

Themed Section: Opioids: New Pathways to Functional Selectivity

RESEARCH PAPER

δ -Opioid receptors in the accumbens shell mediate the influence of both excitatory and inhibitory predictions on choice

Vincent Laurent, Felix L Wong and Bernard W Balleine

Behavioural Neuroscience Laboratory, Brain and Mind Research Institute, The University of Sydney, Camperdown, NSW 2006, Australia

Correspondence

Bernard W Balleine, Brain and Mind Research Institute, University of Sydney, 100 Mallett St, Camperdown, NSW 2050, Australia. E-mail: bernard.balleine@sydney.edu.au

Received

30 October 2013

Revised

5 February 2014

Accepted

18 March 2014

BACKGROUND AND PURPOSE

Stimuli that predict rewarding events can control choice between future actions, and this control could be mediated by δ -opioid receptors in the nucleus accumbens shell (NAc-S). Stimuli predicting the absence of important events can also guide choice, although it remains unknown whether they do so via changes in an accumbal δ -opioid receptor-related process.

EXPERIMENTAL APPROACH

δ -opioid receptor-eGFP mice were trained to perform two instrumental actions that delivered different food outcomes. Choice between the two actions was then tested in the presence of stimuli paired with either the delivery or the non-delivery of each of the two outcomes. Bilateral infusions of the δ -opioid receptor antagonist naltrindole into the NAc-S were used to determine the role of these receptors at the time of choice and δ -opioid receptor expression in the NAc-S used to assess functional activity.

KEY RESULTS

A stimulus predicting a specific outcome biased choice performance towards the action previously earning that same outcome. In contrast, a stimulus signalling the absence of that outcome biased performance away from the action that delivered that outcome towards actions associated with the absence of that outcome. Both effects were associated with increased δ -opioid receptor expression on the membrane of cholinergic interneurons within the NAc-S. Furthermore, both effects were blocked by naltrindole infused into the NAc-S.

CONCLUSIONS AND IMPLICATIONS

These findings suggest that δ -opioid receptors in the NAc-S were involved in the effects of predictive learning on choice between actions, whether those predictions involve the presence or absence of specific rewarding events.

LINKED ARTICLES

This article is part of a themed section on Opioids: New Pathways to Functional Selectivity. To view the other articles in this section visit <http://dx.doi.org/10.1111/bph.2015.172.issue-2>

Abbreviations

ChAT, choline acetyltransferase; CIN, cholinergic interneuron; NAc-S, nucleus accumbens shell; PIT, Pavlovian-instrumental transfer

Introduction

Much evidence suggests that the endogenous opioid system plays a critical role in decision making involving choice between different courses of goal-directed action (Laurent *et al.*, 2012; Bertran-Gonzalez *et al.*, 2013; Lutz and Kieffer, 2013). One variable that regulates this choice is the presence of stimuli sharing a predictive history with the consequences of available actions; for example, Pavlovian stimuli that predict a particular outcome bias choice towards actions that earn that outcome (Colwill and Rescorla, 1988; Dickinson and Balleine, 1994; Holmes *et al.*, 2010). This phenomenon, known as outcome-specific Pavlovian-instrumental transfer (PIT), requires activation of δ -opioid receptors (receptor nomenclature follows Alexander *et al.*, 2013) in the nucleus accumbens shell (NAc-S) (Laurent *et al.*, 2012). Recent work in our laboratory suggests that δ -opioid receptors located on the membrane of cholinergic interneurons (CINs) in the NAc-S are particularly important for outcome-specific PIT (Bertran-Gonzalez *et al.*, 2013). We found that successful PIT is associated with an increase in δ -opioid receptor expression on CINs in the NAc-S but not other regions of the striatum. This increase appeared to be induced by Pavlovian predictive learning and conditioned responding was strongly correlated with δ -opioid receptor translocation to the membrane of CINs in the NAc-S. More importantly, this accumulation was also strongly correlated with the size of the PIT effect, suggesting that predictive learning triggered the translocation of δ -opioid receptors to generate stimulus-based choice between actions.

The finding that outcome-specific PIT involves a long-term change in δ -opioid receptor expression represents a significant advance in our understanding of the cellular mechanisms underlying choice. To date, however, we have only investigated the role of δ -opioid receptors in the control exerted by excitatory stimuli; that is, those predicting the occurrence of biologically significant events such as the delivery of food. Choice can, however, also be influenced by inhibitory stimuli; that is, stimuli predicting the absence of a biologically significant event. For example, Delamater *et al.* (2003) demonstrated that a stimulus predicting the absence of a specific food outcome reduced the performance of actions that previously earned that outcome, relative to actions that did not, the reverse of the usual PIT effect. The neural mechanisms underlying this reduction remain unexplored. Here, we have tested the possibility that inhibitory stimuli influence choice between actions through a mechanism similar to excitatory stimuli. More specifically, we investigated whether the training of Pavlovian inhibitory stimuli produced a change in δ -opioid receptor expression on the membrane of NAc-S CINs and whether blockade of these receptors removed the influence of this inhibitory learning on choice in tests of PIT. The results clearly suggest that

δ -opioid receptors in the NAc-S modulate the effects of predictive learning on choices between actions, whether those predictions involve the presence or absence of biologically significant events.

Methods

Animals

All animal care and experimental procedures were approved by the Animal Ethics Committee at the University of Sydney. All studies involving animals are reported in accordance with the ARRIVE guidelines for reporting experiments involving animals (Kilkenny *et al.*, 2010; McGrath *et al.*, 2010). A total of 56 animals were used in the experiments described here.

