Neuroscience 163 (2009) 770–780

DISRUPTION OF ENDOGENOUS OPIOID ACTIVITY DURING INSTRUMENTAL LEARNING ENHANCES HABIT ACQUISITION

K. M. WASSUM,[a,c]* I. C. CELY,[a] N. T. MAIDMENT[a,c] AND B. W. BALLEINE[a,b,c,d]

[a]*Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA*

[b]*Department of Psychology, UCLA, Los Angeles, CA 90024, USA*

[c]*Brain Research Institute, UCLA, Los Angeles, CA 90024, USA*

[d]*BMRI, University of Sydney, Australia*

Abstract—Considerable evidence suggests that in instrumental conditioning rats learn the relationship between actions and their consequences, or outcomes. Such goal-directed actions are sensitive to changes in outcome value. The present study assessed the role of the endogenous opioid system in goal-directed reward learning. In two experiments, rats were trained to lever press for food pellets either under vehicle or naloxone-induced opioid receptor blockade. Specific satiety procedures were used for outcome devaluation, and the effect of this devaluation on instrumental responding was then tested in extinction. In Experiment 1 outcome devaluation resulted in a reduction in lever pressing in rats that were trained after vehicle injections, indicating that actions in these rats were goal-directed. In contrast, actions in rats trained under naloxone were insensitive to outcome devaluation when tested off drug, suggesting that lever pressing had become habitual in these rats. Interestingly, in Experiment 2 naloxone-induced habitual behavior was shown to be specific to the context in which the training occurred under naloxone; rats showed normal sensitivity to outcome devaluation when tested in an alternate vehicle-trained context. Additionally, in Experiment 2 we found that the acute administration of naloxone on test had no effect in itself, indicating that opioid receptor-related processes contribute to the acquisition of goal-directed actions and not to their general performance. These data suggest that an intact endogenous opioid system is necessary for normal goal-directed learning and more importantly, reveal that a compromised endogenous opioid system during learning enhances the habitual control of actions. **© 2009 IBRO. Published by Elsevier Ltd. All rights reserved.**

Key words: opioid receptors, naloxone, learning, goal-directed action, habit, reward.

Considerable evidence suggests that the performance of goal-directed instrumental actions requires the integration of two learning processes; one through which animals acquire and represent the consequences of their actions, the action–outcome association, and a second through which they assign reward or incentive value to those consequences (Balleine, 2001; Dickinson and Balleine, 2002). It is the dependency on this integrative process between action–outcome learning and outcome value that allows animals rapidly to alter their performance when the experienced value of the outcome changes (Balleine, 2001; Dickinson and Balleine, 2002). Indeed, recent evidence suggests that altering this process by disrupting the corticostriatal network involved in action–outcome learning can result in actions becoming habitual, i.e. insensitive to changes in value and more dependent on antecedent stimuli than their consequences (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Ostlund and Balleine, 2005; Yin et al., 2005). Likewise, disruption of the basolateral amygdala has been found to alter reward-value-related processing and to render instrumental performance insensitive to changes in outcome value (Balleine et al., 2003; Wang et al., 2005; Wellman et al., 2005; Ostlund and Balleine, 2008). It appears, therefore, that disrupting the specific circuitry governing either action–outcome or reward learning processes can significantly compromise the ability of animals to exert goal-directed executive control over instrumental actions.

Although there is now considerable information regarding the neural circuits that control action–outcome learning and the shift to habitual control, it is unclear, at present, whether neurochemical disruption of reward-related learning affects the shift in control from a goal-directed to a stimulus–response habit learning process. From this perspective evidence implicating the endogenous opioid receptor system in reward processing is of considerable interest. The endogenous opioid receptor system has long been implicated in the hedonic aspects of reward processing (Skoubis et al., 2001, 2005; Pecina et al., 2006). However, opioid peptide–containing neurons, their terminals and opioid receptors are present in multiple basal forebrain regions that are involved in motivational, rather than hedonic, aspects of reward processing (Ding et al., 1996; Poulin et al., 2006). Indeed, there is dense mu opioid receptor localization in the dorsomedial striatum and nucleus accumbens core (Daunais et al., 2001), regions which are implicated in the learning and performance of goal-directed actions, respectively (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Ostlund and Balleine, 2005; Yin et al., 2005). Moreover, the opioid peptides, beta-endorphin and enkephalin, have been shown to positively contribute to the motivation for rewards (Hayward et al., 2002). More recently, opioid receptors in the basolateral amygdala have been shown to be critical in encoding changes in the incentive value of an instrumental outcome

*Correspondence to: K. M. Wassum, Department of Psychiatry and Biobehavioral Sciences, UCLA Semel Institute for Neuroscience and Human Behavior, 675 Charles E Young Drive South, Los Angeles, CA 90024, USA. Tel: +1-310-206-7980.
E-mail address: kwassum@ucla.edu (K. M. Wassum).
Abbreviation: RR, random ratio.

0306-4522/09 $ - see front matter © 2009 IBRO. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuroscience.2009.06.071

(Wassum et al., 2009). Therefore, in the present study we tested the hypothesis that blockade of the endogenous opioid system during learning would compromise goal-directed reward learning and so accelerate the control of instrumental performance by a habit learning process.

