Establishing a probabilistic reversal learning test in mice: Evidence for the processes mediating reward-stay and punishment-shift behaviour and for their modulation by serotonin

Christian Ineichen [a], Hannes Sigrist [a], Simona Spinelli [a], Klaus-Peter Lesch [b], Eva Sautter [c], Erich Seifritz [d], Christopher R. Pryce [a],*

[a] Preclinical Laboratory for Translational Research into Affective Disorders, Clinic for Affective Disorders & General Psychiatry, Psychiatric University Hospital Zurich, August Forel-Strasse 7, CH-8008 Zurich, Switzerland
[b] Molecular Psychiatry, Laboratory of Translational Neuroscience, Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany
[c] TSE Systems GmbH, Bad Homburg, Germany
[d] Clinic for Affective Disorders & General Psychiatry, Psychiatric University Hospital Zürich, Zürich, Switzerland

ARTICLE INFO

ABSTRACT

Valid animal models of psychopathology need to include behavioural readouts informed by human findings. In the probabilistic reversal learning (PRL) task, human subjects are confronted with serial reversal of the contingency between two operant stimuli and reward/punishment and, superimposed on this, a low probability (0.2) of punished correct responses/rewarded incorrect responses. In depression, reward-stay and reversals completed are unaffected but response-shift following punished correct response trials, referred to as negative feedback sensitivity (NFS), is increased. The aims of this study were to: establish an operant spatial PRL test appropriate for mice; obtain evidence for the processes mediating reward-stay and punishment-shift responding; and assess effects thereon of genetically- and pharmacologically-altered serotonin (5-HT) function. The study was conducted with wildtype (WT) and heterozygous mutant (HET) mice from a 5-HT transporter (5-HTT) null mutant strain. Mice were mildly food deprived and reward was sugar pellet and punishment was 5-s time out. Mice exhibited high motivation and adaptive reversal performance. Increased probability of punished correct response (PCR) trials per session ($p = 0.1, 0.2$ or 0.3) led to monotonic decrease in reward-stay and reversals completed, suggesting accurate reward prediction. NFS differed from chance-level at p PCR $= 0.1$, suggesting accurate punishment prediction, whereas NFS was at chance-level at $p = 0.2–0.3$. At p PCR $= 0.1$, HET mice exhibited lower NFS than WT mice. The 5-HTT blocker escitalopram was studied acutely at p PCR $= 0.2$: a low dose (0.5–1.5 mg/kg) resulted in decreased NFS, increased reward-stay and increased reversals completed, and similarly in WT and HET mice. This study demonstrates that testing PRL in mice can provide evidence on the regulation of reward and punishment processing that is, albeit within certain limits, of relevance to human emotional-cognitive processing, its dysfunction and treatment.

1. Introduction

Altered emotional-cognitive processing of negative and positive events is central to depression psychopathology, as indicated by the two core diagnostic symptoms of depressed mood (sadness, emptiness) and loss of interest or pleasure (anhedonia) (APA, 2000). Depressed mood reflects high focus on past, current or potential future negative (aversive) events, which elicits feelings and states such as sadness, frustration and catastrophisation (Abramson et al., 1989; Eshel and Roiser, 2010; Nandrino et al., 2004). Anhedonia reflects low focus on past, current, and potential future positive events (rewards), which leads to reduced pleasure and interest (Berridge and Robinson, 2003; Henriques and Davidson, 2000; Sloan et al., 2001). Automated psychological (operant) tasks that require development of cognitive associations between stimuli and outcomes allow for quantitative assessment of

affective responsiveness to negative events ("error" feedback) and positive events ("correct" feedback). It has been reported that depression is characterised by high sensitivity to negative feedback: on memory and planning tasks, when patients receive error feedback on a trial they are more likely to also make another incorrect decision on the next trial, relative to healthy subjects. That is, depression is associated with increased emotional-cognitive reactivity to error feedback, possibly due to exaggerated punishment expectancy/prediction, manifested as increased likelihood of additional errors on subsequent trials (Elliott et al., 1996, 1997).

The probabilistic reversal learning (PRL) task is specifically designed to assess cognitive-emotional ability to develop appropriate expectations/predictions about stimulus–feedback associations on the basis of a combination of accurate and misleading feedback (Chamberlain et al., 2006; Cools et al., 2002; Evers et al., 2005; Jocham et al., 2009). Using two-way operant-stimulus or spatial-response discrimination on a computer touch-screen, the subject is instructed to select one stimulus on each trial in order to maximize correct feedback in the form of rewards and to minimize error feedback in the form of punishers. Rewards and punishers can be symbolic or monetary. The identity of the correct stimulus is reversed when a criterion of consecutive correct responses is attained. The reversal criterion varies within certain limits and is to an extent predictable, therefore. In addition to the accurate correct and accurate error feedback that the subject experiences, at a certain probability (e.g. 0.2) correct and incorrect responses receive misleading error and correct feedback, respectively. The subject is informed that at any one time one stimulus will *usually* be correct. Of particular interest are the subject's decisions on trials that follow misleading error feedback i.e. when the subject is punished for a correct response. The proportion of such trials on which the subject shifts on the next trial to the incorrect stimulus i.e. punishment-shift-punishment responses, provides a measure of punishment expectancy/prediction and is referred to as negative feedback sensitivity (NFS). High NFS is indicative of a cognitive-emotional over-estimation of punishment expectancy/prediction for that stimulus. Depressed patients, whilst being largely unimpaired in the acquisition and reversal of probabilistic reversal in the PRL task, exhibit 2–3 times higher NFS than do healthy controls (Murphy et al., 2003; Taylor Tavares et al., 2008).

There is a current deficiency of animal tests for detailed study of depression-relevant emotional-cognitive processes (Pryce and Seifritz, 2011). This is despite the clinical and therapeutic importance of altered emotional-cognitive processing in depression psychopathology, the availability of non-verbal emotional-cognitive tasks such as PRL to quantify such states, and the potential for these tasks to be adapted to animal species. One exception is indeed the PRL task, a version of which has been recently described for the rat (Bari et al., 2010). This rat automated PRL test is based on operant responding (nose poking) in a two-way spatial (rather than stimulus) discrimination, and a 0.2 probability for both correct and incorrect responses being followed by inaccurate (i.e. non-reversal) negative and positive feedback, respectively. Subjects were food deprived to induce high motivation for correct responding with sugar pellet reinforcement, and incorrect responses received error feedback in the form of no reward plus a time-out delay until the next trial. Compared to human performance, rats exhibited relatively low reward-stay behaviour ($p = 0.6–0.8$ versus close to 1.0 in human) and relatively high NFS ($p = 0.4–0.6$ versus 0.1 in human) and across 200 trials achieved 2–3 reversals on a reversal contingency of eight consecutive responses (Bari et al., 2010).

The first aim of the present study was to attempt to establish a mouse automated PRL test based on nose-poke responding in a two-way spatial discrimination as used in the rat. Mice were trained to a high level of reversal learning performance prior to

introducing PRL conditions. Then the effects of different per session probabilities of punished correct responding on reward-stay, punishment-shift and reversals completed, were studied. It was reasoned that these data would provide insights into the emotional-cognitive processes underlying PRL behaviour in mouse. One assumption underlying the human PRL task is that subjects are cognitively able to acquire accurate reward and punishment expectancy/prediction and thereby to exhibit both (1) accurate rule reversal learning and (2) a probabilistic strategy of maximising correct responses without frequent shifting from the just-punished stimulus (Evers et al., 2005; Murphy et al., 2003). Compared to the human data, relatively low reward-stay probability and relatively high NFS were observed in the rat study (Bari et al., 2010). These findings suggest that applying human PRL test parameters one-to-one to rodents could run counter to the assumption that subjects are cognitively able to maintain accurate reward and punishment expectancy. Extrapolating these important rat findings to the current mouse study, in an attempt to maximize the likelihood of establishing a mouse automated PRL test where subjects are able to acquire accurate reward and punishment expectancy/prediction, the protocol was *a priori* designed to be cognitively less demanding (i.e. less inaccurate/misleading feedback) than the human task. Therefore, misleading feedback was given exclusively in the form of punished correct responding and no rewarded incorrect response trials were used. It is punished correct responding (and not rewarded incorrect responding) which is the basis of one of the major PRL parameters, NFS. In fact, there is at least one example of a human PRL task that is also without rewarded incorrect responses (Ersche et al., 2008). A recent study of a mouse manual PRL test that used a spatial maze and probability of 0.2 for both correct and incorrect responses being followed by inaccurate feedback, reports that reward-stay was only $p = 0.5$, i.e. chance level (Amodeo et al., 2012), thereby providing further evidence for the utility of reducing cognitive demand. Therefore, whilst not using rewarded incorrect response trials is a proviso that needs to be taken into account when comparing the findings of the current study with existing rat and mouse PRL data, these existing data indicate the need for this adjustment in order to better satisfy the assumption of accurate reward/punishment expectancy underlying the human PRL test and thereby to increase the relevance of mouse PRL findings to human.

