

# Acute Tryptophan Depletion in Healthy Volunteers Enhances Punishment Prediction but Does not Affect Reward Prediction

Roshan Cools<sup>\*1</sup>, Oliver J Robinson<sup>2</sup> and Barbara Sahakian<sup>2</sup>

<sup>1</sup>Department of Experimental Psychology, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK;

<sup>2</sup>Department of Psychiatry, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK

Central serotonin (5-HT) has been implicated in emotional and behavioral control processes for many decades, but its precise contribution is not well understood. We used the acute tryptophan depletion procedure in young healthy volunteers to test the hypothesis that central 5-HT is critical for predicting punishment. An observational reversal-learning task was employed that provided separate measures of punishment and reward prediction. Under baseline, subjects made more prediction errors for punishment-associated stimuli than for reward-associated stimuli. This bias was abolished after central 5-HT depletion, which enhanced the ability to predict punishment while not affecting reward prediction. The selective potentiation of punishment prediction concurs with recent theorizing, suggesting that central 5-HT carries a prediction error for future punishment, but not for future reward (Daw *et al*, 2002). Furthermore, the finding highlights the importance of central 5-HT in resilience to adversity and may have implications for a variety of neuropsychiatric disorders including depression and anxiety.

*Neuropsychopharmacology* (2008) **33**, 2291–2299; doi:10.1038/sj.npp.1301598; published online 17 October 2007

**Keywords:** serotonin; depression; tryptophan; reward; punishment; learning

## INTRODUCTION

Adequate adaptation to our constantly changing environment requires the anticipation of biologically relevant events by learning signals of their occurrence, that is prediction. Models of reinforcement learning use a temporal difference prediction error signal, representing the difference between expected and obtained events, to update their predictions based on states of the environment (Sutton and Barto, 1998). A putative neuronal mechanism of the temporal prediction error signal for future reward is the fast-phasic firing of dopamine cells in the ventral tegmental area (Montague *et al*, 1996; Schultz *et al*, 1997). According to this proposal, positive prediction error of reward, that is unexpected reward, produces a burst in the firing of dopamine neurons, whereas negative prediction error of reward, that is unexpected omission of reward, produces a pause in the firing of dopamine neurons.

Recently, Daw *et al* (2002) have highlighted that existing models of reward prediction cannot easily account for the prediction of future punishment. They have extended the proposal that dopamine subserves the reward prediction

error by suggesting a way in which central serotonin (5-HT), released by the dorsal raphe nucleus, could act as a motivational opponent to dopamine in prediction learning. According to this theoretical model, the phasic release of 5-HT mirrors the phasic release of dopamine, and reports a prediction error for future punishment.

The hypothesis that 5-HT is involved in the prediction of aversive signals is plausible, but not yet proven. There is abundant empirical evidence implicating 5-HT in controlling aversion and potentiating anxiety-induced avoidance (Tye *et al*, 1977; Gray, 1982; Graeff *et al*, 1996). The hypothesis concurs with the observation that serotonergic neurotransmission is implicated in a range of mood and anxiety disorders (Young *et al*, 1985; Anderson *et al*, 1990; Deakin *et al*, 1990; Blier and de Montigny, 1999) that are characterized by enhanced anticipation of, and sensitivity to threat-related stimuli, punishment, and negative feedback (Beats *et al*, 1996; Elliott *et al*, 1997; Mathews and Mackintosh, 2000; Steffens *et al*, 2001; Richards *et al*, 2002; Murphy *et al*, 2003). Studies with healthy human volunteers have demonstrated potentiated processing of punishment-related signals after dietary depletion of the 5-HT precursor tryptophan (TRP; acute tryptophan depletion, ATD), particularly in vulnerable individuals. For example, ATD enhanced the amygdala response to fearful faces (Cools *et al*, 2005b; Van der Veen *et al*, 2007), decreased the impact of positively valenced words (Murphy *et al*, 2002), and increased the impact of negatively valenced emotional words in Stroop-like tasks (Evers *et al*, 2006a). In addition,

\*Correspondence: Dr R Cools, Department of Experimental Psychology, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Downing street, Cambridge CB2 3EB, UK, Tel: +44 1223 333587, Fax: +44 1223 333564, E-mail: roshan.cools@gmail.com

Received 6 August 2007; revised 6 September 2007; accepted 7 September 2007

ATD potentiated neural activity during negative feedback in a probabilistic reversal-learning task (Evers *et al*, 2005), which was also sensitive to acute administration of the selective serotonin reuptake inhibitor citalopram (Chamberlain *et al*, 2006). Finally, a processing bias in favor of aversive signals is seen in healthy individuals who carry one or two copies of the short allele of the 5-HT transporter polymorphism, which is associated with reduced expression of the 5-HT transporter (Hariri *et al*, 2002; Heinz *et al*, 2005; Pezawas *et al*, 2005) and possibly reduced 5-HT function (Bethea *et al*, 2004).

However, despite the existence of abundant evidence for a role of 5-HT in aversive processing, there is currently no direct experimental evidence supporting the specific hypothesis that central 5-HT mediates the prediction of future punishment, but not that of future reward. Here, we test this hypothesis by investigating the effects of ATD, a well-known procedure to reduce central nervous system 5-HT (Nishizawa *et al*, 1997; Carpenter *et al*, 1998), on performance of an observational learning paradigm that allowed the independent assessment of reward and punishment prediction. Reward- and punishment-prediction trials were matched in terms of response inhibition demands, and learning demands were maximized by repeatedly reversing contingencies at unpredictable intervals.

## METHODS

### Subjects

Procedures were approved by the Norfolk Research Ethical Committee (06/Q0101/5) and were in accord with the Helsinki Declaration of 1975.

Twelve subjects participated in this study. They were screened for psychiatric and neurological disorders, gave written informed consent, and were compensated for participation. Exclusion criteria were any history of cardiac, hepatic, renal, pulmonary, neurological, psychiatric or gastrointestinal disorders, medication/drug use, and personal or family history of major depression or bipolar affective disorder. Following the screening interview, subjects were assigned in a double-blind approximately counterbalanced fashion to the 'first-TRP-' ( $n = 5$ ) or the 'first-BAL' group ( $n = 7$ ) (mean age (years) 22.4, SD = 4.0; four males).

