ARTICLE OPEN

Withdrawal from escalated cocaine self-administration impairs reversal learning by disrupting the effects of negative feedback on reward exploitation: a behavioral and computational analysis

Peter Zhukovsky[1,2], Mickael Puaud[1,2], Bianca Jupp[1,2], Júlia Sala-Bayo[1,2], Johan Alsiö[1,2], Jing Xia[1,2], Lydia Searle[1,2], Zoe Morris[1,2], Aryan Sabir[1,2], Chiara Giuliano[1,2], Barry J. Everitt[1,2], David Belin [1,2], Trevor W. Robbins[1,2] and Jeffrey W. Dalley[1,2,3]

Addiction is regarded as a disorder of inflexible choice with behavior dominated by immediate positive rewards over longer-term negative outcomes. However, the psychological mechanisms underlying the effects of self-administered drugs on behavioral flexibility are not well understood. To investigate whether drug exposure causes asymmetric effects on positive and negative outcomes we used a reversal learning procedure to assess how reward contingencies are utilized to guide behavior in rats previously exposed to intravenous cocaine self-administration (SA). Twenty-four rats were screened for anxiety in an open field prior to acquisition of cocaine SA over six daily sessions with subsequent long-access cocaine SA for 7 days. Control rats ($n = 24$) were trained to lever-press for food under a yoked schedule of reinforcement. Higher rates of cocaine SA were predicted by increased anxiety and preceded impaired reversal learning, expressed by a decrease in lose-shift as opposed to win-stay probability. A model-free reinforcement learning algorithm revealed that rats with high, but not low cocaine escalation failed to exploit previous reward learning and were more likely to repeat the same response as the previous trial. Eight-day withdrawal from high cocaine escalation was associated, respectively, with increased and decreased dopamine receptor D2 (DRD2) and serotonin receptor 2C (HTR2C) expression in the ventral striatum compared with controls. Dopamine receptor D1 (DRD1) expression was also significantly reduced in the orbitofrontal cortex of high cocaine-escalating rats. These findings indicate that withdrawal from escalated cocaine SA disrupts how negative feedback is used to guide goal-directed behavior for natural reinforcers and that trait anxiety may be a latent variable underlying this interaction.

INTRODUCTION

Despite considerable research, the psychological mechanisms underlying the maladaptive behavior of individuals addicted to drugs remain poorly understood, in particular the propensity of such individuals to continue taking drugs despite mounting negative impacts. A disregard for harmful consequences implies an innate or acquired imbalance in how positive and rewarding outcomes are perceived and processed relative to punishment signals. Consistent with this view, rats exposed to cocaine fail to utilize outcome value to guide behavior [1] and continue to seek drugs despite their devaluation [2–4] or the risk of punishment [5–7]. The neural mechanisms underlying this maladaptive behavior are not well understood but may underlie compulsive forms of drug seeking [8].

Neural activity in the orbitofrontal cortex (OFC) is broadly acknowledged to represent outcome value and expectation used to guide value-based decision-making [9–15]. The OFC also plays a key role in behavioral flexibility, the capacity to rapidly track changing stimulus-response contingencies in reversal learning procedures [16–19], and structural and functional changes in the

OFC are present in individuals addicted to drugs [20–26]. Consistent with these findings, reversal learning is impaired in rats and monkeys exposed to cocaine [17, 27–29]. Thus, drug-induced abnormalities in OFC networks that include the amygdala and striatum [11] may underlie the inflexibility and insensitivity to outcomes associated with drug exposure [1].

The ability to respond flexibly to changing stimulus-response contingencies requires an animal to learn about the prospective values of the responses using both positive and negative feedback. Reversal learning tasks require animals to optimize their choice strategy to maximize the rewards they obtain, while at the same time occasionally exploring alternate reward options. Using the representation of predictive relationships in the environment, acquired through trial and error, reversal learning requires animals to switch responding to a now correct stimulus while ignoring the interference of a recently rewarded, but now no-longer-correct stimulus. Reinforcement learning has been proposed as a tractable computational process underlying trial-and-error learning [30, 31] with utility in modelling aspects of addiction [32, 33] and stimulant administration in rodents [34, 35].

[1]Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK; [2]Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, UK and [3]Department of Psychiatry, University of Cambridge, Cambridge CB2 2QQ, UK
Correspondence: Jeffrey W. Dalley (jwd20@cam.ac.uk)

2164

Indeed, computational psychiatry has become an increasingly popular translational methodology to investigate mental health [36, 37], especially if the computational models are constrained by neurobiological data [38].

In this study we therefore used behavioral and computational methods to define the nature of reversal learning deficits in rats with a history of escalated cocaine self-administration, compared with food-reinforced, cocaine-naïve rats. We hypothesized that cocaine-exposed rats would be less sensitive to negative feedback in a spatial-discrimination reversal learning task than drug-free controls. We further evaluated possible modulatory effects of trait anxiety given that this predicts the individual propensity to escalate cocaine SA under long-access conditions [39–41], as well as response perseveration in a spatial reversal learning task [42]. Finally, we measured ex vivo gene transcript levels of dopamine (DA) and serotonin (5-HT) receptors in the OFC and striatum as neural correlates of cocaine-induced impairments in reversal learning.

MATERIALS AND METHODS
Subjects
Subjects were male Lister-hooded rats ($n = 48$) weighing 280–300 g at the beginning of the experiments (Charles River, Kent, UK). Rats were maintained at 85–95% of free-feeding weights. Each animal received 18 g of food chow once a day within 2 h after behavioral testing and had ad libitum access to water. When no behavioral training or testing took place, rats received 20 g of chow per day. Rats were either housed in groups of four or singly after catheter implantation and during the cocaine self-administration experiment under a reversed 12 h light/dark cycle (lights off 07:00 h until 19:00 h). Two cohorts of rats were trained and tested on a spatial-discrimination serial reversal learning task [42, 43] prior to the assessment of anxiety on an open field test (Fig. 1a). Rats in cohort 1 ($n = 24$) were trained to self-administer intravenous (i.v) cocaine (six daily short-access sessions; seven daily long-access sessions) whereas cohort 2 rats ($n = 24$) were trained to lever-press for food pellets (Noyes dustless pellets, 45 mg, Sandown Scientific, UK) over an equivalent period of days. Rats from both cohorts were re-tested on the reversal learning task 8 days after cocaine or food self-administration. All experiments complied with the statutory requirements of the Animals (Scientific Procedures) Act 1986 following local ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (PPL 70/8072).

Behavioral assessment
Reversal learning. Spatial-discrimination reversal learning was assessed using 12 5-choice operant chambers placed in ventilated, sound-attenuating cubicles (Med Associates, Georgia, VT), as previously described [42, 43]. Subjects were initially habituated to the apparatus over 2 days, with each session lasting 20 min. They were then trained to enter the magazine to trigger the illumination of a single stimulus light (left or right) and to respond in one of the two illuminated apertures for food delivery under a fixed-ratio (FR) 1 schedule of reinforcement. Once rats had achieved 50 correct responses, food reward was successively delivered under FR2 and FR3 schedules to the same criterion within a 30 sec limited hold period. Failure to respond within the 30 sec period resulted in a 5 sec time-out. Once rats were able to achieve criterion under a 5 sec inter-trial interval, they were tested for spatial discrimination followed the next day by a reversal of the stimulus-reward contingency. Firstly, rats were given a maximum of 1 h to complete the discrimination task by achieving nine correct trials across the previous 10 trials. Once rats achieved this criterion and consistently responded at the rewarded (left or right) aperture, the session ended. On the following day, rats were given a retention test of the

discrimination learned on the previous day, in which the same aperture was rewarded. Once rats achieved criterion (9/10 correct), they then completed three reversals (Fig. 3g). This test session lasted for ~1 h. Following each contingency reversal, responses in the previously incorrect aperture were signaled as correct (and reinforced with a food pellet) whereas responses in the previously correct aperture were signaled as incorrect (and not reinforced with food). Two rats (one in the cocaine and one in the control group) who failed to achieve the criterion of three successful reversals were excluded from the analysis. In addition, five rats in the cocaine group were excluded from the study due to suspected catheter failure. Following cocaine or food self-administration, rats were re-trained over five sessions to respond for food on the spatial-discrimination task under a FR3 schedule of reinforcement. On the test day, rats were given a retention test prior to completing three reversals, identical to the procedure described above.

Anxiety assessment. A black, matte arena of 150 cm diameter and 50 cm high walls was used to assess behavior in the open field [44] under white lights (70 lux). The central area of the arena was defined as a circle with a diameter of 75 cm. Exploratory behavior in the maze was recorded and monitored on a ceiling-mounted Yi Action Camera (Xiaomi, Japan) connected to a computer via Wi-Fi and analyzed using Icarus V2.09 (University of Manchester, UK 2002–2003) software. Rats were placed in the center of the arena with behavior recorded for 8 min. The arena was cleaned with water between each animal. An anxiety score was calculated as the proportion of time spent in the center of the arena in the total time of 8 min.