In Experiment 1 we used 45 homozygous male C57Bl/6 δ -opioid receptor-eGFP knock-in transgenic mice (δ -opioid receptor-eGFP^{hi}) in which a functional δ -opioid receptor gene (*Oprd1*) fused to enhanced green fluorescent protein gene (*EGFP*) was inserted in the wild-type *Oprd1* locus. The initial colony was generously provided by the laboratory of Prof. B.L. Kieffer (CNRS, Illkirch, France). Mice were housed in plastic boxes (two to five mice per box) located in a climate-controlled colony room and were maintained on a 12 h light/dark cycle. Experiment 2 used 11 experimentally naive Long-Evans rats (aged 7–12 weeks) obtained from the Monash University Animal Research Platform. Rats were housed in plastic boxes (two or three rats per box) located in another colony room. Five days before the behavioural procedures, all animals were handled daily and were put on a food deprivation schedule to maintain them at ~85% of their *ad libitum* feeding weight.

Drug treatments

The δ -opioid receptor antagonist naltrindole hydrochloride (Tocris Bioscience, Ellisville, MO, USA) was dissolved in 0.9% (w/v) saline containing 5% DMSO to obtain a final concentration of 5 $\mu\text{g}\cdot\mu\text{L}^{-1}$. This dose was selected as it had previously been shown to impair outcome-specific PIT (Laurent *et al.*, 2012; Bertran-Gonzalez *et al.*, 2013) in a relatively specific manner. Indeed, NAc-S infusion of a μ -opioid receptor antagonist failed to produce the same impairment, consistent with other findings revealing that mice with genetic deletion of δ - but not that of μ -opioid receptors, fail to exhibit outcome-specific PIT (Laurent *et al.*, 2012). Infusion of 0.9% (w/v) saline containing 5% DMSO (vehicle) was used to control for any effect of the infusion procedure *per se*.

Surgery and microinjections

At the time of surgery, rats weighed between 290 and 360 g. Continuous flow of a mixed isoflurane and oxygen gas solution was used to anaesthetize rats that were then placed in a stereotaxic frame (Kopf Instruments, Tujunga, CA, USA) with

the incisor bar set at -3.3 mm. The scalp was retracted to expose the skull, and 26 gauge guide cannulae (Plastics One, Roanoke, VA, USA) were bilaterally implanted through holes drilled in the skull in the shell region of the nucleus accumbens at the following coordinates, relative to bregma: AP, $+1.7$; ML, ± 0.75 ; DV, -6.4 . The guide cannulae were maintained in position with dental cement and dummy cannulae were kept in each guide at all times except during microinjections. Immediately after the surgical procedure, rats were injected i.p. with a prophylactic (0.4 mL) dose of $300 \text{ mg}\cdot\text{kg}^{-1}$ solution of procaine penicillin. Rats were allowed 3 days to recover from surgery, during which time they were handled and weighed daily.

Naltrindole and vehicle were infused into the shell region of the nucleus accumbens by inserting a 33 gauge infusion cannula into the guide. The infusion cannulae were connected to a $25 \mu\text{L}$ glass syringe connected to an infusion pump (kdScientific, SDR Clinical Technology, Australia) and projected 1 mm ventral to the tip of the guide. A total volume of $0.2 \mu\text{L}$ was delivered at a rate of $0.1 \mu\text{L}\cdot\text{min}^{-1}$. The infusion cannula remained in place for a further 1 min after the infusion and then removed. On the day before the first infusion, the dummy cannula was removed and the infusion pump was turned on for 2 min in order to familiarize the rats with the procedure and thereby minimize any stress produced by this procedure, when infusions were made.

Behavioural apparatus

Training and testing took place in 32 MED Associates operant chambers (St Albans, VT, USA) (16 for mice and 16 for rats) enclosed in sound- and light-resistant shells. Each chamber was equipped with a pump fitted with a syringe that could deliver 0.1 mL of a 20% sucrose solution into a recessed magazine in the chamber. Each chamber was also equipped with two pellet dispensers that could individually deliver either grain food pellets (20 mg for mice and 45 mg for rats; Bioserve Biotechnologies, Flemington, NJ, USA) or chocolate food pellets (20 mg for mice) when activated. The chambers contained two retractable levers that could be inserted to the left and right sides of the magazine. An infrared photobeam crossed the magazine opening, allowing for the detection of head entries. A 3 W, 24 V house light provided illumination of the operant chamber, and each chamber contained a Son-alert that, when activated, delivered a 3 kHz pure tone, a 28 V DC mechanical relay that was used to deliver a 2 Hz clicker stimulus and a white noise generator (80 dB). A set of four microcomputers running MED Associates proprietary software (Med-PC) controlled all experimental events and recorded magazine entries and lever presses.

Behavioural procedures

Experiment 1: backward conditioning. Training and testing in Experiment 1 was conducted on the δ -opioid receptor-eGFP^{ki} mice with the house light on. All mice initially received 8 days of instrumental training, during which the two levers were made available consisting of training two responses (R1 and R2, left and right lever presses) earning two distinct outcomes (O1 and O2; grain and chocolate pellets) in separate daily sessions. The order of the sessions was counterbalanced, as were the response–outcome relationships. Each

session ended when 20 outcomes were earned or when 30 min had elapsed. For the first 2 days, lever pressing was continuously reinforced (i.e. each response earned an outcome). Then, the probability of the outcome given a response was gradually shifted over days using increasing random ratio schedules: a RR5 schedule ($P = 0.2$) was used on days 3–5 and a RR10 schedule ($P = 0.1$) was used on days 6–8.

The mice then received 14 daily sessions of Pavlovian training, during which the levers were retracted. In group Forward ($n = 21$), each session consisted of presenting two stimuli S1 and S2 (clicker and noise) with each terminating with the delivery of one of the food outcomes (O1 or O2). Mice in group Backward ($n = 24$) received a similar procedure, except that the two outcomes were delivered 10 s before presentation of the stimuli. The stimulus–outcome relationships were fully counterbalanced within the groups and were also counterbalanced with the response–outcome relationships previously established. In each session, the stimuli lasted between 2 and 58 s with an average time of 30 s. There were 12 presentations of each stimulus in a pseudo-random manner with an intertrial interval that varied between 50 and 170 s with an average of 120 s.

