

ACCEPTED ARTICLE PREVIEW

Accepted Article Preview: Published ahead of advance online publication

**The Protective Action Encoding of Serotonin Transients in the Human Brain**

Rosalyn J Moran, Kenneth T Kishida, Terry Lohrenz, Ignacio Saez, Adrian W Laxton, Mark R Witcher, Stephen B Tatter, Thomas L Ellis, Paul EM Phillips, Peter Dayan, P Read Montague

Cite this article as: Rosalyn J Moran, Kenneth T Kishida, Terry Lohrenz, Ignacio Saez, Adrian W Laxton, Mark R Witcher, Stephen B Tatter, Thomas L Ellis, Paul EM Phillips, Peter Dayan, P Read Montague, The Protective Action Encoding of Serotonin Transients in the Human Brain, *Neuropsychopharmacology* accepted article preview 3 January 2018; doi: [10.1038/npp.2017.304](https://doi.org/10.1038/npp.2017.304).

This is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication. NPG are providing this early version of the manuscript as a service to our customers. The manuscript will undergo copyediting, typesetting and a proof review before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Received 26 June 2017; revised 1 October 2017; accepted 28 November 2017;
Accepted article preview online 3 January 2018

a. Title page

The protective action encoding of serotonin
transients in the human brain.

Rosalyn J. Moran^{*1,2}, Kenneth T. Kishida^{*9,10}, Terry Lohrenz^{*1}, Ignacio Saez³, Adrian W. Laxton⁴, Mark R. Witcher⁴, Stephen B. Tatter⁴, Thomas L. Ellis⁴⁺, Paul E.M. Phillips⁵, Peter Dayan⁶, P. Read Montague^{1,7,8}

¹Virginia Tech Carilion Research Institute, Virginia Tech, Roanoke, VA 24016, USA

²Bradley Department of Electrical & Computer Engineering, Virginia Tech, Blacksburg, VA 24061, USA

³Helen Wills Neuroscience Institute, University of California, Berkeley, 175 Li Ka Shing Center, Berkeley, CA 94720, USA

⁴Department of Neurosurgery, Wake Forest Health Sciences, Winston-Salem, NC 27157, USA

⁵Departments of Psychiatry & Behavioral Sciences and Pharmacology, University of Washington, Seattle, WA 98195

⁶The Gatsby Computational Neuroscience Unit, University College London, 25 Howland St. London W1T 4JG, UK

⁷Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA

⁸Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London WC1N 3BG, UK

⁹Department of Physiology and Pharmacology, Wake Forest Health Sciences, Winston-Salem, NC 27157, USA

¹⁰Department of Neurosurgery, Wake Forest Health Sciences, Winston-Salem, NC 27157, USA

Running Title: Protective action encoding of serotonin

Correspondence: Rosalyn J. Moran,
Merchant Venturers Building,
4 Woodland Road,
Clifton,
Bristol BS81UB,
UK
Email: rm16003@bristol.ac.uk
Telephone: +44 (0) 117 95 45468
Fax: +44 (0) 117 331 7391

Keywords: Serotonin, Fast-Scan Cyclic Voltammetry, Reward Prediction Error, Counterfactual Reward Prediction Error, Striatum

b. Abstract

The role of serotonin in human brain function remains elusive due, at least in part, to our inability to measure rapidly the local concentration of this neurotransmitter. We used fast-scan cyclic voltammetry to infer serotonergic signaling from the striatum of fourteen brains of human patients with Parkinson's disease. Here we report these novel measurements and show that they correlate with outcomes and decisions in a sequential investment game. We find that serotonergic concentrations transiently increase as a whole following negative reward prediction errors, while reversing when counterfactual losses predominate. This provides initial evidence that the serotonergic system acts as an opponent to dopamine signaling, as anticipated by theoretical models. Serotonin transients on one trial were also associated with actions on the next trial in a manner that correlated with decreased exposure to poor outcomes. Thus, the fluctuations observed for serotonin appear to correlate with the inhibition of over-reactions and promote persistence of ongoing strategies in the face of short-term environmental changes. Together these findings elucidate a role for serotonin in the striatum, suggesting it encodes a protective action strategy that mitigates risk and modulates choice selection particularly following negative environmental events.

c. Introduction

The neurotransmitter serotonin influences a broad range of brain functions, including mood, sleep, learning, and decision-making (Portas, Bjorvatn et al. 2000, Must, Juhász et al. 2007, Diekhof, Falkai et al. 2008), and is thus implicated in a diverse range of diseases, including obsessive compulsive disorder (Hu, Lipsky et al. 2006), anorexia nervosa (Kaye, Gendall et al. 1998) and depression (Whittington, Kendall et al. 2004, Risch, Herrell et al. 2009). However, there is much debate as to what it actually encodes. Various data and arguments favor aspects of it as being opponent to dopamine, thus suggesting a role in punishment and loss (Deakin 1983, Graeff and Deakin 1991, Daw, Kakade et al. 2002, Cools, Roberts et al. 2008, Crockett, Clark et al. 2009, Tanaka, Shishida et al. 2009, Boureau and Dayan 2011). Other evidence suggests a role in patience while waiting for a reward (Miyazaki, Miyazaki et al. 2012, Worbe, Savulich et al. 2014, Fonseca, Murakami et al. 2015, Li, Zhong et al. 2016), disengagement (Tops, Russo et al. 2009), motor activity (Jacobs 1994), or even reward (Dölen, Darvishzadeh et al. 2013, Liu, Zhou et al. 2014, Li, Zhong et al. 2016). Recent optogenetically tagged recordings from serotonergic neurons in the raphé have not resolved these debates (Cohen, Amoroso et al. 2015), perhaps partly because of heterogeneity among serotonin neurons (Lowry 2002). A voltammetric approach in awake, behaving, humans may offer new insights into the role of serotonin at a temporal resolution that reveals dynamic events several orders of magnitude faster than positron emission tomography

(Yao, Li et al. 2009) whilst also observing concentrations of the neurotransmitter at downstream structures directly, rather than assuming linearity with firing rates in the raphé (Montague, McClure et al. 2004).

Recently, we developed a fast-scan cyclic voltammetry procedure to identify dopamine fluctuations in patients with Parkinson's disease who were undergoing surgery for deep brain stimulator implantation (Kishida, Saez et al. 2016). These unique multi-subject recordings provided millisecond-resolved estimates of relative fluctuations in the concentration of this neurotransmitter whilst patients played an investment task that was designed to elicit adaptive behaviour in the face of rewards and punishments. Serotonin also produces redox reactions around this range; thus, were it possible to distinguish serotonin from dopamine, we could measure relative serotonergic responses from exactly these same signals, thus addressing its involvement in the same task. There are approaches to perform this separation or boost serotonin signals in slices or in vivo in other animals such as picking peaks in the voltammetric recordings visually (John and Jones 2007), using drugs (John and Jones 2007, Hashemi, Dankoski et al. 2011) or even optogenetics (Xiao, Privman et al. 2014). However, these do not readily extend to studies in humans.

Here we derived a principled supervised learning method to apply to these same human recordings to measure trial-by-trial fluctuations in extracellular concentrations of serotonin, distinguished from dopamine. We used this to measure serotonin transients in 16 hemispheres from patients as they performed the investment task (12 of these recordings were used in the previous dopamine assessment). In the task, patients selected an investment level on each trial as market prices unfolded according to real historical financial markets (Lohrenz, McCabe et al. 2007). Dopamine transients in this task were found to encode not the typical 'reward prediction error' but rather a signal that combined actual and counterfactual reward prediction errors (the latter when subjects invested little, and so experienced regret or rejoicing when the market rose or fell respectively) (Kishida, Saez et al. 2016). We hypothesized that serotonin may function in a similar way but in the loss domain, with the potential to represent counterfactual losses from foregone gains akin to a regret signal (a positive reward prediction error when investments are low). Moreover, given the theorized role of serotonin in active avoidance (Dayan and Huys 2008), we investigated whether fluctuations in serotonin correlated with avoidance behavior in the task (i.e. reducing one's level of exposure) to actual or counterfactual losses.

d. Materials and Methods

Participants & Fast-scan cyclic voltammetry

Fourteen patients (aged 61.1 ± 9.7 yrs., 2 female) undergoing surgery for the implantation of deep brain stimulation (DBS) electrodes participated in the experiment.

Two patients participated on separate days during right and left implantation to give a total of $n = 16$ recordings. All patients had a diagnosis of Parkinson's disease. Participants provided informed consent and were instructed that they could opt out of the experiment at any time. Procedures were approved by the Virginia Tech and Wake Forest Baptist Medical Center Institutional Review Boards (IRB #11-078). No adverse effect of the extended surgical procedure was reported. Eleven of our participants (12 hemispheres) formed part of the cohort reported in (Kishida, Saez et al. 2016). In Supplementary Table 1 we provide details of disease duration, medication and comorbidities. During the surgery and during our test period, all patients were off their dopamine-related medications but remained on all other pharmacological treatments. For details of the surgical and voltammetry procedure see Supplementary Materials and Methods. Our voltammetric electrochemical assays relied upon redox reactions of serotonin along a carbon fiber (at 10 Hz), which we assumed would produce measurable current changes in proportion to the concentration of the chemical species in the extracellular space (for full details of the carbon fiber and FSCV protocol see Supplementary Materials and Methods).

Investment Task

Participants played a decision making game in a simulated 'stock market'. Participants were first endowed with 100 points, and, on each trial they had to decide an amount to invest in the stock market. This amount could be 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100% of their points. At the beginning of a game with a single market, participants were first shown a trajectory of previous market moves indicating changes in the value of the market (Figure 3A). They were then asked to make their first investment decision (submit bet) and following a delay (840 ± 12 msec), the market value change was revealed and players lost or gained points in accordance with market returns. Following this outcome, participants then submitted a further 19 investment or 'bet' decisions in a self-timed manner. On average, players waited 4.5 secs to submit their next decision after the reveal event. In all 6 markets of 20 decisions were played. Our markets were consistent with real historical markets (e.g. mimicking events before and after 1929 Wall St. crash and 1987's 'black Monday').

We chose this decision-making task since it is engaging and it could be used to interrogate the serotonergic response to 'what is' as well as 'what could have been'. The latter involve so-called counterfactual or fictive reward prediction errors (RPEs). These errors are ecologically important as learning signals and have been previously shown to activate the human striatum (Lohrenz, McCabe et al. 2007). For the details of how we computed reward prediction error on each trial see Supplementary Materials and Methods

The game was designed with 6 markets of 20 decisions (120 decisions). In 14 of our 16 experiments, the participants completed the full game. For two patients, recordings

were stopped prior to experimental completion with 5 markets (100 decisions) and 2 markets (40 decisions) completed. Thus a total of 1820 trials were played. From those 1820 trials, a total of 1729 trials could be associated with prospective actions within a market (i.e. where $[\text{bet}(t+1) - \text{bet}(t)]$ can be computed within a given market). In figure 4 we present the concentration differences in serotonin for negative compared to positive reward prediction errors at high [60 – 100%] and low [0 – 50%] betting levels. Here 1724 trials are included as 5 trials resulted in RPEs equal to zero. In figures 5 and 6 we examine these 1724 bet decisions parametrically and in terms of the adjustments in bet made on the next trial. We show the transients in serotonin associated with lowering, holding or raising one's bet, from investment levels of [0%], [10%, 20%], [30%, 40%], [50%, 60%], [70%, 80%] and [90%, 100%].

Testing Reward Prediction Errors and Action Control

For each voltammogram we applied the model's serotonin coefficients to predict concentration levels of serotonin at 100 msec temporal resolution. For each trial we collected transient responses of 700 msec duration around the time of market reveal (0 msec) from -100 to 600 msec. Data were z-scored over an individual market and baseline corrected within their cell to zero at either -100 msec (Figure 4) or 0 msec (Figures 5 and 6).

To test for responses that corresponded with reward prediction errors (Figure 4), we tested the fluctuations in serotonin transients at the time of market reveal (the outcome of their decision) from 0 – 600 msec. For this analysis we split our trials according to the outcome-related reward prediction error and according to the bet size (high or low). To test whether these prediction-error related-responses (Figure 4) further corresponded with prospective action encoding (Figures 5 & 6) we performed a second analysis. Specifically we tested, again, the fluctuations in serotonin transients at the time of market reveal from 0 – 600 msec. For this analysis however, we split our trials according to the outcome-related reward prediction error (negative; Figure 5 and positive; Figure 6), and according to the decision to 'raise' 'hold' or 'lower' the current bet from levels (0%, [10-20%], [30-40%], [50-60%], [70-80%], [90-100%]). We aimed to examine differences in serotonin transients associated with negative compared to positive reward prediction errors. Given our previous findings for dopamine (Figure 2), we were also interested in the bet dependence on the prediction error since, for example, positive reward prediction errors might be coded as a poor outcome if the bet decision on that trial was low. We also aimed to determine whether these transients predicted decision making on the next trial – i.e. what adjustment might be applied from current investment levels. For full details of the statistical analysis applied to test for reward prediction error-related serotonin responses and to test for action control-related responses please see Supplementary Materials and Methods.

e. Results

Serotonin Concentrations: Estimation Results

We followed a supervised-learning approach for extracting serotonin signals (Kishida, Saez et al. 2016). We created a training set of voltammograms taken from a set of probes constructed identically to those used in the patient recordings. The probes were exposed to analytes in a flow cell containing a known range of serotonin concentrations (from 0 – 8000 nM) confounded with various concentrations of dopamine and pH levels. The concentration dependence of the shape and magnitude of the serotonin redox current was directly apparent (Figure 1A). We therefore trained a multivariate penalized regression model (Tibshirani 1996) to extract serotonin concentration estimates from all points of each trace (1000 samples over 10 msec, Figure 1A). We aimed to produce a dual-transmitter model that could predict both serotonin and dopamine, since we had previously observed task-related dopamine fluctuations in these samples.