Intravenous cocaine self-administration. Twenty-four operant chambers (31.8 cm long × 25.4 cm width × 34.3 cm high), constructed of Plexiglas and a metal grid floor, were each placed in ventilated, sound-attenuating cubicles (Med Associates, Georgia VT). Whisker Control software (Second Order, Cardinal and Aitken, 2010) controlled the apparatus. Two retractable levers and a white light emitting diode located above each lever were placed along one wall of the chamber, with a house-light positioned on top of the opposite wall. Cocaine infusions were delivered via implanted intravenous-dwelling catheters connected to a syringe-driven infusion pump (Semat Technical, Herts, UK) and Tygon tubing. Infusions were delivered at a rate of 20 µl/sec. Each infusion contained 0.25 mg cocaine hydrochloride.

A single-lumen catheter (CamCath®, Cambridge, UK, inner diameter 0.28 mm; outer diameter 0.61 mm; dead volume 12 µl) was implanted in the right jugular vein under ketamine hydrochloride (100 mg/kg, intraperitoneal, Ketaset) and xylazine (9 mg/kg, i.p., Rompun) anesthesia. The proximal end of the silastic catheter was inserted in the right atrium and the distal end was sutured subcutaneously between the scapulae. To prevent infection, rats were treated with a subcutaneous injection of 10 mg/kg Baytril® (Genus Express, Bury St. Edmunds, UK) on the day before surgery and were then given 10 mg/kg Baytril in mashed pellets for 5 days post-surgery. Following surgery, catheters were flushed daily with saline-heparin (100 IU/ml), with a recovery period of 10 days.

Rats acquired cocaine SA during six daily 1 h long sessions (short-access; ShA), under a fixed-ratio (FR)-1 schedule of reinforcement, and each 0.05 ml infusion containing 0.25 mg cocaine hydrochloride (MacFarlan, UK) was delivered over 5.7 s. Thereafter, rats were given long-access (LgA) exposure to cocaine over seven daily 6 h long sessions. Catheters were flushed with saline-heparin before and after each session. Each session started with the chamber being illuminated and the 2 levers inserted. Active lever presses resulted in a cocaine infusion and a white cue light cue for 5 sec followed by a 20 sec time-out period, during which both levers were retracted. Inactive lever presses had no
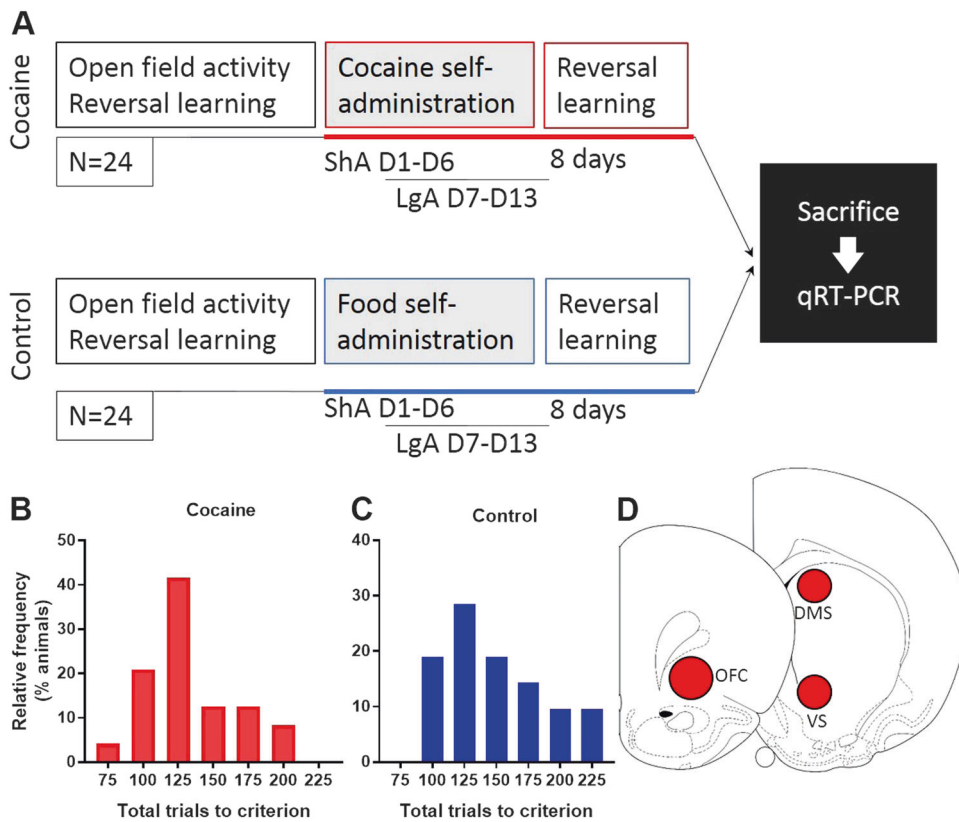
Withdrawal from escalated cocaine self-administration impairs reversal…
P Zhukovsky et al.

2165

Fig. 1 **a** Experimental timeline. Two cohorts of rats (each $n = 24$) were assessed for open field activity as a measure of anxiety followed by spatial-discrimination reversal learning. Rats in cohort 1 were trained to intravenously self-administer cocaine under short- and long-access schedules (ShA; LgA) while rats in cohort 2 (control group) responded for food reinforcement under identical schedules. Finally, rats in both cohorts were re-assessed for reversal learning prior to sacrifice and post mortem qRT-PCR. **b**, **c** Frequency distribution plots of 'total trials to criterion' for the cocaine and control rats. **d** qRT-PCR was used to assess gene expression in the OFC, dorsal and ventral striatum

scheduled consequences. Active and inactive levers were randomly assigned to the 24 rats.

Food reinforcement. Twelve operant chambers of the same configuration and manufacturer as the cocaine SA chambers were used. These only differed by the presence of food pellet dispenser and magazine. Rats ($n = 24$) were trained to make a lever-press response for a single food pellet (Noyes dustless pellets, 45 mg, Sandown Scientific, UK) under an FR-1 schedule for the first 6 daily 1 h sessions. Thereafter, rats responded under an FR5 schedule for the remaining 7 days. Rats in this group did not have a surgically-implanted i.v. catheter. In order to maximize the time spent in the testing context, and in accordance with the cocaine SA experiment, the post-reinforcement time-out period was set to 60 sec. Inactive lever responses were recorded but had no programmed consequences. The maximum number of pellets available was adjusted to match the number of lever-press responses made by the cocaine SA rats. Since rats consumed the food pellets whenever they became available, the maximum number of pellets determined the session duration.

Computational modeling
Several learning models were used to simulate the reversal learning data, including three variants of the Q-learning model [45], defined below and the three parameters: α, β, and κ. Model parameters were fitted to each animal's reversal data individually and then compared using analysis of variance (ANOVA). The learning rate α determines how quickly the model adjusts to the expected value of a response following positive or negative feedback. High α values allow the agent to increase

(or decrease) the expected Q-value placed on that response if the response is followed by a reward (or not). The inverse temperature parameter β regulates how much an agent explores by responding randomly or exploits what the agent learned about the responses to date. A low β value would lead an agent to rely on the expected Q-values of the responses and hence exploit what they have learnt about the responses already. A high β value would lead to exploration that under some circumstances may lead to higher rewarded outcomes. However, in the present reversal task, with deterministic outcomes, a high β value would result in fewer rewards. Finally, the choice autocorrelation parameter κ is a measure of "stickiness", or how likely an animal will perform the same response again regardless of reward outcome. Values of κ close to 1 reflects an agent "sticking" to the previous response while κ values close to -1 reflects choice alternation.

Model-free Q-learning: model 1. Simple Q-learning is equivalent to Rescorla-Wagner learning [30] whereby an agent assigns an expected Q-value to each choice available; presently a left or right response (L or R) at each trial t. The expected Q-value is updated on each trial according to the following:

$$P(c_t = L | Q_t(L), Q_t(R)) = \frac{\exp(Q_t(L)/\beta)}{\exp(Q_t(L)/\beta) + \exp(Q_t(R)/\beta)}$$

where $0 \le \alpha \le 1$ is a learning parameter, $Q_t(c_t)$ is the value of the choice c_t at trial t and r takes the value of 1 if the choice was rewarded and a value of 0 if not. A large α implies faster updating of the expected Q-values of a response after a trial is completed. The probability of making the choice c_t at trial t was calculated