After the final day of Pavlovian training, mice were given two instrumental reminder sessions followed by a single Pavlovian-instrumental transfer test. For the test, both levers were inserted into the box, but no outcomes were delivered. Responding was extinguished on both R1 and R2 for 8 min to reduce the rate of baseline performance. All mice then received presentations of the two Pavlovian conditioned stimuli in the following order: noise, clicker, clicker, noise, clicker, noise, noise, clicker. Each stimulus lasted for 1 min and was separated by a 3 min fixed ITI. Performance on the two instrumental actions was recorded in the presence of either stimulus and in their absence.

Experiment 2: conditioned inhibition. As quantification of δ -opioid receptor expression in δ -opioid receptor-eGFP^{ki} mice would have been inaccurate due to the damage induced by the surgical procedure, Experiment 2 used rats as subjects. Indeed, we have previously reported that, in rats, infusion of naltrindole into the NAc-S impairs specific PIT (Laurent *et al.*, 2012) and that Pavlovian training triggers similar changes in δ -opioid receptor expression on NAc-S CINs as it does in mice (Laurent *et al.*, 2014). In this experiment, the house light was off during all stages because it was used as a Pavlovian stimulus. All rats received 20 sessions of Pavlovian training once a day, during which the levers were retracted. Each session consisted of the presentation of two excitatory stimuli, S1 and S2 (the noise and the clicker), each paired with two distinct food outcomes O1 and O2 (food pellets and sucrose solution). Both S1 and S2 were also presented in compound with one of two other stimuli, S3 and S4 (house light and tone), S1 with S3 and S2 with S4. No outcomes were delivered during presentation of the compound stimuli. S3 and S4 were, therefore, trained as conditioned inhibitors predicting the absence of O1 and O2 respectively. The stimulus–outcome and stimulus–stimulus relationships were fully counterbalanced. The individual stimuli and the compound stimuli lasted 1 min and were each presented four times in a pseudorandom order with an inter-trial interval of 5 min. The two outcomes, sucrose solution or food pellets, were

delivered on a random time 20 s schedule during the appropriate individual stimulus.

Following Pavlovian training, all rats received 8 days of instrumental training in the manner described previously, except for outcome identity (sucrose solution and grain pellets) and the number of training days. After the third day of RR10 training, rats were given *ad libitum* access to food and water for 5 consecutive days before surgery (see above). After recovery from surgery, rats were returned to the food deprivation schedule previously used and received 2 additional days of instrumental training on a RR10 schedule.

After the final day of training, rats received two Pavlovian reminder sessions followed by two Pavlovian-instrumental transfer tests conducted on consecutive days. The procedure was similar to that described earlier except that responding was extinguished on both R1 and R2 for 4 min to establish a low rate of baseline performance. Half of the rats in each of the conditions established during training received presentations of the various compound stimuli in the following order: noise-light, noise-tone, clicker-tone, clicker-light. The remaining rats received the compound presentations in the order: clicker-tone, clicker-light, noise-light, noise-tone.

Tissue preparation to control for cannulae placements

At the end of Experiment 1, the rats received a lethal dose of sodium pentobarbital (300 mg·kg⁻¹; Virbac Pty. Ltd., Regents Park, NSW, Australia). The brains were removed, frozen and sectioned coronally with a cryostat (Leica Microsystems Australia, North Ryde, NSW, Australia) at 40 μm through the core or the shell region of the nucleus accumbens. Every third section was collected on a slide, and the sections were stained with cresyl violet. The location of cannula tips was determined under a microscope by a trained observer, who was unaware of the treatment groups, using boundaries defined using the Paxinos and Watson atlas (Paxinos and Watson, 2006). Animals with inaccurate cannulae placements or with extensive damage at the infusion site were excluded from the statistical analysis.

Transcardial fixation and brain sectioning for immunofluorescence

After the test, mice were rapidly anaesthetized with sodium pentobarbital (500 mg·kg⁻¹, i.p.) and transcardially perfused with cold 4% paraformaldehyde in 0.1 M sodium phosphate buffer (pH 7.5). Brains were post-fixed in the same solution at 4°C overnight. Coronal 30-μm-thick sections (+1.3 from bregma in mice) were cut with a vibratome (Leica Microsystems VT1000 Australia, North Ryde, NSW, Australia) and stored at -20°C in a solution containing 30% ethylene glycol, 30% glycerol and 0.1 M sodium phosphate buffer, until they were processed for immunofluorescence.

Immunofluorescence

Individualized free-floating sections were rinsed with Tris-buffered saline (TBS; 0.25 M Tris, 0.5 M NaCl, pH 7.5), incubated for 5 min in TBS containing 3% H₂O₂ and 10% methanol, and then rinsed three times for 10 min with TBS. After 20 min of incubation in 0.2% Triton X-100 (Sigma-Aldrich, St Louis, MO, USA) in TBS, sections were rinsed three times with TBS again. δ-opioid receptor-eGFP signal was

amplified through incubation with polyclonal rabbit anti-eGFP primary antibody (1:300, #A11122; Life Technologies, Carlsbad, CA, USA) diluted in TBS. Choline acetyltransferase (ChAT) was simultaneously detected using polyclonal goat anti-ChAT (1:300, #AB144P; Millipore, Billerica, MA, USA) (4°C, overnight). Sections were then rinsed 10 min with TBS three times and incubated 60 min at room temperature with compatible sets of fluorescent secondary antibodies diluted in TBS: donkey anti-rabbit Alexa 488 (Life Technologies, Mulgrave, VIC, Australia) (1:400; eGFP amplification and donkey anti-goat CY3 (1:400; ChAT). Sections were rinsed three times for 10 min in TBS, mounted in Superfrost Plus coated slides (Thermo Scientific) and let dry for 10 min before coverslipped in Vectashield fluorescence medium (Vector Laboratories, Burlingame, CA, USA).