In order to test this hypothesis we first compared the effects of instrumental learning under naloxone, a general opioid receptor antagonist, to the effects of over-training, a manipulation known to produce habitual responding, in inducing insensitivity to specific satiety-induced outcome devaluation (Experiment 1). Next we assessed whether naloxone enhanced the stimulus control of instrumental performance by differential training in a naloxone-paired and vehicle-paired context and compared these effects with the effects of acutely administered naloxone on outcome devaluation (Experiment 2). We demonstrate that the blockade of opioid receptors during instrumental training reduces goal-directed control of instrumental performance and increases stimulus control consistent with the claim that actions trained under naloxone rapidly become habitual.

EXPERIMENTAL PROCEDURES

Experiment 1: comparing the effects of over-training and opioid-receptor blockade during under-training on outcome devaluation

Subjects. Male Sprague–Dawley rats (200–220 g, Charles River Laboratories, Wilmington, MA, USA, $n=20$) were housed four per cage and were handled daily for 1 week prior to instrumental training. Training and testing took place during the light phase of the 12-h light/dark cycle. Rats were maintained on a food-deprived schedule whereby they received 10–12 g of their maintenance diet daily in order to maintain approximately 85% free-feeding body weight. All rats were fed after each day's training session. Rats had free access to tap water in the home cage. We used only the number of animals needed for sufficient power to conduct these studies and made every attempt to minimize their discomfort and pain throughout the experiment. All procedures were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of California, Los Angeles Institutional Animal Care and Use Committee.

Apparatus. Training and testing took place in four Med Associates (East Fairfield, VT, USA) operant chambers housed within sound- and light-resistant shells. Each chamber contained two retractable levers that could be inserted to the left and right of the magazine. With the exception of the reinstatement test, only one lever was inserted during training and testing sessions. A 3-W, 24-V house light mounted on the top of the center wall opposite the magazine provided illumination when necessary. Each chamber was equipped with a pellet dispenser that when activated delivered a single 45 mg precision purified sucrose or grain-based pellet (Bio-Serv, Frenchtown, NJ, USA) into a recessed magazine. Chambers provided one of two contexts. One context was the standard Med Associates chamber: grid floor, square in shape, with woodchip bedding and the house light illuminated during training and testing. A second context was altered with compressed paper bedding, a solid sandpaper-textured floor, a curved white textured wall and no house light illumination. The type of pellet reinforcement and lever (either to the left or right of the magazine) was counter-balanced by context and in subsequent training.

Training. The basic training paradigm was adapted from Killcross and Coutureau (2003). A summary of the training and testing procedures can be found in Table 1. In total, training consisted of three components: magazine training, over-training (approximately 500 action–outcome pairings) and relatively less training (i.e. under-training—approximately 50 action–outcome pairings). For each rat magazine training and over-training took place in one context while, concurrently, additional magazine training and under-training took place in an alternate context. Over-training and under-training were conducted on different levers and for different reward outcomes. The contexts, lever position and outcome were counter-balanced between training types.

Table 1. Design of Experiment 1

Training		Test	
Phase I	Phase II	Devalue	Extinction
C_1: $R_1 \rightarrow O_1$	*Group Vehicle* C_1: $R_1 \rightarrow O_1$ C_2: (VEH) $R_2 \rightarrow O_2$	O_1 or O_2	C_1: $R_1 \rightarrow$ ø
C2: O2	*Group Naloxone* C_1: $R_1 \rightarrow O_1$ C_2: (NLX) $R_2 \rightarrow O_2$		C_2: $R_2 \rightarrow$ ø

Rats were initially trained on one lever (R_1) for one outcome (O_1) in one context (C_1) on a random interval 60 s schedule and given a second outcome (O_2) freely in a different context (C_2). They then received further training on a RR-20 schedule in C_1 while in C_2 receiving concurrent training on a second lever (R_2) on RR-20 to earn O_2. This training in C_2 was conducted either under vehicle (VEH) or naloxone (NLX, 4 mg/kg 1 ml/kg i.p.). Two sets of non-rewarded tests on the levers were then conducted; one set after specific satiety-induced devaluation of O_1 and another after devaluation of O_2 (counter-balanced), each set comprising tests in C_1 and C_2 (also counter-balanced).

Magazine training. Rats received 3 days of magazine training in which they were exposed to non-contingent delivery of either the 45 mg sucrose or grain pellet (20 outcomes over 30 min) in the operant chamber with levers retracted. The pellet they received in this training was that delivered on the over-trained lever (see below).

Over-training. After magazine training all rats were given a single session of continuous reinforcement training on a single lever in a single context, in which they could earn 30 outcomes. Subsequently, rats were given 14 additional training sessions, one per day, on a random interval 60 s schedule in which they could earn 30 outcomes per session. Concurrently, all rats received an additional magazine training session for the other reward outcome each day in the other, to be under-trained, context in order to equalize the number of outcomes consumed and the context–outcome pairings. At this point rats had received approximately 450 action–outcome pairings in the over-trained context. After the 15th total training session, rats were given an additional five training sessions, one per day, in the over-trained context in which they could earn a maximum of 10 outcomes on a random ratio 20 (RR-20) schedule. This allowed for an additional 50 action–outcome pairings, for a total of 500 action–outcome pairings in the over-training context.

Under-training and drug administration. Concurrent with the five sessions of RR-20 training in the over-training context, rats were also trained in the other context, on the other lever for the other instrumental outcome. In the first session rats could earn a maximum of 10 outcomes on a RR-10 schedule, then for the remaining four sessions the requirement was increased to a RR-20 schedule. This under-training condition provided 50 action–outcome pairings over 5 days. In addition, prior to each of these under-training sessions rats were given one of two treatments: one group of rats received an injection of vehicle (1 ml/kg), while the other received an injection of naloxone (4 mg/kg; 1 ml/kg i.p.)