The experimental paradigm of interest in the current paper was administered as part of a larger study (data to be published separately by OJR and BJS). One subject vomited immediately after consuming the drink and her data were excluded from analysis. Three subjects failed to comply with the task instructions as revealed by error rates at chance or worse than chance levels in one or more blocks (eg one subject made zero correct responses in one block with a mean reaction time of 305 ms). One subject did not return for a second visit and three subjects encountered technical difficulties with the task.

### General Procedure

Subjects were assessed on a neuropsychological battery on two test sessions, separated by at least 1 week. Volunteers were asked to abstain from alcohol, caffeine, and food from

midnight before each session. During the test days, they followed a low-protein diet. In the morning of a test day (between 0830 and 1030 hours), volunteers arrived at the research center, where a blood sample was taken, and a nutritionally balanced (BAL) or a TRP-free (TRP-) amino-acid drink was ingested. Testing started after a resting period of approximately 5.5 h to ensure stable and low TRP- levels. After a second blood sample, the task was completed.

### Amino-Acid Mixtures

Central TRP was depleted by ingesting an amino-acid load that did not contain TRP but did include other large neutral amino acids (LNAA) (Reilly *et al*, 1997). The quantities of amino acids in each drink were based on those used by Young *et al* (1985), though a 75 g mixture was employed to minimize nausea. Amino-acid mixtures (prepared by SHS international, Liverpool, UK) were as follows:

BAL: L-alanine, 4.1 g; L-arginine, 3.7 g; L-cystine, 2.0 g; glycine, 2.4 g; L-histidine, 2.4 g; L-isoleucine, 6 g; L-leucine, 10.1 g; L-lysine, 6.7 g; L-methionine, 2.3 g; L-proline, 9.2 g; L-phenylalanine, 4.3 g; L-serine, 5.2 g; L-threonine, 4.9 g; L-tyrosine, 5.2 g; L-valine, 6.7 g; and L-tryptophan, 3.0 g—total 78.2 g.

TRP-: L-alanine, 4.1 g; L-arginine, 3.7 g; L-cystine, 2.0 g; glycine, 2.4 g; L-histidine, 2.4 g; L-isoleucine, 6 g; L-leucine, 10.1 g; L-lysine, 6.7 g; L-methionine, 2.3 g; L-proline, 9.2 g; L-phenylalanine, 4.3 g; L-serine, 5.2 g; L-threonine, 4.9 g; L-tyrosine, 5.2 g; and L-valine, 6.7 g—total 75.2 g.

For female participants, the same ratios of amino acids were used, but with a 20% reduction in quantity to take into account lower body weight. The drinks were prepared by stirring the mixture into approximately 200 ml tap water. Subjects were given the choice of adding either lemon-lime or grapefruit flavoring to compensate for the unpleasant taste. They reported no side effects apart from transient nausea following ingestion of the drink.

### Self-Report Measurements

The Positive and Negative Affect Scale (PANAS; Watson *et al*, 1988) was administered on nine occasions during the test day (with the first measure administered before the drink). We analyzed the difference in positive and negative affect scores obtained from the PANAS between the following two time points: (1) immediately before drink ingestion and (2) immediately before test administration. Analysis of these difference scores revealed that ATD did not significantly affect time-related changes in positive affect ( $F_{1,11} = 2.1$ ,  $P = 0.2$ ) or negative affect ( $F_{1,11} = 0.08$ ,  $P = 0.8$ ).

Subjects also completed a number of questionnaires: the behavioral inhibition system/behavioral activation system (BIS/BAS) scales (Carver and White, 1994), the Eysenck Personality Questionnaire (EPQ), the Impulsiveness Venturesomeness Empathy questionnaire (IVE-7; Eysenck and Eysenck, 1978), the Beck Depression Inventory (BDI; Beck *et al*, 1961), and the Barratt Impulsiveness Scale (BIS; Patton *et al*, 1995). Scores are reported in Table 1.

## Task Design

**General description.** The paradigm was previously described by Cools *et al* (2006) and the reader is referred to that manuscript for additional details (Table 2).

Subjects were presented with a series of two stimuli. The two stimuli were the same throughout the experiment. At any one point in time, one of the stimuli was associated with reward, while the other was associated with punishment. On each trial, one of the two stimuli was highlighted and subjects had to predict, based on trial and error learning,

**Table 1** Demographic and Trait Characteristics

Questionnaire	Range	Mean (SD)
BIS	13–22	19 (2.3)
BAS	29–48	37.6 (5.6)
EPQ—psychoticism	1–7	3.6 (2.4)
EPQ—extraversion	3–13	9.0 (2.7)
EPQ—neuroticism	0–9	3.6 (3.0)
EPQ—lie	1–8	3.8 (2.5)
IVE—impulsiveness	2–13	8.4 (3.5)
IVE—venturesomeness	3–15	10.3 (3.7)
IVE—empathy	7–16	12.7 (2.5)
BDI	2–10	4.8 (2.4)
Barratt impulsiveness scale	47–78	65.7 (10.4)

Abbreviations: BAS, behavioral activation system score; BDI, Beck Depression Inventory; BIS, behavioral inhibition system score from the BIS/BAS scale; EPQ, Eysenck Personality Questionnaire; IVE, Impulsiveness Venturesomeness Empathy questionnaire.

**Table 2** Schematic of Sample Trial Sequences from the Two Conditions

Condition	Unexpected punishment							
Highlighted stimulus	A	A	B	A	A	B	B	A
Correct response	P	P	R	P	P	R	P	R
Outcome	P	P	R	P	P	P	<b>P</b>	R
Switch/non-switch	ns	ns	ns	ns	ns	ns	sw	ns

Condition	Unexpected reward							
Highlighted stimulus	A	A	B	A	A	B	B	A
Correct response	R	R	P	R	R	P	R	P
Outcome	R	R	P	R	R	<b>R</b>	R	P
Switch/non-switch	ns	ns	ns	ns	ns	ns	sw	ns

Abbreviations: ns, non-switch trial; sw, switch trials.