We used L1-penalized regression to create a generalizable dual-transmitter regression model that estimated the concentrations of DA independent of the ambient levels of 5-HT and pH, and 5-HT, independent of the level of DA and pH (Supplemental Information, Figure S1). We found that regression coefficients were distributed throughout the trace indicating that concentration levels were best predicted by considering not only the peaks of oxidation and reduction but also other points distributed across the voltage sweep (Figure 1A). Crucially we were able to estimate the concentrations of serotonin in each mixture independent of dopamine (Figure 1B) and we could predict dopamine concentration levels independent of serotonin (Figure 1C). Moreover, our predictions for serotonin were trained to ignore altering pH levels (Supplemental Information, Figure S2). This model estimated the true serotonin levels within 90% confidence intervals of the estimated levels in absolute terms (Figure 1D), even in the presence of dopamine and at differing pH. We also tested whether at very low concentrations we could differentiate 5-HT concentrations and found a resolution of ~100 nM (Supplemental Information, Figure S2).

Figure 1A shows that determining the low contamination of the models is difficult to observe by visual inspection, since the voltammograms for changing dopamine concentrations appear similar to those for serotonin (Figure 1A). For additional validation of our procedure, we compared the dopamine predictions to our previously published findings (Figure 2). We confirmed that, on the identical datasets to those previously published (17 recordings in total), we could replicate previous results on transient fluctuations in dopamine from the dorsal striatum (Kishida, Saez et al. 2016) (Figure 2). We also show in a supplemental analysis (Figure S2), the correlation structure

amongst our DA and 5 HT estimates where small positive correlations were found to exist.

Behavior on the Sequential Investment Task

Figure 3A shows the sequence of events in the investment game (Lohrenz, McCabe et al. 2007) which participants played during voltammetric recordings from dorsal striatum. The game was designed to elicit prediction-, prediction error-, reward- and future investment-related signals associated with revelation of market price movements on 120 separate trials over 6 historical markets (20 moves per market). On each trial participants chose a level of investment for their current endowment with possible choices from 0-100% with 10% increments (Figure 3A) and submitted their choices. Then participants were shown the market move (its change in value Figure 3A), to end a trial. Our behavioral data showed that over all subjects, bets were distributed bimodally across these 11 possible investment choices (Figure 3B), with investment levels distributed around 50% and also peaking at 100%. Reward prediction errors measure the difference between the return on a trial and a prediction. We defined the return as the fractional change in wealth (combining the current bet size and market change), and the prediction as an average of recent previous returns. We also scaled this difference by the standard deviation of those previous returns (see Experimental Procedures). Across the cohort, this led to a spectrum of positive and negative reward prediction errors (Figure 3C). Further, based on our previous results (Kishida, Saez et al. 2016) we considered counterfactual as well as real outcomes depending on the current betting level. This study suggested that negative outcomes could be experienced in two ways: the first were those outcomes where negative reward prediction errors were experienced and so events were 'worse than expected' (since bets were high). The second were counterfactual negative events, in which positive reward prediction errors occurred when bets were low and thus regret on a foregone gain. We correlated the whole collection of events and choices with relative fluctuations in transmitter concentrations.

Serotonin encodes loss prediction errors

We assessed serotonin responses in voltammograms at a repetition frequency of 10 Hz using the penalized regression models developed above. We examined fluctuations in estimated concentrations at the time of trial outcomes (as the market move is revealed) and tested for the serotonergic encoding of prediction errors. Figure 4A displays the serotonin transients associated with positive and negative reward prediction errors. Remarkably, when considering all betting levels, serotonin displayed an upward fluctuation to negative prediction errors and a downward fluctuation to positive prediction errors (Figure 4A and see also Supplemental Information, Figure S3). Given the potential difference in response to negative reward prediction errors at high and low betting levels (loss and counterfactual loss/foregone gain respectively), we examined serotonin fluctuations across a median split of bet levels. Figure 4B shows that this encoding reversed for the lower half of bets with upward serotonin

fluctuations encoding positive errors and downward fluctuations encoding negative errors. The inversion of the encoding can be understood as the presence of a counterfactual term for serotonin; which responds to negative outcomes both in the context of a surprising loss when one was highly invested in the market and a surprising gain when one was not. The difference between Figures 4B and Figure S3 is the baseline normalization of the signals to 100ms before revelation of the outcome (Figure 4B) or at the time of revelation (Figure S3). We present both as a form of exploratory result. They suggest that the dynamics of how the prediction component of the prediction error is represented in serotonin concentration would be worth exploring in future studies; in particular, higher temporal resolution in the voltammetric signal could elucidate early dynamics that alter baseline properties (Schmidt et al. 2013, Figure S3). Here the interaction of reward prediction error and bet amount was significant for the earlier baseline (Figure 4B). In our Supplemental Information (Figure S3) we also include a random effects analysis across patients to ensure that our results are not driven by only a few subjects. These additional analyses support our findings when accounting for individual differences.

To allow for duration of serotonergic signaling to be altered in response to reward prediction errors, we computed the area under the curve to indicate ‘cumulative serotonin’ responses. For this analysis the interaction of prediction error (positive or negative) and bet invested (high or low) was significant (Figure 4C). In particular, the response to positive reward prediction errors seemed to induce a depression in serotonin that was more prolonged than in the low bet condition. Further, a parametric analysis revealed a small but significant negative correlation between the serotonin response and RPE at high bets and a small but significant positive correlation between serotonin response and RPE at low bets (Figure 4D). To examine the subjective effects of actual and counterfactual gains and losses, rather than to reward prediction error per se, we conducted a further supplemental analysis (Figure S4). This revealed a lack of a parametric effect in gains or losses (Figure S4).

Serotonin protects investors from loss

Given these bet-dependent prediction error transients, we sought to establish serotonin’s influence on investment decisions. The effect of counterfactual outcomes on both dopamine (Kishida, Saez et al. 2016) and serotonin (Figure 4) suggests that it is crucial to perform the analysis of action encoding (betting more or less) at different bet levels since a bet of 0% could result in large foregone gains (i.e., counterfactual losses), while a bet of 100% could result in large actual losses. In other words, in the context of this task, one’s next move carries two distinct risks of loss on the upcoming trial. We tested whether fluctuations in 5-HT could be used to predict the bet level on the next trial. Using a multiple linear regression we tested for serotonin and game factors in predicting the next decision. Specifically, our independent variables included the area under the curve of the 5-HT transient from 100 – 600 msec, the bet level, the polarity of the RPE at trial (t), as well as their interactions. Our dependent variable was the change

in bet at trial ($t+1$). Our regression model revealed significant predictive power in upcoming decision (F-statistic vs. constant model: 32.4, p -value < 0.00001 , Supplementary Table 2). Importantly the regressor describing the interaction of serotonin and current bet level was a significant predictor of the upcoming decision ($p = 0.04$). This was a negative interaction indicating that for large serotonin responses and large bets, participants tended to decrease their bet and for large serotonin responses and small current bets, participants tended to increase their bets. We found a that the three-way interaction of serotonin, bet level and RPE sign was at trend level significance ($p = 0.13$), Supplementary Table 2.

To examine and illustrate these regression effects, we first separated out ranges of current bet levels (Figure 5A), and examined how serotonin transients were associated with decisions following a negative reward prediction error (Figure 5). We tested the relationship between serotonin responses and current bet levels at trial (t) for decisions to 'lower' and for decisions to 'hold or raise' the bet on trial ($t+1$) (Figure 5B). These analyses are a recapitulation of the negative RPE responses in Figure 4 but separated according to what the subject decides to do next. We found that under the conditions of the decision to withdraw from the market following negative reward prediction errors there was a strong positive correlation between 5-HT and current betting levels (Figure 5C). This is important given that withdrawal from the market (i.e. lowering one's bet) is consistent with the hypothesized role for serotonin in forms of avoidance (Dayan and Huys 2008). This striking parametric effect is indicated in the serotonin time courses of figure 5B. We can see that reducing the bet from a high amount implies reducing the risk of actual loss, and is associated with positive serotonin fluctuations (Figures 5B). Reducing the bet from an already low amount implies increasing the risk of counterfactual losses, and is associated with negative fluctuations. A trend towards a significant positive correlation for serotonin and decisions to hold-or-raise one bets was also observed. In the time-courses we can see that particularly at 10-20% bet levels, serotonin rises following a negative RPE and is associated with a subsequent raise-or-hold bet decision. This direction is again consistent with serotonin protecting against counterfactual losses on the next trial.

In figure 6 we explore the same dependencies but following positive reward prediction errors. Here at low bets, again from 10-20% levels we observe a upgoing serotonin transient (Figure 6B) that dominates the low bet regime (Figure 4B). At this betting level both 'lower' and 'raise-hold' decisions are associated with a positive transient. Significant negative-going fluctuations are observed at the lowest betting level of 0%. No parametric effects are observed for either decision following positive reward prediction errors (Figure 6C and D).

In order to investigate the timing of these decision-related transients further, we extracted the peak 5-HT response from every trial and tested whether a faster time to peak corresponded with decisions from investment on the next trial. For responses to negative reward prediction errors we found no timing effects in an analysis of decision x

bet-level. However, in response to positive reward prediction errors we saw a significant effect of time to peak on the decision to lower, hold or raise ones' bets following the outcomes. Specifically, the decision to raise ones' bets was associated with slower 5-HT transient peaks as compared to decisions to hold or reduce current betting levels. No effect of bet level or interaction was observed (Supplemental Information Figure S5). This may suggest that fast serotonin signals are associated with withdrawal from market-investment, even when the 'going is good'.

f. Discussion

We used a modern statistical method to extract serotonin signals from fast-scan cyclic voltammetric data. We showed *in vitro* using a flow cell that we could extract separate 5-HT and dopamine signals from a single voltammogram, even at variable pH, and then applied our method to recordings taken *in vivo* from human Parkinson patients playing an investment game. We found that at the time that the outcome of a round was revealed, serotonin encoded a prediction error for actual reward when subjects were substantially invested, and an inverted prediction error for counterfactual reward (i.e., regret) when subjects had failed to invest. Moreover, these serotonin concentration fluctuations on a trial were positively correlated with protective choices made by subjects in the subsequent trial following negative reward prediction errors.

Our findings provide novel evidence that serotonin encodes loss-related prediction errors. This finding ratifies and extends previous theoretical accounts which hypothesized a role for serotonin in aversive prediction and learning (Deakin 1983, Daw, Kakade et al. 2002). For the high bet case, our findings demonstrate the opposite of standard accounts of the activity of dopamine neurons (Schultz, Dayan et al. 1997, Kishida, Saez et al. 2016) or transient fluctuations of dopamine concentrations (Flagel, Clark et al. 2011) recorded in conventional Pavlovian or instrumental paradigms in animals other than humans. For the low bet case, our results for serotonin are the mirror image of those for dopamine on the same task, showing a sensitivity to counterfactual as well as actual outcomes (Kishida, Saez et al. 2016).

Our second major finding was an emergent action code in serotonin that could be used to predict the change in bet following negative reward prediction errors. In particular our findings related to betting less or holding on the next trial are consistent with computational accounts of serotonin in active avoidance (Dayan and Huys 2008, Dayan and Huys 2009). However, here too the signaling was not unidirectional. Positive fluctuations to lower bets and negative ones to holding predominated when current bet levels were high. However, they flipped polarity at 50%, suggesting that serotonin drops when avoidance is overall detrimental, and likely to expose the player to foregone gains. Colloquially this might be deemed as a signal that both 'secures in place' when risk is already low and 'retreats' from dangerous high-risk investments. The dependencies on prior sensory experience and its sensitivity to context (previous bet amounts) may help

clarify some formerly confusing and seemingly contradictory findings with respect to learning signals and reward error processing (Yacubian, Gläscher et al. 2006). (Figures 5B & C, 6A).

Recently serotonergic firing in rodents has been associated with patience (Miyazaki, Miyazaki et al. 2012). They found that optogenetic stimulation of rodent serotonergic neurons in the dorsal raphé enables waiting for delayed reward (Miyazaki, Miyazaki et al. 2014). Similarly, fibre photometry recordings from the dorsal raphé nucleus have recently demonstrated increased tonic firing of serotonergic neurons during a reward-related anticipatory period, and phasic firing on reward acquisition (Li, Zhong et al. 2016). In future work, an experimental manipulation of the time from bet-submission to outcome in our task would enable us to formally test the role of serotonin in signaling patience for anticipated reward. With regard the outcome-acquisition related activity (Li, Zhong et al. 2016), our results also find increased 5-HT to positive reward prediction errors but only in the context of low bets (when outcomes could have been better). Studies that systematically vary actual and counterfactual gains and losses might be required to unravel these effects in animals.

Our findings are more directly comparable to studies which investigated lose-shift and win-stay behaviors. For example, Bari et al (Bari, Theobald et al. 2010) have shown that an acute-dose of SSRIs in rodents can increase lose-shift behavior but that longer-term chronic administration increased win-stay behaviors. Similar effects on lose-shift behaviors in humans have been associated with genetic polymorphisms in the serotonin transporter gene (SERT), which was dissociated from dopamine transporter polymorphisms on perseveration (den Ouden, Daw et al. 2013) in the same task. We show that serotonin estimates from our participants are associated with lose-shift behaviours. But interestingly the associated shift behavior is not a simple 'withdrawal' from the game. Rather, the shift behaviors associated with serotonin increases (Figure 5B) is towards the centre of our betting levels. These effects are observed following negative reward prediction errors and might mean that the player is not exposed to 'too much' risk while also ensuring that they do not 'miss out' on future gains.

Pharmacological manipulations using tryptophan depletion in humans supports this idea of behavioral effects of 5-HT in the face of aversive outcomes with reports of both Pavlovian and instrumental predictions of negative events reduced following dietary depletion (Crockett, Clark et al. 2012), as well as reports of a specific model-based or goal-directed deficit (Worbe, Palminteri et al. 2016). Overall, our methodology provides a unique opportunity to understand the role of serotonin in the human brain in computational terms. Relating these findings to our previous work we provide evidence for serotonin's loss-opponency to dopamine's gain-dependent signals (Kishida, Saez et al. 2016) and further extend this valence dependent activity with evidence for a role in subsequent action selection. Our analysis could be further extended to examine, for example, individual differences in reward prediction errors or in decision making. This may reflect altered model-based goals. Testing such a proposal is beyond the scope of

the present work, but would start from developing a quantitative parametrized computational account of the full task. For example, the proposed role of serotonin in regulating temporal discounting and impulsivity (Doya, 2007) could be explored with a temporal difference model of the task (and manipulations of cue-timings for example). Our results, at least informally, are consistent with a role for serotonin in controlling impulsivity. In particular, following negative reward prediction errors, the impulsive choice may be to lower further one's investment. However, at low bet levels, serotonin increases are associated with decisions to 'hold' or even 'raise' the bet. This may be a protective signal that guards against over-reactions to negative outcomes. Further in our supplementary analysis (Figure S5), we show that after outcomes that elicit positive reward prediction errors, the serotonin transient is faster for lower, or hold decisions, perhaps preventing an impulsive raise in betting levels. Our reward-prediction error and action encoding transients might also be considered in light of models of risk assessment. Risk, typically modelled as predicted outcome variance, has been mapped to serotonin function in the basal ganglia (e.g. in Balasubramani et al. 2014). Our transient serotonin increases to negative RPEs at high bets and low bets are associated with decisions to lower the bet and to hold or raise bets respectively. Thus 5-HT may seek a mid-point where the potential variability in both actual and counterfactual losses are balanced (e.g. minimizing maximal losses).