15 min prior to training in the under-training context. To ensure that the effects of drug were limited to training in the under-training context this treatment and training occurred in the afternoon, after the over-training session was completed.

Drug administration. The non-specific opioid receptor antagonist naloxone (4 mg/kg, 1 ml/kg, Sigma, St. Louis, MO, USA) was dissolved in sterile saline and injected i.p. 15 min prior to each training session. The same volume of sterile saline was injected i.p. as the vehicle control. No drug was administered on test day.

Outcome devaluation and testing. After the instrumental training phase, each rat received 2 days of specific-satiety outcome devaluation and testing separated by a day of retraining. Testing was conducted sequentially in both the over-trained and under-trained context in extinction on both days, counter-balanced for order. The 2 days of testing differed only in the outcome that was devalued. The testing procedures are illustrated in Table 1. On the day after the final day of training all rats were individually given free access to one of the two outcomes, counter-balanced for whether over-trained or under-trained, for 1 h in a chamber identical to their home cage, but devoid of bedding. (Prior to the onset of the experiment (i.e. before magazine training) rats were allowed experience with the pre-feeding chamber to ensure appropriate outcome consumption at test.) The amount of the outcome consumed during the pre-feeding period was recorded. Immediately thereafter rats were given a 10 min extinction test in each of the over-training and under-training contexts, during which lever press responses were recorded in the absence of outcome delivery. The order of testing was counter-balanced with respect to training and pre-fed outcome type. After the second extinction test, rats were returned to the home cage–like feeding chambers and presented with each of the outcomes for 15 min in counter-balanced for order. The amount of the devalued and non-devalued outcome consumed was measured. This test allowed for evaluation of the effectiveness of the pre-feeding procedure to devalue the specific outcome. The following day rats received a single retraining session in both the over-training and under-training contexts on the RR-20 schedule. Naloxone or vehicle injections, as appropriate, were given 15 min prior to the under-training session. The next day, a second test was conducted identical to the first, except rats were pre-fed on the alternate outcome to that presented in test 1 (e.g. if given grain pellets prior to test 1, they received sucrose pellets prior to test 2). As in test 1, the order of testing was counter-balanced with respect to training and pre-fed outcome type and a post-test feeding assessment was conducted as in test 1.

Data analysis. All responding in extinction after the devaluation tests was normalized to baseline prior to analysis. The baseline for each rat was taken as the average response rate on the lever, calculated separately for each context, in the last two training sessions prior to the test. For the first devaluation test this included the last two training sessions prior to the onset of the first devaluation test, while for the second test this included the last training session prior to the onset of the first devaluation test and the single retraining session in between the two devaluation tests. Response rate/percentage baseline data were analyzed using two-tailed two- or three-way ANOVAs followed by Bonferroni multiple comparison post hoc tests where appropriate (GraphPad Prism, San Diego, CA, USA and SPSS, Chicago, IL, USA). For all tests the alpha level for significance was set to $P<0.05$.

Experiment 2: contextual control of the effects of opioid receptor blockade during learning on outcome devaluation

Subjects and apparatus. Male Sprague–Dawley rats (200–220 g, Charles River Laboratories, Wilmington, MA, USA, $n=24$) were housed, handled and maintained as in Experiment 1. Training and testing took place in the apparatus and the contexts described in Experiment 1. Drug delivery was also conducted using the procedures described in Experiment 1.

Training. Experiment 2 differed from the first experiment in that a within-subjects training design with respect to both context and drug treatment was employed in which each rat was trained on one action–outcome relationship in one context after an injection of vehicle (1 ml/kg) and a second action–outcome relationship in a second context after an injection of naloxone (4 mg/kg, 1 ml/kg i.p.). In this experiment, all rats were under-trained. The contexts, levers and outcomes were counter-balanced between drug conditions.

Magazine training. Rats received 3 days of magazine training (both contexts each day) in which they were exposed to non-contingent delivery of either the 45 mg sucrose or grain pellets (20 outcomes over 30 min) in the operant chamber with the lever retracted.

Training. After magazine training all rats were given instrumental training. Fifteen minutes prior to each training session in the morning rats received an injection of vehicle that matched the volume of naloxone (4 mg/kg, 1 ml/kg i.p.) given 15 min prior to the afternoon training session. Approximately 2.5 h separated each training session. The context and type of outcome earned were counter-balanced across sessions. Training was otherwise identical across these sessions. The first four training sessions consisted of continuous reinforcement training on the single lever in both contexts, in which rats could earn a maximum of 10 outcomes. Subsequently, rats were moved to an RR-5 schedule for 2 days, followed by 2 days of RR-10 and 4 days of RR-20 training. In each training session rats could earn up to 10 outcomes, allowing for approximately 120 action–outcome pairings in both the context trained under vehicle and that trained under naloxone-induced opioid receptor blockade.