The four top rows represent a sample trial sequence from the unexpected punishment condition. The four bottom rows represent a sample trial sequence from the unexpected reward condition. Subjects were presented with two stimuli (A and B). One of these two stimuli was highlighted (highlighted stimulus: A or B). Subjects pressed a red or green button depending on whether they predicted reward (R) or punishment (P) (correct response: R or P). Following the response, the outcome was presented (outcome: R or P). The stimulus-outcome contingencies were deterministic and reversed after a certain number of consecutively correct responses. These reversals were signaled by unexpected outcomes (highlighted in boxes), following which subjects needed to update their predictions on switch trials (sw).

whether the highlighted stimulus would lead to reward or punishment. The outcome was presented after subjects made their prediction. Outcomes were not response contingent, but depended on which stimulus was highlighted. Thus, the outcome did not provide performance feedback. To minimize confusion regarding the task instructions, we provided performance feedback in an indirect fashion, by highlighting the same stimulus after error trials. This procedure was identical for punishment- and reward-prediction trials and allowed us to track whether subjects adhered to the task instructions. During the task, the stimulus-outcome contingencies reversed multiple times provided attainment of learning criteria.

**Trial details.** On each trial subjects were presented two vertically adjacent stimuli, one scene and one face (location randomized), at about 19-inch viewing distance (subtending about 3° horizontally and 3.5° vertically). One of the stimuli was highlighted with a black border surrounding the stimulus. Subjects indicated their predictions by pressing, with the index or middle finger, one of two colored buttons (corresponding to keys 'b' and 'n' depending on the response-outcome mapping) on a laptop keyboard. They pressed the green button for reward and the red button for punishment. The outcome-response mappings were counter-balanced between subjects. The (self-paced) response was followed by an interval of 1000 ms, after which the outcome was presented for 500 ms. Reward consisted of a green smiley face, a '+ \$100' sign and a high-frequency jingle tone. Punishment consisted of a red sad face, a '−\$100' sign and a single low-frequency tone. After the outcome, the screen was cleared for 500 ms, after which the next two stimuli were presented.

**Task procedure.** Each subject performed one practice block and four experimental blocks. Each practice block consisted of one acquisition stage and one reversal stage. Each experimental block consisted of one acquisition stage and a variable number of reversal stages. The task proceeded from one stage to the next following a number of consecutive correct trials, as determined by a preset learning criterion. Learning criteria (that is the number of consecutive correct trials following which the contingencies changed) varied between stages (mean = 6.9, SD = 1.8, range from 5 to 9), to prevent predictability of reversals. The maximum number of reversal stages per experimental block was 16, although the block terminated automatically after completion of 120 trials (~6.6 min), so that each subject performed 480 trials (four blocks) per experimental session.

The task consisted of two conditions (two blocks per condition). A schematic of sample trial sequences for each condition is shown in Table 2. In the unexpected reward condition, reversals were signaled by unexpected reward. Specifically, on reversal trials of the unexpected reward condition, the previously punished stimulus was highlighted and followed unexpectedly by reward. In the unexpected punishment condition, reversals were signaled by unexpected punishment. Thus, on reversal trials of this condition, the previously rewarded stimulus was highlighted and followed unexpectedly by punishment. The order of conditions was counterbalanced between

groups (six subjects received the unexpected punishment condition first).

The stimulus that was highlighted on the first trial of each reversal stage (on which the unexpected outcome was presented) was always highlighted again on the second trial of that stage (ie the switch trial on which the subject had to implement the reversed contingencies and switch their predictions) (Table 2). For example, if the previously rewarded stimulus A was highlighted on the first trial of a reversal stage and followed by unexpected punishment, then stimulus A was highlighted again on the second trial of that reversal stage.

## Data Analysis

**Biochemical measures.** Blood (venous) samples (10 ml) were taken immediately before ingestion of the amino-acid drink and after the testing session, approximately 5.5 h after administration, to determine the level of total and free TRP in plasma, and the TRP/ $\sum$ LNAA ratio. This ratio was calculated from the serum concentrations of total TRP divided by the sum of the LNAAs (tyrosin, phenylalanine, valine, isoleucine, and leucine) and is important, because the uptake of TRP in the brain is strongly associated with the amounts of other competing LNAAs (TRP and the other LNAA share the blood-brain barrier). Venous samples were taken in lithium heparin tubes and stored at  $-20^{\circ}\text{C}$ . Plasma TRP concentrations were determined by an isocratic high-performance liquid chromatography (HPLC) method of analysis. Plasma proteins were removed by precipitation with 3% trichloroacetic acid and centrifugation at 3000 revolutions,  $4^{\circ}\text{C}$  for 10 min, and then pipetted into heparin aliquots. An aliquot was diluted in mobile phase before injection onto the HPLC analytical column. Fluorescence end-point detection was used to identify TRP.

**Behavioral measures.** Data were analyzed in three steps. First, we assessed the effects of ATD on the mean number of errors on the task as a whole, regardless of outcome type (that is regardless of whether subjects predicted reward or punishment). Errors were square-root transformed to stabilize variances and decrease skewness ( $\sqrt{x}$ ; as is usual when data are in the form of counts (Howell, 1997, p327)) and submitted to an ANOVA with condition (which contrasted the unexpected punishment condition with the unexpected reward condition) and drink (TRP- vs BAL) as within-subject factors.

Second, the data were decomposed according to outcome type. This trial-by-trial analysis included only those trials that followed correct responses. (We excluded trials following errors, because errors on these error +1 trials probably reflected a failure to maintain the task instructions. This assumption was based on the fact that the same stimulus was highlighted on trials following errors and likely provided no significant cognitive challenge on non-switch trials.) Errors were transformed into proportional scores, given that the number of data points varied per trial type as a function of performance. Mean proportions of errors were arcsine transformed ( $2 \times \arcsine(\sqrt{x})$ ; as is appropriate when the variance is proportional to the mean (Howell, 1997; p328)) and analyzed using

repeated-measures ANOVAs (SPSS 11, Chicago, IL) with condition, drink, and outcome type as within-subject factors.

Finally, we separately analyzed trials after unexpected outcomes that required a behavioral switch (switch trials) and trials after expected outcomes that did not require such switching (non-switch trials) (Table 2). This analysis allowed us to assess whether ATD differentially affected switching as a function of the valence of the unexpected outcomes. For these analyses we excluded trials from the first acquisition stage of each block, which did not differ between conditions.