The second aim of the study was to investigate the effects of manipulation of serotonergic function on mouse behaviour in the PRL test. Aversive stimuli (punishments) induce 5-HT release in cortico-limbic-striatal regions (Millan, 2003). Accordingly, 5-HT has been proposed as a major mediator of adaptive behavioural responses to punishing stimuli, including those experienced in tasks such as PRL (Boureau and Dayan, 2011; Cools et al., 2011, 2008;). Genetic polymorphisms that impact on 5-HT signalling could also impact on aversive stimulus processing. For example, the "short" form of the 5-HT transporter-linked polymorphic region of the gene (SLC6A4, 5-HTT) is associated with decreased 5-HTT/increased 5-HT activity, increased reactivity to aversive stimuli, and is a risk factor for depression (Canli and Lesch, 2007; Caspi et al., 2003; Hariri et al., 2002). One of the major hypotheses for depression pathophysiology is 5-HT deficiency (Sharp and Cowen, 2011), and altered sensitivity to rewarding and punishing feedback could mediate the effects of 5-HT deficiency on depression psychopathology. Accordingly, a mouse translational PRL test for the study of effects of 5-HT manipulations would be beneficial. The present study was conducted with wildtype (WT) and heterozygous mutant (HET) mice from a 5-HTT (Slc6a4) null mutant strain. Relative to WT, the 5-HTT HET mouse exhibits reduced 5-HT clearance (Montanez et al., 2003) and increased extracellular 5-HT levels (Mathews et al., 2004), but otherwise normal 5-HT

transmission (Jennings et al., 2010). It provides a model for human 5-HTT polymorphisms that are associated with reduced 5-HTT function (Murphy and Lesch, 2008). The selective 5-HTT blocker citalopram, when administered acutely as a tool compound, was found to increase NFS in healthy humans (Chamberlain et al., 2006) and to increase and decrease NFS at low and high doses, respectively, in rat (Bari et al., 2010). In the present study the potent 5-HTT blocker escitalopram was used as a tool compound in order to investigate the effects of acute, specific blocking of 5-HTT in WT and HET mice, using the acute doses demonstrated to be effective in mouse antidepressant- and anxiolytic-screening tests (Sanchez et al., 2003). Therefore, the effects of genetically- and pharmacologically-induced reduced 5-HTT function were studied in a novel mouse PRL test. It is important to note that the present study did not aim to produce a mouse model of depression-relevant deficits in PRL behaviour, but rather to establish and validate a mouse PRL test. The major application of such a test would then be to investigate whether environmental and genetic manipulations induce depression-relevant deficits in PRL behaviour, as observed in human depression, and to use such a model to study their neuropharmacological reversal.

2. Materials and methods

2.1. Animals

Male and female mice of a 5-HTT null mutant strain on a C57BL/6J background (>20 backcross generations) were transferred from the University of Würzburg (Bengel et al., 1998) and breeding was established in-house with WT dams and HET sires. Male offspring were weaned at age 4 weeks and caged as brother pairs throughout the study. Mice were maintained on a reversed 12:12 h light–dark cycle (white lights off at 07:00 h) in an individually-ventilated cage system, with temperature at 20–22 °C and humidity at 50–60%. Each cage measured 43 (L) × 27 (W) × 19 (H) cm, with wood chip bedding and an igloo. The standard diet was complete pellet (Provimi, Kliba Ltd, Kaiseraugst, Switzerland) and water, both available continuously and *ad libitum* until age 10 weeks. The study was conducted with 15 WT and 15 HET mice, born to seven different breeding pairs. All procedures were conducted under a permit for animal experimentation issued by the Veterinary Office, Zurich, Switzerland, in accordance with the Animal Protection Act (1978) Switzerland. All efforts were made to minimise the number of mice used and any suffering of those mice that were used. Whilst it will be important to study PRL behaviour in both male and female mice, the present study was conducted with males only. As the drug study used a cross-over design (see below) it was essential to exclude any potential confounding effects of the oestrous cycle on PRL behaviour. The rat PRL study was also conducted with males specifically (Bari et al., 2010).

2.2. Restricted feeding protocol

Mice were reduced to 90% of their free-feeding adult body weight. To achieve this, 100% food intake (g) per brother-pair was measured daily for two weeks. Mice were reduced to 75% of their average daily intake and were also fed sugar pellets either in the home cage prior to onset of the training phase or earned during training/testing. On days without training/testing, mice were fed 100% free-feeding food intake. Initial body weight was 24.5 ± 2.5 g (mean ± SD) and experimental body weight was 23.1 ± 1.3 g.

2.3. Operant apparatus

Each operant box (TSE Systems GmbH, Bad Homburg, Germany) had inner dimensions of 20 × 17 × 18 cm. The back and side walls and ceiling were out of stainless steel and the floor was transparent Plexiglas with perforations to allow urine to flow into a waste tray. The front wall was transparent Plexiglas and was also a forward-opening door. Situated in one side wall were two nose-poke ports (∅ = 20 mm × depth 30 mm) that detected nose pokes via an infra-red beam. A white lamp was set into the recess of each port and could be illuminated to indicate it was active. The nose-poke ports were situated left and right of a central feeder port. Sugar pellets (14 mg, Dustless Precision Pellets, TSE Systems GmbH) were delivered singly into the feeder port (∅ = 20 mm × depth 35 mm). Pellet delivery was signalled by a tone (conditioned stimulus (CS)) from a speaker located above the feeder port. Pellet retrieval was detected via infra-red beam. The centre-to-centre distance between the nose-poke ports was 107 mm and the centre-to-centre distance between each nose-poke port and the central feeder port was 55 mm. A house light provided 30 lux during training and test sessions. All elements were

removable for cleaning. Four such operant boxes were run in parallel from one control PC and interface. Dedicated programs controlled and recorded all experimental events and output.

2.4. Operant training

Mice were taken out of the colony room into the adjacent testing room, under dim illumination. Each mouse was allocated to a specific operant box. The consecutive training stages are described below. Autoshaping and operant learning were conducted in daily sessions Monday–Sunday and thereafter training was conducted in daily sessions Monday–Friday.

Autoshaping. At an inter-trial-interval (ITI) of 45 s, a 1-s tone CS signalled delivery of a sugar pellet into the feeder port. The first session was 20 min and 24 pellets were delivered; thereafter sessions were 30 min and 39 pellets. Criterion to proceed to operant training was 2 sessions with ≥30 pellets retrieved and eaten. For all subsequent stages the following parameters were applied. After each correct operant response: the tone CS was presented; if the nose-poke port light(s) was in use it was switched off; subsequent nose pokes were registered but were without consequence; the mouse had to retrieve the pellet from the feeder to initiate the ITI of 2.5 s; during the ITI, nose-poke port light(s) remained off; onset of the subsequent trial was indicated by switching on nose-poke port light(s). After each incorrect nose poke response: if the nose-poke port light(s) was in use it was switched off; a punishment ITI of 5.0 s was initiated; nose pokes were registered but were without consequence; onset of the subsequent trial was indicated by switching on nose-poke port light(s).

Nose-poke learning. Sugar pellets (5, 3 or 1, across sessions 1–3) were placed in the nose-poke ports. One or more nose pokes into either left or right port triggered the tone CS and 1 pellet delivery. Session duration was 30 min or 40 reinforcements maximum. Criterion to proceed to the next stage was 1 session with ≥20 pellets earned and eaten.

Nose-poke learning with light. Both nose-poke ports were illuminated. One or more nose pokes into either port triggered the tone CS and 1 pellet delivery. Criterion to proceed to the next stage was 2 consecutive sessions with ≥30 pellets retrieved and eaten.

Reversal learning (Fig. 1A). Reversal learning comprised four consecutive stages:

(1) Only the active ("correct") nose-poke port was illuminated, so that the light served as a discriminatory stimulus (S^D). A nose poke at the active port triggered the tone CS and 1 pellet delivery, and into the inactive ("incorrect") port triggered extinguishing of the S^D in the correct port and a 5.0 s punishment time out i.e. light off used as $S\Delta$. After 5 correct responses on 5 (not necessarily consecutive) trials, reversal took place, i.e. position of correct and incorrect ports was reversed. Session duration was 30 min or 40 reinforcements, so that a maximum of 8 reversals was possible. Criteria to proceed to the next stage were 2 consecutive sessions where the mouse obtained all 40 reinforcements and reward-shift errors (i.e. trials with response to the incorrect port immediately after a correct trial/total trials immediately after a correct trial) set to maximum 5 per port ($p \leq 0.2$).