Fluorescence analysis

For each neuron located in the ventromedial extension of the nucleus accumbens shell (approximate coordinates AP, + 1.3; ML, ± 0.8; DV, -5), a single focal plane with optimal ChAT immunoreactivity was determined in channel 2 (Ch02, HeNe green laser). Sequential 58.93 μm² single confocal images (optical magnification: 60×; digital zoom: 4×; resolution: 17.378 pixels·μm⁻¹) were obtained for ChAT signal (Ch02, HeNe green laser) and corresponding δ-opioid receptor-eGFP signal (Ch01, Ar) with a Kaplan filter (5 averaging scans). On each image, two different regions of interest were subsequently defined in the ChAT image of each neuron: ROI 1 comprised the somatic region (located in the intracellular-extracellular inter-phase defined by the ChAT staining), whereas ROI 2 was used as a background correction, and comprised the nuclear region (as defined by the central region devoid of ChAT staining). Mean grey value for each ROI was then collected from the overlapped eGFP image and expressed, for each neuron, as ROI1-ROI2. One single fluorescence value was finally obtained per animal (average of all the quantified neurons). In all cases, an experimenter unaware of the behavioural score underlying the samples performed microscope acquisitions, and all image files in each experiment were randomly renumbered using a MS Excel plug-in (Bio-excel2007 by Romain Bouju, France) prior to all quantifications.

Data analysis

Statistical analyses were conducted using within-subject or mixed-model ANOVA depending upon the experimental design. For all analyses, significance was assessed against a type I error rate of 0.05. ANOVAs were followed by simple main effects analyses to establish the source of any significant interactions. Due to the level of counterbalancing employed in the reported experiments, performance across the various counterbalancing conditions (i.e. stimuli, compound stimuli and outcome identity) was averaged during all analyses.

Results

Pavlovian inhibitory training produces δ-opioid receptor accumulation on the membrane of NAc-S CINs

Two groups of food-deprived δ-opioid receptor-eGFP^{hi} mice initially learned that two instrumental responses (R1 and R2)

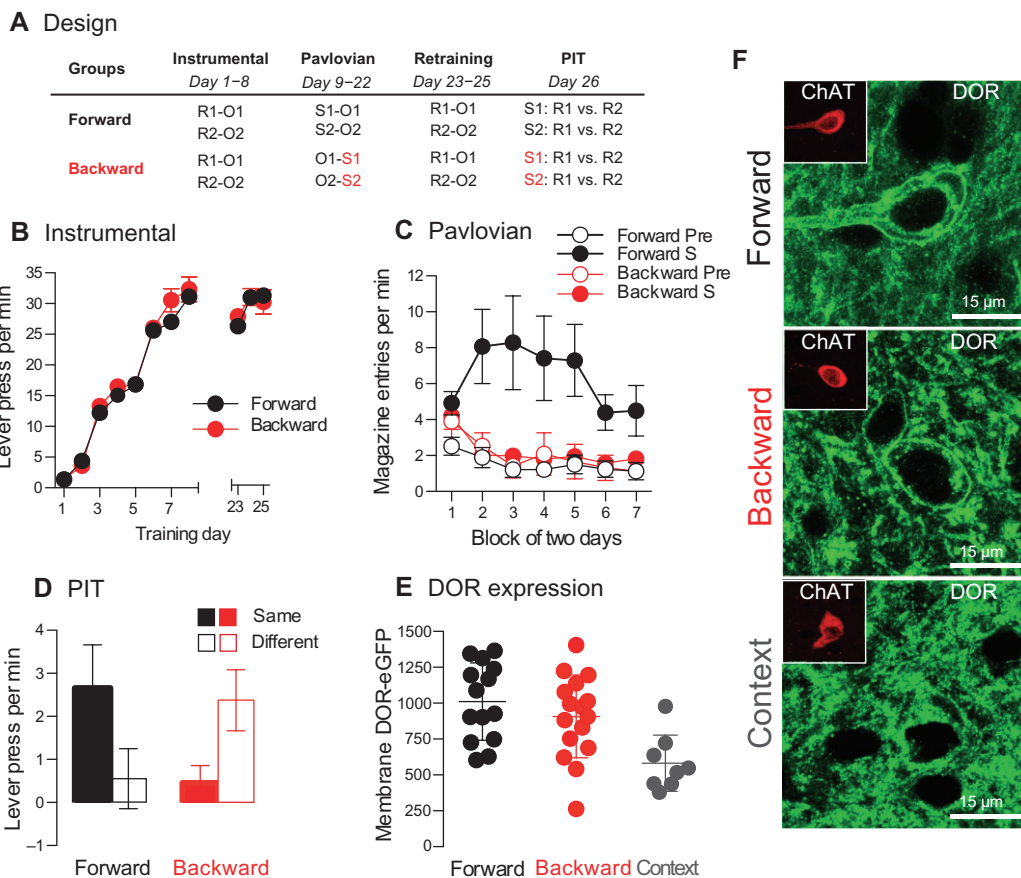


Figure 1

Pavlovian inhibitory training produces δ -opioid receptor accumulation on the membrane of NAC-S CINs. (A) Two groups of δ -opioid receptor-eGFP^{ki} mice were given instrumental followed by Pavlovian conditioning. The Pavlovian conditioning was conducted in a forward manner for one group (i.e. S→O or Forward) and in a backward fashion for the other group (i.e. O→S or Backward). All mice were then given a PIT test. (B) The number of lever presses per minute gradually increased across instrumental training and retraining. (C) Pavlovian conditioning only produced a gradual increase in magazine entries per minutes in the group Forward. (D) Group Forward exhibited outcome-specific PIT as the lever press per minute was higher when the stimulus predicted the same outcome as the response (Same) than when the stimulus predicted a different outcome (Different). Group Backward displayed the opposite pattern: a stimulus predicting the absence of a particular outcome favoured performance on the Different response. (E) Quantification of membrane δ -opioid receptor (DOR) expression in NAC-S CINs revealed no difference between groups Forward and Backward. Both groups displayed higher levels of expression than a control group exposed to the conditioning chambers (grey). (F) Confocal micrographs showing δ -opioid receptor (DOR) distribution in NAC-S ChAT-immunoreactive neurons in mice that received forward or backward training and mice exposed to the context.