We report two-tailed *P*-values. Greenhouse–Geisser corrections were applied when the sphericity assumption was violated (Howell, 1997). The data in the figures represent raw data.

## RESULTS

### Biochemical Measures

Repeated-measures ANOVA revealed significant two-way interactions of drink by time of blood test, due to significant reductions in total TRP levels ( $F_{1,11} = 94.3$ ,  $P < 0.0001$ ), free TRP levels ( $F_{1,11} = 28.2$ ,  $P < 0.0001$ ), and the critical ratio TRP/ $\sum$ LNAA measure ( $F_{1,11} = 64.6$ ,  $P < 0.0001$ ) approximately 5.5 h after TRP- relative to BAL (Table 3).

Analyses of simple effects for the critical ratio data revealed a significant main effect of time for the TRP-drink ( $T_{11} = 12.7$ ,  $P < 0.0001$ ), but not for the BAL drink ( $T_{11} = -1.2$ ,  $P = 0.25$ ). Thus, the ratio of TRP/ $\sum$ LNAA was significantly reduced after the TRP- drink, but remained unaltered after the BAL drink.

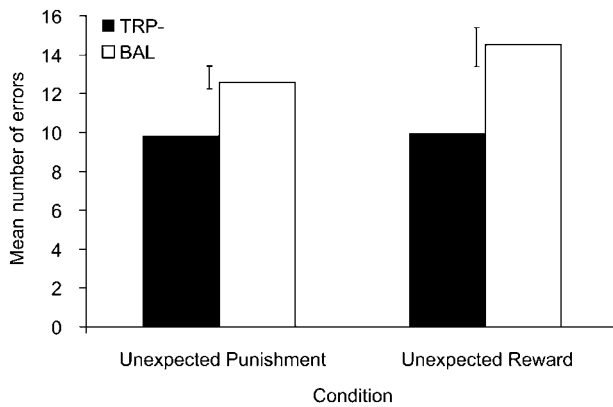
### Behavioral Data: Block Analysis

In Figure 1 we present the total number of errors made on the task as a whole as a function of condition and drink. Subjects made on average 5% errors on the task (chance = 50%). Significantly fewer errors were made after the TRP- drink than after the BAL drink (main effect of drink:  $F_{1,11} = 9.5$ ,  $P = 0.01$ ). The effect of ATD did not differ between the unexpected punishment and unexpected reward condition (drink by condition interaction:  $F_{1,11} = 0.1$ ,  $P = 0.7$ ; main effect of condition:  $F_{1,11} = 0.5$ ,  $P = 0.5$ ). Separate paired-sample *t*-tests confirmed that ATD improved performance in both the unexpected reward

**Table 3** Biochemical Measures as a Function of Time of Test and Drink

	TRP- $T_0$	TRP- $T_{5.5}$	BAL- $T_0$	BAL- $T_{5.5}$
Ratio TRP/ $\sum$ LNAA	0.15 (0.01)	0.01 (0.001)	0.15 (0.01)	0.17 (0.02)
Total TRP	42.4 (2.0)	8.6 (0.7)	42.1 (3.4)	95.1 (9.9)
Free TRP	2.4 (0.3)	0.5 (0.04)	3.4 (0.9)	7.9 (1.7)

Abbreviations: TRP, tryptophan-depleting drink; BAL, balancing drink;  $T_0$ , test at baseline;  $T_{5.5}$ , test approximately 5.5 h after drink. Values represent means (SEM).



**Figure 1** The mean number of errors as a function of condition and drink. Error bars represent standard errors of the difference as a function of drink.

( $T_{11} = 2.2$ ,  $P = 0.05$ ) and the unexpected punishment condition ( $T_{11} = 2.6$ ,  $P = 0.03$ ).

In keeping with the reduced number of errors, subjects completed more stages within the maximum of 120 trials after the TRP– drink than after the BAL drink and this main effect of drink on the number of completed stages across both conditions was marginally significant ( $F_{1,11} = 3.8$ ,  $P = 0.08$ ).

### Behavioral Data: Trial-by-Trial Analysis

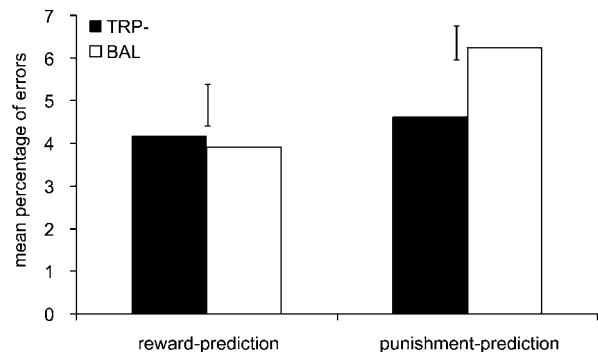
We assessed whether the improvement depended on outcome type by comparing reward- and punishment-prediction trials using a second ANOVA. This analysis revealed that the effect of ATD was not equally distributed between punishment- and reward-prediction trials (drink by outcome interaction:  $F_{1,11} = 5.6$ ,  $P = 0.04$ ). In keeping with our hypothesis, the effect of ATD was restricted to punishment-prediction trials and did not extend to reward-prediction trials. Simple effects analyses confirmed that subjects made significantly fewer punishment-prediction errors after the TRP– drink than after the BAL drink ( $F_{1,11} = 8.3$ ,  $P = 0.015$ ), whereas there was no drink effect on reward-prediction errors ( $F_{1,11} = 0.1$ ,  $P = 0.7$ ). There was also a significant interaction between condition and outcome ( $F_{1,11} = 10.2$ ,  $P = 0.009$ ), due to more punishment- than reward-prediction errors in the unexpected punishment condition, but more reward- than punishment-prediction errors in the unexpected reward condition. However, as mentioned above, the condition factor did not interact with drink.

The next set of analyses assessed switch and non-switch trials separately. An ANOVA on switch trials with condition and drink as within-subject factors revealed that there was no main effect of drink on switch trials (Table 4;  $F_{1,11} = 0.2$ ,  $P = 0.7$ ), nor a drink by condition interaction ( $F_{1,11} = 0.005$ ,  $P = 0.9$ ). Note that on switch trials (ie trials immediately after unexpected outcomes), the condition factor overlaps with the outcome factor, because unexpected punishment was always followed by a punishment-prediction trial and unexpected reward was always followed by a reward-prediction trial. Thus, there was no punishment prediction improvement on switch trials.