Withdrawal from escalated cocaine self-administration impairs reversal...
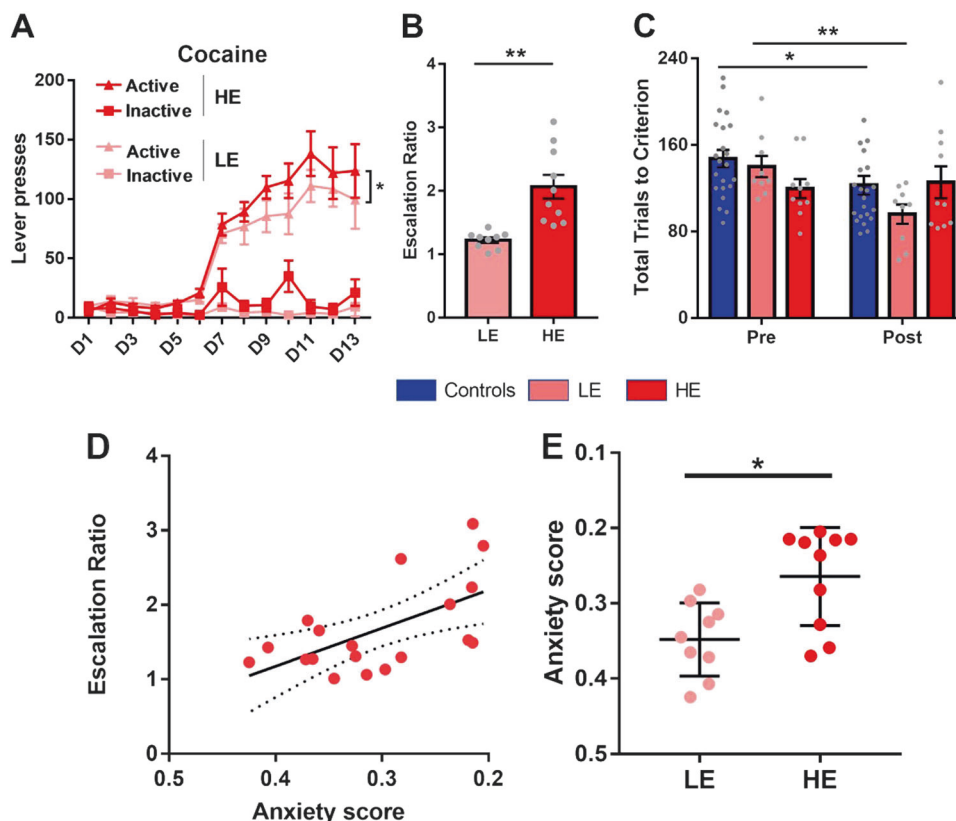P Zhukovsky et al.

2166

Fig. 2 **a** Active and inactive lever-press responses of rats trained to self-administer cocaine. Data shown are means ± SEM. Since rats responded on a fixed-ratio 1 schedule, the number of lever presses was equivalent to the number of infusions received. Rats were divided into two groups: high escalation (HE) and low escalation (LE), based on a median split of escalation ratios. The escalation ratio was calculated as the ratio of the average number of active lever responses on days 12 and 13 to the number of lever responses on day 7 (D7—the first long-access session). During the first 6 days rats were given short-access to cocaine (1 h daily sessions) under a fixed-ratio-1 (FR-1) schedule of reinforcement. On days 7–13 inclusive, access to cocaine was increased to 6 h under an FR-1 schedule. **b** Escalation ratios for each animal in the high and low escalation groups, based on a median split (independent samples $t_{17} = 4.2$, $p = 0.0006$). **c** Individual reversal learning scores (total trials to criterion) before and after cocaine exposure in LE and HE rats compared with control rats. Data are means ± SEM. *$p < 0.05$. **$p < 0.01$. Relationships between anxiety and escalation of intravenous cocaine self-administration are shown in plots (**d**, **e**), including a line of best of fit with 95% confidence intervals in dotted lines. A lower anxiety score equates to increased anxiety in the open field arena. **d** Significant positive relationship between escalation ratio during the first hour of cocaine self-administration and anxiety scores ($r^2 = 0.29$, $p < 0.05$), consistent with significant group differences in anxiety scores between LE and HE rats (**e**)

using the *softmax* rule:

$$P(c_t = L|Q_t(L), Q_t(R)) = \frac{\exp(\beta * Q_t(L))}{\exp(\beta * Q_t(L)) + \exp(\beta * Q_t(R))}$$

where β is the inverse temperature parameter, with larger β values leading to more exploration of the responses with lower Q-values. On the other hand, smaller β values result in exploitation of the response with higher Q-values.

Model-free Q-learning: model 2. Model 1 was extended to include a separate α for learning from rewards and losses, α_{REWARD} and α_{NO} $_{REWARD}$, depending on whether the animal received a reward on trial t. The decision probability was updated in the same way as in Model 1.

Model-free Q-learning: model 3. A different variation of Model 1 included only one learning parameter α as in Model 1, but an additional autocorrelation parameter in the observational part of the model:

$$P(c_t = L|Q_t(L), Q_t(R), L_{t-1}, R_{t-1})$$
$$= \frac{\exp(Q_t(L)/\beta + \kappa * L_{t-1})}{\exp(Q_t(L)/\beta + \kappa * L_{t-1}) + \exp(Q_t(R)/\beta + \kappa * R_{t-1})}$$

whereby a larger κ results in greater probability of the choice c_t at trial t being the same as the choice c_t at trial $t-1$. The same approach was applied to the right sided choice.

Model fitting. The probability of *Data D* (a sequence of choices and rewards) is the product of the individual probabilities of making a choice c_t at trial t:

$$P(\text{Data } D|\text{Model } M, \text{parameters } \theta) = P(D|M, \theta) = \prod P(c_t|Q_t(L), Q_t(R))$$

Model space was treated as discrete, using the following range of parameters: $0.001 \leq \alpha \leq 1$ with a step size of 0.08; $0.005 \leq \beta \leq 5$ with a step size of 0.08 and $-1 \leq \kappa \leq 1$ with a step size of 0.08. Parameter range was chosen based on the *a priori* expectations regarding α and κ, as well as empirical information about best fit β parameters. Best fit parameters ($\hat{\theta}_M$) were chosen to maximize the log-likelihood of the observed data for each participant over all parameter sets (θ) by finding the maximum of the probability density function, $\arg\max_{\theta} P(D|M, \theta)$.

Model comparison. Nested models were compared using the likelihood ratio test that contrasts the log-likelihood of the data

Withdrawal from escalated cocaine self-administration impairs reversal…
P Zhukovsky et al.

2167

given the best fit parameters ($\widehat{\theta}_M$):

$$d = 2 * \left[\log P\left(D \middle| M_2, \widehat{\theta}_{M_2}\right) - \log \log P\left(D \middle| M_1, \widehat{\theta}_{M_1}\right) \right]$$

As d follows the χ-square distribution, the difference in data likelihood associated with increasing the number of parameters from two (α, β) to three (α_{REWARD}, $\alpha_{NO\ REWARD}$, β or α, β, κ) is significant at $p = 0.05$ for $d > 3.842$. An example of model predicted probability of choosing left or right (for $\widehat{\theta}_M$) together with the sequence of observed responses and rewards is shown in Fig. 3g. A biased measure of model fit, *pseudo r^2*, was computed as follows:

$$\text{pseudo } r^2 = \frac{\log P\left(D \middle| M, \widehat{\theta}_M\right) - 0.5^n}{0.5^n}$$

where n represents the number of trials and the probability of observing the data when the best fit parameters are contrasted against the probability of observing the data at random (0.5^n). Although *pseudo r^2* will increase with the number of parameters fitted and does not penalize overfitting, it can be useful in linking the modelling results to more traditional statistical methods such as linear regression. Finally, the Bayesian Information Criterion (BIC) provided an alternative measure of model fit:

$$\text{BIC} = \log(P(D|M)) \approx \log\left(P\left(D \middle| M, \widehat{\theta}_M\right)\right) - \frac{n}{2} \log m$$

where n = number of free parameters and m = number of observations. We implemented this analysis using in-house Matlab scripts (R2016a), which can be found in the following link: https://github.com/peterzhukovsky/reversal_learning).

Postmortem gene expression
Aliquots of brain tissue (diameter 1.0 mm) were extracted from 150 μm frozen slices. Their location is shown in Fig. 1d. miRNeasy Mini kit (Qiagen, UK) with additional DNAse digestion was used to extract RNA from the frozen samples. RNA yields were quantified using a Nanodrop 2000 spectrophotometer (Thermo Fisher, UK). First-strand cDNA was synthesized from 5 ng total RNA using random hexamer primers from the RevertAid First-Strand cDNA Synthesis Kit (Thermo Scientific, UK) and diluted to 2.5 ng per μl. SYBR green-based quantitative real-time polymerase chain reaction (qRT-PCR) was performed on the CFX96 Touch Thermal Cycler (Bio-Rad, UK). PCR on duplicates was performed using 0.25 mM of each primer. Efficiencies were calculated using linregPCR and the ΔΔCt method [46], normalizing against two reference genes (Tubulin and B-Actin) and the mean of the food control group. Primer pairs were purchased from Sigma–Aldrich, as detailed previously [43]. PCR runs were set up as follows: 95 °C for 5 min; 40 cycles at 95 °C for 10 s; 60 °C for 10 s, and 72 °C for 1 min.