Limitations of the current study are related to both the human brains from which the signals were acquired and the specificity of voltammetric signal extraction. Our cohort comprised 14 patients with Parkinson's disease. This pathology has been associated with aberrations in some decision-making parameters – including impulsivity (Voon, Irvine et al. 2014). Though our paradigm did not address learning *per se*, the effect of valence on learning and subsequent decision-making has been shown to be affected by dopamine medication. In patients off medication positive outcomes tend to affect learning more prominently, while patients off dopamine medication show a greater sensitivity to negative outcomes – effects thought to be controlled by a high dopamine 'tone' or 'floor' on medication (Frank, Seeberger et al. 2004). Though in the absence of learning, a similar off/on study showed that decision making tends to be improved by dopamine medications in patients generally (Shiner, Seymour et al. 2012), though here the loss domain was not explicitly investigated. Other studies have corroborated (Frank, Seeberger et al. 2004) showing reduced sensitivity to negative feedback when patients are on medications (Euteneuer, Schaefer et al. 2009). Our task did not require learning. In other decision-making studies without learning, it has been shown that patients with Parkinson's disease and age-matched controls perform comparably (unlike Parkinson's patients with dementia, who are not in our cohort) (Delazer, Sinz et al. 2009). Hence, the type of game we chose for our participants has been shown to engender near-normal performance. Nevertheless, the pattern of choice selection may not fully reflect the statistics of choice in people without neurological disease. Secondly, patients with Parkinson's disease may exhibit cross-loading of these particular neurotransmitters into the alternate axonal terminals. Specifically, levodopa is thought to induce a 'false-transmission' of dopamine via serotonin axons, and may contribute to the dyskinesias

associated with long term L-Dopa use (Mathur and Lovinger 2012, Politis, Wu et al. 2014). This cross loading would reduce the distinct effects of dopamine and serotonin in the striatum (Montague, Kishida et al. 2016). These caveats deserve study in their own right using this type of protocol. Here we show the feasibility of dissociating dopamine from serotonin and thus the procedures may be extended to test particular cross-talk hypotheses that might contribute to movement and decision-making impairments. Our measurements were restricted to the caudate and putamen – according to each patient’s pre-planned surgical trajectory for the eventual placement of the DBS electrode (which followed after our recordings had been completed along the same guide tube). These regions have been shown to activate in response to real and counterfactual reward prediction errors in previous fMRI studies of this task: caudate (Lohrenz, McCabe et al. 2007), putamen (Chiu, Lohrenz et al. 2008). However, we could not access the ventral striatum with our probe, where reward prediction errors are pronounced (Pagnoni, Zink et al. 2002) and thus cannot rule out a role for 5-HT in modulating this or other regions such as the orbitofrontal cortex (Knutson and Cooper 2005), during this task. Furthermore, our extraction model only accounted for dopamine, serotonin and pH changes, but will not account for systematic voltammetric changes induced by other neuromodulators or metabolites. For example, serotonin’s metabolite, as 5-hydroxyindole acetic acid (5H-1AA) has been observed at levels higher than serotonin in voltammetric recordings in vivo in rodents and with similar oxidation and reduction characteristics. Though we cannot directly rule out the role of a metabolite, its systematic (and speeded) fluctuation in concert with decision variables would still suggest a role for the serotonergic system in reacting to negative outcomes and pose important questions for future experimental and theoretical study.

g. Funding & Disclosure

The authors declare no conflict of interest. This work was supported by a Wellcome Trust Principal Research Fellowship to PRM.

h. Acknowledgments

We would like to thank the patients who took part in this study.

i. References

- Anderson, M. J. (2001). "Permutation tests for univariate or multivariate analysis of variance and regression." Canadian journal of fisheries and aquatic sciences **58**(3): 626-639.
- Bari, A., D. E. Theobald, D. Caprioli, A. C. Mar, A. Aidoo-Micah, J. W. Dalley and T. W. Robbins (2010). "Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats." Neuropsychopharmacology **35**(6): 1290-1301.
- Balasubramani, P.P., Chakravarthy, V.S., Ravindran, B. and Moustafa, A.A. (2014). An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. Frontiers in computational neuroscience, 8.
- Boureau, Y.-L. and P. Dayan (2011). "Opponency revisited: competition and cooperation between dopamine and serotonin." Neuropsychopharmacology **36**(1): 74-97.
- Broderick, P. A. (1988). "Distinguishing in vitro electrochemical signatures for norepinephrine and dopamine." Neuroscience letters **95**(1-3): 275-280.
- Chiu, P. H., T. M. Lohrenz and P. R. Montague (2008). "Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task." Nature neuroscience **11**(4): 514-520.
- Cohen, J. Y., M. W. Amoroso and N. Uchida (2015). "Serotonergic neurons signal reward and punishment on multiple timescales." Elife **4**: e06346.
- Cools, R., A. C. Roberts and T. W. Robbins (2008). "Serotonergic regulation of emotional and behavioural control processes." Trends in cognitive sciences **12**(1): 31-40.
- Crockett, M. J., L. Clark, A. M. Apergis-Schoute, S. Morein-Zamir and T. W. Robbins (2012). "Serotonin modulates the effects of Pavlovian aversive predictions on response vigor." Neuropsychopharmacology **37**(10): 2244-2252.
- Crockett, M. J., L. Clark and T. W. Robbins (2009). "Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans." The Journal of neuroscience **29**(38): 11993-11999.
- Daw, N. D., S. Kakade and P. Dayan (2002). "Opponent interactions between serotonin and dopamine." Neural Networks **15**(4): 603-616.
- Dayan, P. and Q. J. Huys (2008). "Serotonin, inhibition, and negative mood." PLoS Comput Biol **4**(2): e4.
- Dayan, P. and Q. J. Huys (2009). "Serotonin in affective control." Annual review of neuroscience **32**: 95-126.
- Deakin, J. (1983). "Roles of serotonergic systems in escape, avoidance and other behaviours." Theory in psychopharmacology **2**: 149-193.
- Delazer, M., H. Sinz, L. Zamarian, H. Stockner, K. Seppi, G. Wenning, T. Benke and W. Poewe (2009). "Decision making under risk and under ambiguity in Parkinson's disease." Neuropsychologia **47**(8): 1901-1908.
- den Ouden, H. E., N. D. Daw, G. Fernandez, J. A. Elshout, M. Rijpkema, M. Hoogman, B. Franke and R. Cools (2013). "Dissociable effects of dopamine and serotonin on reversal learning." Neuron **80**(4): 1090-1100.

- Diekhof, E. K., P. Falkai and O. Gruber (2008). "Functional neuroimaging of reward processing and decision-making: a review of aberrant motivational and affective processing in addiction and mood disorders." Brain research reviews **59**(1): 164-184.
- Dölen, G., A. Darvishzadeh, K. W. Huang and R. C. Malenka (2013). "Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin." Nature **501**(7466): 179-184.
- Doya, K. (2007) Reinforcement learning: Computational theory and biological mechanisms, HFSP Journal, 1:1, 30-40
- Euteneuer, F., F. Schaefer, R. Stuermer, W. Boucsein, L. Timmermann, M. T. Barbe, G. Ebersbach, J. Otto, J. Kessler and E. Kalbe (2009). "Dissociation of decision-making under ambiguity and decision-making under risk in patients with Parkinson's disease: a neuropsychological and psychophysiological study." Neuropsychologia **47**(13): 2882-2890.
- Fang, H., M. L. Pajski, A. E. Ross and B. J. Venton (2013). "Quantitation of dopamine, serotonin and adenosine content in a tissue punch from a brain slice using capillary electrophoresis with fast-scan cyclic voltammetry detection." Analytical Methods **5**(11): 2704-2711.
- Flagel, S. B., J. J. Clark, T. E. Robinson, L. Mayo, A. Czuj, I. Willuhn, C. A. Akers, S. M. Clinton, P. E. Phillips and H. Akil (2011). "A selective role for dopamine in stimulus-reward learning." Nature **469**(7328): 53-57.
- Fonseca, M. S., M. Murakami and Z. F. Mainen (2015). "Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing." Current Biology **25**(3): 306-315.
- Frank, M. J., L. C. Seeberger and R. C. O'reilly (2004). "By carrot or by stick: cognitive reinforcement learning in parkinsonism." Science **306**(5703): 1940-1943.
- Göthert, M. (2016). "5-HT Receptors Mediating Pre-synaptic Autoinhibition in Central Serotonergic Nerve." Serotonin: 56.
- Graeff, F. and J. Deakin (1991). "5-HT and mechanisms of defence." Journal of Psychopharmacology **5**(4): 305-315.
- Hashemi, P., E. C. Dankoski, K. M. Wood, R. E. Ambrose and R. M. Wightman (2011). "In vivo electrochemical evidence for simultaneous 5 - HT and histamine release in the rat substantia nigra pars reticulata following medial forebrain bundle stimulation." Journal of neurochemistry **118**(5): 749-759.
- Hu, X.-Z., R. H. Lipsky, G. Zhu, L. A. Akhtar, J. Taubman, B. D. Greenberg, K. Xu, P. D. Arnold, M. A. Richter and J. L. Kennedy (2006). "Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder." The American Journal of Human Genetics **78**(5): 815-826.
- Jacobs, B. L. (1994). "Serotonin, motor activity and depression-related disorders." American Scientist **82**(5): 456-463.
- John, C. E. and S. R. Jones (2007). "Fast scan cyclic voltammetry of dopamine and serotonin in mouse brain slices."
- John, C. E. and S. R. Jones (2007). "Voltammetric characterization of the effect of monoamine uptake inhibitors and releasers on dopamine and serotonin uptake in

- mouse caudate-putamen and substantia nigra slices." *Neuropharmacology* **52**(8): 1596-1605.
- Kaye, W., K. Gendall and M. Strober (1998). "Serotonin neuronal function and selective serotonin reuptake inhibitor treatment in anorexia and bulimia nervosa." *Biological psychiatry* **44**(9): 825-838.
- Kishida, K. T., I. Saez, T. Lohrenz, M. R. Witcher, A. W. Laxton, S. B. Tatter, J. P. White, T. L. Ellis, P. E. Phillips and P. R. Montague (2016). "Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward." *Proceedings of the National Academy of Sciences* **113**(1): 200-205.
- Kishida, K. T., S. G. Sandberg, T. Lohrenz, Y. G. Comair, I. Sáez, P. E. Phillips and P. R. Montague (2011). "Sub-second dopamine detection in human striatum." *PloS one* **6**(8): e23291.
- Knutson, B. and J. C. Cooper (2005). "Functional magnetic resonance imaging of reward prediction." *Current opinion in neurology* **18**(4): 411-417.
- Li, Y., W. Zhong, D. Wang, Q. Feng, Z. Liu, J. Zhou, C. Jia, F. Hu, J. Zeng and Q. Guo (2016). "Serotonin neurons in the dorsal raphe nucleus encode reward signals." *Nature communications* **7**.
- Liu, Z., J. Zhou, Y. Li, F. Hu, Y. Lu, M. Ma, Q. Feng, J.-e. Zhang, D. Wang and J. Zeng (2014). "Dorsal raphe neurons signal reward through 5-HT and glutamate." *Neuron* **81**(6): 1360-1374.
- Lohrenz, T., K. McCabe, C. F. Camerer and P. R. Montague (2007). "Neural signature of fictive learning signals in a sequential investment task." *Proceedings of the National Academy of Sciences* **104**(22): 9493-9498.
- Lowry, C. (2002). "Functional subsets of serotonergic neurones: implications for control of the hypothalamic - pituitary - adrenal axis." *Journal of neuroendocrinology* **14**(11): 911-923.
- Mathur, B. N. and D. M. Lovinger (2012). "Serotonergic action on dorsal striatal function." *Parkinsonism & related disorders* **18**: S129-S131.
- Miyazaki, K., K. W. Miyazaki and K. Doya (2012). "The role of serotonin in the regulation of patience and impulsivity." *Molecular neurobiology* **45**(2): 213-224.
- Miyazaki, K. W., K. Miyazaki and K. Doya (2012). "Activation of dorsal raphe serotonin neurons is necessary for waiting for delayed rewards." *The Journal of Neuroscience* **32**(31): 10451-10457.
- Miyazaki, K. W., K. Miyazaki, K. F. Tanaka, A. Yamanaka, A. Takahashi, S. Tabuchi and K. Doya (2014). "Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards." *Current Biology* **24**(17): 2033-2040.
- Montague, P. R., K. T. Kishida, R. J. Moran and T. M. Lohrenz (2016). "An efficiency framework for valence processing systems inspired by soft cross-wiring." *Current Opinion in Behavioral Sciences* **11**: 121-129.
- Montague, P. R., S. M. McClure, P. Baldwin, P. E. Phillips, E. A. Budygin, G. D. Stuber, M. R. Kilpatrick and R. M. Wightman (2004). "Dynamic gain control of dopamine delivery in freely moving animals." *The Journal of Neuroscience* **24**(7): 1754-1759.
- Must, A., A. Juhász, Á. Rimanóczy, Z. Szabó, S. Kéri and Z. Janka (2007). "Major depressive disorder, serotonin transporter, and personality traits: why patients use