**Table 4** The Mean Percentage of Errors on Switch Trials as a Function of Condition and Drink

	Unexpected punishment	Unexpected reward
TRP	7.2 (2.8)	10.0 (3.1)
BAL	7.8 (3.0)	9.3 (2.3)

Abbreviations: TRP, tryptophan-depleting drink; BAL, balancing drink. Values represent means (SEM).



**Figure 2** The mean percentage of errors on non-switch trials as a function of drink and outcome trial type. Error bars represent standard errors of the difference as a function of drink.

Conversely, an ANOVA on non-switch trials with condition, drink, and outcome as within-subject factors confirmed again a significant drink by outcome interaction ( $F_{1,11} = 10.6$ ,  $P = 0.008$ ; Figure 2; Table 5). Simple effects analyses revealed that subjects made significantly fewer errors on punishment-prediction trials after the TRP– drink than after the BAL drink ( $F_{1,11} = 7.6$ ,  $P = 0.02$ ), while there was no drink effect on reward-prediction trials ( $F_{1,11} = 0.07$ ,  $P = 0.4$ ). Further simple effects analyses of these non-switch trials revealed that subjects made significantly more errors on punishment- than reward-prediction trials after the BAL drink ( $F_{1,11} = 4.8$ ,  $P = 0.05$ ). By contrast, there was no difference between punishment- and reward-prediction errors after the TRP– drink ( $F_{1,11} = 0.003$ ,  $P = 0.96$ ). Thus, ATD abolished a disproportionate difficulty with punishment prediction, but did not affect reward prediction. This effect was restricted to non-switch trials, and did not extend to switch trials.

The order of drink administration could not account for the data, as revealed by additional analyses of non-switch trials, evidencing that the significant interaction between drink and outcome remained significant when drink order was inserted as a between-subject variable ( $F_{1,10} = 10.9$ ,  $P = 0.008$ ). In addition, there was no evidence for an interaction between drink, outcome, and drink order ( $F_{1,10} = 0.07$ ,  $P = 0.4$ ).

In summary, ATD improved performance by abolishing a disproportionate difficulty with punishment prediction relative to reward prediction. The effect of ATD was present only on non-switch trials, when subjects found punishment prediction more difficult than reward prediction after the BAL drink. There was no effect of ATD on switch trials. These findings indicate that ATD increased the prediction

**Table 5** The Mean Percentage of Errors on Non-switch Trials as a Function of Condition, Drink, and Outcome

	Unexpected punishment		Unexpected reward	
	Reward prediction	Punishment prediction	Reward prediction	Punishment prediction
TRP	4.8 (1.6)	4.3 (1.2)	3.6 (0.8)	3.5 (0.6)
BAL	4.1 (1.1)	5.9 (1.0)	5.2 (1.7)	6.6 (1.6)

Abbreviations: TRP, tryptophan-depleting drink; BAL, balancing drink. Values represent means (SEM).

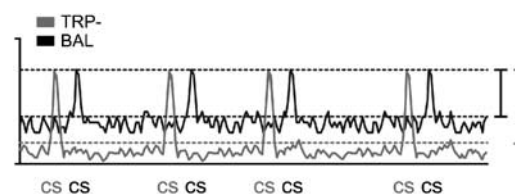
of punishment, but left unchanged the prediction of reward. Furthermore, ATD did not affect the ability to flexibly alter responding based on unexpected outcomes.

## DISCUSSION

The observation that ATD increased punishment prediction concurs with classic and more recent findings indicating that 5-HT controls the processing of aversive signals (Tye *et al*, 1977; Iversen, 1984; Deakin and Graeff, 1991; Daw *et al*, 2002; Cools *et al*, 2005b; Pezawas *et al*, 2005; Harmer *et al*, 2006). More specifically, the selective effect of ATD on punishment prediction is consistent with a recent theoretical model, suggesting that in prediction learning 5-HT acts as a motivational opponent to dopamine, which is commonly implicated in the prediction of future reward (Daw *et al*, 2002).

In this model, learning to predict punishment depends on a transfer, with learning, of a high-amplitude transient-phasic 5-HT response from an aversive stimulus to a conditioned stimulus that predicts it. We demonstrate that a modest reduction in 'background' levels of tonic 5-HT increased the ability to predict punishment. One possibility is that the depletion of tonic 5-HT increased the dynamic range and thus the impact of changes in phasic 5-HT, thus shifting the system from a tonic mode of neurotransmission to a phasic mode of neurotransmission, effectively reducing the signal to noise ratio (Figure 3). Similar antagonistic interactions between phasic and tonic neurotransmission have been proposed for dopamine, where tonic levels regulate the phasic dopamine responses to biologically relevant stimuli (Grace, 1991). Although definitive confirmation of the pharmacological mechanism underlying our selective effect requires electrophysiological recording from serotonergic neurons and voltammetric data during punishment and reward prediction, our findings provide the first direct evidence in support of the hypothesis that 5-HT is critical for the prediction of punishment.

As with most neurochemical manipulations available for human research, we cannot fully exclude the possibility that manipulation of TRP levels did not also affect levels of dopamine, due to known interactions between 5-HT and dopamine (Millan *et al*, 1998). However, it should be noted that direct manipulation of dopamine by withdrawal of the dopamine precursor L-DOPA and dopamine receptor agonists had diametrically opposite effects on this same

**Figure 3** Schematic representation of the hypothetical effect of ATD (in red) on phasic 5-HT neuronal activity. ATD is hypothesized to increase the dynamic range of the phasic 5-HT burst. Time series for the balanced (BAL) and the TRP-free (TRP-) condition are shifted in time to facilitate visualization. The bars on the right represent the height of the phasic burst. CS, conditioned stimulus predictive of punishment.

learning paradigm from those reported here. Specifically, we demonstrated that withdrawal of dopaminergic medication in patients with mild Parkinson's disease selectively improved the ability to switch predictions based on unexpected punishment, while not affecting the ability to predict punishment (or reward) on non-switch trials (Cools *et al*, 2006). Thus, the effects of ATD dissociated from the effects of withdrawal of dopaminergic drugs, likely reflecting neurochemically specific effects of central 5-HT and dopamine, respectively.