Outcome devaluation and test. After instrumental training each rat received 4 days of specific-satiety outcome devaluation and testing separated by retraining days. Testing was conducted sequentially in both the vehicle- and naloxone-trained contexts in extinction each day, counter-balanced for order. The 4 days of testing differed in the outcome that was devalued and the drug administered (naloxone or vehicle) prior to satiety treatment and testing. The testing procedures are illustrated in Table 2. On the

Table 2. Design of Experiment 2

Training	Test	
	Devalue	Extinction
C_1: (VEH) $R_1 \rightarrow O_1$	(VEH): O_1	C_1: $R_1 \rightarrow ø$ C_2: $R_2 \rightarrow ø$
C_2: (NLX) $R_2 \rightarrow O_2$	(VEH): O_2	C_1: $R_1 \rightarrow ø$ C_2: $R_2 \rightarrow ø$
	(NLX): O_1	C_1: $R_1 \rightarrow ø$ C_2: $R_2 \rightarrow ø$
	(NLX): O_2	C_1: $R_1 \rightarrow ø$ C_2: $R_2 \rightarrow ø$

Rats were trained on one lever (R_1) for one outcome (O_1) in one context (C_1) and a on a second lever (R_2) for a second outcome (O_2) in a second context (C_2). Training in C_1 was conducted after an injection of vehicle (VEH) and in C_2 after an injection of naloxone (NLX, 4 mg/kg 1 ml/kg i.p.). Four sets of non-rewarded tests on the levers were then conducted, each set incorporating tests in C_1 and C_2. Two sets were conducted after an injection of vehicle and two after an injection of NLX. The two sets within each drug treatment differed in the outcome that was devalued. Devalued outcome and drug were counter-balanced for order.

day after the final day of training all rats were injected with either vehicle (1 ml/kg) or naloxone (4 mg/kg, 1 ml/kg i.p.) and individually given free access to one of the outcomes, counter-balanced for whether the outcome was earned in the vehicle or naloxone context, for 1 h in a chamber identical to their home cage, but devoid of bedding and water. The amount of the outcome consumed during this period was recorded. Immediately after this pre-feeding session rats were given a 10 min extinction test in each of the vehicle- and naloxone-trained contexts, during which lever press responses were recorded in the absence of outcome delivery. The order of testing was counter-balanced with respect to drug condition during training and pre-fed outcome type. Following the second extinction test rats were returned to the home cage-like feeding chambers and presented with each of the outcomes for 15 min (counter-balanced for order). The amount of each outcome consumed was measured allowing us to evaluate the effectiveness of the pre-feeding procedure to devalue the specific outcome.

The next day rats received a single retraining session in each context on the RR-20 schedule after the appropriate injection of either vehicle or naloxone. A second devaluation test was then conducted the next day prior to which rats were injected with the same drug (either vehicle or naloxone) that they received on the first test day. This test was identical to the first, except rats were pre-fed on the alternate outcome to that used in test 1, e.g. if they were fed grain pellets prior to test 1, they were fed sucrose pellets prior to test 2. As in test 1, the order of testing was counter-balanced with respect to drug on training condition and pre-fed outcome type. After this test the post-feeding test was conducted as in test 1.

Rats were then retrained for 2 days in both the vehicle and naloxone contexts on the RR-20 schedule following the appropriate injections. The set of two satiety tests was then conducted again, with an intervening day of retraining, exactly as described above except that rats received the alternate drug treatment on testing, i.e. if they received vehicle prior to the previous tests, they received naloxone prior to these tests.

Outcome specific reinstatement test. After the devaluation tests, a subset of rats ($n=8$) was given an additional test to check for their ability to discriminate between the two outcomes in the naloxone-trained context. For this test, conducted off drug, half of the rats were placed in the vehicle-trained and the other half in the naloxone-trained context. Rats were given at least 5 min of extinction responding with both levers inserted; extinction continued until their response rate dropped below one response per 15 s at which point the levers were removed. An average of 60 s after the end of the extinction period either the sucrose or grain pellet was presented non-contingently, counter-balanced across context. After the rat retrieved the pellet from the magazine both levers were inserted into the chamber and the rat's outcome-specific reinstatement-induced responding was assessed. When the rats met the above extinction criterion again, the other pellet was delivered (if previously a grain pellet was delivered then now a sucrose pellet was delivered). This reinstatement assessment was repeated for each pellet in the same order. For this test responding on the levers for the 2 min after pellet delivery was normalized to a baseline that consisted of an average of the rats' response rate in the 1 min before each pellet delivery.

Data analysis. All responding in extinction after the devaluation tests was normalized to baseline prior to analysis. The baseline for each rat was taken as the average response rate on the lever, calculated separately for each context, in the last two training sessions prior to the test. Response rate/percentage baseline data were analyzed using two-tailed two- or three-way ANOVAs followed by Bonferroni multiple comparison post hoc tests where appropriate (GraphPad Prism and SPSS). For all tests the alpha level for significance was set to $P<0.05$.

In several cases a manipulation, such as devaluation, was found to have no effect on lever pressing actions in a particular condition. These findings proved critical to the interpretation of the data. As such, we computed Bayes factors for use in supporting the null hypothesis in select cases (Gallistel, 2009; Rouder et al., 2009). These analyses were conducted using a freely-available Bayes factor calculator (http://pcl.missouri.edu/bayesfactor) (Rouder et al., 2009), and have recently been suggested as an appropriate method for expressing a preference for the null hypothesis (Gallistel, 2009; Rouder et al., 2009).

RESULTS

Experiment 1: over-training and opioid-receptor blockade during under-training results in insensitivity to outcome devaluation

Experiment 1 was designed to compare the effects of an over-training procedure known to produce habitual behavior in rats to the effects of systemic opioid receptor blockade during under-training (see Table 1) on action performance after outcome devaluation.