delivered two distinct food outcomes (O1 and O2) (see Figure 1A). One group of mice (group Forward; $n = 21$) then received forward Pavlovian training, during which a stimulus, S1, predicted the delivery of O1, whereas S2 predicted the delivery of O2. The other group of mice (group Backward; $n = 24$) received similar training, except that O1 and O2 were delivered 10 s prior to presentation of S1 and S2 respectively. The insertion of this time delay is critical because in previous work, it has been shown to allow the development of an inhibitory relationship between the outcome and the stimulus that follows; that is, it allows S1 to predict the absence of subsequent O1 delivery and S2 to predict the absence of subsequent O2 delivery; that is, stimuli trained in such manner exhibit properties (such as retardation) identical to those display by other conditioned inhibitors (Rescorla, 1969; Maier *et al.*, 1976; Delamater *et al.*, 2003). After Pavlovian

training, both groups of mice received a single PIT test, during which choice between R1 and R2 was assessed both in the absence of S1 and S2 and in their presence.

During instrumental training (Figure 1B), all mice acquired the lever press responding, which increased as the ratio parameters increased across days [$F_{(7,359)} = 307.7$; $P < 0.001$]. The two groups of mice did not differ ($F < 0.7$). The data obtained during Pavlovian training (Figure 1C) revealed no difference in conditioned responding (i.e. magazine entries) between the two groups when the stimuli were absent (period Pre; $F < 0.3$). However, group Forward exhibited higher levels of magazine entries than group Backward in the presence of the stimuli [period S; $F_{(1,258)} = 8.2$; $P < 0.01$]. Further analysis showed that group Forward entered the magazine more in the presence of the stimuli than in their absence [$F_{(1,120)} = 13$; $P < 0.001$], and that the difference

between the two periods increased across days [$F_{(6,293)} = 3.2$; $P < 0.01$]. In contrast, group Backward did not discriminate between the period Pre and the period S ($F < 0.1$).

The data from the PIT test are presented in Figure 1C. The performance is plotted as the mean number of lever presses per minute when the stimulus predicted the delivery, or absence, of the same outcome as the response (Same) and when the stimulus predicted the delivery, or absence, of the different outcome from the response (Different). This test was conducted in extinction, and as such, no outcome was delivered. The levels of responding in the absence of the stimuli were subtracted from these scores to show the elevation over this baseline, as there was no difference in responding in the absence of the stimuli between the two groups ($F < 1$). The Pavlovian training procedure had a clear effect on choice between actions [$F_{(1,89)} = 13.2$; $P < 0.001$]. Group Forward showed the classical PIT effect, a stimulus-biased choice toward the action with which it shared the same outcome [$F_{(1,20)} = 5.2$; $P < 0.05$]; that is, S1:R1 > R2 and S2:R1 < R2. Interestingly, the exact opposite pattern of behaviour was observed in group Backward. Indeed, a stimulus predicting the absence of a particular outcome favoured performance on the action that signalled the delivery of the other food outcome [$F_{(1,23)} = 8.7$; $P < 0.01$]; that is, S1:R1 < R2 and S2:R1 > R2. Thus, inhibitory and excitatory stimuli exerted an opposite influence over choice between actions.

Next, we conducted a confocal analysis to quantify δ -opioid receptor expression on the membrane of NAc-S CINs. We focused upon NAc-S CINs for three reasons. The first is that CINs have been shown to play a critical role in associative learning (Apicella, 2007; Stocco, 2012). The second reason is that we have previously reported a modest and unspecific expression of δ -opioid receptors on striatum medium spiny neurons that contrasted with a clear and somatic expression of these receptors on CINs (Bertran-Gonzalez *et al.*, 2013). The third reason is that we have also found that Pavlovian-induced δ -opioid receptor translocation occurs in the NAc-S but not in other regions of the striatum, such as the dorsal areas and the nucleus accumbens core. Thus, our analysis was conducted on randomly selected subgroups of the mice in groups Forward ($n = 14$) and Backward ($n = 16$), as well as control mice (group control; $n = 8$). These control mice had received the same behavioural procedure as the experiment mice, except that Pavlovian training had been omitted. Instead, unrewarded exposure to the conditioning chambers was given. The results (Figure 1E–F) showed that the mice in groups Forward and Backward displayed higher levels of δ -opioid receptor-eGFP expression in the membrane of CINs in the NAc-S than the control mice [$F_{(1,20)} = 15.6$; $P < 0.001$ and $F_{(1,22)} = 8.1$; $P < 0.01$]. More importantly, membrane δ -opioid receptor accumulation was similar in groups Forward and Backward and did not differ significantly ($F < 1.1$), suggesting that this accumulation may be necessary for both excitatory and inhibitory stimuli to influence choice between actions.