In temporal difference models, the prediction error for future punishment is largest when events are unexpected. At first sight, one may thus argue that the effect of ATD should be most pronounced following unexpected punishment. In fact, the improvement was not present on such switch trials, but only surfaced on non-switch trials. This finding may be reconciled with the above-described model, by assuming that the prediction error due to unexpected punishment was too large and too robust to be sensitive to the small reduction in central 5-HT.

An important implication of the lack of effect on switch trials is that ATD did not modulate attention to punishment *per se*. Thus, the matched performance following unexpected punishment indicates that regardless of drink subjects attended equally to the unexpected punishment and were equally able to implement the changed contingency on the next trial.

An alternative account of our effect on punishment prediction is that it does not reflect a modulation of learning *per se*, but rather that of the memory of specific stimulus-punishment contingencies. These two alternative learning and memory hypotheses can be disentangled in future study by assessing the effect of ATD on slow learning curves after reversals of more difficult (eg probabilistic) contingencies than those presented in the present paradigm (where learning curves reached asymptote on the second trial following reversal).

The increased tendency to learn and/or memorize stimulus-punishment contingencies was not a result of a nonspecific, generalized increase in punishment anticipation, because subjects did not predict punishment more often for reward-associated stimuli. This finding indicates that our effect reflects enhanced learning and/or memory of specific stimulus-punishment contingencies and concurs with results from studies with experimental animals indicating an important role for 5-HT in fear conditioning and fear memory (Inoue *et al*, 1993, 1996, 2004; Wilkinson *et al*, 1995; Burghardt *et al*, 2004).

After the BAL drink, subjects found punishment prediction significantly more difficult than reward prediction. It is unlikely that this reflects an effect of the BAL drink for two reasons. First, the BAL drink did not affect the critical ratio of TRP/ $\sum$ LNA. Second, we observed similar disproportionate difficulty with punishment prediction in elderly volunteers who did not take any substance in our previous study with this paradigm (Cools *et al*, 2006). Therefore, the selective difficulty with punishment prediction may reflect a protective bias in subjects under baseline. Suppression of the learning and/or memory of stimulus-punishment contingencies may be adaptive in this task, where the punishment is uncontrollable. Critically, the difference between punishment and reward prediction was abolished by ATD. Thus, after TRP— subjects exhibited a form of depressive realism (Alloy and Abramson, 1979), ATD did not induce a negative bias but rather abolished a protective bias against punishment anticipation. This observation concurs with previous suggestions that depressed individuals, who exhibit low 5-HT levels, do not show an attentional bias to negative information, but rather fail to demonstrate the protective bias that is evident in non-depressed individuals (McCabe and Gotlib, 1995).

The protective bias under baseline, that is the impairment in punishment relative to reward prediction may reflect resilience to aversive signals (Amat *et al*, 2005; Yehuda *et al*, 2006; JV Taylor Tavares, L Clark, ML Furey, GB Williams, BJ Sahakian, WC Drevets, unpublished observations). Resilience protects subjects from the detrimental consequences of exposure to adversity and enables them to quickly recover from negative experiences. In the present task, resilience may have resulted in a paradoxical impairment in the ability to anticipate punishment given specific predictive stimuli. Resilience has been hypothesized to result from cortical, top-down control (from the prefrontal cortex, PFC) over subcortical brain regions that mediate aversive conditioning (eg the amygdala and the dorsal raphe nucleus) (Quirk and Gehlert, 2003; Amat *et al*, 2005; Pezawas *et al*, 2005; Urry *et al*, 2006; Yehuda *et al*, 2006). In keeping with this hypothesis, recent neuroimaging observations suggest that the PFC controls amygdala activity when subjects are presented with negatively valenced stimuli (Ochsner *et al*, 2002; Phelps and LeDoux, 2005). Based on previous suggestions that 5-HT conveys resilience to adversity (Deakin and Graeff, 1991; Deakin, 1991; Richell *et al*, 2005), we hypothesize that ATD disrupts PFC-mediated control over subcortical brain regions, such as the amygdala and/or the dorsal raphe nucleus (Amat *et al*, 2005; Heinz *et al*, 2005; Pezawas *et al*, 2005). Such a top-down control failure may interact with reductions in 'background' levels of tonic 5-HT to bias the system toward anticipation of adversity (by increasing prediction errors for future punishment). This hypothesis can be tested using event-related functional neuroimaging.

Trial-by-trial analyses showed that the effect of ATD was present only on non-switch trials and there was no evidence for a similar effect on switch trials. This is consistent with the finding that effects of systemic serotonergic manipulations in human volunteers are not specific to the reversal stage of discrimination learning tasks, but extend to simple and compound discrimination learning stages of such tasks (Park *et al*, 1994; Rogers *et al*, 1999; Murphy *et al*, 2002;

Chamberlain *et al*, 2006). In keeping with these findings, ATD enhanced the BOLD response to punishment in a probabilistic reversal-learning paradigm regardless of whether punishment led to switching (Evers *et al*, 2005). Thus, ATD in human volunteers does not selectively alter behavioral flexibility, but rather has a more generalized effect on the learning and/or memory of contingencies via effects on punishment processing.