- suboptimal decision-making strategies?" Journal of affective disorders **103**(1): 273-276.
- Pagnoni, G., C. F. Zink, P. R. Montague and G. S. Berns (2002). "Activity in human ventral striatum locked to errors of reward prediction." Nature neuroscience **5**(2): 97-98.
- Politis, M., K. Wu, C. Loane, D. J. Brooks, L. Kiferle, F. E. Turkheimer, P. Bain, S. Molloy and P. Piccini (2014). "Serotonergic mechanisms responsible for levodopa-induced dyskinesias in Parkinson's disease patients." The Journal of clinical investigation **124**(3): 1340-1349.
- Portas, C. M., B. Bjorvatn and R. Ursin (2000). "Serotonin and the sleep/wake cycle: special emphasis on microdialysis studies." Progress in neurobiology **60**(1): 13-35.
- Risch, N., R. Herrell, T. Lehner, K.-Y. Liang, L. Eaves, J. Hoh, A. Griem, M. Kovacs, J. Ott and K. R. Merikangas (2009). "Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis." Jama **301**(23): 2462-2471.
- Schultz, W. (2007). "Multiple dopamine functions at different time courses." Annu. Rev. Neurosci. **30**: 259-288.
- Schmidt, R., Leventhal, D.K., Mallet, N., Chen, F., Berke, J. (2013) Canceling actions involves a race between basal ganglia pathways. Nature Neurosci. **16**, 1118–1124.
- Schultz, W., P. Dayan and P. R. Montague (1997). "A neural substrate of prediction and reward." Science **275**(5306): 1593-1599.
- Selvaraju, T. and R. Ramaraj (2003). "Simultaneous determination of ascorbic acid, dopamine and serotonin at poly (phenosafranine) modified electrode." Electrochemistry communications **5**(8): 667-672.
- Shankar, S. S. and B. K. Swamy (2014). "Detection of epinephrine in presence of serotonin and ascorbic acid by TTAB modified carbon paste electrode: a voltammetric study." Int J Electrochem Sci **9**(3): 1321-1339.
- Shiner, T., B. Seymour, K. Wunderlich, C. Hill, K. P. Bhatia, P. Dayan and R. J. Dolan (2012). "Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease." Brain: aw083.
- Tanaka, S. C., K. Shishida, N. Schweighofer, Y. Okamoto, S. Yamawaki and K. Doya (2009). "Serotonin affects association of aversive outcomes to past actions." The Journal of Neuroscience **29**(50): 15669-15674.
- Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso." Journal of the Royal Statistical Society. Series B (Methodological): 267-288.
- Tops, M., S. Russo, M. A. Boksem and D. M. Tucker (2009). "Serotonin: modulator of a drive to withdraw." Brain and cognition **71**(3): 427-436.
- Voon, V., M. A. Irvine, K. Derbyshire, Y. Worbe, I. Lange, S. Abbott, S. Morein-Zamir, R. Dudley, D. Caprioli, N. A. Harrison, J. Wood, J. W. Dalley, E. T. Bullmore, J. E. Grant and T. W. Robbins (2014). "Measuring "waiting" impulsivity in substance addictions and binge eating disorder in a novel analogue of rodent serial reaction time task." Biol Psychiatry **75**(2): 148-155.
- Whittington, C. J., T. Kendall, P. Fonagy, D. Cottrell, A. Cotgrove and E. Boddington (2004). "Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data." The Lancet **363**(9418): 1341-1345.

- Worbe, Y., S. Palminteri, G. Savulich, N. Daw, E. Fernandez-Egea, T. Robbins and V. Voon (2016). "Valence-dependent influence of serotonin depletion on model-based choice strategy." *Molecular psychiatry* **21**(5): 624-629.
- Worbe, Y., G. Savulich, V. Voon, E. Fernandez-Egea and T. W. Robbins (2014). "Serotonin depletion induces 'waiting impulsivity' on the human four-choice serial reaction time task: cross-species translational significance." *Neuropsychopharmacology* **39**(6): 1519-1526.
- Wu, K., J. Fei and S. Hu (2003). "Simultaneous determination of dopamine and serotonin on a glassy carbon electrode coated with a film of carbon nanotubes." *Analytical biochemistry* **318**(1): 100-106.
- Xiao, N., E. Privman and B. J. Venton (2014). "Optogenetic control of serotonin and dopamine release in *Drosophila* larvae." *ACS chemical neuroscience* **5**(8): 666-673.
- Yacubian, J., J. Gläscher, K. Schroeder, T. Sommer, D. F. Braus and C. Büchel (2006). "Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain." *The Journal of Neuroscience* **26**(37): 9530-9537.
- Yao, H., S. Li, Y. Tang, Y. Chen, Y. Chen and X. Lin (2009). "Selective oxidation of serotonin and norepinephrine over eriochrome cyanine R film modified glassy carbon electrode." *Electrochimica Acta* **54**(20): 4607-4612.

j. **Figure Legends**

Figure 1. Serotonin concentration prediction from dual transmitter model

(A) An illustration of voltammograms acquired for varying levels of serotonin concentration (left) and dopamine concentration (right) in the flow cell. We see that low

to high concentration levels produce changes in current magnitude around the oxidation potentials (insets). Concentration is denoted by []. (B) The flow cell predictions are illustrated for serotonin under varying concentrations of dopamine in the mixture. Serotonin was sampled at each level of concentration from 0.1 - 8 μ M in 0.1 μ M increments. For each of these 80 concentrations we computed the serotonin prediction over different levels of dopamine. Not all concentration mixtures for dopamine and serotonin were acquired and acquired mixtures are denoted by the asterisk. Over this test grid we interpolated (using linear triangulation) across the acquired tests samples to produce a 3-dimensional heat map of serotonin predictions. This plot show that serotonin predictions do not vary systematically with increasing dopamine levels. We illustrate one outlying serotonin prediction which was observed but not included in the interpolated plot for visualization purposes. (C) Flow cell predictions for dopamine in the mixture, plotted as a function of increasing dopamine and increasing serotonin as per (B). Again, dopamine predictions using our model do not appear to be systematically affected by the level of serotonin in the sample. (D) We tested 200 out-of-sample voltammograms for each concentration level to quantify the error in generalizability. Illustrated in grey are predictions for different concentration ranges and their 99.99% confidence intervals. Red crosses denote the mean of the (correct) test range. For each range we randomly selected 20 predictions from the full test set.

Figure 2. Dopamine replication in mixture model

We performed an *in vivo* validation by replicating the previous dopamine findings using our new multivariate mixture model. We split trials into low ([0-50%]), medium ([60-80%]) and high ([90-100%]) bets and examined dopamine transients at these different bet levels in response to positive and negative reward prediction errors. As per the findings reported previously for a univariate model (Kishida, Saez et al. 2016) dopamine estimates from the dual transmitter model predict dopamine encoding of prediction errors. Using two-way ANOVAs with factors RPE (negative RPE, positive RPE) and bet levels (high, medium, low) we found a significant interaction at 200 msec ($p = 0.005$), 300 msec ($p = 0.0001$), 400 msec ($p = 0.0016$), 500 msec ($p = 0.0079$) and 600 msec ($p = 0.02$). Post-hoc 2 sample *t test* results for each bet level and time point are illustrated (* $p < 0.05$, *** $p \leq 0.001$). For validation purposes, we report here, those measurements from the original cohort of 17 patients reported in (Kishida, Saez et al. 2016). Data were baseline corrected to zero at 0 msec, and bar graphs depict the mean and s.e.m.

Figure 3. Investment game and distributions of bets and RPEs over trials

(A) In this figure we provide an illustration of the overall task design. To investigate the role of serotonin we used an investment game (Lohrenz, McCabe et al. 2007) where participants were endowed with an initial 100 'points' and were instructed to invest a percentage of this amount for investment into a stock market (historic markets, e.g. the 1929 Wall St. crash). Participants could choose to invest 0 – 100% (color bar) in 10%

increments (blue arrow, $Bet(t)$). On each trial participants submitted their investment (upper panel) and 840 msec later (± 12 msec std.) were shown the market return (middle panel). On the outcome, participants either lost or gained in accordance with their investment. From these market moves we calculated the reward prediction error on that trial. Following this outcome, participants submitted their next investment (blue arrow, $Bet(t+1)$) at their own pace (lower panel). (B) Distribution of investment choices over all participants. (C) Distribution of reward prediction errors, calculated over each market move over all participants.

Figure 4. Serotonin encodes negative reward prediction errors at high bets

(A) Testing across all outcomes and separating according to either a concomitant positive or negative reward prediction error, we found that serotonin fluctuated significantly more positively for negative (black line) compared to positive (cyan line) reward prediction errors. Six 2-sample t tests were performed over temporal bins [100 – 600 msec] comparing concentration levels; significant effects of RPE were observed at 300 and 400 msec ($*p < 0.05$). Here we baseline corrected at -100msec. (B) Two-way analyses of variance of serotonin's transient response at presentation of the outcome or market move were performed for six temporal bins [100-600 msec], with factors reward prediction error polarity; positive and negative and bet level; low (0 – 50%), and high (60 – 100%). These revealed a significant interaction of reward prediction error and bet level at 100 msec ($F = 14.34$, $p = 0.0002$) and 500 msec ($F = 4.89$, $p = 0.027$). Post-hoc 2-sample t tests were performed using permutation testing to assess within bet-range differences in the response to negative compared to positive reward prediction errors. For the high bet range (60-100% invested), serotonin transients were significantly greater for negative compared to positive reward prediction errors at 100 msec; $p = 0.001$, 300 msec; $p = 0.011$, 400 msec; $p = 0.005$ and 500 msec; $p = 0.01$. While for the low bet range (0 – 50% invested) responses were significantly greater for positive compared to negative reward prediction errors at 100 msec; $p = 0.016$. Only the differential response at 100 msec in the high bet case survived FWE-correction $p = 0.005$ ($**p \leq 0.005$, $*p \leq 0.05$, $(**)$ FWE-corrected). We also applied one-sample, 2-sided t tests in order to investigate the effects of RPE and bet size on 5-HT responses as compared to baseline. We find that the difference is driven by significant decreases in 5-HT following positive reward prediction errors at high bets, and to negative reward prediction errors at low bets ($<<p < 0.005$, $< p < 0.05$). Bar graphs depict the mean and s.e.m. Comparisons of transients with an alternate baseline is presented in Supplemental Information figure S3. (C) The area under the curve in B, revealed a significant interaction ($F = 7.13$, $p = 0.0077$) of RPE and bet level with larger (in time and amplitude) negative-going transients for positive reward prediction error responses in the high bet condition. (D) We tested the serotonin response at 100msec and its correlation with the sign and polarity of the RPE. After omitting 65 outliers ($\sim 3\%$ of trials) that may drive the effect (outliers defined as RPEs with an absolute magnitude greater than 3 and z-scores with an absolute magnitude greater than 5) we see a small

but significant correlation for the different bet levels. Serotonin transients are negatively correlated with the RPE for high bets ($R = -0.0714$; $p = 0.0113$) and positively correlated with the RPE for low bets ($R = 0.0653$; $p = 0.0494$). To explore these results more granularly, we examined individual bins. We found that the only significant individual bins were at [20%, 30%], [60%, 70%], and [80%, 90%] with correlation coefficients and p-values of ($R = 0.19$; $p = 0.01$), ($R = -0.08$; $p = 0.06$) and ($R = -0.14$; $p = 0.009$) respectively. This suggests that a putative ‘indifference point’ for counterfactual and actual losses occurs around 40-50%.

Figure 5. Serotonin and active avoidance following negative reward prediction errors