Training. Training progressed smoothly in both contexts for all rats (see supplementary data for statistical analysis of training data). On the last day of training the mean response rates for each group of rats were very similar: in the over-trained context, rats that received vehicle in the under-trained context responded at an average rate of 26.8 presses/min (SEM=3.9), whereas those that received naloxone in the under-trained context responded at an average rate of 18.71 presses/min (SEM=2.2). In the under-trained context, rats in the vehicle group responded at an average response rate of 24.1 presses/min (SEM=4.2), whereas those in the naloxone group responded at an average response rate of 20.0 presses/min (SEM=3.4).

Over-training and under-training after opioid receptor blockade results in similar insensitivity to outcome devaluation. The results of the outcome devaluation tests are presented in Fig. 1. The effect of outcome devaluation on action performance in the over-trained context in Groups Under-trained Vehicle and Under-trained Naloxone are presented in the left panels of Fig. 1a and 1b, respectively. Inspection of these figures suggests that, irrespective of under-trained context drug treatment, rats were insensitive to devaluation in the over-trained context responding at similar rates on the lever when the outcome associated with responding was devalued as when it was not. In contrast, performance in the under-trained context depended on drug treatment. Inspection of the right panels of Fig. 1a and 1b suggests that rats in Group Vehicle, that received vehicle injections prior to each under-training session, were sensitive to outcome devaluation and reduced performance on the lever appropriately when its training outcome was devalued relative to when it was not. This was not true of rats in Group Naloxone (Fig. 1b) who, in similar fashion to the effects of over-training, were, again, demonstrably insensitive to the effects of outcome devaluation.

A three-way ANOVA with a between subjects factor of drug (naloxone vs. vehicle) and within-subjects factors of

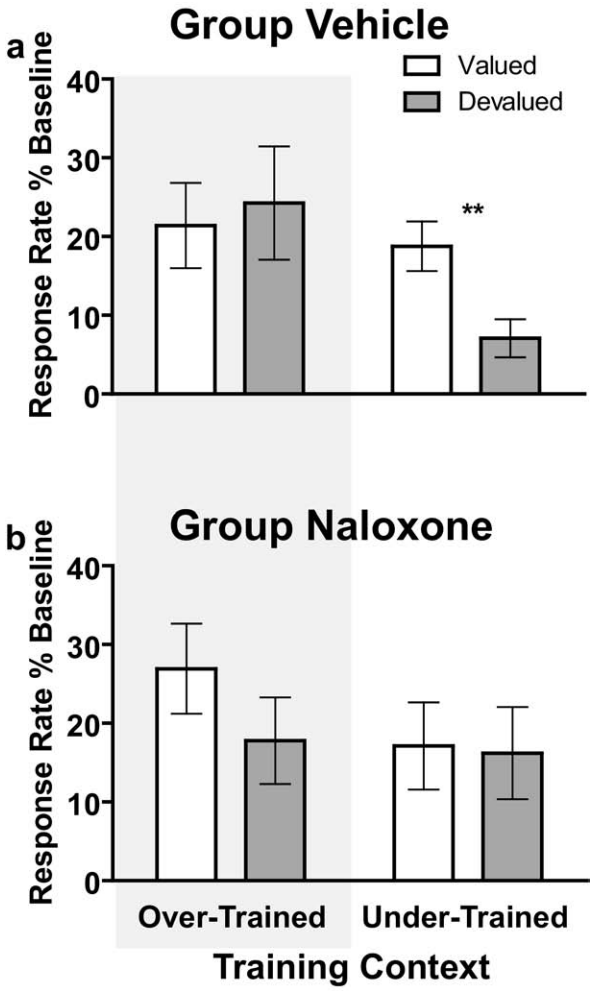

Fig. 1. Over-training and naloxone, but not vehicle, during under-training resulted in insensitivity to outcome devaluation. In Experiment 1 rats were trained on two outcomes, one in each of two contexts, one over-trained (500 action-outcome pairings) and one under-trained (50 action-outcome pairings) following an injection of naloxone or vehicle. At test rats were pre-fed on one of the outcomes for 1 h and then allowed to respond in extinction in both contexts. (a) Extinction response rates in both contexts, plotted as a percentage of baseline, following specific-satiety–induced outcome devaluation for rats that were given vehicle prior to under-training sessions. Devalued refers to the outcome which was pre-fed to the rats for 1 h prior to the test. (b) Extinction response rates in both contexts, plotted as a percentage of baseline, following specific-satiety-induced outcome devaluation for rats that were given naloxone prior to under-training sessions (n=20). ** $P<0.01$.