(2) The maximum available reinforcements were increased to 48. Criteria to proceed to the next stage were 2 consecutive sessions where the mouse obtained all 48 reinforcements and reward-shift errors set to maximum 6 per port ($p \leq 0.2$).

(3) Both nose-poke ports were illuminated. Criteria to proceed to the next stage was 2 consecutive sessions where the mouse obtained all 48 reinforcements and reward-shift errors set to maximum 6 per port ($p \leq 0.2$).

(4) Reversal criterion was increased to 8 correct consecutive trials, so that a maximum of 6 reversals were possible. Criteria to proceed to the next stage were 1 session where the mouse obtained all 48 reinforcements, reward-shift errors set to maximum 6 per port ($p \leq 0.2$), and a minimum of 3 reversals completed. Mice then progressed to the PRL test.

2.5. Probabilistic reversal learning test (Fig. 1B)

Both nose-poke ports were illuminated. Session duration was 30 min or 60 reinforcements. Reversal criterion was 8 correct consecutive trials, so that a maximum of 7 reversals were possible. A proportion of correct responses was punished by 5.0 s time out. Additional parameters were: the first correct response per session was always rewarded; the maximum number of consecutive punished correct responses was 2. For PRL test validation, test sessions with different overall probability of punished correct response (PCR) trials were used. Thus each mouse was given 4 PRL tests, one test each at a PCR trial probability of 0.0, 0.1, 0.2 and 0.3, e.g. at $p = 0.1$ PCR, 6 of 60 correct responses were punished. A latin-square cross-over design was used to control for any effect of test order. For the drug study, escitalopram was tested at $p = 0.2$ PCR. Subjects were tested daily Monday–Friday, with 2 PRL tests per week (Tuesdays, Fridays) and with intervening sessions of reversal learning. After each PRL test, mice had to attain criterion (see above) on at least one reversal learning session to proceed to the next PRL test.
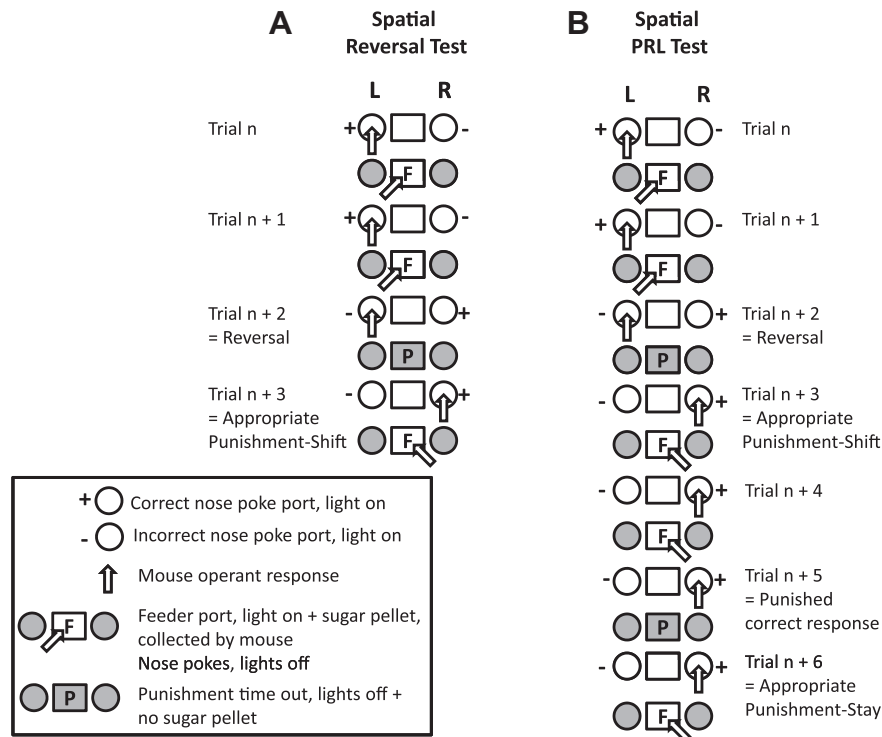
Fig. 1. Schematic for the layout and procedure for (A) the spatial reversal test and (B) the spatial probabilistic reversal learning test. L and R indicate left and right nose-poke ports, respectively.

Measures of interest were: probability of reward-stay responding (p (trials with response to nose-poke port that was correct on previous trial/total trials immediately following a correct trial)); negative feedback sensitivity (p (trials with response to the opposite nose-poke port to that at which a correct response was punished on previous trial/total trials immediately following a punished correct response)); and number of reversals completed.

2.6. Progressive ratio schedule test

To assess whether HET and WT mice differed in their incentive-motivation for the sugar pellet reward, when mice had completed PRL testing, *ad libitum* feeding was given for 7 days and mice were then tested once on a progressive ratio schedule (PRS) test. For the PRS test, only one nose-poke port was deployed. The maximum session duration was 45 min. Session parameters were start ratio (required number of nose-pokes on first trial) = 1, repetition factor (number of consecutive trials for which the ratio remained constant) = 10, and step-wise linear fixed ratio increase (number of nose-pokes by which the ratio increased per increment) = 2, e.g. reward ratio on trials 1–10 = 1 response, on trials 11–20 = 3 responses, on trials 21–30 = 5 responses, and on trial 51–60 = 11 responses. The break point, was 600 s. The PRS measures of interest were total nose pokes committed, total reinforcements obtained, final ratio reached, and the session duration.

2.7. Drug

Mice were adjusted to ip injection by saline injection on 3 consecutive days prior to the onset of drug testing. They were also injected ip with: (i) saline prior to a reversal learning session to ensure that performance was not affected; (ii) escitalopram at 0.5 mg/kg at 1 week prior to the drug study so that any general effects of the compound on vegetative or psychological state had been experienced prior to the experiment. Escitalopram oxalate (ESC) was supplied by Lundbeck (H. Lundbeck A/S, Valby, Denmark). It was dissolved in sterile physiological saline and stored at 4 °C for 5 days maximum, protected from light. Solutions were prepared in order to give a base concentration of 0.5, 1.5, 4.5 or 7.5 mg/kg/5 ml. Effects of acute ESC on PRL behaviour were studied, with a wash-out period of 3–4 days i.e. the interval between successive PRL tests. Escitalopram was administered ip at 30 min prior to PRL testing. All tests were conducted between 08:30 and 11:30 h. Mice were administered vehicle (2×), 0.5, 1.5 and 4.5 mg/kg ESC (1× per dose) using a latin-square cross-over design, and 7.5 mg/kg (1×) was the final dose to be administered in all mice. Two vehicle (VEH) sessions were included in the latin square design in order that it would be possible to conduct a dose-range analysis that would

accommodate inter-individual variation in dose–response effects of ESC. That is, the better PRL performance (higher p reward-stay, lower p NFS) at 0.5 versus 1.5 mg/kg ESC ("low ESC dose") and at 4.5 versus 7.5 mg/kg ("medium–high ESC dose") was compared with the better PRL performance of the two VEH sessions.

2.8. Experimental design

Mice were trained at age 12–16 weeks and included in the PRL test validation at age 17–19 weeks, the study of acute ESC effects in the PRL test at age 21–24 weeks, and in the PRS test at age 26 weeks. Following each PRL test, re-baselining to criterion-level performance on the reversal learning test was required before proceeding to the next PRL test, as were at least three days between consecutive PRL tests for ESC wash-out. One WT mouse had to be excluded from the study during the validation experiment (skeletal injury in the home cage), and one WT mouse had to be excluded from the drug study (acute inactivity induced by ESC).

2.9. Data analysis

Statistical analysis was conducted using SPSS (version 17, SPSS Inc., Chicago IL, USA). Mixed model analysis of variance (ANOVA) was conducted for each dependent measure. Depending on experiment, fixed effects were genotype, p PCR, drug dose, and random effect was mouse-subject. In the case of significant interactions including genotype, a posteriori re-analysis was conducted using t-test or one-way ANOVA. *Post hoc* testing was conducted using the Bonferroni correction for multiple comparisons. Dunnett's test was run to assess whether negative feedback sensitivity was significantly different from the chance probability of NFS = 0.5. Significance was set at $p < 0.05$ (two-tailed). All data are presented as mean ± standard deviation (SD).