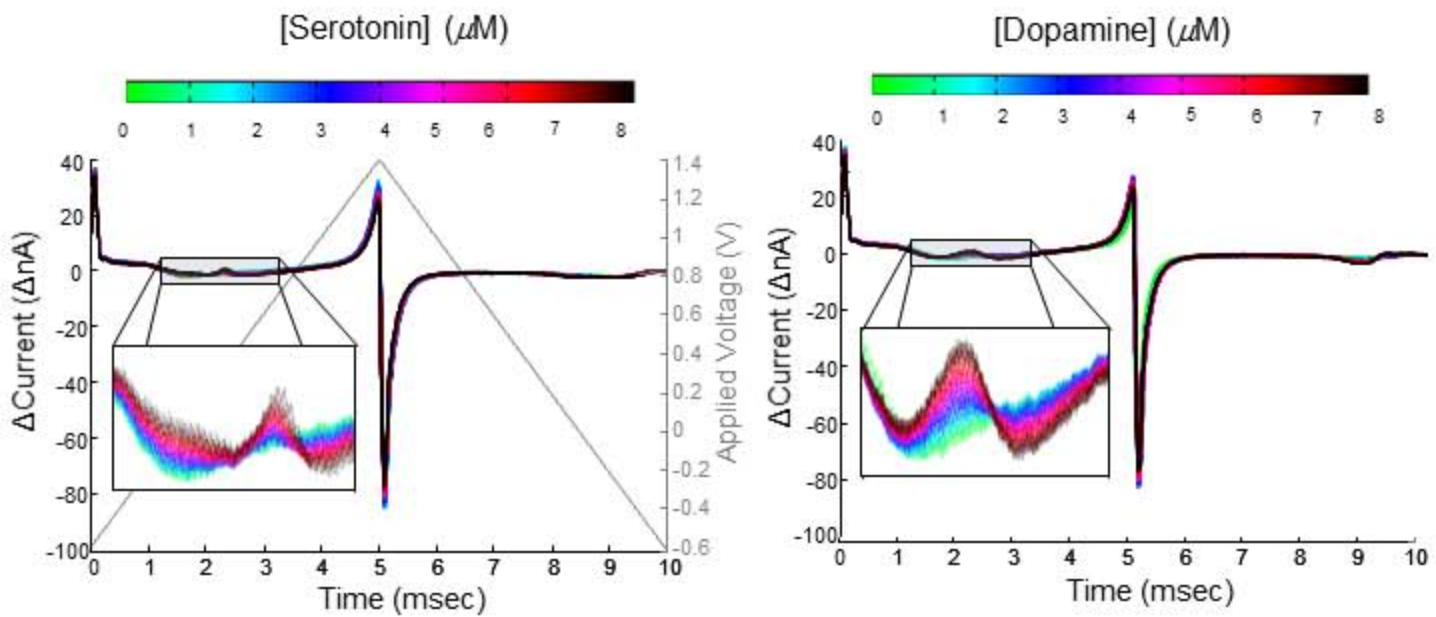
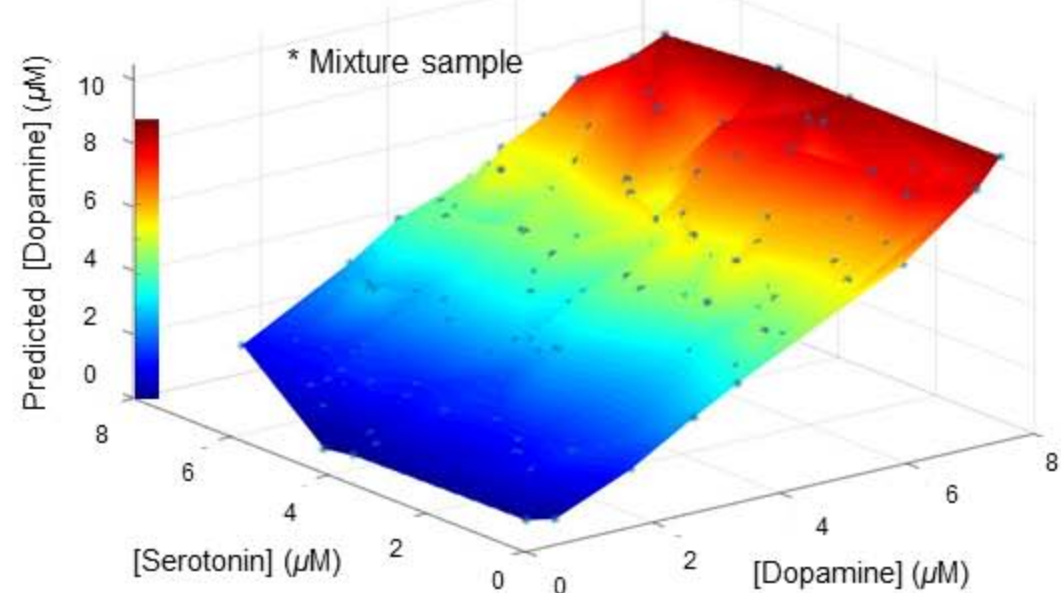
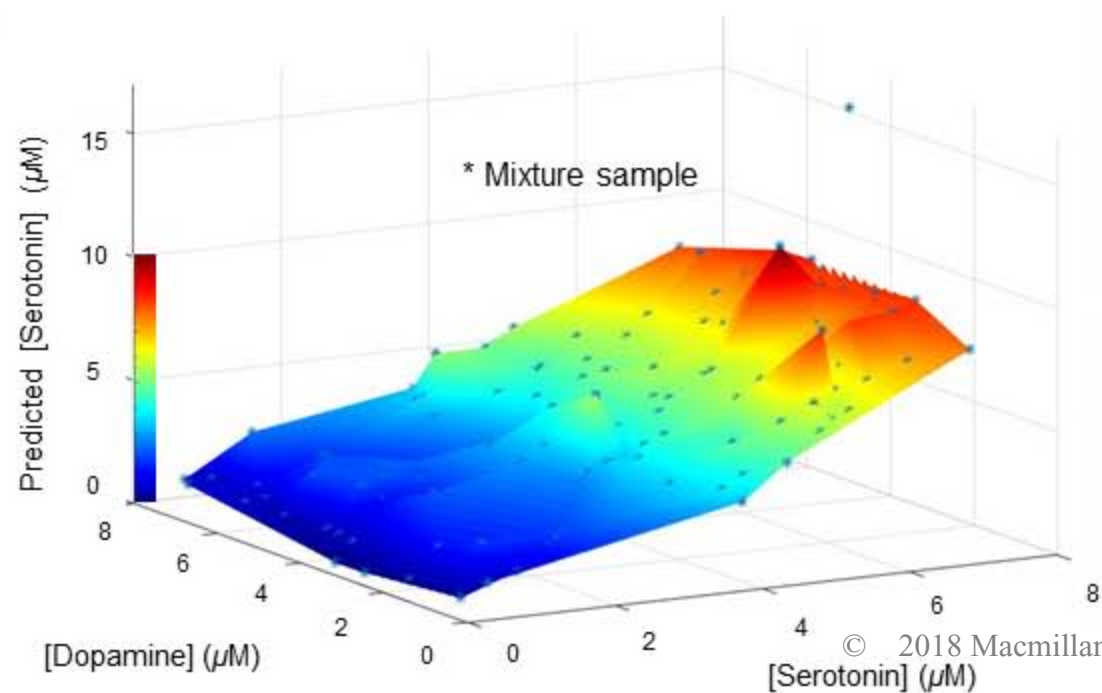
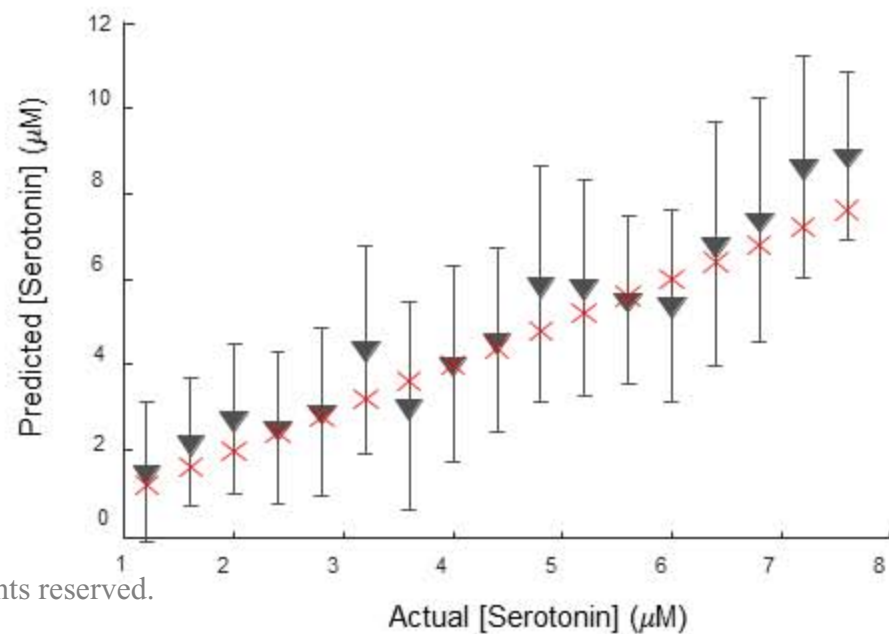
(A) Depiction of ‘next actions’. Responses to trial (t) were analysed for $\text{bet}(t) = \{0\%, [10-20\%], [30-40\%], [50-60\%], [70-80\%] \text{ and } [90-100\%]\}$. For each of these six levels we examined ‘lower bet’ next actions (black arrows) and ‘hold-or-raise’ bet next actions (grey arrows). (B) The $n=842$ negative RPE transients presented in Figure 4A and 4B are represented here but separated according to next-bet decision and current betting level. These results are provided in order to explore the significant negative interaction between serotonin and current bet on predicting change in bet (Supplementary Table 2). Consistent with this negative interaction in the regression analysis, we observe that large positive 5-HT transients at large bets predict a lowering of the bet at trial $t+1$ (black line, right panels). While dips in 5-HT are associated with reducing one’s bet at low bet levels to even lower levels (black line, left panels). The opposite effect is observed for holding or raising one bets, grey lines (with 0% not showing any significant transient effects). Significance here is indicated for uncorrected t tests against zero ($*p < 0.05$, $**p < 0.01$, $***p < 0.005$). (C) Applying a correlation analyses to examine the relationship between serotonin and current bet levels when the next decision is to lower one’s bet. We find that over all ‘lower bet’ decisions, serotonin and the current bet level were positively correlated ($R = 0.3$; $p < 0.00001$). This indicates that market withdrawal is affected by serotonin following poor outcomes. More specifically it indicates that serotonin may prevent further withdrawal (when investment is already low) and promote withdrawal (when investment is high). (D) Correlating decisions to hold or raise bets suggests the opposite effect (but at only trend-level significance).

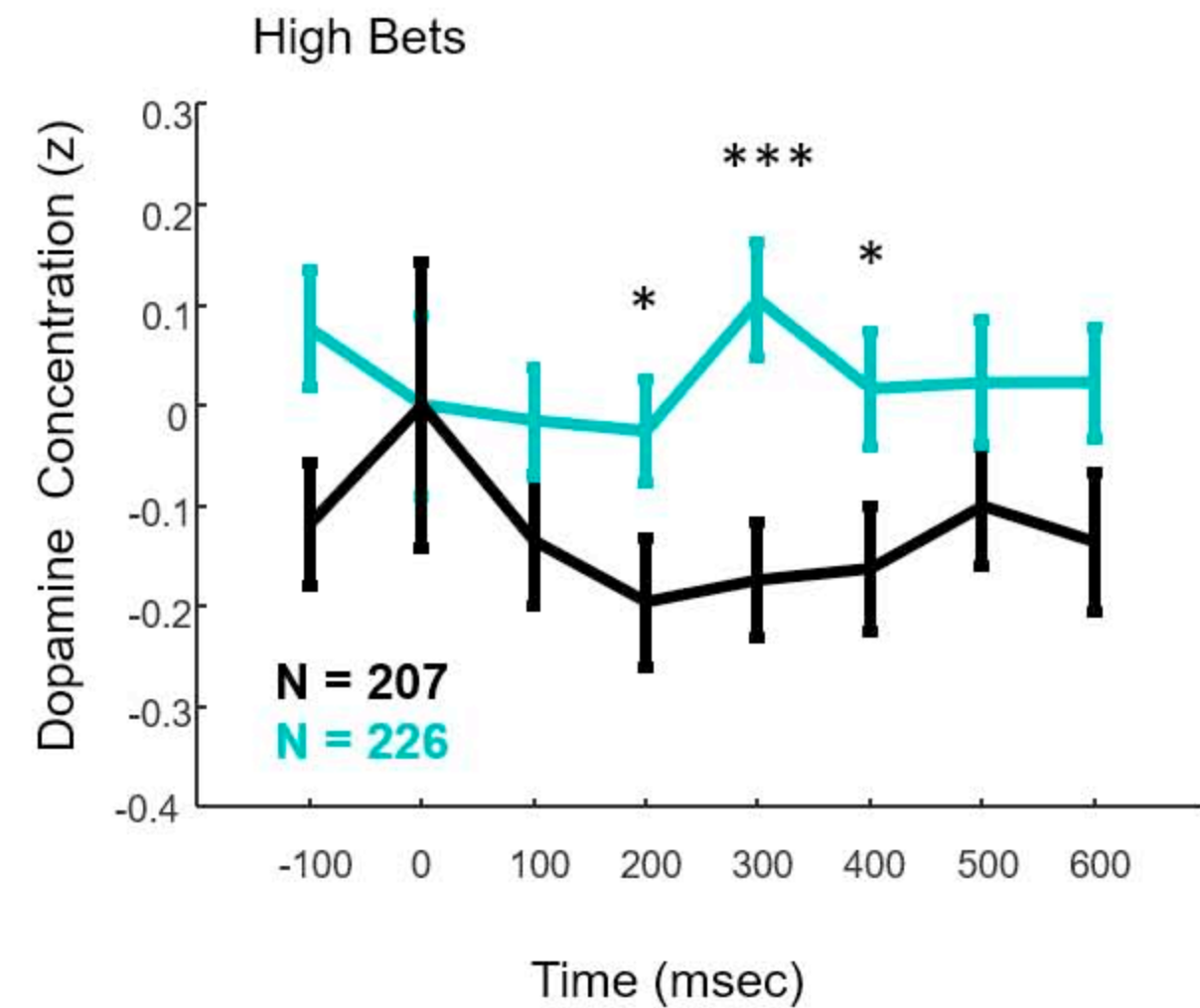
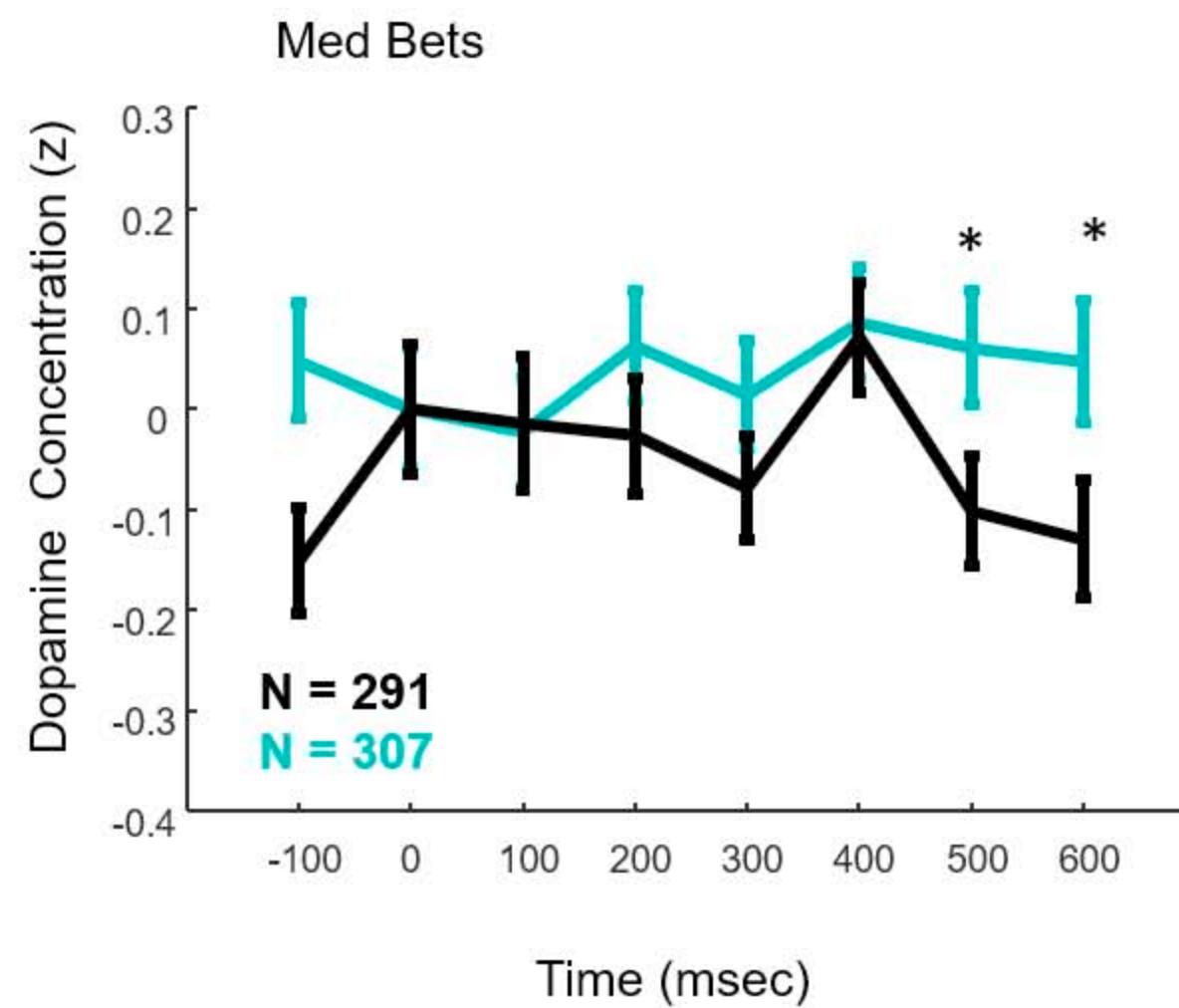
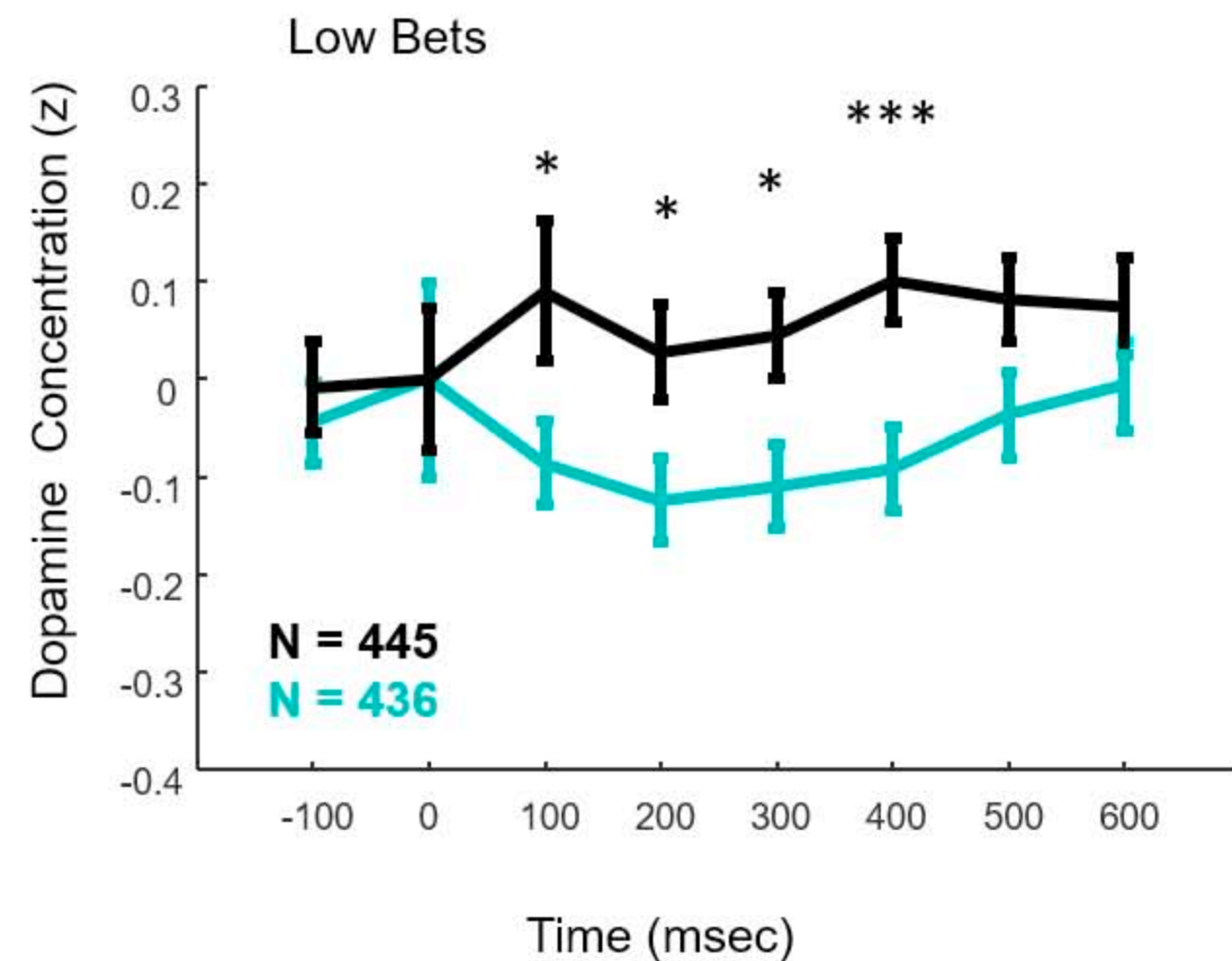
Figure 6. Serotonin and active avoidance following positive reward prediction errors