training context (over-trained vs. under-trained) and devaluation (devalued vs. not devalued) revealed main effects of training context ($F_{1,17}$=4.35, P=0.05) and devaluation ($F_{1,17}$=5.11, P=0.04), but no overall main effect of drug ($F_{1,17}$=0.09, P=0.77). Although this analysis found no training context by drug interaction ($F_{1,17}<0.001$, P=0.98), devaluation×drug interaction ($F_{1,17}$=0.47, P=0.50), or training context×devaluation interaction ($F_{1,17}$=0.41, P=0.53), it did, importantly, reveal a significant training context×devaluation×drug three-way interaction ($F_{1,17}$=8.38, P=0.01). To clarify these effects individual two-way repeated measures ANOVAs were conducted separately for extinction responding in the over-trained and under-trained contexts, represented in the left and right panels of Fig. 1, respectively. A two-way ANOVA conducted on responding in the over-trained context (Fig. 1a and b left panels) with under-training drug treatment and devaluation as factors, revealed neither an effect of drug treatment ($F_{1,17}$=0.003, P=0.95), nor devaluation ($F_{1,17}$=0.93, P=0.35), and found no drug×devaluation interaction ($F_{1,17}$=3.39, P=0.08). This analysis indicates that, regardless of under-training drug treatment, rats responding in the over-trained context were insensitive to specific-satiety induced outcome devaluation. A similar analysis of responding in the under-trained context (Fig. 1a and b, right panels), revealed no overall main effect of drug ($F_{1,17}$=0.41 P=0.53), but did find a significant main effect of devaluation ($F_{1,17}$=6.07, P=0.02), and, importantly, a significant drug×devaluation interaction ($F_{1,17}$=4.43, P=0.04). Bonferroni post hoc analysis of these data, correcting for multiple comparisons, revealed that the effect of devaluation was restricted to the group that was under-trained after vehicle injections ($P<$ 0.01); outcome devaluation had no significant effect on extinction responding in the under-trained context in Group Naloxone ($P>0.05$). Indeed, bayesian analysis (Gallistel, 2009; Rouder et al., 2009) supported this claim; the null hypothesis that, in the under-trained context, Group Naloxone showed no difference in response rate when the outcome was devalued was found to be 3.15 times more probable than the alternative hypothesis.

These results indicate that naloxone-induced opioid receptor blockade during learning results in insensitivity to outcome devaluation at test, in a manner comparable to the effects of over-training. To confirm that the effects of naloxone were not significantly different from the effects of the over-training manipulation two-way ANOVAs were conducted separately for Group Vehicle and Group Naloxone for the data in represented in Fig. 1a and b, respectively. A two-way ANOVA conducted on the data from Group Vehicle using within subjects factors of training context, either over or under, and devaluation found neither a main effect of devaluation ($F_{1,17}$=3.61, P=0.07), nor training ($F_{1,17}$=2.27, P=0.15), but did reveal a significant devaluation×training context interaction ($F_{1,17}$=9.77, P= 0.006). Bonferroni post hoc analysis of these data again revealed that the effect of devaluation was restricted to the group of rats responding in the context that was under-trained ($P<0.01$). Similar two-way ANOVA conducted on the data from Group Naloxone (Fig. 1b) found neither a significant main effect of training context ($F_{1,17}$=0.62, P=0.44), nor devaluation ($F_{1,17}$=2.13, P=0.16) with no training×devaluation interaction ($F_{1,17}$=1.43, P=0.25). Again, these data indicate that rats that were under-trained after naloxone-induced blockade of opioid receptors were insensitive to outcome devaluation in a manner that was similar to that induced by over-training. For an analysis of the devaluation data over the time course of each extinction test see supplementary data.

Although it is possible that naloxone affected outcome taste discriminability or the ability of the outcome to actu-

ally lose value, both Group Vehicle and Group Naloxone decreased their preference for the devalued outcome relative to the non-devalued outcome during a consumption test on both outcomes conducted after the extinction tests in feeding cages separate from the operant box. A two-way ANOVA with a between-subjects factor of drug and a within-subjects factor of devaluation, revealed a significant main effect of devaluation ($F_{1,37}=66.56$, $P<0.001$) on post-feeding consumption, with no main effect of drug ($F_{1,37}=2.13$, $P=0.15$), and no drug×devaluation interaction ($F_{1,37}=0.02$, $P=0.62$). These data indicate that, in both groups, the specific satiety pre-feeding did in fact result in a reliable decrease in the relative value of the outcome on which rats were sated prior to the extinction test.

Taken together, the data from Experiment 1 indicate that, as has been shown previously (Coutureau and Killcross, 2003; Yin et al., 2004), using a within-subjects training procedure over-training in one context resulted in habitual control of actions, as revealed by insensitivity to specific satiety-induced outcome devaluation, in specifically the over-trained context. As might also be expected, in the same group of rats, under-training that was restricted to 50 action–outcome pairings produced goal-directed performance that remained sensitive to changes in outcome value when tested in the under-trained context. Importantly, rats under-trained under naloxone-induced blockade of opioid receptors were insensitive to outcome devaluation consistent with the claim that, when in the naloxone-trained context, their actions were insensitive to changes in value and so under habitual control.

Experiment 2: chronic naloxone during learning, but not acute naloxone on test, results in insensitivity to outcome devaluation that is context specific

One explanation of the effect observed in Group Naloxone in Experiment 1 is that the naloxone injection may have affected the ability of the rats to discriminate between the two training contexts, i.e. rats that received naloxone during under-training may not have been able to discriminate between this context and the over-trained context, possibly accounting for their habitual performance at test. Experiment 2 was designed to assess this interpretation of the data and, further, to test the context specificity of the effect of naloxone on the goal-directed control of instrumental performance using a within-subjects drug treatment design. Additionally, at test rats were given either an injection of naloxone or vehicle prior to the specific satiety procedure and extinction tests to test the state-dependency of naloxone's effect as well as to assess the hypothesis that the effect of naloxone was limited to chronic opioid receptor blockade during learning and would not result from acute drug treatment. The design of Experiment 2 is laid out in Table 2.

Training. Training progressed smoothly for all rats in both contexts. On the last day of training the average response rate was similar in both contexts with rats responding on average 17.3 presses/min (SEM=1.9) in the vehicle context and 16.5 presses/min (SEM=1.9) in the naloxone context (see supplementary data for statistical analysis of training data).