Statistical analyses
All statistical analyses were carried out using SPSS (IBM version 23). Rats assigned to the cocaine SA experiment were segregated into two groups ($n_1 = 9$; $n_2 = 10$) using a median split based on the escalation ratio, defined as the proportion of infusions taken on the last 2 days of LgA to the infusions taken on the first day of LgA. A mixed-effects ANOVA with session (13 levels) and cocaine escalation group (High vs Low) as within- and between-subject factors, respectively, was used to confirm the different cocaine self-administration profiles. Further, two-way ANOVAs were used to assess the effect of group (controls vs high vs low cocaine escalation) on reversal performance, including the total number of trials to reach criterion, the number of perseverative errors (7/10 incorrect) to criterion; lose-shift and win-stay probabilities; alpha, beta and kappa model parameters. While group was used as a between-subject factor, time of testing (at baseline or post-

cocaine/food SA) was used as within-subject factor. A mixed-effects two-way ANOVA was used to test for the between-subject effects of group (controls, HE, and LE) and the within-subject effect of region (OFC, VS, and DMS) on mRNA expression of each mRNA receptor subtype (DRD2, DRD1, HT2AR, and HT2CR). LSD tests were chosen for post hoc comparisons due to the increased power. If sphericity was violated as indicated by Mauchly's test, a Greenhouse-Geisser correction was used. Linear regressions were used to test for associations between reversal learning, cocaine escalation, and anxiety scores. Statistical significance threshold was set at $p < 0.05$.

RESULTS
High escalation of cocaine SA impairs reversal learning following 8 days of withdrawal
Following the assessment of reversal learning and anxiety, rats acquired i.v. cocaine SA over 6 consecutive days (D1-D6), as shown in Fig. 2a. Over the 13 days of cocaine SA rats responded differentially on the active and inactive levers and in response to increased cocaine availability (D7-D13) increased their responding for cocaine as shown by a significant increase in the number of active lever-press responses during this period ($F_{2.93,132} = 5.0$, $p = 0.004$, $\eta^2 = 0.19$). Two groups of rats were subsequently formed— low escalation (LE) and high escalation (HE)—based on a median split of escalation ratio, calculated as the mean number of infusions during days 12 and 13 divided by the number of infusions on day 7 (i.e., the first day of long-access exposure). The mean (±1SEM) escalation ratio for LE and HE rats was 1.2 ± 0.04 and 2.1 ± 0.19, respectively (Fig. 2b). However, the groups did not differ in terms of the total amount of cocaine taken during the LgA sessions (total of 165.4 mg/kg/rat and 176.4 mg/kg/rat for LE and HE rats, respectively, $p > 0.6$). A separate group of control rats ($n = 23$) responded for food reinforcement and were matched to the cocaine group in terms of the maximum number of lever-press responses they could make.

We next assessed whether variation in reversal learning predicted cocaine escalation and, in turn, what effect long-access cocaine exposure had on reversal learning itself, measured 8 days after the end of self-administration. As shown in Fig. 2c, rats were generally faster to reverse when assessed for the second time on the reversal learning task (i.e., made fewer trials to criterion). Thus, a mixed-effects ANOVA with exposure time (pre- versus post-food/cocaine) and group (control, LE, and HE) as within- and between subjects factors, respectively, revealed a significant main effect of exposure time ($F_{1,38} = 7.74$, $p = 0.008$) and a trend for an interaction between group and exposure time ($F_{2,38} = 3.1$, $p = 0.056$). Post-hoc LSD tests revealed that while the number of trials to criterion significantly decreased during the second ('post') assessment in control and LE rats, this was not the case in HE rats. No significant differences between the three groups were found at baseline or post-cocaine (post hoc LSD, $p > 0.05$).

An analysis of perseverative errors revealed an interaction between time and group ($F_{2,38} = 3.3$, $p = 0.047$). The low escalation group improved over time (LSD, $p = 0.007$), resulting in significant group differences between LE and HE (LSD, $p = 0.035$) and LE and controls (LSD, $p = 0.049$) during the second assessment. However, baseline (i.e., 'pre') levels of perseverative responding were not significantly different between control, LE, and HE rats. These findings indicate that rats with a history of escalated cocaine intake (HE) failed to show the expected improvement in behavioral flexibility after repeated testing on the reversal learning task.

Anxiety but not reversal learning predicts cocaine escalation
Figure 2 summarizes the dimensional relationships of anxiety with cocaine escalation. Anxiety was positively related to the escalation

Withdrawal from escalated cocaine self-administration impairs reversal…
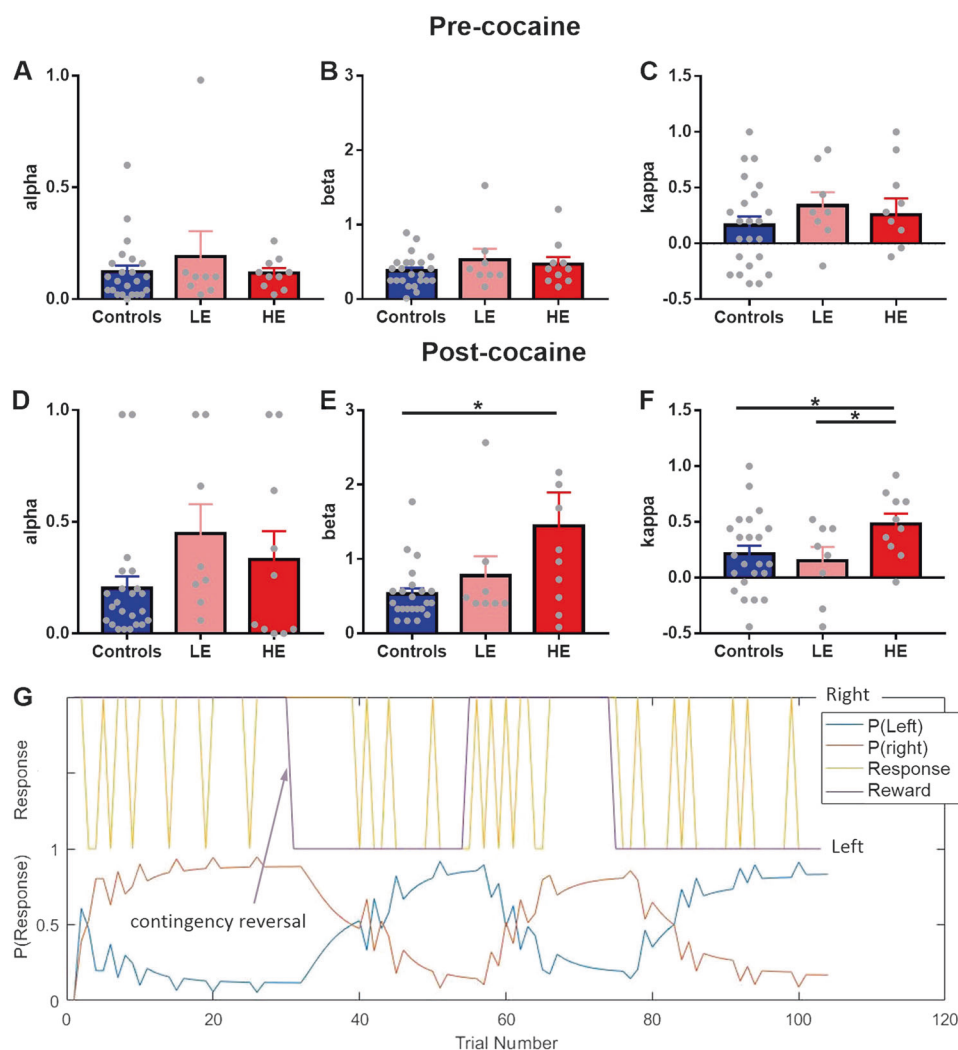P Zhukovsky et al.

2168

Fig. 3 Modelling variables of learning and response flexibility on the spatial reversal learning task before and after intravenous cocaine self-administration compared with control rats. Data are means ± 1SEM. No significant differences in α, β and κ were observed in future LE and HE rats compared with control rats (a–c, respectively). Whereas the rate of learning of a response after the completion of each trial (α) was not significantly affected by cocaine exposure (d), a significant increase in β (e) and κ (f) was observed in HE rats (*$p < 0.05$; **$p < 0.01$). Thus, HE rats failed to exploit what they had previously learnt (increased β) and showed an increased tendency to make the same response as the previous trial (increased κ). An example of the model fit is shown in the lower panel (g) with individual left and right responses in the upper yellow traces alongside the rewarded side (violet trace) and in the lower trace the modelled probabilities of the same animal making a left or right response using the modelled values of α, β and κ

of cocaine SA ($F_{1,17} = 9.3$, $r^2 = 0.354$, $p = 0.007$, Fig. 2d). Thus, anxiety scores were significantly different between future LE and HE rats (Fig. 2e). However, using linear regression models, we found neither a relationship between baseline behavioral flexibility (total trials to criterion) and escalation ratio ($r^2 = 0.01$, $p > 0.05$, supplementary fig. 1A) nor a significant relationship between anxiety and behavioral flexibility ($r^2 < 0.06$, $p > 0.05$, supplementary fig. 1B).