(A) Depiction of ‘next actions’ as per negative RPE analysis in Figure 5. (B) Serotonin transients ($n=882$) following positive reward prediction errors as per figure 4A and 4B but shown here separated according to decision on next trial and current bet level (lower bet on trial ($t+1$): cyan line, hold or raise bet on trial ($t+1$): dark green line). Only at low bets (where counterfactual losses dominate) did we observe large transients – the direction of the transient was not discriminative however in terms of next-bet decision. (C) Unlike following negative RPES, no parametric effect, in terms of current

bet level and serotonin response, was observed for the decision to lower bet following positive reward prediction errors. (D) Similarly, no parametric effects were observed for serotonin responses preceding a decision to raise or increase bet levels (following a positive reward prediction error).

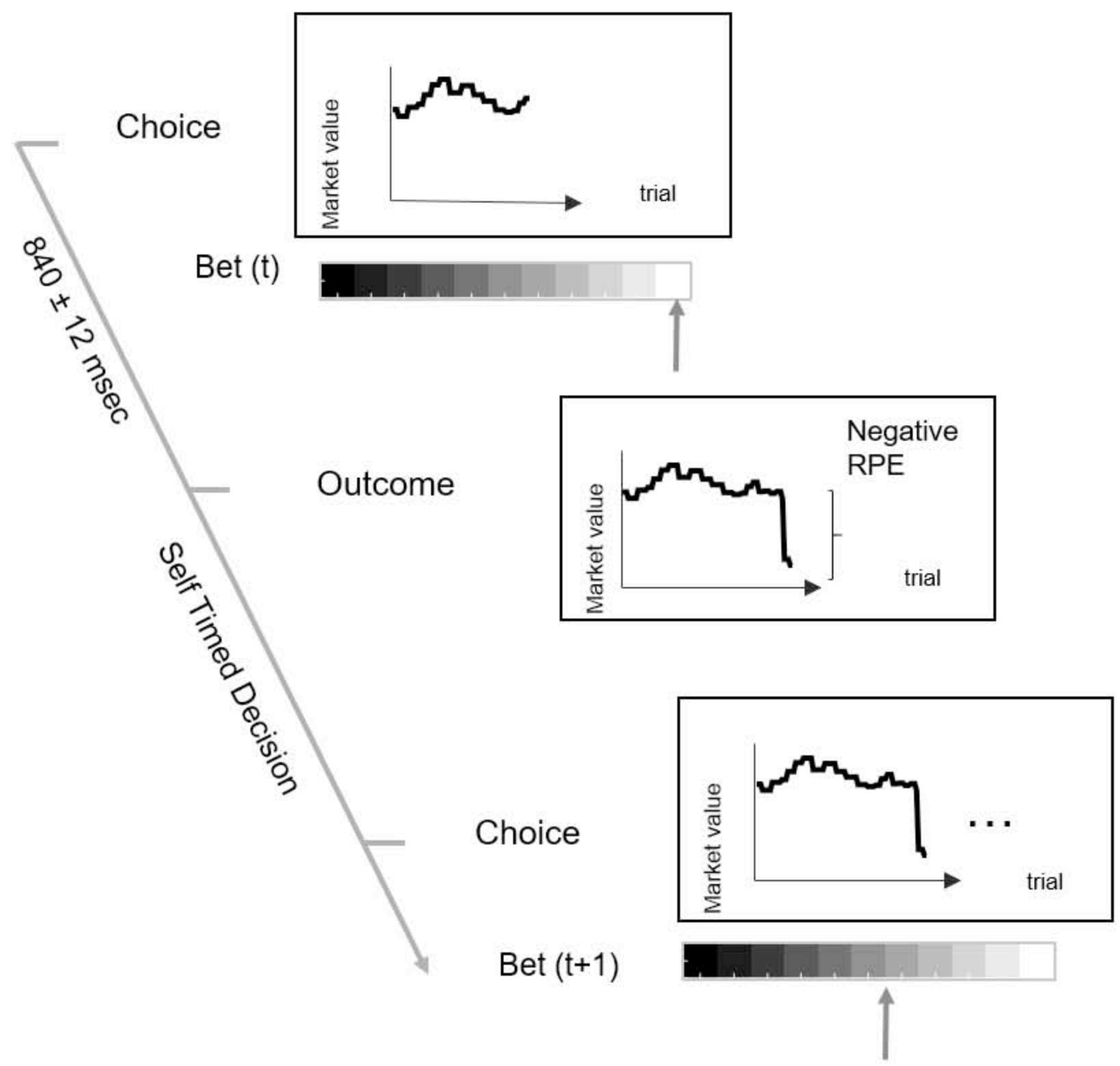
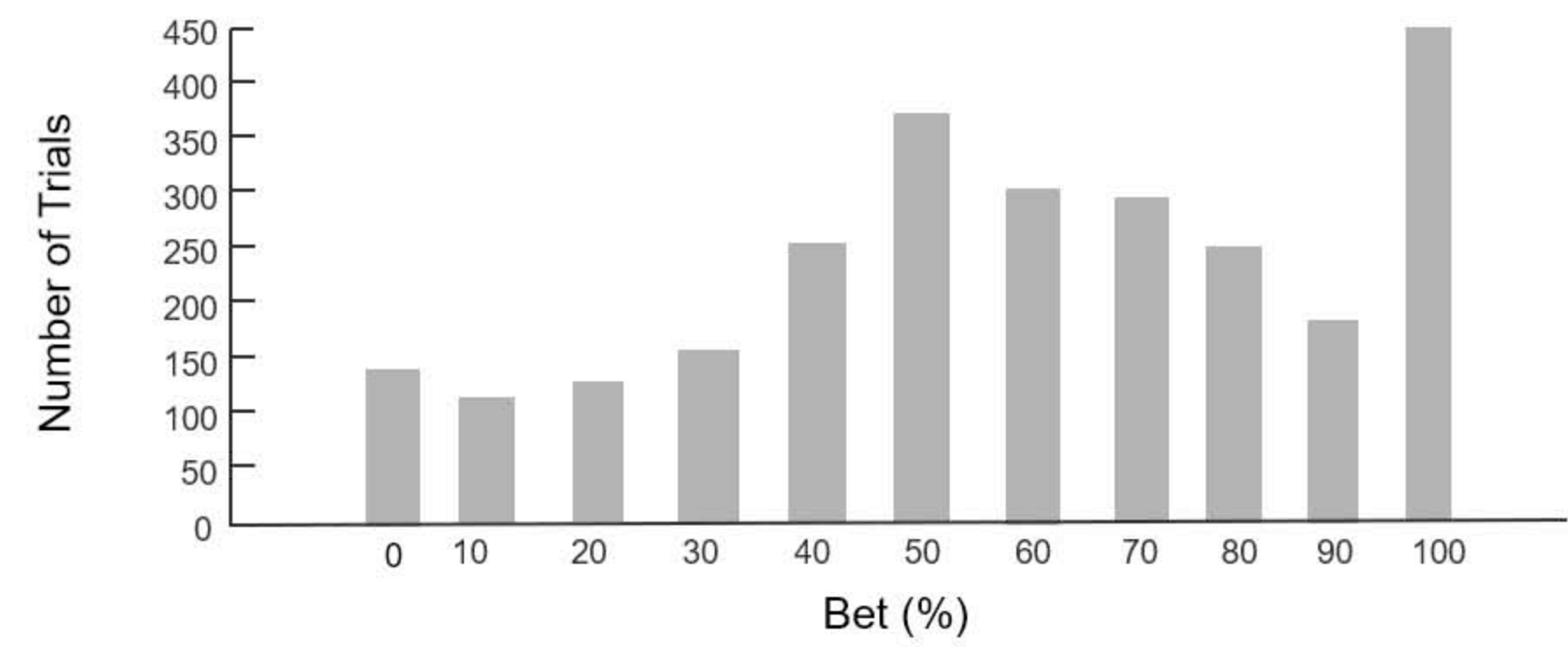
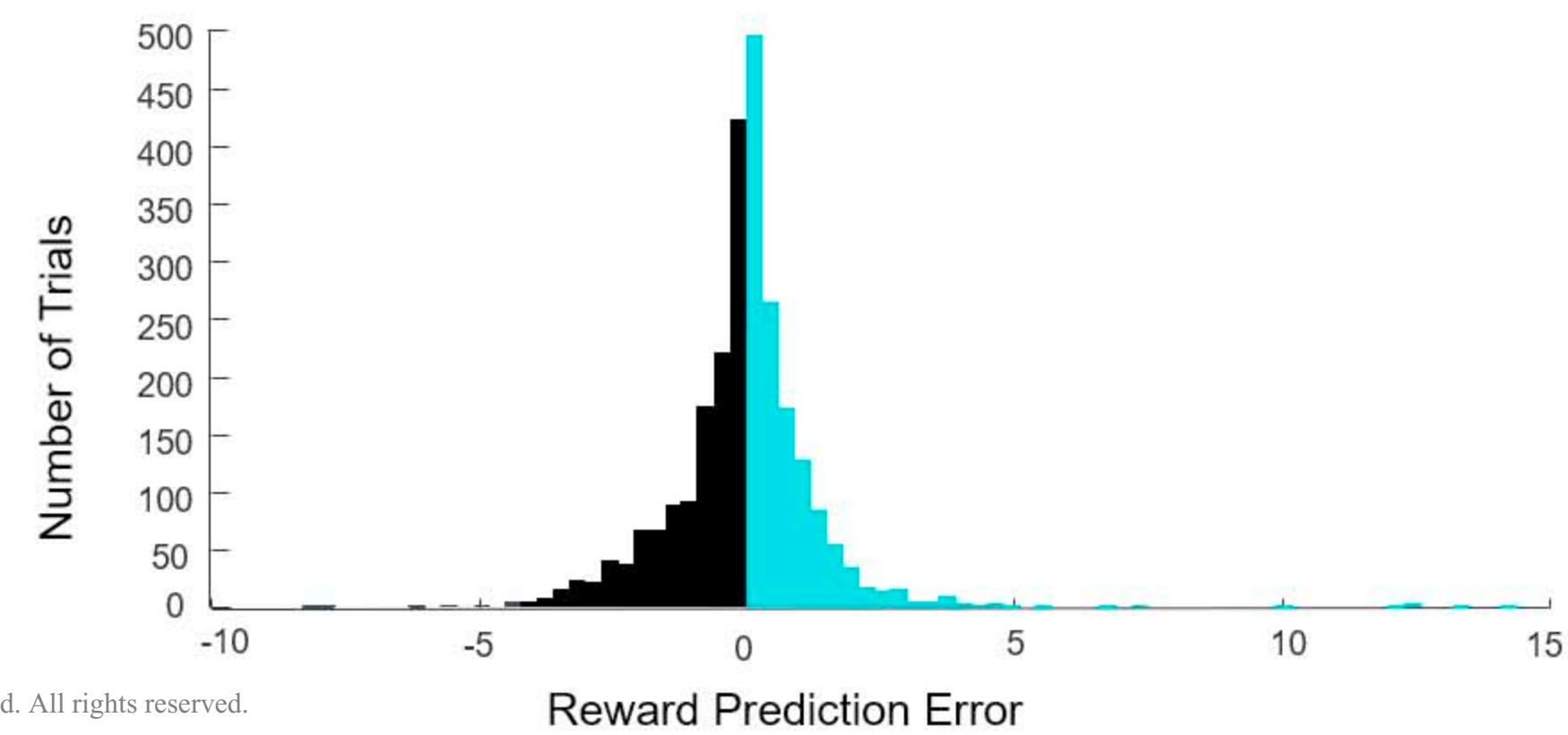
Accepted manuscript