Naloxone during training produces insensitivity to outcome devaluation in the naloxone-paired, but not vehicle-paired context, irrespective of drug treatment on test. Fig. 3a and 3b illustrates the results from the devaluation tests in Experiment 2. Irrespective of whether they were tested under acute systemic injection of naloxone or vehicle, the rats showed a clear outcome devaluation effect when they were tested in the vehicle trained context (Fig. 2a). However, when tested in the naloxone-trained context under both acute systemic injection of naloxone and vehicle there was no apparent outcome devaluation effect; rats responded at a similar rate when the outcome was devalued relative to when it was not (Fig. 2b).

A three-way repeated measures ANOVA using within-subjects factors of training drug (vehicle vs. naloxone), drug given on test (vehicle vs. naloxone), and devaluation revealed a main effect of devaluation ($F_{1,19}=19.65$, $P<0.0001$), but neither a main effect of training drug ($F_{1,19}=3.01$, $P=0.1$), nor test drug ($F_{1,19}=0.36$, $P=0.58$). Importantly, this ANOVA did reveal a significant devaluation× training drug interaction ($F_{1,19}=5.08$, $P=0.04$), but not a devaluation×test drug interaction ($F_{1,19}=0.30$, $P=0.60$), a training drug×test drug interaction ($F_{1,19}=0.57$, $P=0.47$), or a training drug×test drug×devaluation three-way interaction ($F_{1,19}=0.20$, $P=0.66$). Although the graph in Fig. 2b appears to show a tendency for lower response rates in the naloxone-trained context, the lack of a main effect of training drug indicates that this small tendency is not significant. Indeed, analysis of rats' responding in the naloxone-trained context in the first 4 min of the extinction test, when their response rates were higher and comparable to their response rates in the vehicle context shows only a main effect of time (vehicle on test: $F_{3,17}=7.96$, $P=0.001$, naloxone on test: $F_{3,17}=6.05$, $P=0.005$), with no effect of devaluation (vehicle on test: $F_{1,19}=0.194$, $P=0.67$, naloxone on test: $F_{1,19}=3.56$, $P=0.08$), or devaluation×time interaction (vehicle on test: $F_{3,17}=0.58$, $P=0.64$, naloxone on test: $F_{3,17}=0.87$, $P=0.47$) (Fig. 2b, insets). Additionally, multiple extinction testing, as conducted in this experiment, has the possibility alone of affecting response sensitivity to changes in goal value by disrupting the perception of the response–reward contingency. For this reason the testing order was counterbalanced with respect to drug treatment on test. Moreover, this multiple extinction testing did not observably affect the results; the same pattern of effects was seen in the first set of devaluation extinction tests as in the second in all conditions.

To further clarify the effect of naloxone during training individual two-way ANOVAs were conducted separately for the extinction data collected in the vehicle- and naloxone-trained contexts shown in Fig. 2a and b, respectively. Two-way ANOVA of the data from rats responding in the vehicle trained context (Fig. 2a), with within-subjects variables drug at test (vehicle vs. naloxone) and devaluation revealed a main effect of devaluation ($F_{1,15}=18.46$,

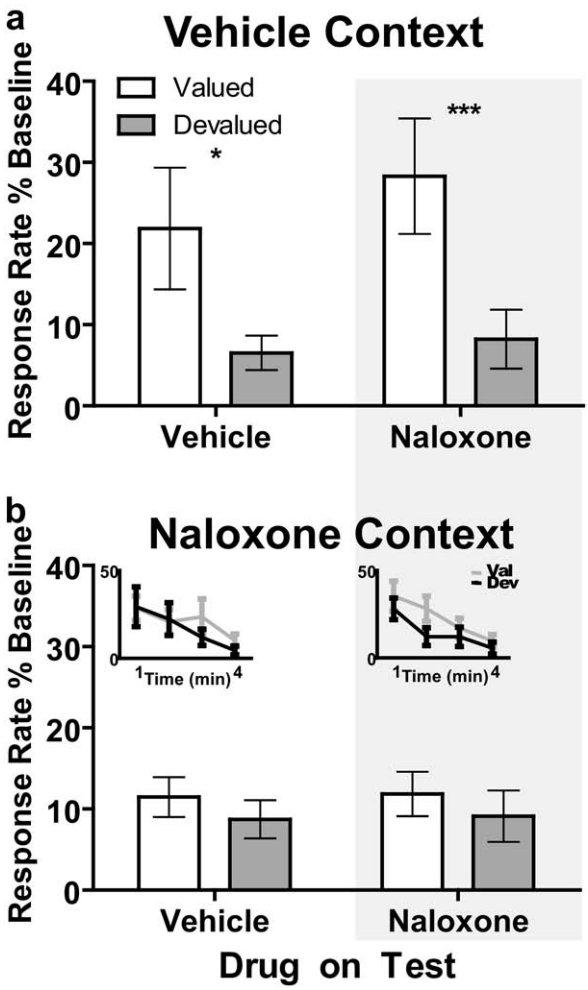

Fig. 2. Rats showed insensitivity to outcome devaluation when tested under naloxone or vehicle conditions in naloxone- but not vehicle-trained contexts. In Experiment 2 rats were under-trained in two contexts (120 action–outcome pairings). In one context training occurred after an injection of vehicle and in the alternate context training occurred 15 min after an injection of naloxone. At test rats were pre-fed on one of the outcomes for 1 h and then allowed to respond in extinction in both contexts. This test either occurred after an injection of vehicle or naloxone prior to the specific satiety treatment. (a) Extinction response rates in the vehicle-trained context, plotted as a percentage of baseline, following an injection of vehicle or naloxone (shaded panel) prior to the specific-satiety induced outcome devaluation. Devalued refers to the outcome which was pre-fed to the rats for 1 h prior to the test. (b) Extinction response rates in the naloxone-trained context, plotted as a percentage of baseline, following an injection of vehicle or naloxone (shaded panel) prior to the specific-satiety-induced outcome devaluation. The insets show response rate as a percentage of baseline on the y-axis over the first 4 min of the extinction test conducted in the naloxone-trained context and tested after an injection of vehicle or naloxone (shaded panel) (n=24). * P<0.05, *** P<0.001.