The influence of inhibitory stimuli on choice between actions requires δ -opioid receptor activation in the NAc-S

In Experiment 2, food-deprived rats initially learned that two stimuli (S1 and S2) predicted two distinct food outcomes (O1

and O2; see Figure 2A). Two other stimuli (S3 and S4) were simultaneously trained as conditioned inhibitors, predicting the absence of O1 and O2 respectively. This was achieved through a feature negative conditioned inhibition procedure (Rescorla, 1969), during which two excitatory stimuli, S1 and S2, predicted O1 and O2, respectively, whereas two compound stimuli, one composed of S1 and S3 and S2 and S4, were repeatedly presented in the absence of any outcome. Thus, whereas S1 predicted O1, S3 predicted the absence of O1 and, similarly, whereas S2 predicted O2, S4 predicted the absence of O2. Following Pavlovian training, rats received instrumental training as before and were then given two consecutive PIT tests. These tests occurred after NAc-S infusion (Figure 2G) of either vehicle or the δ -opioid receptor antagonist naltrindole and consisted of presenting two types of compound stimuli: congruent and incongruent. Congruent compounds were those that had been trained during Pavlovian conditioning (i.e. S1S3 and S2S4), whereas incongruent compounds involved presenting a predictor of one outcome with the inhibitor of the other outcome; that is, S1S4 and S2S3. The rationale for comparing the effects of these stimulus compounds is quite straightforward; as a result of Pavlovian training, the congruent compounds S1S3 and S2S4 should act as inhibitors, predicting the absence of O1 and O2 respectively. In contrast, each incongruent compound is a good predictor of one of the two outcomes; for example, presenting S1 – which predicts O1 – with S4 – which signals the absence of O2 – is equivalent to presenting S1 alone. The same logic applies to the compound S2S3, which acts as a good predictor of O2.

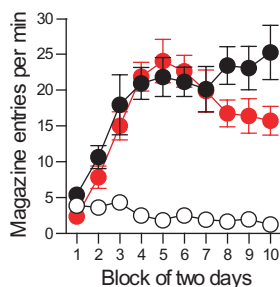
The data from Pavlovian training are presented in Figure 2B. Rats spent significantly more time in the magazine in the presence of the stimuli than in their absence [$F_{(1,90)} = 83.5$; $P < 0.001$ for S1/S2 and $F_{(1,90)} = 67.8$; $P < 0.001$ for the stimulus compounds], and this difference grew larger across days [$F_{(10,219)} = 8.5$; $P < 0.001$ for S1/S2 and $F_{(1,219)} = 19.3$; $P < 0.001$ for the stimulus compounds]. Importantly, rats exhibited higher responding during S1 and S2 than during the stimulus compounds [$F_{(1,90)} = 18.7$; $P < 0.001$], indicating successful inhibitory training. Subsequent instrumental training (Figure 2C) was successful and lever press responding increased over the course of instrumental training [$F_{(9,90)} = 31.8$; $P < 0.001$].

Performance across the PIT test (Figures 2E–F) is plotted for each type of compound (congruent and incongruent) as the number of lever presses per minute when the stimuli predicted the same outcome as the response (Same), when the stimuli predicted a different outcome from the response (Different) or when there was no stimulus (Baseline). Baseline responding was assessed to ensure that any drug-related effect was due to a modulation of choice rather than a change in the animals' ability to perform the task due to a motor impairment. The data of primary interest are those relating to the congruent compounds (Figure 2E). Inspection of the figure clearly suggests that choice between actions was affected by naltrindole and, indeed, the statistical analysis revealed a significant *drug* \times *action* interaction [$F_{(2,20)} = 4.3$; $P < 0.05$]. Thus, vehicle-treated rats elevated their responses on the different response relative to both baseline [$F_{(1,10)} = 7.8$; $P < 0.05$] and the same response [$F_{(1,10)} = 6.4$; $P < 0.05$], whereas the latter did not differ from baseline ($F < 0.4$). These results

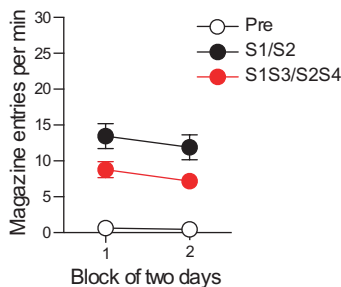
A Design

Pavlovian Day 1–20	Instrumental Day 21–30	Retraining Day 31–34	PIT Day 35–36	
S1-O1 S2-O2 S1S3-∅ S2S4-∅	R1-O1 R2-O2	S1-O1 S2-O2 S1S3-∅ S2S4-∅	Incongruent	S1S4: R1 vs. R2 S2S3: R1 vs. R2
			Congruent	S1S3: R1 vs. R2 S2S4: R1 vs. R2
Vehicle or Naltrindole				