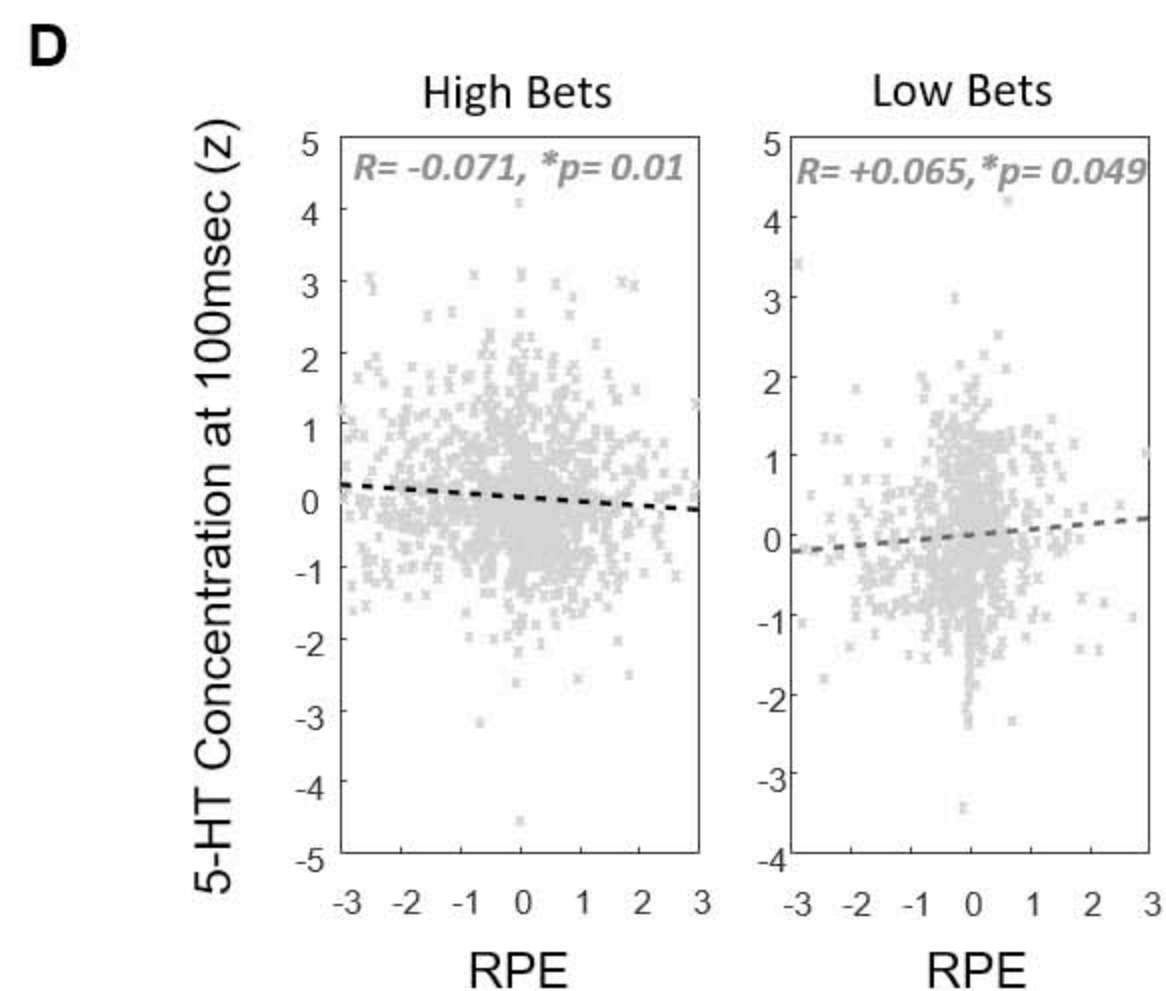
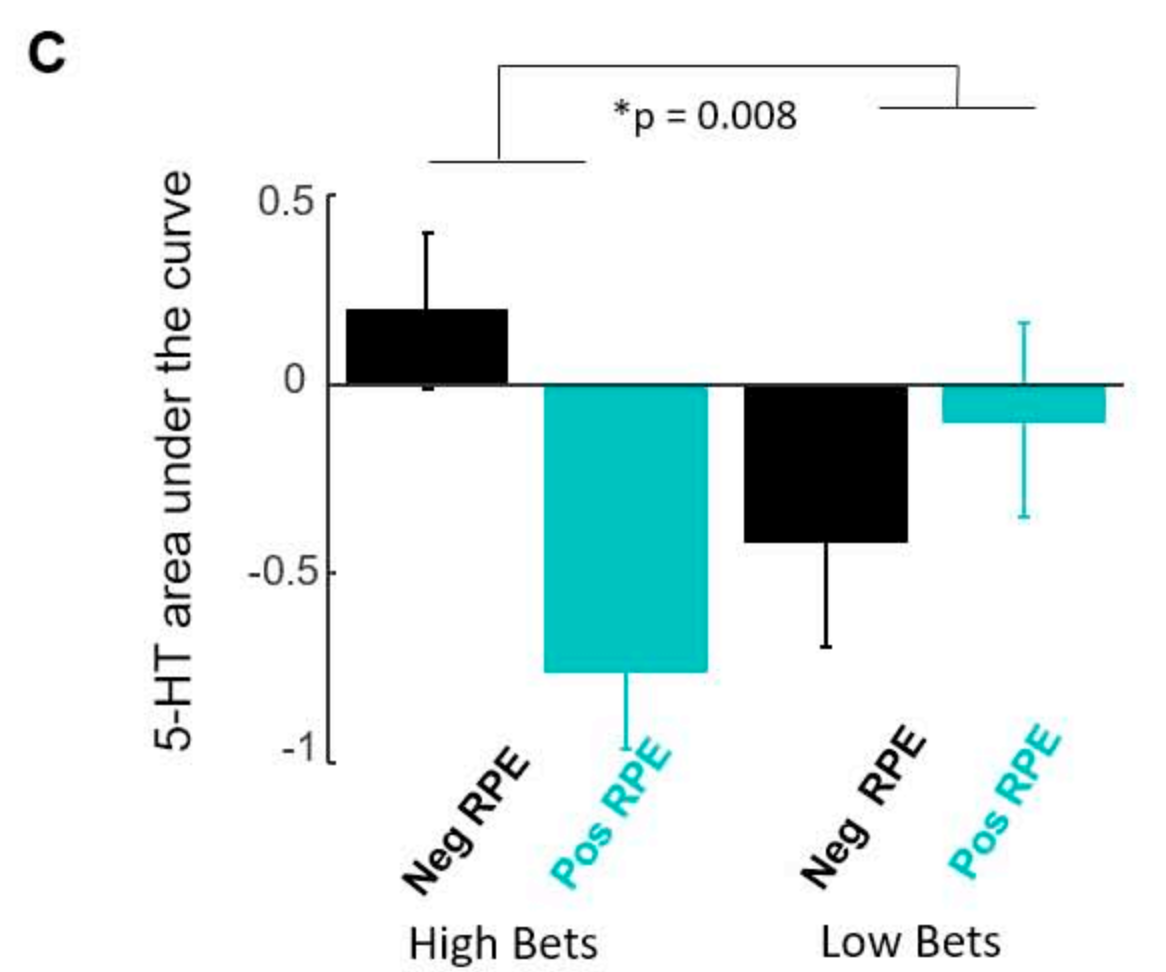
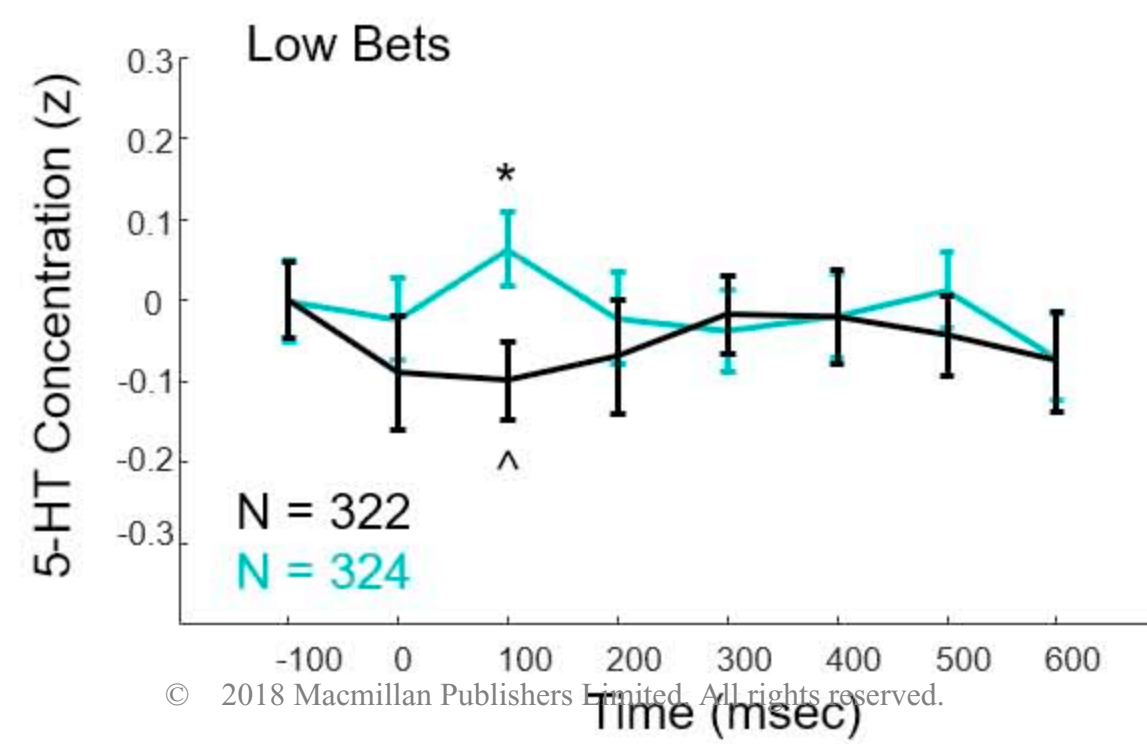
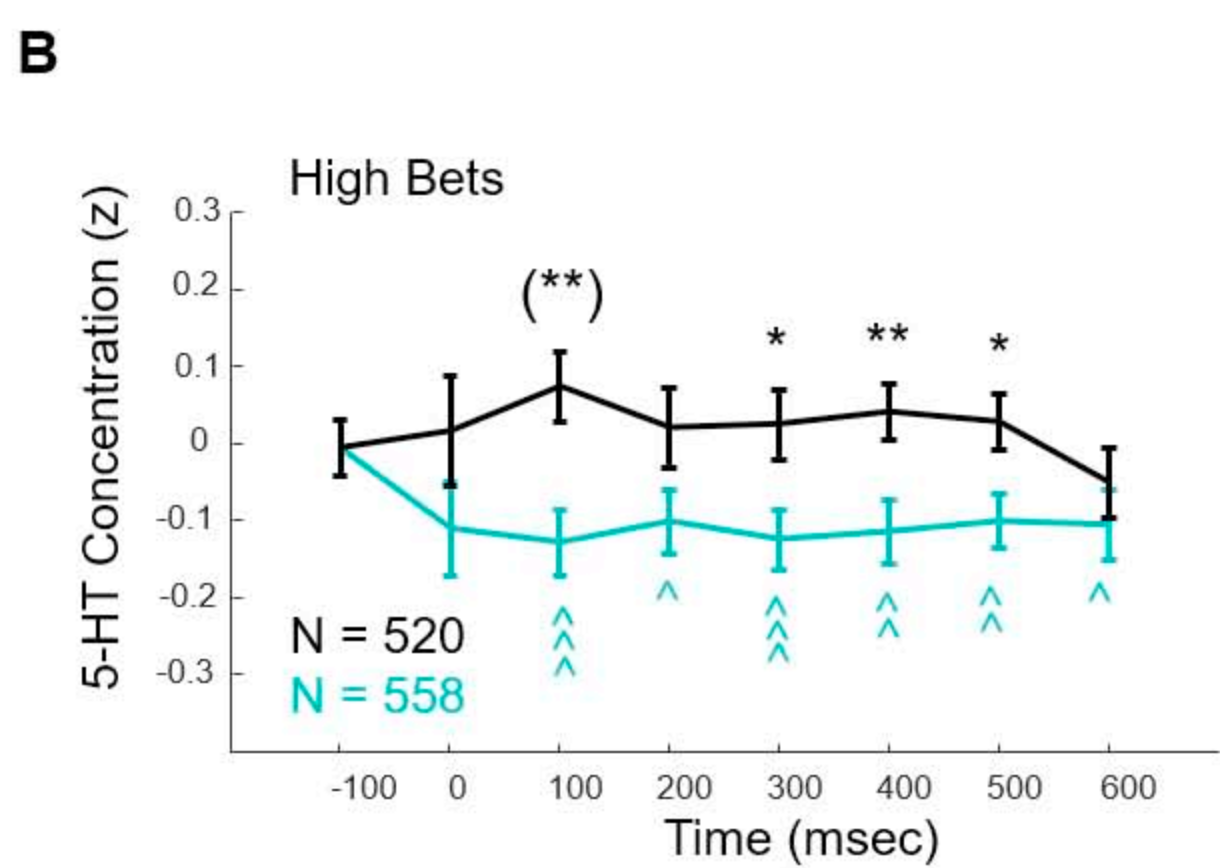
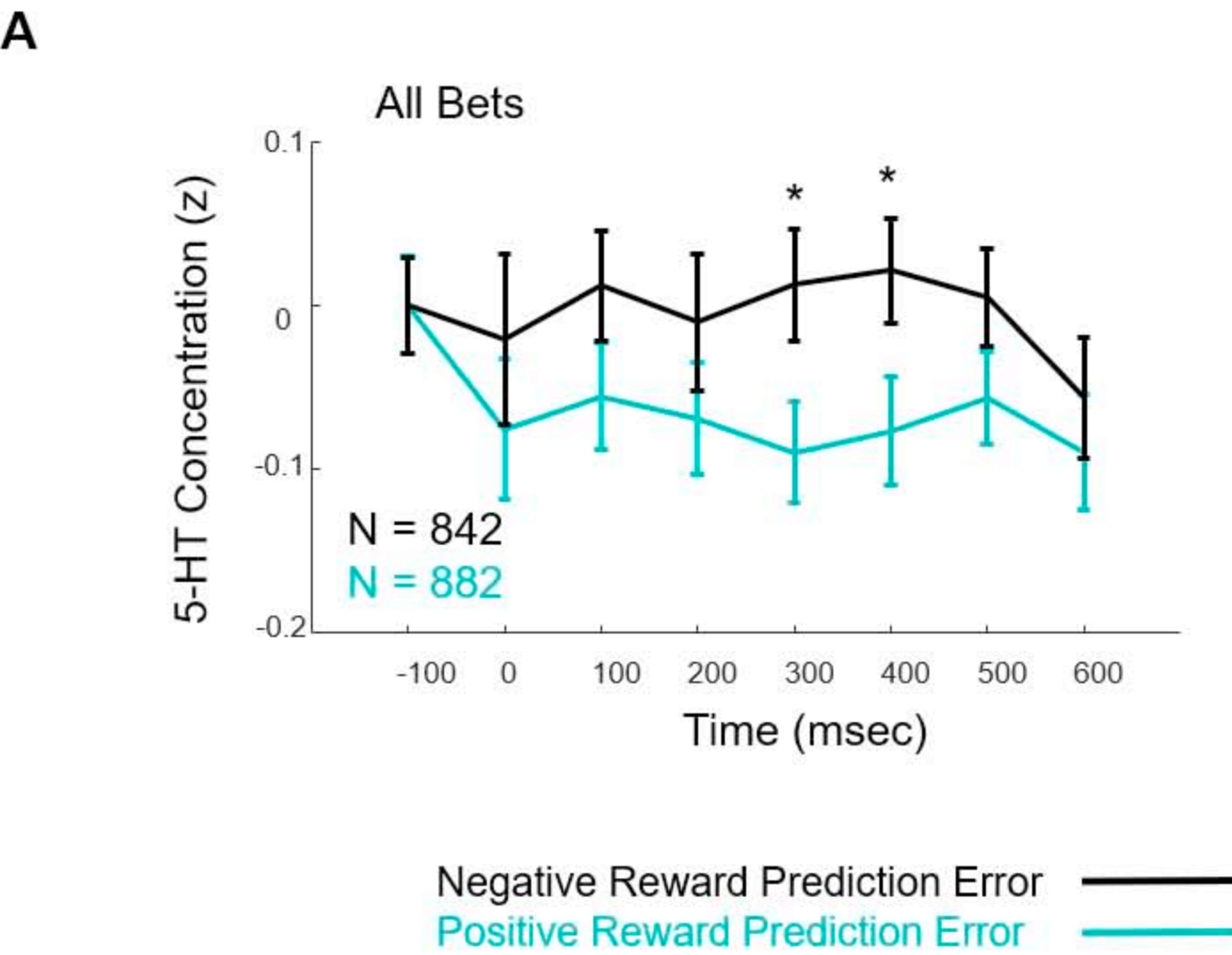
A**C****B****D**