High cocaine escalation decreases exploitation of previously learnt choice values and increases choice autocorrelation
Adding a third parameter in models 2 and 3 significantly improved the model fit compared with model 1. Model 3 provided a better fit of the data derived from the cocaine group than model 2 (*average pseudo r^2 = 0.16* compared to *pseudo r^2 = 0.14*, and average BIC = 66.2 compared to average BIC = 67.6, respectively) while model 2 provided a better fit of the data derived from the control group than model 3 (*pseudo r^2 = 0.20* compared to *pseudo r^2 = 0.21*, and average BIC = 69.8 compared to average BIC = 70.3, respectively). Model 3 was therefore chosen as the preferred

model given its superiority in modeling the post-cocaine data, the main dataset of interest, and as a means to assess choice autocorrelation. A fourth model was also tested that included four parameters: a reward learning rate, a non-reward learning rate, beta, and kappa. This model failed to improve upon the fit of model 3 and hence was not included in the analysis (Supplementary Table 1).

Figure 3 reports individual modeled parameters for control, LE, and HE rats before and after cocaine SA. In addition to a significant main effect of time ($F_{1,38} = 9.5$, $p = 0.004$) and group ($F_{1,38} = 4.7$, $p = 0.015$), a significant interactive effect of group (controls vs LE vs HE) and time (pre vs post-cocaine) was found on beta ($F_{2,38} = 3.3$, $p = 0.048$) but not on alpha and kappa ($F_{2,38} = 1.2$, $p = 0.33$, $F_{2, 38} = 2.2$, $p = 0.13$). Post hoc comparisons revealed no significant group differences in α, β, or κ prior to cocaine exposure (all $p > 0.3$, Fig. 3a–c). However, following cocaine self-administration, HE rats showed a significantly increased β value (lower exploitation, Fig. 3e) compared with controls (LSD, $p = 0.0002$) and LE rats ($p = 0.024$) together with a significantly increased κ value (an increased tendency to repeat the last response, Fig. 3f) compared with
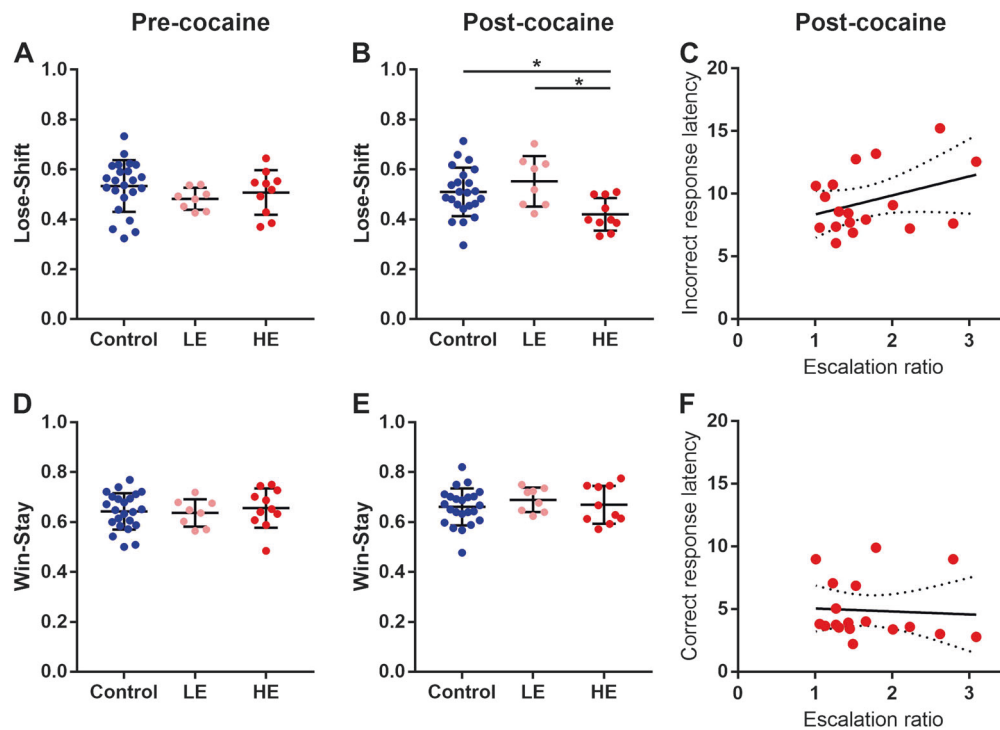
Withdrawal from escalated cocaine self-administration impairs reversal...
P Zhukovsky et al.

2169

Fig. 4 Lose-shift and win-stay probabilities on the spatial reversal learning task before and after intravenous cocaine self-administration compared with control rats. Data are means ± SEM. Prior to cocaine exposure there were no significant differences in lose-shift and win-stay probabilities between any of the groups (**a**, **d**). However, in rats exhibiting high escalation, lose-shift probabilities significantly decreased compared with low escalation and control rats (**b**) unlike win-stay probabilities (**e**). Escalation ratios did not significantly correlate with incorrect (**c**) or correct (**f**) response latencies. Shown are the lines of best fit (solid lines) and 95% confidence intervals (dotted lines)

control and LE rats ($p = 0.049$). Importantly, neither LE nor HE rats differed from controls in the rate of learning parameter, a, after cocaine SA ($p > 0.1$, Fig. 3d).

Cocaine exposure has differential effects on lose-shift and win-stay behavior
A mixed effect ANOVA revealed a significant interactive effect of time ('pre' vs 'post') and group (controls vs LE vs HE) on lose-shift probability ($F_{2,38} = 5.2$, $p = 0.01$, Fig. 4), but not on win-stay probability. Post hoc LSD tests revealed that this effect was driven by a significant decrease in lose-shift probability in HE rats (LSD, $p = 0.014$) compared with control and LE groups (LSD, $p = 0.004$, Fig. 4b) and notably was not present before the rats were exposed to cocaine (Fig. 4a). In contrast, win-stay probability was unaffected by cocaine exposure (Fig. 4e) and was no different between control, LE, and HE rats prior to cocaine SA (Fig. 4d). Using linear models, we found no significant relationship between escalation ratio, assessed over 6 h sessions, and incorrect response latencies, defined as time to initiate a new trial after the end of the previous trial, ($r^2 = 0.13$, $p = 0.14$, Fig. 4c) or correct response latencies ($r^2 = 0.06$, $p = 0.8$, Fig. 4f).

Differential effects of cocaine on the expression of genes encoding DA and 5-HT receptors
Figure 5 shows gene transcription levels of candidate DA and 5-HT receptors in the OFC, ventral striatum (VS), and dorsomedial striatum (DMS). Two-way ANOVA with group (control, LE, and HE) and region (OFC, VS, and DMS) revealed significant interactions between region and group for *DRD2* ($F_{4,113} = 4.6$, $p = 0.002$, Fig. 5a) and *HT2CR* ($F_{4,113} = 3.2$, $p = 0.017$, Fig. 5d), but not *DRD1* ($F_{4,82} = 2.3$, $p = 0.06$, Fig. 5b) and *HT2AR* ($F_{4,82} = 1.5$, $p = 0.20$, Fig. 5c). Post hoc LSD contrasts revealed significant increases in *DRD2* expression in the VS of the HE group and in the DMS of LE and HE groups compared with the control group. *DRD1* expression

in the OFC significantly decreased in the HE groups compared with the control group, whereas *HTR2A* expression increased significantly in both escalation groups in the OFC compared with controls. *HTR2C* expression was significantly decreased in the VS of HE rats compared with controls.

DISCUSSION
Our findings demonstrate several features and consequences of long-access cocaine self-administration that selectively affect how negative and positive feedback signals are processed to guide behavior in a reversal learning task. In agreement with our previous findings [41], we found that rats exhibiting high baseline trait anxiety showed greater escalation of cocaine. These rats were also more likely to perseverate with their previous response regardless of whether the outcome was rewarded or not. Importantly, high cocaine escalation rats learned as quickly as control and low cocaine escalation rats from the outcome of each trial but were unable to exploit this information flexibility to adjust behavior when the stimulus-reward contingencies were reversed. These findings support and extend previous findings that cocaine impairs insight and makes actions less sensitive to response outcomes [1, 47] by showing that high rates of cocaine self-administration, associated with trait anxiety, cause a selective disruption in the way negative feedback is used to guide behavior to a food incentive.