3. Results

3.1. Operant training and reversal learning performance

For autoshaping, mice required 3.5 ± 1.6 sessions to eat ≥30 sugar pellets in two consecutive sessions. There was no significant effect of genotype on autoshaping e.g. total number of sessions to attain criterion ($F (1, 28) < 1$, $p < 0.58$). For operant nose-poke learning, mice required 344 ± 80 total trials over 7.6 ± 2.0 sessions to achieve ≥30 rewards in each of two consecutive

sessions. There was no significant effect of genotype on operant learning e.g. total number of trials to criterion (F (1, 28) < 1, p < 0.51). For reversal learning, mice required 15.3 ± 4.0 sessions (Fig. 2A) and a total of 977 ± 276 trials (Fig. 2B) to reach criterion on the final training stage. There was no significant effect of genotype on reversal learning e.g. total number of reversal errors (WT: 266 ± 70, HET: 320 ± 122; F (1, 28) = 2.22, p < 0.15; Fig. 2C). Fig. 3 gives acquisition data for reversal learning by individual mice, illustrating both the acquisition curves and the inter-individual variation observed.

When mice had attained criterion on reversal learning they proceeded to PRL test sessions on 2 days per week. Reversal learning sessions continued to be conducted three days per week to ensure that baseline reversal learning was consistent across the experiment, as required by the cross-over/latin-square design. For the 8 weeks of PRL testing, Fig. 4 gives the weekly mean reversal session scores for reward-stay responses and number of reversals completed. Using ANOVA for main effects of week and genotype and their interaction, there was no significant main or interaction effect of week on p reward-stay (main effect: F (7, 169) < 1, p < 0.52) or number of reversals completed (main effect: F (7, 171) = 2.22, p < 0.15). There was also no significant effect of genotype on these two measures.

3.2. PRL testing at different probabilities of punished correct responding

The behavioural data for WT and HET mice in PRL test sessions conducted at different specific probabilities of punished correct responding (p PCR) are presented in Table 1 and Fig. 5. Data analysis was conducted using ANOVA to investigate effects of p PCR and genotype and their interaction.

There was a significant main effect of p PCR on the number of trials required to attain the 60 available reinforcements (F (3, 82) = 142.29, p < 0.0005), on session duration (F (3, 82) = 5.47, p < 0.002), and on the total number of nose pokes into the feeder port (F (3, 82) = 27.39, p < 0.0005). As indicated by the significant post hoc pair-wise comparisons reported in Table 1, these effects reflected, respectively, a monotonic increase in number of trials, session duration and total nose pokes, as p PCR increased. There was a significant main effect of p PCR on p reward-stay responses (F (3, 82) = 56.2, p < 0.0005); p reward-stay decreased significantly at p PCR = 0.1 relative to 0.0 (p < 0.002) and at 0.2/0.3 relative to 0.0/0.1 (p < 0.0005) (Table 1, Fig. 5A). There was a significant main effect of p PCR on number of reversals completed (F (3, 82) = 12.5, p < 0.0005); the number of reversals completed decreased significantly at p PCR = 0.2/0.3 relative to 0.0 (p ≤ 0.003) and at 0.3

relative to 0.1 (p < 0.02) (Table 1, Fig. 5B). There was no significant effect of p PCR on number of perseverative errors per reversal (Table 1). Each of these significant effects of p PCR was similar in WT and HET mice i.e. there was no main effect or interaction effect of genotype on any of the above behavioural measures (Table 1).

For p negative feedback sensitivity (NFS), there was a significant interaction effect of p PCR × genotype (F (2, 55) = 12.86, p < 0.0005) and a significant main effect of genotype (F (1, 28) = 16.84, p < 0.0005) (Fig. 5C). Based on the significant interaction term, a posteriori one-way ANOVA was conducted for each genotype. For WT mice, there was a significant main effect of p PCR (F (2, 27) = 10.02, p < 0.001); p NFS was increased at p PCR = 0.1 versus 0.2/0.3 (p ≤ 0.008; Fig. 5C). At p PCR = 0.1, WT mice exhibited p NFS (0.81) that was significantly higher than chance-level i.e. p = 0.5 (Dunnett's test critical distance at p < 0.01 = 0.143), whereas at p PCR = 0.2 and 0.3, p NFS was not significantly different from chance-level. For HET mice, there was also a significant main effect of p PCR (F (2, 28) = 6.47, p < 0.005); p NFS was decreased at p PCR = 0.1 versus 0.2/0.3 (p ≤ 0.05; Fig. 5C). At none of the p PCR's studied did HET mice exhibit p NFS that was significantly different from chance-level (Dunnett's test critical distance at p < 0.01 = 0.158). Also based on the significant interaction term, a posteriori t-tests of genotype differences in p NFS were conducted at specific p PCR's: p NFS was significantly lower in HET relative to WT mice at p PCR = 0.1 (t = 2.05, df = 27, p < 0.0005), with no genotype effect at other p PCR's (Fig. 5C).

3.3. Effects of escitalopram in PRL test

Effects of escitalopram (ESC) on PRL test behaviour were studied at p PCR = 0.2: at this probability, WT and HET mice exhibited similar behaviour in terms of the measures p reward-stay and p NFS, and both measures had values commensurate with detection of ESC-induced increase or decrease (Fig. 5). Two analyses of ESC effects on PRL behaviour were conducted: Firstly, individual dose effects, i.e. behaviour at VEH, 0.5, 1.5, 4.5 and 7.5 mg/kg were analysed using ANOVA for main effects of dose and genotype and their interaction. There were no significant main or interaction effects of individual doses of ESC on PRL test behaviour, e.g. main effect of Dose: reward-stay (F (4, 107) = 2.08, p < 0.09); NFS (F (4, 107) = 0.54, p < 0.71); number of reversals completed (F (4, 107) = 1.41, p < 0.23). There was also no significant effect of genotype. Second, dose-range effects were analysed i.e. comparison of PRL behaviour at VEH, 0.5–1.5 mg/kg (low ESC dose) and 4.5–7.5 mg/kg (medium–high ESC dose) sessions. The dose-range data are given in Fig. 6. ANOVA was conducted for main effects of dose-range, genotype and their interaction: For p reward-stay there
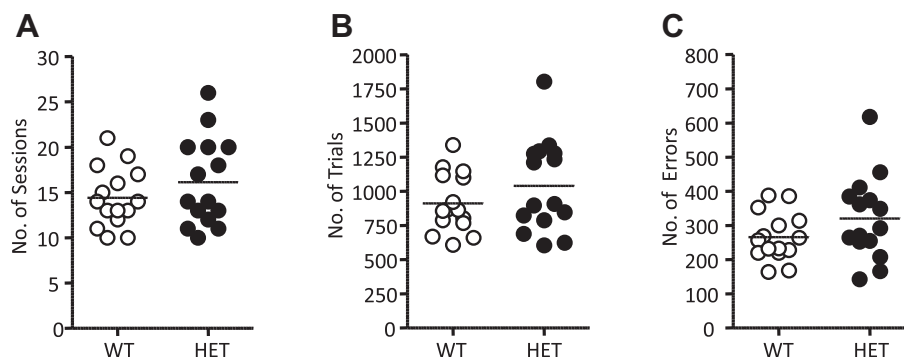