A**B****C**

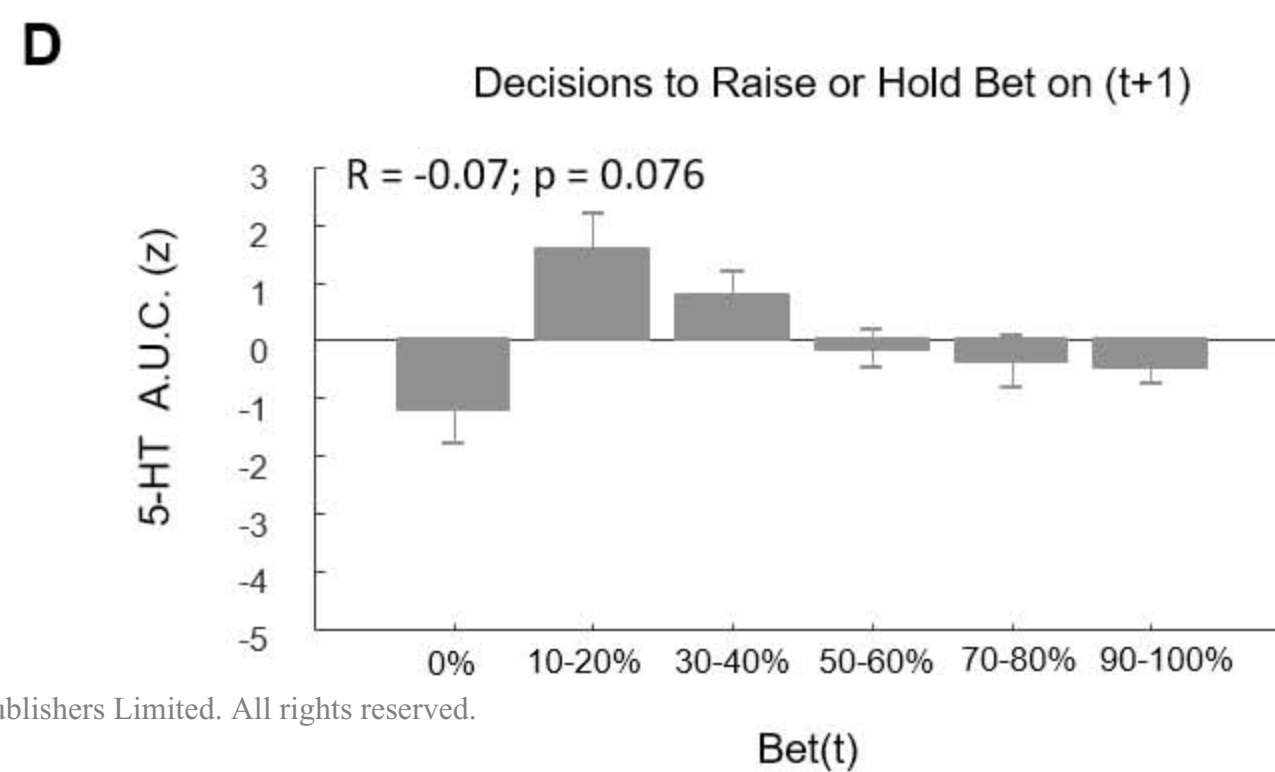
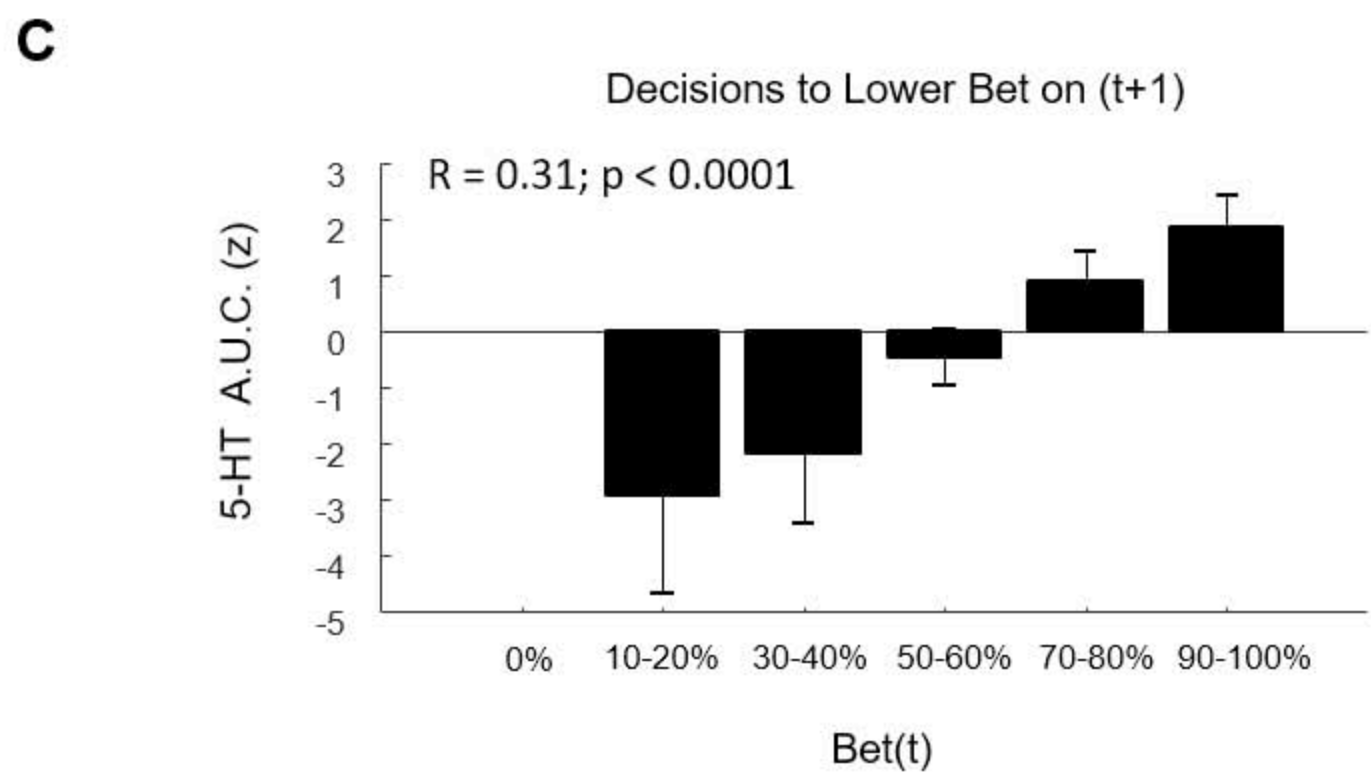
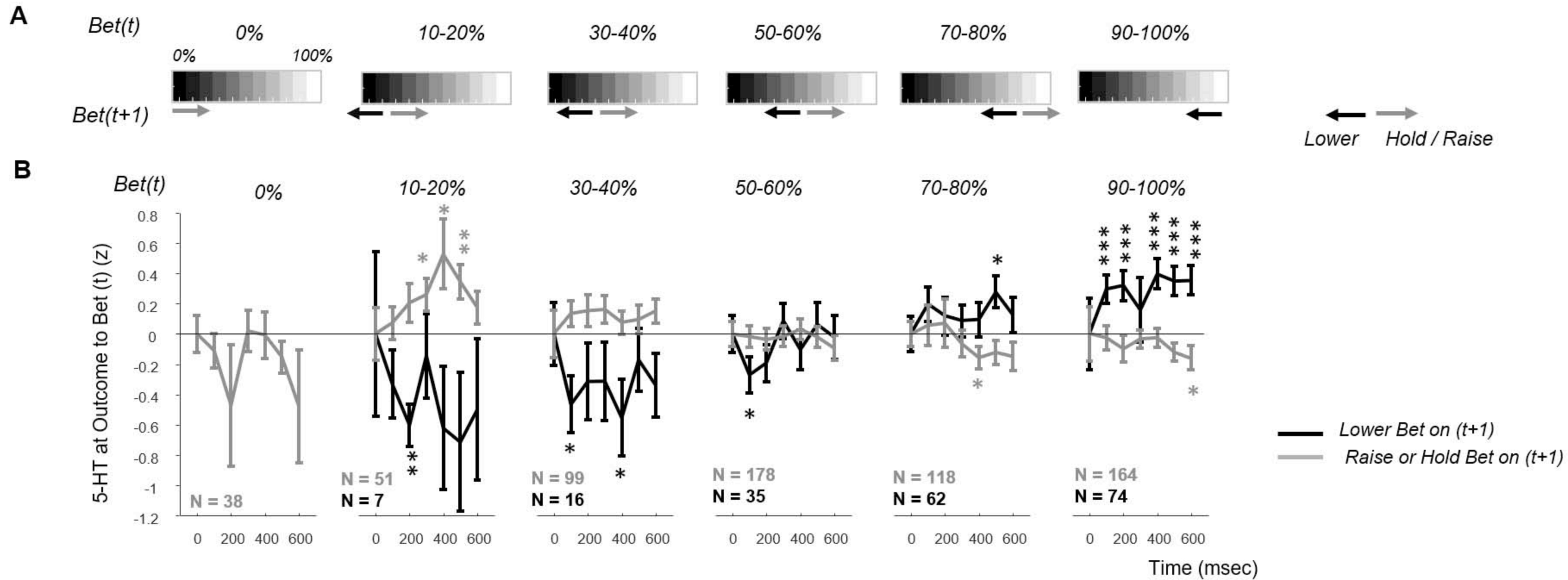
Negative Reward Prediction Error —

Positive Reward Prediction Error —

A**B****C**



Decision-making & 5-HT responses to **Negative** RPEs



Decision-making & 5-HT responses to **Positive** RPEs