Substantial evidence suggests that anxiety can be both a precursor and consequence of drug abuse [48–51], with the perpetuation of drug use possibly reflecting the self-medication of chronic anxiety states [52]. Increased anxiety in rats predicts the propensity to develop a conditioned place preference for cocaine [53], increased oral and intravenous cocaine escalation [40, 41], and increased motivation to self-administer cocaine [39]. These findings were supported by the present study with increased
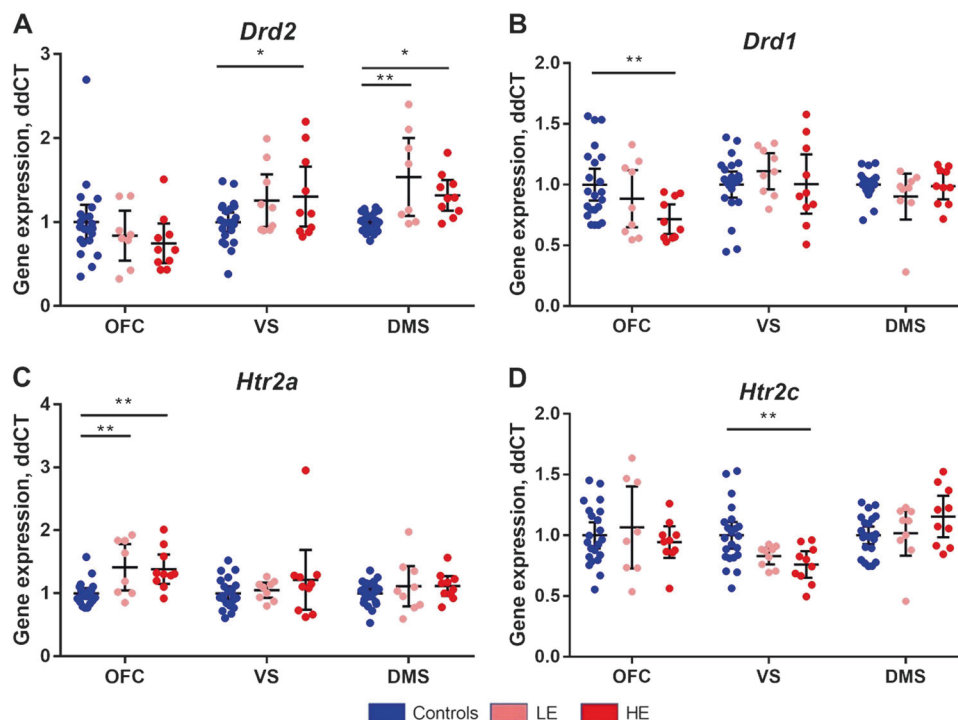
Withdrawal from escalated cocaine self-administration impairs reversal…
P Zhukovsky et al.

2170

Fig. 5 mRNA expression of *DRD2* (**a**) *DRD1* (**b**), *HTR2A* (**c**), and *HTR2C* (**d**) in the orbitofrontal cortex (OFC), ventral striatum (VS), and dorsomedial striatum (DMS) of control ($n = 23$), LE ($n = 9$), and HE (n = 10) rats. *p < 0.05 versus controls. Data are means ± 95% CIs

cocaine escalation rates in highly anxious rats particularly during the first hour of each session. However, unlike our previous study [42], where anxiety was assessed using an elevated plus maze rather than an open field, we found no relationship between trait anxiety and behavioral flexibility. This discrepancy may reflect the different measures of anxiety used in each case and that only 15% of the variance in perseverative errors was explained by trait anxiety in our earlier study. Whereas trait anxiety in humans has long been associated with a preferential bias toward negative external cues [54, 55], and impaired set shifting [56, 57], deficits in task-switching reportedly only clearly manifest when attentional control is challenged in highly anxious individuals [58–60]. Thus, the low attentional load of serial spatial reversal learning involving intra-dimensional rather than extra-dimensional shifting [61] may have impeded the expected relationship between trait anxiety and behavioral flexibility to have been reliably detected in the present study.

An important objective of this research was to investigate the nature of the widely reported impairing effects of cocaine on the flexibility of goal-directed behavior [17, 28, 62–65]. Our finding that lose-shift behavior is decreased in HE rats who had some reversal experience is consistent with findings in rats exposed to methamphetamine [34, 35] and in human addicts [66–68]. Rats self-administering methamphetamine have also been reported to show impaired learning from unrewarded outcomes, resulting in reduced model-free learning after stimulant treatment [34] and after non-contingent methamphetamine administration [35]. In addition, model-based impairments have been reported in rats during reversal [34] and habitual behavior on reinforced learning tasks has been reported in humans [67]. In the present study we used a model-free learning algorithm to explain performance on a spatial serial reversal task. By assessing reversal learning before and after response-contingent cocaine administration and using the reinforcement learning framework of Q-learning [45], we were able to define the effects of prior history of escalated cocaine intake on behavioral flexibility measured 8 days after the end of cocaine treatment. Our results demonstrate that rats in the

control, low, and high cocaine escalation groups learnt from negative and positive feedback on any given trial and appropriately updated internal representations of choice values, as revealed by no significant change in the alpha modelling parameter, both before and after food or cocaine exposure. Nevertheless, important differences became evident in the way the different groups of rats exploited the value assigned to each choice. This was particularly the case for HE rats, which were more likely to perseverate with their previous choice regardless of the received outcomes. As this deficit was not present prior to drug exposure it was likely the consequence of cocaine itself.

Our analysis of lose-shift and win-stay probabilities revealed that those rats more prone to escalate cocaine self-administration subsequently were less likely to switch behavior on trials that were not rewarded in the reversal learning task. This deficit was clearly the consequence of prior cocaine exposure and did not extend to trials with rewarded outcomes. Maladaptive exploration, as indexed by increased beta and kappa values in HE rats, was also significantly associated with decreased lose-shift probabilities (see supplementary table 2), providing a behavioral validation of the modeling parameters. The relationship between lose-shift probabilities and kappa in cocaine and control rats (supplementary table 3) is mathematically plausible since both measures attempt to capture the association between an agent's choice on a given trial with their choice on the previous trial. These findings reveal a hitherto unreported deficit in behavioral flexibility caused by cocaine that was restricted to rats with a greater propensity to escalate cocaine intake. Notably, this deficit was present 8 days after the last cocaine session, suggesting it may be caused by relatively long-lasting neural changes, consistent with other studies [28, 62, 69].

Previous research has shown that stimulant addiction in humans is associated with increased perseveration on reversal learning tasks [70, 71]. The present findings go some way to explaining the nature of this deficit whilst building on the earlier finding that cocaine affects the utilization of expected reward value to guide behavior [1], possibly due to impaired executive

Withdrawal from escalated cocaine self-administration impairs reversal…
P Zhukovsky et al.

2171

control over action selection. Specifically, our computational analysis revealed that rewarded and non-rewarded trials were differentially exploited in rats with a history of high-escalation cocaine self-administration. The finding that rats that more readily escalate cocaine self-administration do so because they become insensitive to the anxiogenic properties of cocaine is consistent with this account [72–74]. However, it should be noted that whilst anxiety predicted increased rates of cocaine self-administration, anxiety per se did not predict the failure of HE rats to exploit negative feedback to guide behavior. Our findings suggest therefore that interactive effects between trait anxiety and cocaine exposure were somehow responsible for the inability of HE rats to exploit reward value during the reversal session.

Reversal learning has been widely shown to depend on monoaminergic mechanisms in the OFC and striatum [75–80] with substantial evidence implicating D2 receptors in the striatal indirect pathway [81–85]. However, rather than decreasing *DRD2* expression in the striatum, as predicted from prior positron emission tomography imaging studies in humans and other rats, [86–88], striatal *DRD2* expression increased significantly in the DMS after 8 days of abstinence from cocaine in LE and HE rats. This effect has been reported before after-cocaine exposure and may reflect a delayed compensatory rebound in D2 receptor regulation [89–91]. Since *DRD2* expression in the DMS increased in both HE and LE rats, this was presumably the consequence of prior cocaine exposure rather than a contributing factor to the failure of HE rats to utilize outcome value during reversal. However, without additional studies to measure protein levels it is unclear whether increased *DRD2* expression resulted in increased D2 receptor density. By contrast, HE rats exhibited increased *DRD2* and decreased *HTR2C* expression in the ventral striatum, with a corresponding reduction in *DRD1* expression in the OFC. However, one should be cautious about linking these differentially-expressed genes for OFC-striatal circuit function and specifically whether they contributed to the failure of HE rats to exploit previously learnt outcome value, especially as qualitatively similar but statistically non-significant effects were also observed in LE rats. Nevertheless, highly-impulsive rats that subsequently developed persistent cocaine-taking in the face of aversive outcomes [92] also exhibited reduced *HTR2C* expression in the ventral striatum after long-access cocaine SA [93]. Since the 5-HT2C receptor has been shown to modulate learning from negative feedback in the context of reversal learning [94, 95], impaired 5-HT2C receptor transmission may have contributed to the failure of HE rats to process negative feedback in the present study.

Synthesis and conclusions
The present findings add to the growing body of evidence that cocaine impairs how negative feedback is used to guide behavior. Using traditional and computational methods of analysis, we report that rats exposed to response-contingent cocaine, and which more rapidly escalate intake, were able to learn the value of changing reward contingencies but were compromised in exploiting this knowledge to guide appropriate actions on a serial reversal task. Previous research has shown that the encoding of expected outcomes to acquired values depends on interactions between the OFC and basolateral amygdala [96] and that cocaine disrupts insight into the consequences of behavior by OFC-dependent mechanisms [47]. Abnormalities within this circuitry may thus be relevant to understanding why individuals addicted to drugs persist with drug consumption despite adverse consequences of continued drug use.