It may be noted that the effect of systemic serotonergic manipulations on human reversal learning differs from that of selective 5-HT depletion following injection of the neurotoxin 5,7-dihydroxytryptamine (5,7-DHT) in the nonhuman primate OFC. This manipulation dramatically impairs the ability to inhibit responding to the previously rewarded stimulus, while not affecting the initial acquisition of a discrimination (Clarke *et al*, 2004, 2005, 2007). To explain this discrepancy, we must take into account three factors: (1) the method of depletion, (2) the extent to which the task depends on inhibitory control, and (3) the neural site of action of 5-HT. First, injection of 5,7-DHT leads to almost complete removal of brain 5-HT levels, whereas ATD in humans reduces central 5-HT levels only modestly. These methods may well have different effects on the hypothetical equilibrium between tonic and phasic modes of neurotransmission. Second, the (serial) reversal-learning tasks in studies with humans, particularly the one employed here, do not load on inhibitory control as much as do the paradigms used in studies with nonhuman primates, for whom reinforcement is more salient and thus, habit formation more pronounced (Clarke *et al*, 2007). Finally, neuroimaging studies with human volunteers have shown particularly pronounced effects of ATD on the (dorso)medial PFC during cognitive performance (Cools *et al*, 2005a; Evers *et al*, 2005, 2006a, b; Talbot and Cooper, 2006; Van der Veen *et al*, 2007). Conversely, the disinhibitory effects resulted from selective 5-HT depletion in the OFC (Clarke *et al*, 2004, 2005, 2007). Thus, 5-HT depletion may have different functional consequences depending on the extent of depletion, task demands and on the neural site of action (medial PFC vs OFC).

## ACKNOWLEDGEMENTS

This work was conducted within the Behavioural and Clinical Neuroscience Institute, which is cofunded by the Medical Research Council and the Wellcome Trust. We are grateful to Trevor W Robbins, Molly Crocket, and Luke Clark for helpful discussion. RC holds a Royal Society University Research Fellowship and OJR holds an MRC Research Studentship.

## DISCLOSURE/CONFLICT OF INTEREST

We declare that except for income received from our primary employers no financial support or compensation has been received from any individual or corporate entity over the past 3 years for the current research and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest.

## REFERENCES

- Alloy L, Abramson L (1979). Judgement of contingency in depressed and nondepressed students: sadder but wiser? *J Exp Psychol Gen* 108: 441–485.
- Amat J, Baratta MV, Paul E, Bland ST, Watkins LR, Maier SF (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat Neurosci* 8: 365–371.
- Anderson I, Parry-Bilings M, Newsholme E, Poortmans J, Cowen P (1990). Decreased plasma tryptophan concentration in major depression: relationship to melancholia and weight loss. *J Affect Disord* 20: 185–191.
- Beats B, Sahakian B, Levy R (1996). Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. *Psych Med* 26: 591–603.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. *Arch Gen Psychiatry* 11: 561–571.
- Bethea CL, Streicher JM, Coleman K, Pau FK, Moessner R, Cameron JL (2004). Anxious behavior and fenfluramine-induced prolactin secretion in young rhesus macaques with different alleles of the serotonin reuptake transporter polymorphism (5HTTLPR). *Behav Genet* 34: 295–307.
- Blier P, de Montigny C (1999). Serotonin and drug-induced therapeutic responses in major depression, obsessive-compulsive and panic disorders. *Neuropsychopharmacology* 21: 91S–98S.
- Burghardt N, Sullivan G, McEwen B, Gorman J, LeDoux JE (2004). The selective serotonin reuptake inhibitor citalopram increases fear after acute treatment but reduces fear with chronic treatment: a comparison with tianeptine. *Biol Psychiatry* 55: 1171–1178.
- Carpenter L, Anderson G, Pelton G, Gudin J, Kirwin P, Price L et al (1998). Tryptophan depletion during continuous CSF sampling in healthy human subjects. *Neuropsychopharmacology* 19: 26–35.
- Carver C, White T (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J Pers Soc Psychol* 67: 319–333.
- Chamberlain SR, Muller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006). Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* 311: 861–863.
- Clarke H, Dalley J, Crofts H, Robbins T, Roberts A (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304: 878–880.
- Clarke HF, Walker SC, Crofts HS, Dalley JW, Robbins TW, Roberts AC (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J Neurosci* 25: 532–538.
- Clarke HF, Walker SC, Dalley JW, Robbins TW, Roberts AC (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cereb Cortex* 17: 18–27.
- Cools R, Altamirano L, D'Esposito M (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* 44: 1663–1673.
- Cools R, Calder AJ, Lawrence AD, Clark L, Bullmore E, Robbins TW (2005a). Individual differences in threat sensitivity predict serotonergic modulation of amygdala response to fearful faces. *Psychopharmacology* 180: 670–679.
- Cools R, Calder AJ, Lawrence AD, Clark L, Bullmore E, Robbins TW (2005b). Individual differences in threat sensitivity predict serotonergic modulation of amygdala response to fearful faces. *Psychopharmacology (Berl)* 180: 670–679.
- Daw N, Kakade S, Dayan P (2002). Opponent interactions between serotonin and dopamine. *Neural Netw* 15: 603–616.
- Deakin J, Graeff F (1991). 5-HT and mechanisms of defence. *J Psychopharmacol* 5: 305–315.
- Deakin J, Pennell I, Upadhyaya A, Lofthouse R (1990). A neuroendocrine study of 5HT function in depression: evidence for biological mechanisms of endogenous and psychosocial causation. *Psychopharmacology* 101: 85–92.
- Deakin JF (1991). Depression and 5HT. *Int Clin Psychopharmacol* 6(Suppl 3): 23–28; discussion 29–31.
- Elliott R, Sahakian B, Herrod J, Robbins T, Paykel E (1997). Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *J Neurol Neurosurg Psychiatry* 63: 74–82.
- Evers EA, Cools R, Clark L, van der Veen FM, Jolles J, Sahakian BJ et al (2005). Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacology* 30: 1138–1147.
- Evers EA, van der Veen FM, Jolles J, Deutz NE, Schmitt JA (2006a). Acute tryptophan depletion improves performance and modulates the BOLD response during a Stroop task in healthy females. *Neuroimage* 32: 248–255.
- Evers EA, van der Veen FM, van Deursen JA, Schmitt JA, Deutz NE, Jolles J (2006b). The effect of acute tryptophan depletion on the BOLD response during performance monitoring and response inhibition in healthy male volunteers. *Psychopharmacology (Berl)* 187: 200–208.
- Eysenck S, Eysenck H (1978). Impulsiveness and venturesomeness: their position in a dimensional system of personality description. *Psychol Rep* 43: 1247–1255.
- Grace AA (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41: 1–24.
- Graeff F, Guimaraes F, De Andrade T, Deakin JFW (1996). Role of 5-HT in stress, anxiety and depression. *Pharmacol Biochem Behav* 54: 129–141.
- Gray J (1982). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford University Press: Oxford.
- Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D et al (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297: 400–403.
- Harmer CJ, Mackay CE, Reid CB, Cowen PJ, Goodwin GM (2006). Antidepressant drug treatment modifies the neural processing of nonconscious threat cues. *Biol Psychiatry* 59: 816–820.
- Heinz A, Braus D, Smolka M, Wrase J, Puls I, Hermann D et al (2005). Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter. *Nat Neurosci* 8: 20–21.
- Howell DC (1997). *Statistical Methods for Psychology*. Wadsworth Publishing: Belmont.
- Inoue T, Kyama T, Yamashita I (1993). Effect of conditioned fear stress on serotonin metabolism in the rat brain. *Pharmacol Biochem Behav* 44: 371–374.
- Inoue T, Li X, Abekawa T, Kitaichi Y, Izumi T, Nakagawa S et al (2004). Selective serotonin reuptake inhibitor reduces conditioned fear through its effect in the amygdala. *Eur J Pharmacol* 497: 311–316.
- Inoue T, Tsuchiya K, Koyama T (1996). Serotonergic activation reduces defensive freezing in the conditioned fear paradigm. *Pharmacol Biochem Behav* 53: 825–831.
- Iversen S (1984). 5HT and anxiety. *Neuropsychopharmacology* 23: 1553–1560.
- Mathews A, Mackintosh B (2000). Induced emotional interpretation bias and anxiety. *J Abnorm Psychol* 109: 602–615.
- McCabe SB, Gotlib IH (1995). Selective attention and clinical depression: performance on a deployment-of-attention task. *J Abnorm Psychol* 104: 241–245.
- Millan MJ, Dekeyne A, Gobert A (1998). Serotonin (5-HT) 2c receptors tonically inhibit dopamine (DA) and noradrenaline (NA), but not 5-HT, release in the frontal cortex *in vivo*. *Neuropharmacology* 37: 953–955.
- Montague P, Dayan P, Sejnowski T (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16: 1936–1947.