Fig. 2. Operant spatial reversal learning acquisition data for WT (N = 15) and HET (N = 15) mice of a 5-HTT null mutant strain on a C57BL/6 background. (A) Total number of reversal learning training sessions. (B) Total number of reversal learning training trials. (C) Total number of reversal errors i.e. sum of perseverative responses to the previously correct port following reversal of the response-outcome contingency.
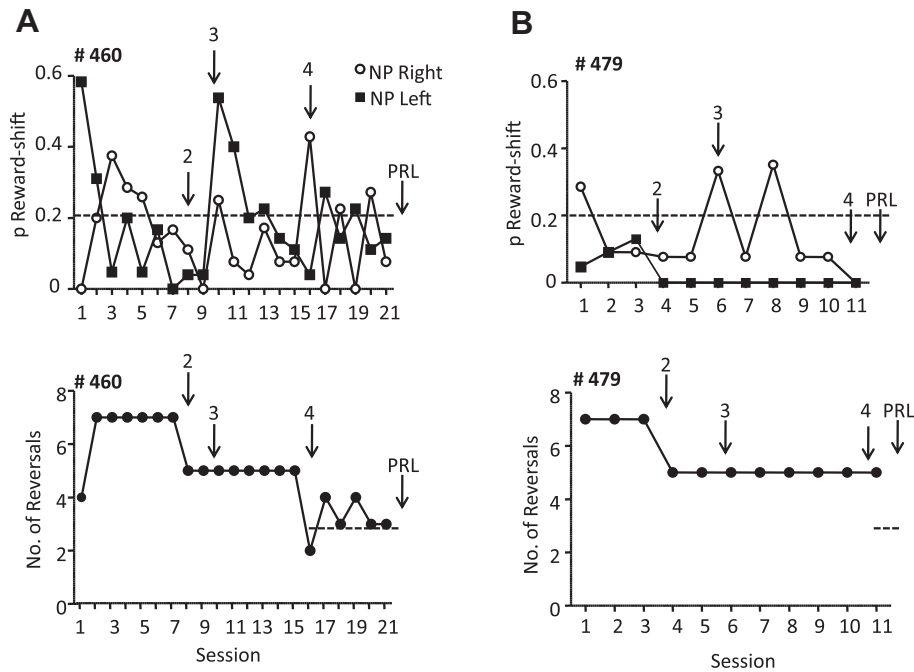
Fig. 3. Examples of reversal learning acquisition data for two individual mice, (A) HET mouse requiring a relatively high number of reversal training sessions, and (B) HET mouse requiring a relatively low number of reversal training sessions. The arrow and number indicate the first session of a stage following the attaining of learning criterion on the previous training stage. The dashed lines depict the threshold criteria for advancement to the next training stage i.e. probability of reward-shift ≤0.2 on each nose poke at stages 1–4 and at least 3 reversals at stage 4. NP = nose-poke port.

was a significant main effect of dose-range (F (2, 55) = 3.35, $p < 0.04$; Fig. 6A); p reward-stay was increased at low-dose ESC relative to VEH ($p < 0.04$) and medium–high dose values were intermediate. For p NFS there was a significant main effect of dose-range (F (2, 55) = 9.69, $p < 0.0005$; Fig. 6B); p NFS was decreased at low-dose ESC relative to VEH ($p < 0.0005$) and medium–high dose values were intermediate. There was no significant effect of geno-type on these two measures. Commensurate with its effects on p reward-stay and p NFS, there was a significant main effect of ESC dose-range on number of reversals completed (F (2, 81) = 4.44,

Fig. 4. Operant spatial reversal test baseline performance across the period of PRL testing in terms of probability of reward-stay responses and number of reversals achieved in WT and HET adult male mice of a 5-HTT null mutant strain. Mice were tested three days per week: the weekly mean score per mouse was calculated and the data shown are the overall weekly means (+/− SD) for all mice per genotype. There was no significant effect of either week or genotype on these two measures. On the two other days mice were tested on probabilistic reversal learning.

$p < 0.02$; Fig. 5C); more reversals were completed at low-dose ESC relative to VEH ($p < 0.02$) and medium–high dose values were intermediate. For number of reversals completed there was also a significant main effect of genotype (F (1, 81) = 4.16, $p < 0.05$) with HET mice completing more reversals than WT mice (Fig. 5C). There was also a significant main effect of dose-range on total number of nose pokes at the feeder port (F (2, 54) = 14.60, $p < 0.0005$), with decreased feeder responses at low-dose ESC (100.2 ± 20.5) relative to VEH (112.3 ± 18.1) ($p < 0.009$) and at medium–high dose ESC (91.5 ± 20.7) relative to VEH ($p < 0.0005$).

3.4. Progressive ratio schedule

Based on PRS test performance there was no significant effect of genotype on incentive-motivation for sugar pellet reward. Thus, WT and HET mice exhibited similar scores for each of the measures, namely: total nose pokes committed (WT: 413 ± 217; HET: 466 ± 281; $t = -0.56$, df = 27, $p < 0.59$), total reinforcements obtained (WT: 58 ± 16, HET: 60 ± 21; $t = -0.26$, df = 27, $p < 0.79$), final ratio reached (WT: 11.6 ± 3.2; HET: 11.9 ± 4.3; $t = -0.25$, df = 27, $p < 0.79$), and session duration, with no subject reaching break point i.e. all continued to nose poke up to the final 5 min of the 45-min session.

4. Discussion

The present study describes an operant two-way spatial discrimination task for the study of probabilistic reversal learning in mice, the responsiveness of behaviour to changes in the proba-bility of punishment of correct responding, and the effects on PRL behaviour of constitutive genetic and acute pharmacological manipulation of serotonin function.

Reversal learning is the basis of the PRL task and it is important to briefly review the evidence for 5-HT modulation of this prior to discussing the PRL data. Reversal learning requires adaptation to
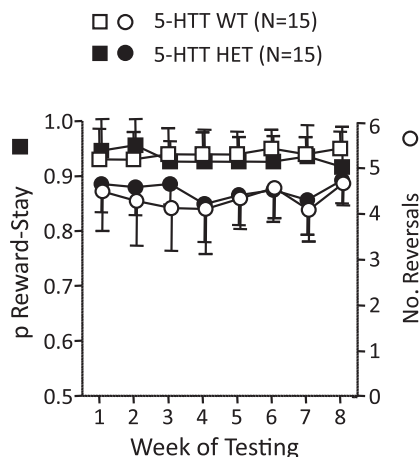
Table 1
Effects of probability of punished correct responding (p PCR) on PRL test scores (mean, SD) in WT and HET 5-HTT null mutant mice.

p PCR session-type measure	WT + HET*				WT				HET				P		
	0.0	0.1	0.2	0.3	0.0	0.1	0.2	0.3	0.0	0.1	0.2	0.3	p PCR	G	G × p PCR
No. trials	75[a]	85[b]	105[c]	121[d]	76	84	106	122	75	85	105	120	0.0005	ns	ns
	5	5	9	13	5	5	9	11	5	6	8	16			
Session duration (min)	12[a]	11[a]	14[ab]	15[b]	12	11	15	15	13	12	14	16	0.002	ns	ns
	6	4	4	5	6	2	3	3	6	5	5	7			
No. rewarded responses	60	60	60	60	60	60	60	60	60	60	60	60	ns	ns	ns
					0	0	0	0	0	0	0	0			
No. feeder responses	89[a]	95[a]	119[b]	125[c]	85	95	118	125	93	96	120	124	0.0005	ns	ns
	13	15	20	27	11	15	23	24	14	16	17	31			
Feeder response latency (ms)	969	906	768	763	856	717	693	660	1075	1081	842	859	ns	ns	ns
	632	644	275	400	541	193	78	79	709	852	372	541			
No. errors after reversal	9[a]	8[ab]	7[b]	5[c]	10	8	7	5	9	9	7	5	0.001	ns	ns
	4	4	3	4	3	4	4	3	5	5	3	4			
No. errors per reversal	2	2	2	1	2	2	2	1	2	2	1	1	ns	ns	ns
	1	1	1	1	1	1	1	1	1	1	0	1			
p reward-stay	0.92[a]	0.85[b]	0.74[c]	0.67[c]	0.91	0.85	0.74	0.67	0.92	0.86	0.74	0.67	0.0005	ns	ns
	0.08	0.07	0.07	0.07	0.08	0.05	0.07	0.06	0.07	0.08	0.06	0.08			
No. reversals completed	5[a]	5[ab]	4[bc]	3[c]	5	5	4	3	5	5	4	4	0.0005	ns	ns
	2	1	1	1	2	1	2	1	1	2	1	1			
p NFS		0.61	0.58	0.63		0.81	0.60	0.64		0.43	0.57	0.61	ns	0.0005	0.0005
		0.28	0.09	0.10		0.18	0.10	0.09		0.24	0.08	0.11			

* For measures with a significant main effect of p PCR session-type, p PCR's that do not share at least one letter were significantly different for all subjects (WT + HET).

regular changes in stimulus-reward contingency and comprises the separate but inter-dependent processes of reward-stay responding, negative feedback detection, inhibition of on-going response, and punishment-shift responding (Clark et al., 2004). In the present study there was no effect of 5-HTT genotype on either reward-stay responding or punishment-shift responding during reversal sessions. Serotonin transporter WT, HET and null mutant mice were studied using a computer touch-screen spatial reversal learning test (Brigman et al., 2010): There was no effect of genotype on reward-stay behaviour, whereas following a single reversal both HET and null mutant mice committed less reversal errors than did WT mice. In 5-HT depleted mice, achieved either via constitutive 5-HT neuronal depletion in Pet-1 null mutant mice or acute 5-HT depletion using para-chlorophenylalanine, there was no impact on reward-stay or reversal behaviour (Brigman et al., 2010). In non-human primates (common marmoset), comprehensive depletion of 5-HT fibres in the prefrontal cortex led to impaired reversal learning due to failure to inhibit on-going responding i.e. perseveration (Clarke et al., 2005, 2007). In healthy humans, partial depletion of 5-HT did not affect reversal learning about reward-related stimuli (Cools et al., 2008). Therefore, with the exception of the common marmoset study, there is little evidence that 5-HTT or 5-HT manipulations impact on reward-stay or punishment-shift responding in reversal learning. It is important to note that the common marmoset study was the only one to-date to deplete 5-HT in the prefrontal cortex specifically, with all other studies involving global manipulation of 5-HTT or 5-HT.

