important implications for our understanding of clinical disorders characterized by both circadian and reward learning disturbances, and future research is needed to confirm whether this diurnal variation has a truly circadian origin.

Declaration of interest

Over the past three years, Dr. Pizzagalli received consulting fees from Akili Interactive Labs, BlackThorn Therapeutics, Boehreinger Ingelheim, Pfizer and Posit Science for activities unrelated to the present study. Dr. Pizzagalli has a financial interest in BlackThorn Therapeutics, which has licensed the copyright to the Probabilistic Reward Task through Harvard University. Dr. Pizzagalli's interests were reviewed and are managed by McLean Hospital and Partners HealthCare in accordance with their conflict of interest policies. No from these entities was used to support the current work, and all views expressed are solely those of the authors. All other authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Contributions

DAP designed the original studies that contributed the data for the current analyses; AEW, MM and MLI performed the data analysis with critical input from DAP; AEW and DAP drafted the manuscript and GM provided expert input on later drafts of the manuscript.

Funding

The current analyses were supported by R01MH068376, R01MH101521, and R37MH068376. AEW is supported by funding from the National Health and Medical Research Council, the Brain and Behavior Research Foundation, McLean Hospital and the National Institute of Mental Health.

References

Alloy LB, Nusslock R, Boland EM. 2015. The development and course of bipolar spectrum disorders: An integrated reward and circadian rhythm dysregulation model. Annu Rev Clin Psychol. 11:213–50. doi:10.1146/annurev-clinpsy-032814-112902.

Barr RS, Pizzagalli DA, Culhane MA, Goff DC, Evins AE. 2008. A single dose of nicotine enhances reward responsiveness in nonsmokers: Implications for development of dependence. Biol Psychiatry. 63:1061–65. doi:10.1016/j.biopsych.2007.09.015.

Bilder RM, Volavka J, Lachman HM, Grace AA. 2004. The catechol-O-methyltransferase polymorphism: Relations to the tonic–phasic dopamine hypothesis and neuropsychiatric phenotypes. Neuropsychopharmacology. 29:1943–61. doi:10.1038/sj.npp.1300542.

Blum ID, Zhu L, Moquin L, Kokoeva MV, Gratton A, Giros B, Storch K-F. 2014. A highly tunable dopaminergic oscillator generates ultradian rhythms of behavioral arousal. Elife. 3:e05105. doi:10.7554/eLife.05105.

Bogdan R, Perlis RH, Fagerness J, Pizzagalli DA. 2010. The impact of mineralocorticoid receptor ISO/VAL genotype (rs5522) and stress on reward learning. Genes Brain Behav. 9:658–67. doi:10.1111/j.1601-183X.2010.00600.x.

Bogdan R, Pizzagalli DA. 2009. The heritability of hedonic capacity and perceived stress: A twin study evaluation of candidate depressive phenotypes. Psychol Med. 39:211–18. doi:10.1017/S0033291708003619.

Byrne JE, Hughes ME, Rossell SL, Johnson SL, Murray G. 2017. Time of day differences in neural reward functioning in healthy young men. J Neurosci. 37:8895–900. doi:10.1523/JNEUROSCI.0918-17.2017.

Byrne JE, Murray G. 2017. Diurnal rhythms in psychological reward functioning in healthy young men: 'Wanting', liking, and learning. Chronobiol Int. 34:287–95. doi:10.1080/07420528.2016.1272607.

Cajochen C, Kräuchi K, Wirz-Justice A. 2003. Role of melatonin in the regulation of human circadian rhythms and sleep. J Neuroendocrinol. 15:432–37. doi:10.1046/j.1365-2826.2003.00989.x.

Clark LA, Watson D, Leeka J. 1989. Diurnal variation in the positive affects. Motiv Emot. 13:205–34. doi:10.1007/BF00995536.

Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature. 482:85–88. doi:10.1038/nature10754.

Der-Avakian A, D'souza M, Pizzagalli D, Markou A. 2013. Assessment of reward responsiveness in the response bias probabilistic reward task in rats: Implications for cross-species translational research. Transl Psychiatry. 3:e297. doi:10.1038/tp.2013.74.

Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. Nature. 525:243–46. doi:10.1038/nature14855.

First MB, Spitzer RL, Gibbon M, Williams JB. 1994. Structured clinical interview for Axis I DSM-IV disorders. New York: Biometrics Research.

Goetz EL, Hariri AR, Pizzagalli DA, Strauman TJ. 2013. Genetic moderation of the association between regulatory focus and reward responsiveness: A proof-of-concept study. Biol Mood Anxiety Disord. 3:3. doi:10.1186/2045-5380-3-3.

Golder SA, Macy MW. 2011. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science. 333:1878–81. doi:10.1126/science.1202775.

Hasler BP, Forbes EE, Franzen PL. 2014. Time-of-day differences and short-term stability of the neural response to monetary reward: A pilot study. Psychiatry Res. Neuroimaging. 224:22–27. doi:10.1016/j.pscychresns.2014.07.005.

Hautus MJ. 1995. Corrections for extreme proportions and their biasing effects on estimated values of d'. Behav Res Meth Instrum Comput. 27:46–51. doi:10.3758/BF03203619.

Hijmans RJ. (2017). Geosphere: spherical trigonometry. R package version 1.5-7. https://CRAN.R-project.org/package=geosphere

Kaiser RH, Treadway MT, Wooten DW, Kumar P, Goer F, Murray L, Beltzer M, Pechtel P, Whitton A, Cohen AL. 2017. Frontostriatal and dopamine markers of individual differences in reinforcement learning: A multi-modal investigation. Cereb Cortex:1–10 doi:10.1093/cercor/bhx281.