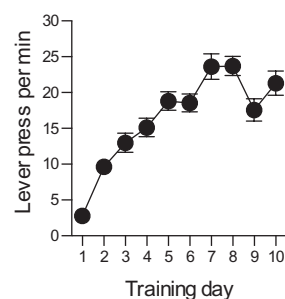
B Pavlovian



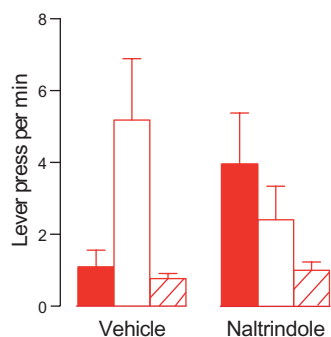
D Retraining



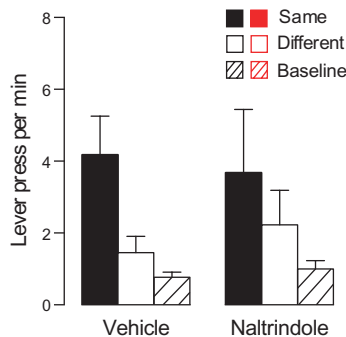
C Instrumental



E PIT Congruent



F PIT Incongruent



G

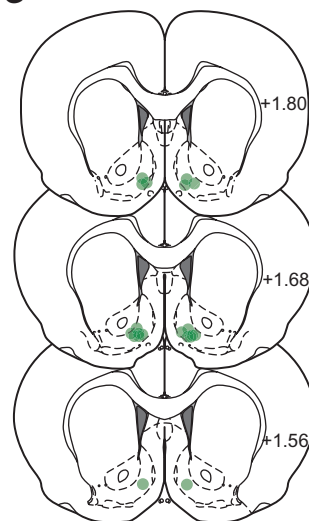


Figure 2

The influence of inhibitory stimuli on choice between actions requires δ -opioid receptor activation in the NAc-S. (A) During Pavlovian training, rats learned that two stimuli predicted distinct outcomes when they were presented alone, but that when presented in compound with two other stimuli, these outcomes were omitted. Following instrumental training and two Pavlovian reminder sessions, the animals received two PIT tests after infusion of either vehicle or naltrindole into the NAc-S. The PIT tests involved presentation of Congruent (red) or Incongruent (black) compounds. The congruent compounds were identical to those presented during Pavlovian training (S1S3 and S2S4), whereas incongruent compounds were composed of novel pairs of stimuli (S1S4 and S2S3). (B and D) The levels of magazine entries per minute were higher when the stimuli were presented alone than when they were presented with the inhibitors. (C) The number of lever presses per minute gradually increased across instrumental training and retraining. (E) Congruent compounds that predicted the absence of particular outcome biased choice away from the action earning that outcome towards the action associated with the absence of this same outcome. This reversal of the usual PIT effect was blocked by NAc-S infusion of naltrindole. (F) Incongruent compounds promoted specific PIT biasing choice towards the action associated with the outcome predicted by the excitatory stimulus. This bias was blocked by NAc-S infusion of naltrindole. (G) Placement of the injection cannula tips in the NAc-S. Distances on the atlas templates are indicated in millimetres relative to bregma.

are consistent with those obtained in Experiment 1 using the backwardly paired stimuli and confirm that an inhibitor of a particular outcome biases choice away from an action associated with that outcome (i.e. same) towards an action associated with the absence of the outcome (i.e. different). Importantly, this effect was abolished by infusion of naltrindole into the NAc-S. Indeed, there was no significant difference between responding the same and different response and between those responses and baseline (all F 's < 4.6; P 's > 0.05). Similar effects were observed for the

incongruent compounds, reflecting an assessment of the effect of conditioned excitors on choice, and which replicated our previous findings (Figure 2E); that is, vehicle-treated rats exhibited the usual PIT effect, elevating responding on the same response relative to both the different response [$F_{(1,10)} = 10.2$; $P < 0.05$] and baseline [$F_{(1,21)} = 10.2$; $P < 0.05$], whereas, consistent with our previous findings (Laurent *et al.*, 2012), naltrindole blocked the ability of the incongruent compounds to guide choice (F 's < 2.5, P 's > 0.05).

Discussion and conclusions

The present experiments examined whether excitatory and inhibitory stimuli influence choice between actions via a δ -opioid receptor-related process in the NAc-S. In both experiments, we found that excitatory stimuli produced outcome-specific PIT and they biased responding towards the action earning the outcome predicted by the stimuli. In contrast, inhibitory stimuli, by predicting the absence of a specific outcome, favoured actions delivering alternative outcomes (i.e. outcomes whose absence was not anticipated). Interestingly, both effects were associated with an increase in δ -opioid receptor expression at the membrane of NAc-S CINs and both effects were prevented by δ -opioid receptor blockade in the NAc-S. These findings suggest that δ -opioid receptor-related processes provide a general mechanism through which both excitatory and inhibitory stimuli influence choice between actions.

The increase in δ -opioid receptor expression observed in the current study is likely to have occurred as a consequence of the Pavlovian training as previous work in our laboratory revealed a similar increase in mice that were only given this training (Bertran-Gonzalez *et al.*, 2013). However, these experiments also demonstrated that exposure to a strong stimulus-outcome contingency was critical; mice receiving uncorrelated stimulus-outcome presentations failed to show δ -opioid receptor accumulation on CINs. This increase in expression was confirmed here in mice given forward pairings of the stimuli and outcome but, importantly, was also observed in mice given backward training. Indeed, the latter exhibited a similar and high degree of δ -opioid receptor accumulation to the forward paired group. Thus, learning about (backward) stimuli that reliably predict the absence of food is as able to promote changes in δ -opioid receptor expression as learning about (forward) stimuli that predict the presence of food. It appears, therefore, that δ -opioid receptors accumulate on NAc-S CINs when stimuli are good predictors of important events, irrespective of whether these events predict the presence or the absence of those events.

Although Pavlovian predictive learning triggers a long-term change in δ -opioid receptor expression in the NAc-S, these receptors do not appear necessary for that learning *per se*. For instance, mice carrying a genetic deletion of δ -opioid receptors show no apparent deficits in Pavlovian training (Laurent *et al.*, 2012), yet they are unable to show outcome-specific PIT. The involvement of a δ -opioid receptor-related process at the time of choice was later confirmed by showing that outcome-specific PIT was impaired by blockade of these receptors in the NAc-S using naltrindole (Laurent *et al.*, 2012). This impairment was reproduced here in the presence of a compound of Pavlovian stimuli. More importantly, the present study found that δ -opioid receptor activity in the NAc-S is also necessary for inhibitory compounds to affect choice between actions. This is entirely consistent with the previous finding that δ -opioid receptors accumulate in the NAc-S as a result of backward/inhibitory training and confirms that this accumulation is functional at the time of testing.