Moving to the findings for probabilistic reversal learning, mice that had received extensive reversal training responded to inclusion of punished correct response trials with decreased reward-stay behaviour. The decrease in reward-stay probability was monotonic relative to the increase in p PCR. Indeed, the observed probability of reward-stay was close to the actual probability of obtaining reward from the correct nose-poke port at any given stage of the session. These data are consistent with mice exhibiting accurate reward expectancy and, furthermore, with reward expectancy determining their reward-stay/-shift decision making. Mice did not apply a trial-by-trial reward-stay strategy, therefore. This interpretation of the current findings is schematized in Fig. 7A. The observed decrease in reversals completed as p PCR increased
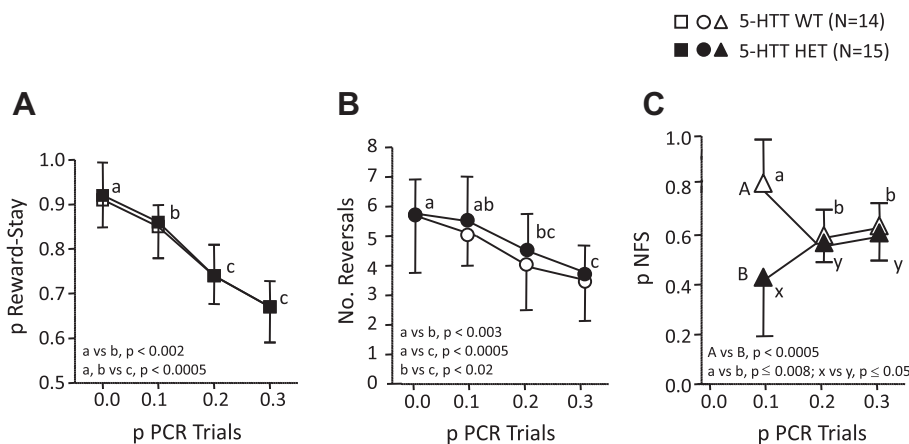
□ ○ △ 5-HTT WT (N=14)
■ ● ▲ 5-HTT HET (N=15)

Fig. 5. Validation of mouse PRL test in terms of effects of varying the probability of punished correct responding (p PCR) on behaviour. For each subject (WT = 14, HET = 15), one session with a maximum of 60 reinforced trials was carried out at each of p PCR = 0.0, 0.1, 0.2 or 0.3. (A) Probability of reward-stay trials. (B) Number of reversals completed (maximum = 7). (C) Negative feedback sensitivity. Values are mean ± SD. Pairs of data points that do not share at least one letter are significantly different, at the probability indicated in the figure.
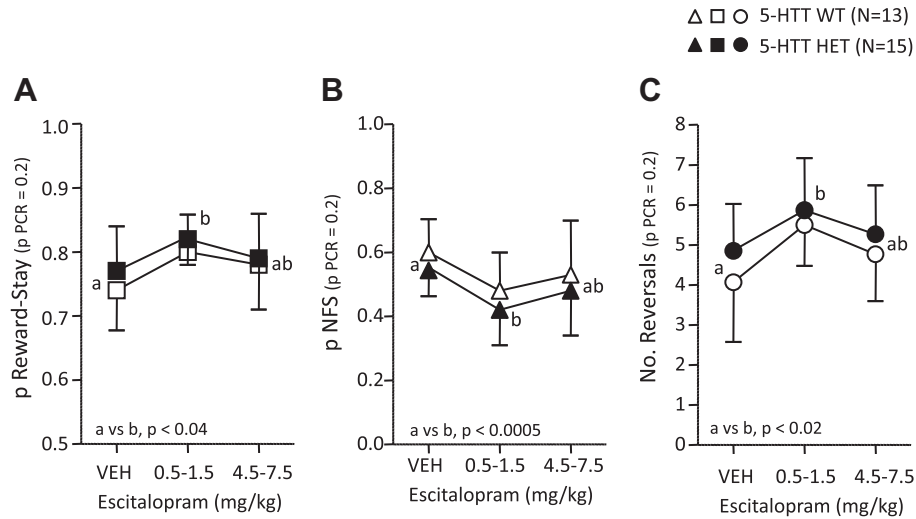
Fig. 6. Effects of escitalopram (ESC) on mouse behaviour in the PRL test. Each subject (WT = 13, HET = 15) was tested in 6 sessions, each with a maximum of 60 reinforced trials and *p* PCR of 0.2. Using a latin-square design, at 30 min pre-test, each mouse received vehicle (2 sessions), 0.5 mg/kg ESC (1), 1.5 mg/kg (1) and 4.5 mg/kg (1), and all mice received 7.5 mg/kg as the final test dose. For 2 vehicle sessions, 0.5–1.5 mg/kg and 4.5–7.5 mg/kg, the higher reward-stay score and lower NFS score were used per subject. (A) Probability of reward-stay trials. (B) Negative feedback sensitivity. (C) Number of reversals completed (maximum = 7). Values are means +/− SD. Pairs of data points that do not share at least one letter are significantly different, at the probability indicated in the figure.

ran parallel to the decrease in reward-stay behaviour, suggesting that reward expectancy was the major determinant of number of reversals completed. Comparison of these mice findings with those reported for other species is important, albeit that caution is required given the inter-study and -species differences. In the rat PRL test, where *p* PCR was 0.2, the reward-stay probability was 0.6–0.8 (Bari et al., 2010), and therefore similar to that observed at *p* PCR of 0.2 for mouse in the present study. A difference between the two studies is that in the rat study probability was set at 0.2 for both punished correct-response (PCR) and rewarded incorrect-response (RIR). In this respect the rat protocol was analogous to most human PRL protocols. In a previous mouse PRL study conducted in a two-way maze, both *p* PCR and *p* RIR of 0.2 were used and the reward-stay probability was 0.5 (Amodeo et al., 2012). The rationale for not including rewarded incorrect-response trials in the present mouse PRL paradigm was to attempt to avoid imposing a too-high cognitive demand on the mice, such that they could maintain accurate reward expectancy and, to some extent, accurate punishment expectancy. Mice completed more reversals in the present study than did rats in the rat PRL study (Bari et al., 2010), and this could be due to the absence of RIR trials, and also to increased reversal training in the present study. Therefore, the mice in the present study exhibited reward-stay and reversal behaviour that reflected accurate reward expectancy and was superior to that reported in previous rat and mouse PRL studies. This provides support for not including RIR trials in mouse PRL tests, so that they satisfy the assumption of human PRL tests that subjects are able to maintain accurate reward expectancy (Evers et al., 2005; Murphy et al., 2003). A study of mouse and rat under identical training/ testing conditions will be required to allow full comparison of the PRL behaviour of these two laboratory species and the effects thereon of including and excluding RIR trials. In human, under conditions of probability set at 0.2 for both punished correct- and rewarded incorrect-response trials, reward-stay probability is maintained at close to 1 (Taylor Tavares et al., 2008). This strongly suggests that human subjects are not using estimated overall reward expectancy but rather a strategy of trial-by-trial reward-stay behaviour, which would render spontaneous errors unlikely (Fig. 7B). This is also the case for depressed subjects (Taylor Tavares et al., 2008). Finally here, there was no effect of mouse serotonin

transporter genotype on the observed relationship between punished correct responding and reward-stay behaviour.