REFERENCES
1. Schoenbaum G, Setlow B. Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. Cereb Cortex. 2005;15:1162–9.
2. Miles FJ, Everitt BJ, Dickinson A. Oral cocaine seeking by rats: action or habit? Behav Neurosci. 2003;117:927–38.
3. Hogarth L, He Z, Chase HW, Wills AJ, Troisi J, Leventhal AM, et al. Negative mood reverses devaluation of goal-directed drug-seeking favouring an incentive learning account of drug dependence. Psychopharmacol (Berl). 2015;232:3235–47.
4. Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol Psychiatry. 2012;72:389–95.
5. Vanderschuren LJMJ, Everitt BJ. Drug seeking becomes compulsive after prolonged cocaine self-administration. Science . 2004;305:1017–9.
6. Belin D, Everitt BJ. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. Neuron. 2008;57:432–41.
7. Pelloux Y, Everitt BJ, Dickinson A. Compulsive drug seeking by rats under punishment: effects of drug taking history. Psychopharmacol (Berl). 2007;194:127–37.
8. Everitt BJ, Giuliano C, Belin D. Addictive behaviour in experimental animals: prospects for translation. Philos Trans R Soc B Biol Sci. 2018;373:pii: 20170027.
9. Rolls ET, Critchley HD, Mason R, Wakeman EA. Orbitofrontal cortex neurons: role in olfactory and visual association learning. J Neurophysiol. 1996;75:1970–81.
10. Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. Nature. 2006;441:223–6.
11. Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, Schoenbaum G. Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat Commun. 2014;5:1–13.
12. Rudebeck PH, Murray EA. The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Urology. 2014;84:1143–56.
13. Wallis JD. Orbitofrontal cortex and its contribution to decision-making. Annu Rev Neurosci. 2007;30:31–56.
14. Blanchard TC, Hayden BY, Bromberg-Martin ES. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron. 2015;85:602–14.
15. Mansouri FA, Buckley MJ, Tanaka K. The essential role of primate orbitofrontal cortex in conflict-induced executive control adjustment. J Neurosci. 2014;34:11016–31.
16. Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. Nature. 1996;380:69–72.
17. Jentsch JD, Olausson P, De La Garza R, Taylor JR. Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. Neuropsychopharmacology. 2002;26:183–90.
18. Fellows LK, Farah MJ. Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. Brain. 2003;126:1830–7.

Withdrawal from escalated cocaine self-administration impairs reversal...
P Zhukovsky et al.

2172

19. McAlonan K, Brown VJ. Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. Behav Brain Res. 2003;146:97–103.

20. London ED, Ernst M, Grant S, Bonson K, Weinstein A. Orbitofrontal cortex and human drug abuse: functional imaging. Cereb Cortex. 2000;10:334–42.

21. Volkow ND, Chang L, Wang G-J, Fowler JS, Ding Y-S, Sedler M, et al. Low level of brain dopamine D2 receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. Am J Psychiatry. 2001;158:2015–21.

22. Dom G, Sabbe B, Hulstijn W, van den Brink W. Substance use disorders and the orbitofrontal cortex. Br J Psychiatry. 2005;187:209–20.

23. Everitt BJ, Hutcheson DM, Ersche KD, Pelloux Y, Dalley JW, Robbins TW. The orbital prefrontal cortex and drug addiction in laboratory animals and humans. Ann N Y Acad Sci. 2007;1121:576–97.

24. Schoenbaum G, Shaham Y. The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. Biol Psychiatry. 2008;63:256–62.

25. Fettes P, Schulze L, Downar J. Cortico-striatal-thalamic loop circuits of the orbitofrontal cortex: promising therapeutic targets in psychiatric illness. Front Syst Neurosci. 2017;11:1–23.

26. Moorman DE. The role of the orbitofrontal cortex in alcohol use, abuse, and dependence. Prog Neuropsychopharmacol Biol Psychiatry. 2018;87:85–107.

27. Stalnaker Ta, Takahashi Y, Roesch MR, Schoenbaum G. Neural substrates of cognitive inflexibility after chronic cocaine exposure. Neuropharmacology. 2009; 56:63–72.

28. McCracken CB, Grace AA. Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization. J Neurosci. 2013;33:17469–82.

29. Izquierdo A, Belcher AM, Scott L, Cazares VA, Chen J, O'Dell SJ, et al. Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: possible involvement of striatal dopamine. Neuropsychopharmacology. 2010;35: 505–14.

30. Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. Class Cond II Curr Res Theory. 1972;21:64–99.

31. Miller RR, Barnet RC, Grahame NJ. Assessment of the Rescorla-Wagner Model. Psychol Bull. 1995;117:363–86.

32. Dayan P. Dopamine, reinforcement learning, and addiction. Pharmacopsychiatry. 2009;42:S56–65.

33. Keramati M, Durand A, Girardeau P, Gutkin B, Ahmed SH. Cocaine addiction as a homeostatic reinforcement learning disorder. Psychol Rev. 2017;124:130–53.

34. Groman SM, Massi B, Mathias SR, Lee D, Taylor JR. Model-free and model-based influences in addiction-related behaviors. Biol Psychiatry. 2019;4:pii: S0006–3223. https://doi.org/10.1016/j.biopsych.2018.12.017

35. Groman SM, Rich KM, Smith NJ, Lee D, Taylor JR. Chronic exposure to methamphetamine disrupts reinforcement-based decision making in rats. Neuropsychopharmacology. 2018;43:770–80.

36. Huys QJM, Maia TV, Paulus MP. Computational psychiatry: from mechanistic insights to the development of new treatments. Biol Psychiatry Cogn Neurosci Neuroimaging. 2016;1:382–5.

37. Adams RA, Huys QJM, Roiser JP. Computational Psychiatry: towards a mathematically informed understanding of mental illness. J Neurol Neurosurg Psychiatry. 2016;87:53–63.

38. Stephan KE, Iglesias S, Heinzle J, Diaconescu AO. Translational perspectives for computational neuroimaging. Neuron. 2015;87:716–32.

39. Homberg JR, Akker M Van den, Raasø HS, Wardeh G, Binnekade R, Schoffelmeer ANM, et al. Enhanced motivation to self-administer cocaine is predicted by self-grooming behaviour and relates to dopamine release in the rat medial prefrontal cortex and amygdala. Eur J Neurosci. 2002;15:1542–50.

40. Walker QD, Schramm-Sapyta NL, Caster JM, Waller ST, Brooks MP, Kuhn CM. Novelty-induced locomotion is positively associated with cocaine ingestion in adolescent rats; anxiety is correlated in adults. Pharmacol Biochem Behav. 2009;91:398–408.

41. Dilleen R, Pelloux Y, Mar AC, Molander A, Robbins TW, Everitt BJ, et al. High anxiety is a predisposing endophenotype for loss of control over cocaine, but not heroin, self-administration in rats. Psychopharmacol (Berl). 2012;222:89–97.

42. Zhukovsky P, Alsiö J, Jupp B, Xia J, Guiliano C, Jenner L, et al. Perseveration in a spatial-discrimination serial reversal learning task is differentially affected by MAO-A and MAO-B inhibition and associated with reduced anxiety and peripheral serotonin levels. Psychopharmacology (Berl). (2017). https://doi.org/10.1007/s00213-017-4569-x

43. Barlow RL, Alsiö J, Jupp B, Rabinovich R, Shrestha S, Roberts AC, et al. Markers of serotonergic function in the orbitofrontal cortex and dorsal Raphé nucleus predict individual variation in spatial-discrimination serial reversal learning. Neuropsychopharmacology. 2015;40:1619–30.

44. Gould TD, Dao DT, Kovacsics CE. Mood and anxiety related phenotypes in mice. Neuromethods. 2009;42:1–20.

45. Daw, N. D. Trial-by-trial data analysis using computational models. Decis. Making, Affect. Learn. Atten. Perform. XXIII (2011). https://doi.org/10.1093/acprof:oso/9780199600434.003.0001

46. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008;3:1101–8.

47. Lucantonio F, Takahashi YK, Hoffman AF, Chang C, Bali-Chaudhary S, Shaham Y, et al. Erratum: orbitofrontal activation restores insight lost after cocaine use. Nat Neurosci. 2014;17:1287–1287.

48. Norton GR. Substance use/abuse and anxiety sensitivity: what are the relationships? Addict Behav. 2001;26:935–46.

49. Sinha R. How does stress increase risk of drug abuse and relapse? Psychopharmacol (Berl). 2001;158:343–59.

50. Lejuez CW, Zvolensky MJ, Daughters SB, Bornovalova MA, Paulson A, Tull MT, et al. Anxiety sensitivity: a unique predictor of dropout among inner-city heroin and crack/cocaine users in residential substance use treatment. Behav Res Ther. 2008;46:811–8.

51. Ersche KD, Turton AJ, Chamberlain SR, Müller U, Bullmore ET, Robbins TW. Cognitive dysfunction and anxious-impulsive personality traits are endophenotypes for drug dependence. Am J Psychiatry. 2012;169:926–36.

52. Khantzian EJ. The self medication hypothesis of addictive disorders: focus on heroin and cocaine dependence. Am J Psychiatry. 1985;142:1259–64.