Knutson B, Gibbs SE. 2007. Linking nucleus accumbens dopamine and blood oxygenation. Psychopharmacology. 191:813–22. doi:10.1007/s00213-006-0686-7.

Liu W-H, Chan RC, Wang L-Z, Huang J, Cheung EF, Gong Q-Y, Gollan JK. 2011. Deficits in sustaining reward responses in subsyndromal and syndromal major depression. Prog Neuropsychopharmacol Biol Psychiatry. 35:1045–52. doi:10.1016/j.pnpbp.2011.02.018.

Macmillan N, Creelman C. 1991. Detection theory: A user's guide. Cambridge UP: Cambridge.

McCarthy D, Davison M. 1979. Signal probability, reinforcement and signal detection. J Exp Anal Behav. 32:373–86. doi:10.1901/jeab.1979.32-373.

McClung CA. 2007. Circadian genes, rhythms and the biology of mood disorders. Pharmacol Ther. 114:222–32. doi:10.1016/j.pharmthera.2007.02.003.

Moore RY, Eichler VB. 1972. Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. Brain Res. 42:201–06. doi:10.1016/0006-8993(72)90054-6.

Morris BH, Bylsma LM, Yaroslavsky I, Kovacs M, Rottenberg J. 2015. Reward learning in pediatric depression and anxiety: Preliminary findings in a high-risk sample. Depress Anxiety. 32:373–81. doi:10.1002/da.22358.

Murray G, Allen NB, Trinder J. 2002. Mood and the circadian system: Investigation of a circadian component in positive affect. Chronobiol Int. 19:1151–69. doi:10.1081/CBI-120015956.

Murray G, Nicholas CL, Kleiman J, Dwyer R, Carrington MJ, Allen NB, Trinder J. 2009. Nature's clocks and human mood: The circadian system modulates reward motivation. Emotion. 9:705–16. doi:10.1037/t06070-000.

Nikolova Y, Bogdan R, Pizzagalli DA. 2012. Perception of a naturalistic stressor interacts with 5-HTTLPR/rs25531 genotype and gender to impact reward responsiveness. Neuropsychobiology. 65:45–54. doi:10.1159/000329105.

Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M. 2008. Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. J Psychiatr. 43:76–87. doi:10.1016/j.jpsychires.2008.03.001.

Pizzagalli DA, Jahn AL, O'Shea JP. 2005. Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. Biol Psychiatry. 57:319–27. doi:10.1016/j.biopsych.2004.11.026.

Ren Z-Y, Zhang X-L, Liu Y, Zhao L-Y, Shi J, Bao Y, Zhang XY, Kosten TR, Lu L. 2009. Diurnal variation in cue-induced responses among protracted abstinent heroin users. Pharmacol Biochem Behav. 91:468–72. doi:10.1016/j.pbb.2008.08.023.

Santesso DL, Evins AE, Frank MJ, Schetter EC, Bogdan R, Pizzagalli DA. 2009. Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. Hum Brain Mapp. 30:1963–76. doi:10.1002/hbm.20642.

Schultz W, Carelli RM, Wightman RM. 2015. Phasic dopamine signals: From subjective reward value to formal economic utility. Curr Op Behav Sci. 5:147–54. doi:10.1016/j.cobeha.2015.09.006.

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science. 275:1593–99. doi:10.1126/science.275.5306.1593.

Sleipness EP, Sorg BA, Jansen HT. 2005. Time of day alters long-term sensitization to cocaine in rats. Brain Res. 1065:132–37. doi:10.1016/j.brainres.2005.10.017.

Starkweather CK, Babayan BM, Uchida N, Gershman SJ. 2017. Dopamine reward prediction errors reflect hidden-state inference across time. Nat Neurosci. 20:581–89. doi:10.1038/nn.4520.

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. 2013. A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci. 16:966–73. doi:10.1038/nn.3413.

Stephan FK, Zucker I. 1972. Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions. Proc Natl Acad Sci. 69:1583–86. Retrieved from http://www.pnas.org/

Tripp G, Alsop B. 1999. Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. J Clin Child Adolesc Psychol. 28:366–75. doi:10.1207/S15374424jccp280309.

Van Dongen HP, Dinges DF. 2005. Sleep, circadian rhythms, and psychomotor vigilance. Clin Sports Med. 24:237–49. doi:10.1016/j.csm.2004.12.007.

Vrieze E, Ceccarini J, Pizzagalli DA, Bormans G, Vandenbulcke M, Demyttenaere K, Van Laere K, Claes S. 2013a. Measuring extrastriatal dopamine release during a reward learning task. Hum Brain Mapp. 34:575–86. doi:10.1002/hbm.21456.

Vrieze E, Pizzagalli DA, Demyttenaere K, Hompes T, Sienaert P, De Boer P, Schmidt M, Claes S. 2013b. Reduced reward learning predicts outcome in major depressive disorder. Biol Psychiatry. 73:639–45. doi:10.1016/j.biopsych.2012.10.014.

Watson D, Weber K, Assenheimer JS, Clark LA, Strauss ME, McCormick RA. 1995. Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. J Abnormal Psychol. 104:3–14. doi:10.1037/0021-843X.104.1.3.

Watson D, Wiese D, Vaidya J, Tellegen A. 1999. The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. J Pers Soc Psychol. 76:820–38. doi:10.1037/0022-3514.76.5.820.

Webb IC, Baltazar RM, Lehman MN, Coolen LM. 2009. Bidirectional interactions between the circadian and reward systems: Is restricted food access a unique zeitgeber? Eur J Neurosci. 30:1739–48. doi:10.1111/j.1460-9568.2009.06966.x.

Whitton AE, Treadway MT, Pizzagalli DA. 2015. Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. Curr Opin Psychiatry. 28:7–12. doi:10.1097/YCO.0000000000000122.