Mouse negative feedback sensitivity i.e. punishment-shift responses to punished correct-response trials, was also sensitive to the probability of such punishment. However, in contrast to reward-stay behaviour, the relationship between *p* PCR and NFS was not monotonic and, furthermore, it was dependent on serotonin transporter genotype (discussed below). At *p* PCR of 0.2 or 0.3, NFS was not significantly different from chance level. Extending the above interpretation that mouse reward expectancy at *p* PCR of 0.2–0.3 was relatively low, then chance-level NFS might reflect simultaneous inability to maintain accurate punishment prediction i.e. to accurately predict when reversal is/is not due (Fig. 7A). In contrast, in WT mice at least, NFS was significantly different from chance level at *p* PCR of 0.1, conditions under which, based on their reward-stay and reversals completed, mice maintained accurate and high reward expectancy. Under these conditions, mean NFS was either 0.8 i.e. relatively high (WT mice) or 0.4 i.e. relatively low (HET mice). These data suggest that at this low probability of PCR, when there were few reward-shift errors and correct negative feedback (i.e. reversals) occurred on a predictable schedule, mice processed misleading negative feedback differently than at higher PCR probabilities. That is, mice appear capable of accurate punishment expectancy under conditions of low punishment experience. In mouse (Amodeo et al., 2012) and rat (Bari et al., 2010), at *p* of 0.2 for both PCR and RIR, mean NFS was 0.5–0.6, and therefore at chance level and similar to that of WT and HET mice at *p* PCR of 0.2–0.3 in the present study. In human subjects, at *p* of 0.2 for PCR and RIR, mean NFS was (only) 0.1 in healthy probands and 0.3 in depressed probands (Taylor Tavares et al., 2008). Thus, healthy humans demonstrate accurate punishment expectancy and have the cognitive ability to inhibit punishment-shift responses to single trials (Fig. 7B). The documented increase in NFS on the PRL task in depression might reflect an increased emotional reactivity to single unexpected punishments (increased punishment prediction error) (Murphy et al., 2003). Given the importance of expectancy/prediction to human PRL behaviour, then a mouse PRL paradigm that promotes both accurate reward and punishment prediction would increase its translational value to the human paradigm. According to the current findings, this would
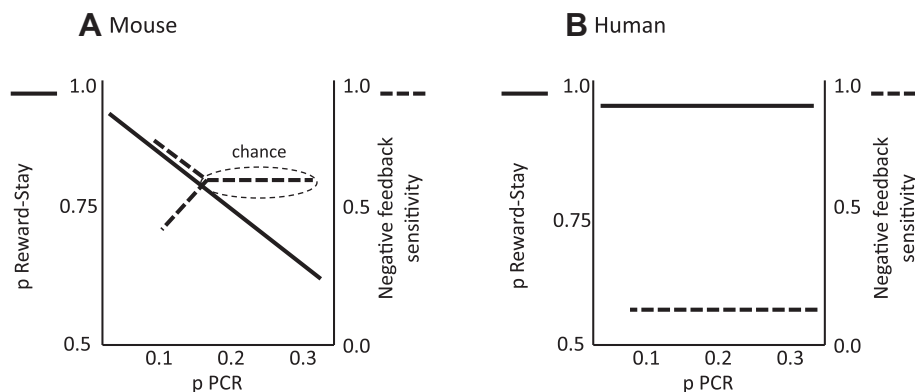
Fig. 7. Schematic representation of the relationships between punished correct responding and reward-stay and punishment-shift behaviour, and interpretation of underlying emotional-cognitive processes, in (A) mouse and (B) human. (A) In mouse, accurate estimation of reward expectancy from the current correct stimulus, combined with cognitive inflexibility to utilize a trial-by-trial reward-stay strategy, is proposed to account for the inverse monotonic relationship between p PCR and p reward-stay behaviour. For example, at p PCR 0.1, there will be approximately 13 punishments of "correct responding" if the subject is behaving optimally: 7 at reversal and 6 PCR's. Given the 60 rewarded CR's, p rewarded CR $= 60/73 = 0.82$, and observed p reward-stay was 0.85–0.86. At p PCR 0.2, there will be approximately 19 punishments of "correct responding" if the subject is behaving optimally: 7 at reversal and 12 PCR's. Given the 60 rewarded CR's, p rewarded CR $= 60/79 = 0.76$, and observed p reward-stay was 0.74. The complex relationship between p PCR and punishment-shift/NFS is interpreted as follows: at low p PCR (0.1), there is accurate, high reward expectancy and NFS reflects responsiveness to unexpected punished correct responses. At high p PCR (0.2–0.3), there is accurate, low reward expectancy and chance level NFS reflects responsiveness to expected punished correct responses. (B) In human, accurate estimation of reward expectancy combined with cognitive ability to utilize a trial-by-trial reward-stay strategy, is proposed to account for the high p reward-stay and lack of relationship between p PCR and reward-stay behaviour. Accurate estimation of punishment expectancy combined with emotional-cognitive ability to inhibit trial-by-trial punishment-shift responses, is proposed to account for the low NFS and lack of relationship between p PCR and NFS.

be at p PCR of around 0.1. However, even at this p PCR mice did not display evidence of applying a trial-by-trial reward-stay strategy or of consistently suppressing trial-by-trial punishment-shift responding. Any future protocol modification that supported such behaviours would be advantageous from a translational perspective.

Under the conditions of low p PCR (0.1) at which mice exhibited behaviour consistent with accurate reward and punishment prediction, a constitutive reduction in 5-HTT led to significantly lower negative feedback sensitivity. Heterozygous 5-HTT mutant mice have been demonstrated to exhibit high tonic synaptic 5-HT titres relative to WT (Mathews et al., 2004). By extrapolation, the reduction in NFS observed in HET mice co-occurred with, and is possibly attributable to, increased synaptic 5-HT levels. Given that 5-HTT polymorphisms predisposing to reduced 5-HTT function are risk factors for depression, it is paradoxical that 5-HTT HET mice exhibited reduced NFS. However, 5-HTT polymorphisms are undoubtedly pleiotropic. For example, 5-HTT HET mice also exhibit increased reactivity to fear-inducing stimuli relative to WT (Pryce et al., 2012), as do healthy human carriers of the "short" 5-HTTLPR polymorphism relative to "long" allele carriers (Hariri et al., 2002).

With respect to the study of the effects of acute pharmacological manipulations of 5-HT, although p PCR of 0.1 is perhaps optimal for mouse PRL (see above), the escitalopram experiment was conducted at p PCR of 0.2. At this p PCR, HET and WT mice performed similarly in terms of both reward-stay and NFS, therefore allowing for direct comparison of acute ESC effects in the two genotypes. There was a lack of consistent effect of ESC on PRL behaviour in both WT and HET mice when the analysis was based on single dose data. To accommodate inter-individual differences in sensitivity to this potent SSRI, the experimental design included two vehicle sessions so that the data for pairs of ESC doses (e.g. 0.5 and 1.5 mg/kg) could be integrated. Inter-individual differences in ESC effects on PRL scores were observed and when the data were re-analysed using the best vehicle, best low dose (0.5 or 1.5 mg/kg) and best medium–high dose (4.5 or 7.5 mg/kg) behavioural scores, low-dose ESC led to a robust decrease in NFS relative to vehicle. Low-dose ESC also increased reward-stay behaviour and number of reversals

completed relative to vehicle. These ESC effects occurred to a similar extent in WT and HET mice. That these effects were specific to lower doses of ESC could be attributable to the net change in 5-HT signalling due to synaptic and somatodendritic effects: ESC has high potency in terms of pre-synaptic 5-HT uptake inhibition. A low dose is already able to increase 5-HT availability in the synapse and a moderate dose might increase 5-HT to an extent that there is also increased binding to somatodendritic 5-HT$_{1A}$ autoreceptors in the DRN leading to reduced 5-HT neuronal firing and less synaptic 5-HT than at low dose ESC (Sanchez et al., 2003). In rat, acute citalopram (the racemic mixture of S[+]- and R [−]-enantiomers) led to increased NFS and decreased reversals at an acute low dose (1 mg/kg) and to decreased NFS and increased reversals at an acute high dose (10 mg/kg), whilst 5 mg/kg was without effect, and none of the doses tested had an effect on reward-stay behaviour (Bari et al., 2010). In healthy humans, a single clinically-relevant dose of citalopram (30 mg) led to increased NFS and impaired learning of the PRL task to criterion (Chamberlain et al., 2006). The authors proposed that a 5HT$_{1A}$ autoreceptor-mediated reduction of 5-HT signalling could underlie this impairing effect of citalopram on human PRL performance (Chamberlain et al., 2006; Evers et al., 2005). Thus, both acute low-dose ESC in mouse and acute high-dose citalopram in rat led to decreased NFS and increased reversals completed and acute low-dose ESC in mouse also led to increased reward-stay behaviour. Given that these effects were obtained at p PCR of 0.2, conditions under which mice (and possibly also rat) appear not to be able to maintain accurate reward and punishment predictions, then a parsimonious interpretation would be that an acute increase in 5-HT synaptic signalling functions to reduce negative feedback sensitivity in mice even on a background of chance-level punishment prediction, but not to an extent that compromises flexibility at reversal. Future studies will expand on these findings by studying drug effects at lower p PCR.