53. Pelloux Y, Costentin J, Duterte-Boucher D. Anxiety increases the place conditioning induced by cocaine in rats. Behav Brain Res. 2009;197:311–6.

54. Fox E, Russo R, Bowles R, Dutton K. Do threatening stimuli draw or hold visual attention in subclinical anxiety? J Exp Psychol Gen. 2001;130:681–700.

55. Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, Van Ijzendoorn MH. Threat-related attentional bias in anxious and nonanxious individuals: ameta-analytic study. Psychol Bull. 2007;133:1–24.

56. Goodwin AH, Sher KJ. Deficits in set-shifting ability in nonclinical compulsive checkers. J Psychopathol Behav Assess. 1992;14:81–92.

57. Caselli RJ, Reiman EM, Hentz JG, Osborne D, Alexander GE. A distinctive interaction between chronic anxiety and problem solving in asymptomatic APOE e4 homozygotes. J Neuropsychiatry Clin Neurosci. 2004;16:320–9.

58. Derakshan N, Smyth S, Eysenck MW. Effects of state anxiety on performance using a task-switching paradigm: an investigation of attentional control theory. Psychon Bull Rev. 2009;16:1112–7.

59. Berggren N, Derakshan N. Blinded by fear? Prior exposure to fearful faces enhances attentional processing of task-irrelevant stimuli. Q J Exp Psychol. 2013; 66:2204–18.

60. Berggren N, Richards A, Taylor J, Derakshan N. Affective attention under cognitive load: reduced emotional biases but emergent anxiety-related costs to inhibitory control. Front Hum Neurosci. 2013;7:1–7.

61. Robbins TW, Arnsten AFT. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. Annu Rev Neurosci. 2009;32:267–87.

62. Calu DJ, Stalnaker, A. T, Franz TM, Singh T, Shaham Y, et al. Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependant reversal learning in rats. Learn Mem. 2007;14:325–8.

63. Ersche KD, Roiser JP, Robbins TW, Sahakian BJ. Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacol (Berl). 2008;197:421–31.

64. Moreno-López L, Perales JC, Son D van, Albein-Urios N, Soriano-Mas C, Martinez-Gonzalez JM, et al. Cocaine use severity and cerebellar gray matter are associated with reversal learning deficits in cocaine-dependent individuals. Addict Biol. 2015;20:546–56.

65. Cox BM, Cope ZA, Parsegian A, Floresco SB, Aston-Jones G, See RE. Chronic methamphetamine self-administration alters cognitive flexibility in male rats. Psychopharmacol (Berl). 2016;233:2319–27.

66. Parvaz MA, Konova AB, Proudfit GH, Dunning JP, Malaker P, Moeller SJ, et al. Impaired neural response to negative prediction errors in cocaine addiction. J Neurosci. 2015;35:1872–9.

67. Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LHE, Luijten M, et al. Carrots and sticks fail to change behavior in cocaine addiction. Science. 2016;352: 1468–71.

68. Ekhtiari H, Victor TA, Paulus MP. Aberrant decision-making and drug addiction—how strong is the evidence? Curr Opin Behav Sci. 2017;13:25–33.

69. Bechard AR, LaCrosse A, Namba MD, Jackson B, Knackstedt LA. Impairments in reversal learning following short access to cocaine self-administration. Drug Alcohol Depend. 2018;192:239–44.

70. Chamberlain SR, Sahakian BJ. The neuropsychology of mood disorders. Curr Psychiatry Rep. 2006;8:458–63.

71. Ersche KD, Roiser JP, Abbott S, Craig KJ, Mller U, Suckling J, et al. Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a D 2/3 receptor agonist. Biol Psychiatry. 2011;70:754–62.

Withdrawal from escalated cocaine self-administration impairs reversal...
P Zhukovsky et al.

2173

72. Rogerio R, Takahashi RN. Anxiogenic properties of cocaine in the rat evaluated with the elevated plus-maze. Pharmacol Biochem Behav. 1992;43:631–3.

73. Ettenberg A, Raven MA, Danluck DA, Necessary BD, Ettenberg A, Raven MA, et al. Evidence for opponent-process actions of intravenous cocaine. Pharmacol Biochem Behav. 1999;64:507–12.

74. David V, Gold LH, Koob GF, Cazala P. Anxiogenic-like effects limit rewarding effects of cocaine in BALB/cByJ mice. Neuropsychopharmacology. 2001;24:300–18.

75. Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC. Cognitive inflexibility after prefrontal serotonin depletion. Science. 2004;304:878–80.

76. Haluk DM, Floresco SB. Ventral striatal dopamine modulation of different forms of behavioral flexibility. Neuropsychopharmacology. 2009;34:2041–52.

77. Roberts AC. The importance of serotonin for orbitofrontal function. Biol Psychiatry. 2011;69:1185–91.

78. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol Psychiatry. 2011;69:e113–25.

79. Clarke HF, Hill GJ, Robbins TW, Roberts AC. Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. J Neurosci. 2011;31:4290–7.

80. Floresco SB. Prefrontal dopamine and behavioral flexibility: shifting from an 'inverted-U' toward a family of functions. Front Neurosci. 2013;7:1–12.

81. Holstein M Van, Aarts E, Schaaf ME Van Der, Geurts DEM, Verkes RJ, Franke B, et al. Human cognitive flexibility depends on dopamine D2 receptor signaling. Psychopharmacol (Berl). 2011;218:567–78.

82. Groman SM, Lee B, Seu E, James AS, Feiler K, Mandelkern MA, et al. Dysregulation of D(2)-mediated dopamine transmission in monkeys after chronic escalating methamphetamine exposure. J Neurosci. 2012;32:5843–52.

83. MacPherson T, Morita M, Wang Y, Sasaoka T, Sawa A, Hikida T. Nucleus accumbens dopamine D2-receptor expressing neurons control behavioral flexibility in a place discrimination task in the IntelliCage. Learn Mem. 2016;23:359–64.

84. Morita M, Wang Y, Sasaoka T, Okada K, Niwa M, Sawa A, et al. Dopamine D2L receptor is required for visual discrimination and reversal learning. Mol Neuropsychiatry. 2016;2:124–32.

85. Linden J, James AS, McDaniel C, Jentsch JD. Dopamine D2 receptors in dopaminergic neurons modulate performance in a reversal learning task in mice. Eneuro. 2018;5:ENEURO.0229-17.2018.

86. Volkow ND, Fowler JS, Wang GJ, Hitzemann R, Logan J, Schlyer DJ, et al. Decreased dopamine-d(2) receptor availability is associated with reduced frontal metabolism in cocaine abusers. Synapse. 1993;14:169–77.

87. Nader MA, Czoty PW. PET imaging of dopamine D2 receptors in monkey models of cocaine abuse: Genetic predisposition versus environmental modulation. Am J Psychiatry. 2005;162:1473–82.

88. Dalley JW, Everitt BJ, Robbins TW. Impulsivity, compulsivity, and top-down cognitive control. Neuron. 2011;69:680–94.

89. Przewłocka B, Lasoń W. Adaptive changes in the proenkephalin and D2 dopamine receptor mRNA expression after chronic cocaine in the nucleus accumbens and striatum of the rat. Eur Neuropsychopharmacol. 1995;5:465–9.

90. Schmidt-Mutter C, Muller C, Zwiller J, Gobaille S, Maitre M. Gamma-hydroxybutyrate and cocaine administration increases mRNA expression of dopamine D1 and D2 receptors in rat brain. Neuropsychopharmacology. 1999;21:662–9.

91. Belin D, Deroche-Gamonet V, Jaber M. Cocaine-induced sensitization is associated with altered dynamics of transcriptional responses of the dopamine transporter, tyrosine hydroxylase, and dopamine D2 receptors in C57Bl/6J mice. Psychopharmacol (Berl). 2007;193:567–78.

92. Belin D, Mar AC, Dalley JW, Robbins TW, Everitt BJ. High impulsivity predicts the switch to compulsive cocaine-taking. Science. 2008;320:1352–5.

93. Besson M, Pelloux Y, Dilleen R, Theobald DE, Lyon A, Belin-Rauscent A, et al. Cocaine modulation of frontostriatal expression of Zif268, D2, and 5-HT2c receptors in high and low impulsive rats. Neuropsychopharmacology. 2013;38:1963–73.

94. Phillips BU, Dewan S, Nilsson SRO, Robbins TW, Heath CJ, Saksida LM, et al. Selective effects of 5-HT2C receptor modulation on performance of a novel valence-probe visual discrimination task and probabilistic reversal learning in mice. Psychopharmacol (Berl). 2018;235:2101–11.

95. Nilsson SRO, Alsiö J, Somerville EM, Clifton PG. The rat's not for turning: dissociating the psychological components of cognitive inflexibility. Neurosci Biobehav Rev. 2015;56:1–14.

96. Schoenbaum G, Setlow B, Ramus SJ. A systems approach to orbitofrontal cortex function: recordings in rat orbitofrontal cortex reveal interactions with different learning systems. Behav Brain Res. 2003;146:19–29.