- Murphy F, Michael A, Robbins TW, Sahakian BJ (2003). Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychol Med* 33: 455–467.
- Murphy FC, Smith K, Cowen P, Robbins TW, Sahakian BJ (2002). The effects of tryptophan depletion on cognitive and affective processing in healthy volunteers. *Psychopharmacology* 163: 42–53.
- Nishizawa S, Benkelfat C, Young S, Leyton M, Mzengeza S, de Montigny C et al (1997). Differences between males and females in rates of serotonin synthesis in human brain. *Proc Natl Acad Sci USA* 94: 5308–5313.
- Ochsner KN, Bunge SA, Gross JJ, Gabrieli JD (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J Cogn Neurosci* 14: 1215–1229.
- Park S, Coull J, McShane R, Young A, Sahakian B, Robbins T et al (1994). Tryptophan depletion in normal volunteers produces selective impairments in learning and memory. *Neuropharmacology* 33: 575–588.
- Patton J, Stanford M, Barratt E (1995). Factor structure of the Barratt impulsiveness scale. *J Clin Psychol* 51: 768–774.
- Pezawas L, Meyer-Lindenberg A, Drabant EM, Verchinski BA, Munoz KE, Kolachana BS et al (2005). 5-HTTLPR polymorphism impacts human cingulate–amygdala interactions: a genetic susceptibility mechanism for depression. *Nat Neurosci* 8: 828–834.
- Phelps EA, LeDoux JE (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48: 175–187.
- Quirk GJ, Gehlert DR (2003). Inhibition of the amygdala: key to pathological states? *Ann NY Acad Sci* 985: 263–272.
- Reilly J, McTavish S, Young A (1997). Rapid depletion of plasma tryptophan: a review of studies and experimental methodology. *J Psychopharmacol* 11: 381–392.
- Richards A, French C, Calder A, Webb B, Fox R, Young A (2002). Anxiety-related bias in the classification of emotionally ambiguous facial expressions. *Emotion* 2: 283–287.
- Richell RA, Deakin JF, Anderson IM (2005). Effect of acute tryptophan depletion on the response to controllable and uncontrollable noise stress. *Biol Psychiatry* 57: 295–300.
- Rogers RD, Blackshaw AJ, Middleton HC, Matthews K, Hawtin K, Crowley C et al (1999). Tryptophan depletion impairs stimulus-reward learning while methylphenidate disrupts attentional control in healthy young adults: implications for the monoaminergic basis of impulsive behaviour. *Psychopharmacology* 146: 482–491.
- Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* 275: 1593–1599.
- Steffens D, Wagner H, Levy R, Horn K, Krishnan K (2001). Performance feedback deficit in geriatric depression. *Biol Psychiatry* 50: 358–363.
- Sutton R, Barto A (1998). *Reinforcement Learning*. MIT Press: Cambridge, MA.
- Talbot PS, Cooper SJ (2006). Anterior cingulate and subgenual prefrontal blood flow changes following tryptophan depletion in healthy males. *Neuropsychopharmacology* 31: 1757–1767.
- Tye N, Everitt B, Iversen S (1977). 5-Hydroxytryptamine and punishment. *Nature* 268: 741–743.
- Urry HL, van Reekum CM, Johnstone T, Kalin NH, Thurow ME, Schaefer HS et al (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. *J Neurosci* 26: 4415–4425.
- Van der Veen FM, Evers EA, Deutz NE, Schmitt JA (2007). Effects of acute tryptophan depletion on mood and facial emotion perception related brain activation and performance in healthy women with and without a family history of depression. *Neuropsychopharmacology* 32: 216–224.
- Watson D, Clark LA, Tellegen A (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 54: 1063–1070.
- Wilkinson LS, Humby T, Robbins TW, Everitt BJ (1995). Differential effects of forebrain 5-hydroxytryptamine depletions on Pavlovian aversive conditioning to discrete and contextual stimuli in the rat. *Eur J Neurosci* 7: 2042–2052.
- Yehuda R, Flory JD, Southwick S, Charney DS (2006). Developing an agenda for translational studies of resilience and vulnerability following trauma exposure. *Ann NY Acad Sci* 1071: 379–396.
- Young S, Smith S, Pihl R, Ervin F (1985). Tryptophan depletion causes a rapid lowering of mood in normal males. *Psychopharmacology* 87: 173–177.