In summary, the present study demonstrates that an automated two-choice operant spatial discrimination with food reinforcement can be used to study probabilistic reversal learning in mice. Evidence is provided that the mouse is sensitive to the probability of punished correct responding. At a low probability, mice appear

able to maintain accurate reward expectancy and punishment expectancy, as evidenced by high reward-stay, high reversals completed and non-chance negative feedback sensitivity. As the PCR probability increases, these desirable features for a translational PRL paradigm attenuate. Under conditions consistent with accurate reward/punishment expectancy, genetic reduction in 5-HTT function led to reduced negative feedback sensitivity, as did acute pharmacological reduction in 5-HTT function, even though this experiment was not conducted under conditions consistent with accurate reward/punishment expectancy. This mouse PRL paradigm can be used to investigate the effects of environmental (stress) and genetic manipulations on reward and punishment prediction and negative feedback sensitivity, and their neurobiological and pharmacological regulation, with the aim of translating findings to development of improved treatments for emotional-cognitive dysfunction in depression.

Acknowledgements

References

Abramson, L.Y., Metalsky, G.I., Alloy, L.B., 1989. Hopelessness depression: a theory-based subtype of depression. Psychol. Rev. 96, 358–372.

Amodeo, D.A., Jones, J.H., Sweeney, J.A., Ragozzino, M.E., 2012. Differences in BTBR T+ tf/J and C57BL/6J mice on probabilistic reversal learning and stereotyped behaviors. Behav. Brain Res. 227, 64–72.

APA, 2000. Diagnostic and Statistical Manual of Mental Disorders. Revision American Psychiatric Association, Washington, DC, fourth ed. American Psychiatric Association, Washington, DC.

Bari, A., Theobald, D.E., Caprioli, D., Mar, A.C., Aidoo-Micah, A., Dalley, J.W., Robbins, T.W., 2010. Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. Neuropsychopharmacology 35, 1290–1301.

Bengel, D., Murphy, D.L., Andrews, A.M., Wichems, C.H., Feltner, D., Heils, A., Mössner, R., Westphal, H., Lesch, K.-P., 1998. Altered brain serotonin homeostasis and locomotor insensitivity to 3,4-methylenedioxymethamphetamine ("ecstasy") in serotonin transporter-deficient mice. Mol. Pharmacol. 53, 649–655.

Berridge, K.C., Robinson, T.E., 2003. Parsing reward. TINS 26, 507–513.

Boureau, Y.-L., Dayan, P., 2011. Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology 36, 74–97.

Brigman, J.L., Mathur, P., Harvey-White, J., Izquierdo, A., Saksida, L.M., Bussey, T.J., Fox, S., Deneris, E., Murphy, D.L., Holmes, A., 2010. Pharmacological or genetic inactivation of the serotonin transporter improves reversal learning in mice. Cereb. Cortex 20, 1955–1963.

Canli, T., Lesch, K.-P., 2007. Long story short: the serotonin transporter in emotion regulation and social cognition. Nat. Neurosci. 10, 1103–1109.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Honalee, H., McClay, J., Mill, J., Martin, J.B., Braithwaite, A., Poulton, R., 2003. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 301, 386–389.

Chamberlain, S.R., Müller, U., Blackwell, A.D., Clark, L., Robbins, T.W., Sahakian, B.J., 2006. Neurochemical modulation of response inhibition and probabilistic learning in humans. Science 311, 861–863.

Clark, L., Cools, R., Robbins, T.W., 2004. The neuropsychology of ventral prefrontal cortex: decision-making and reversal learning. Brain Cogn. 55, 41–53.

Clarke, H.F., Walker, S.C., Crofts, H.S., Dalley, J.W., Robbins, T.W., Roberts, A.C., 2005. Prefronal serotonin depletion affects reversal learning but not attentional set shifting. J. Neurosci. 25, 532–538.

Clarke, H.F., Walker, S.C., Dalley, J.W., Robbins, T.W., Roberts, A.C., 2007. Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. Cereb. Cortex 17, 18–27.

Cools, R., Clark, L., Owen, A.M., Robbins, T.W., 2002. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic imaging. J. Neurosci. 22, 4563–4567.

Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacology 36, 98–113.

Cools, R., Roberts, A.C., Robbins, T.W., 2008. Serotonergic regulation of emotional and behavioural control processes. Trends Cogn. Sci. 12, 31–40.

Elliott, R., Sahakian, B., McKay, A.P., Herrod, J.J., Robbins, T.W., Paykel, E.S., 1996. Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. Psychol. Med. 26, 975–989.

Elliott, R., Sahakian, B.J., Herrod, J.J., Robbins, T.W., Paykel, E.S., 1997. Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. J. Neurol. Neurosurg. Psychiatry 63, 74–82.

Ersche, K.D., Roiser, J.P., Robbins, T.W., Sahakian, B.J., 2008. Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology 197, 421–431.

Eshel, N., Roiser, J.P., 2010. Reward and punishment processing in depression. Biol. Psychiatry 68, 118–124.

Evers, E.A.T., Cools, R., Clark, L., van der Veen, F.M., Jolles, J., Sahakian, B.J., Robbins, T.W., 2005. Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. Neuropsychopharmacology 30, 1138–1147.

Hariri, A.R., Mattay, V.S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M.F., Weinberger, D.R., 2002. Serotonin transporter genetic variation and the response of the human amygdala. Science 297, 400–403.

Henriques, J.B., Davidson, R.J., 2000. Decreased responsiveness to reward in depression. Cognition Emotion 14, 711–724.

Jennings, K.A., Lesch, K.-P., Sharp, T., Cragg, S.J., 2010. Non-linear relationship between 5-HT transporter gene expression and frequency sensitivity of 5-HT signals. J. Neurochem. 115, 965–973.

Jocham, G., Neumann, J., Klein, T.A., Danielmeier, C., Ullsperger, M., 2009. Adaptive coding of action values in the human rostral cingulate cortex. J. Neurosci. 29, 7489–7496.

Mathews, T.A., Fedele, D.E., Coppelli, F.M., Avila, A.M., Murphy, D.L., Andrews, A.M., 2004. Gene dose-dependent alterations in extraneuronal serotonin but not dopamine in mice with reduced serotonin transporter expression. J. Neurosci. Methods 140, 169–181.

Millan, M.J., 2003. The neurobiology and control of anxious states. Prog. Neurobiol. 70, 83–244.

Montanez, S., Owens, W.A., Gould, G.G., Murphy, D.L., Daws, L.C., 2003. Exaggerated effect of fluvoxamine in heterozygote serotonin transporter knockout mice. J. Neurochem. 86, 210–219.

Murphy, D.L., Lesch, K.-P., 2008. Targeting the murine serotonin transporter: insights into human neurobiology. Nat. Rev. Neurosci. 9, 85–96.

Murphy, F.C., Michael, A., Robbins, T.W., Sahakian, B.J., 2003. Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. Psychol. Med. 33, 455–467.

Nandrino, J.-L., Dodin, V., Martin, P., Henniaux, M., 2004. Emotional information processing in first and recurrent major depressive episodes. J. Psychiat. Res. 38, 475–484.

Pryce, C.R., Azzinnari, D., Sigrist, H., Gschwind, T., Lesch, K.-P., Seifritz, E., 2012. Establishing a learned helplessness effect paradigm in C57BL/6 mice: behavioural evidence for emotional, motivational and cognitive effects of aversive uncontrollability per se. Neuropharmacology 62, 358–372.

Pryce, C.R., Seifritz, E., 2011. A translational research framework for enhanced validity of mouse models of psychopathological states in depression. Psychoneuroendocrinology 36, 308–329.

Sanchez, C., Bergqvist, P.B.F., Brennum, L.T., Gupta, S., Hogg, S., Larsen, A., Wiborg, O., 2003. Escitalpram, the S-(+)-enantiomer of citalopram, is a selective serotonin reuptake inhibitor with potent effects in animal models predictive of antidepressant and anxiolytic activities. Psychopharmacology 167, 353–362.

Sharp, T., Cowen, P.J., 2011. 5-HT and depression: is the glass half-full? Curr. Opin. Pharmacol. 11, 45–51.

Sloan, D.M., Strauss, M.E., Wisner, K.L., 2001. Diminshed response to pleasant stimuli by depressed women. J. Abnorm Psychol. 110, 488–493.

Taylor Tavares, J.V., Clark, L., Furey, M.L., Williams, G.B., Sahakian, B., Drevets, W.C., 2008. Neural basis of abnormal response to negative feedback in unmedicated mood disorders. NeuroImage 42, 1118–1126.