It is important to note that the behavioural effects found in the current study are not entirely the same as those previously reported. In the study conducted by Delamater *et al.*

(2003), although a Pavlovian inhibitor of a particular outcome reduced responding on the action earning the same outcome, it left responding on the 'Different' action relatively intact relative to baseline. In contrast, responding on the 'Different' action was clearly elevated in the present experiments. However, Delamater *et al.* (2003) did not extinguish instrumental responding prior to testing the Pavlovian stimuli, leaving open the possibility that observing an increase in responding on the 'Different' action was prevented by a ceiling effect. Whatever the source of this discrepancy between the two studies, the essential finding – that backward conditioning reverses the usual difference between same and different actions – was clearly replicated and suggests that choice between actions can be biased by both excitatory and inhibitory predictions.

Although learning about excitatory and inhibitory stimuli may well depend upon distinct neural substrates, the present experiments indicate that both forms of learning influence choice between actions through similar mechanisms. These mechanisms involve Pavlovian learning-induced translocation of δ -opioid receptors on CINs in the NAc-S and the subsequent activation of these receptors at the time of choice. Clearly, the present research constitutes only a first step in our attempt to understand the role played by δ -opioid receptors in decision-making processes. For instance, it would be interesting to evaluate whether the same mechanisms are implicated in choices involving other rewards than food. It remains also essential to determine the functional consequences of δ -opioid receptors in the NAc-S at the time of PIT, and to guide our assessment of this issue, we have recently described a model of NAc-S function in which δ -opioid receptors modulate CINs to influence the activity of both medium spiny neurons bearing dopamine D₁ receptors and outcome-specific PIT (Laurent *et al.*, 2014). Finally, it is worth noting that the integrative process through which stimuli influence instrumental actions is deficient in many psychiatric conditions, including drug addiction, obesity and psychotic disorders (Hyman, 2005; Simpson *et al.*, 2010; Petrovich, 2011). It is therefore critical to improve our understanding of the cellular mechanisms involved in the stimulus control of actions in order to develop new pharmacological targets for treating these disorders.

Acknowledgements

The authors thank Mr Pietro Vanni for his assistance, Professor Brigitte Kieffer for δ -opioid receptor-eGFP knock-in mice and Professor Mac Christie for helpful discussions of this project. This research was supported by funding from Australian Research Council (ARC), grant DP#130103965, to V.L. and B.W.B., a by Laureate Fellowship, #F0992409, from the ARC to B.W.B.

Author contributions

Vincent Laurent helped design the experiments, implemented the experimental design, analysed the data and

helped write the manuscript. Felix Wong conducted the experiments, helped in the analysis of the data and helped write the manuscript. Bernard Balleine designed the experiments, helped implement the experimental design and wrote the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

- Alexander SPH, Benson HE, Faccenda E, Pawson AJ, Sharman JL, Spedding M *et al.* (2013). The Concise Guide to PHARMACOLOGY 2013/14: G Protein-Coupled Receptors. *Br J Pharmacol* 170: 1459–1581.
- Apicella P (2007). Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci* 30: 299–306.
- Bertran-Gonzalez J, Laurent V, Chieng BC, Christie MJ, Balleine BW (2013). Learning-related translocation of δ -Opioid receptors on ventral striatal cholinergic interneurons mediates choice between goal-directed actions. *J. Neurosci.* 33: 16060–16071.
- Colwill RM, Rescorla RA (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J Exp Psychol Anim Behav Process* 14: 155–164.
- Delamater AR, LoLordo VM, Sosa W (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Learn Behav* 31: 393–402.
- Dickinson A, Balleine B (1994). Motivational control of goal-directed action. *Anim Learn Behav* 22: 1–18.
- Holmes NM, Marchand AR, Coutureau E (2010). Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neurosci Biobehav Rev* 34: 1277–1295.
- Hyman SE (2005). Addiction: a disease of learning and memory. *Am J Psychiatry* 162: 1414–1422.
- Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG (2010). Animal research: reporting *in vivo* experiments: the ARRIVE guidelines. *Br J Pharmacol* 160: 1577–1579.
- Laurent V, Leung B, Maidment N, Balleine BW (2012). μ - and δ -opioid-related processes in the accumbens core and shell differentially mediate the influence of reward-guided and stimulus-guided decisions on choice. *J. Neurosci.* 32: 1875–1883.
- Laurent V, Bertran-Gonzalez J, Chieng BC, Balleine BW (2014). δ -opioid and dopaminergic processes in accumbens shell modulate the cholinergic control of predictive learning and choice. *J Neurosci* 34: 1358–1369.
- Lutz P-E, Kieffer BL (2013). The multiple facets of opioid receptor function: implications for addiction. *Curr Opin Neurobiol* 23: 473–479.
- Maier SF, Rapaport P, Wheatley KL (1976). Conditioned inhibition and the UCS-CS interval. *Anim Learn Behav* 4: 217–220.
- McGrath JC, Drummond GB, McLachlan EM, Kilkenny C, Wainwright CL (2010). Guidelines for reporting experiments involving animals: the ARRIVE guidelines. *Br J Pharmacol* 160: 1573–1576.
- Paxinos G, Watson C (2006) *The Rat Brain in Stereotaxic Coordinates*: Hard Cover Edition. Academic Press: London.
- Petrovich GD (2011). Forebrain circuits and control of feeding by learned cues. *Neurobiol Learn Mem* 95: 152–158.
- Rescorla RA (1969). Pavlovian conditioned inhibition. *Psychol Bull* 72: 77–94.
- Simpson EH, Kellendonk C, Kandel E (2010). A possible role for the striatum in the pathogenesis of the cognitive symptoms of schizophrenia. *Neuron* 65: 585–596.
- Stocco A (2012). Acetylcholine-based entropy in response selection: a model of how striatal interneurons modulate exploration, exploitation, and response variability in decision-making. *Front Neurosci* 6: 18.