P=0.001), with no significant effect of test drug ($F_{1,15}$=3.21, P=0.09), or devaluation×test drug interaction ($F_{1,15}$=2.17, P=0.16). Bonferroni post hoc analysis of these data indicates that, when responding in the vehicle-trained context, rats significantly reduced response rates when the outcome was devalued both when tested acutely under vehicle (P<0.05) and naloxone (P<0.01), indicating that acute naloxone alone does not alter sensitivity to outcome devaluation. This same analysis of the data from rats responding in extinction in the naloxone-trained context (Fig. 2b) revealed neither a main effect of devaluation ($F_{1,38}$=0.037, P=0.55), nor test drug ($F_{1,38}$<0.0001, P=0.99), with no devaluation×test drug interaction ($F_{1,38}$=0.83, P=0.78). Bonferroni post hoc analysis of these data confirmed the lack of effect of devaluation under both test drug conditions (P>0.05), further indicating that the effect of naloxone on responding after outcome devaluation is not state-dependent. Indeed, given the naloxone-trained context response data, bayesian analysis supports this claim showing the null hypothesis, that there was no difference in response rates when the outcome was devalued relative to when it was not, to be 3.43 times more probable than the alternative hypothesis, for both the test conducted on vehicle and that on naloxone (see also the supplementary data for an analysis of the devaluation effect over the course of each extinction test).

Consumption tests conducted outside of the operant box on the devalued and non-devalued outcomes after the devaluation extinction test found, in all situations, that the rats consumed less of the devalued outcome relative to the non-devalued outcome. A three-way repeated measures ANOVA with within-subjects factors of training drug, test drug, and devaluation revealed significant main effects of devaluation ($F_{1,23}$=120.96, P<0.0001) and training drug ($F_{1,23}$=4.61, P=0.04), but no main effect of test drug ($F_{1,23}$=0.003, P=0.96), and no devaluation×training drug ($F_{1,23}$=0.38, P=0.54), devaluation×test drug ($F_{1,23}$=3.73, P=0.07), training drug×test drug ($F_{1,23}$=0.56, P=0.46), or devaluation×test drug×training drug interactions ($F_{1,23}$=0.92, P=0.67). Bonferroni post hoc analyses indi-

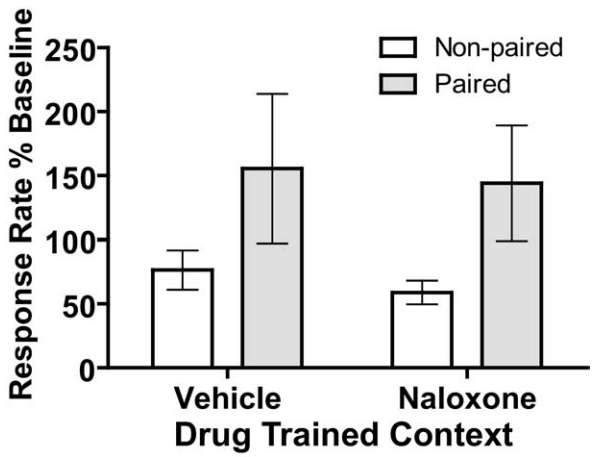

Fig. 3. Rats show outcome-specific discrimination when responding in both the vehicle and naloxone trained contexts. A subset of rats (n=8) was given an outcome-specific reinstatement test in which, after a period of extinction, both of the previously trained pellets were delivered and rats were allowed to respond, in extinction on both levers. This test was conducted in both the vehicle- and naloxone-trained contexts. Response rates as percentage of baseline are plotted on the y-axis for both the vehicle- and naloxone-trained context on the x-axis. Paired refers to responding on the lever that was previously paired with the outcome delivered to induce reinstatement.

cate that, in each condition, there was significantly less consumption of the devalued outcome relative to the non-devalued outcome (vehicle context: vehicle on testing $P<0.05$, naloxone on testing $P<0.001$; naloxone context: vehicle and naloxone on testing $P<0.001$). These data indicate that, while rats did generally consume slightly less of the naloxone-trained outcome, in all conditions, the specific satiety devaluation was effective in significantly changing the value of the outcome, irrespective of training or test condition.

Naloxone during training does not affect outcome discriminability. It is possible that rather than accelerating habit learning, naloxone during training rendered the rats incapable of discriminating between the two trained outcomes when in the naloxone-paired context, thereby resulting in similar response rates in the extinction tests for both devalued and non-devalued outcomes. To test this hypothesis a subset ($n=8$) of rats from Experiment 2 was tested for their ability to discriminate between outcomes in an outcome-specific reinstatement procedure. For this test, the rats' instrumental responding was extinguished prior to counter-balanced presentation of both the vehicle- and naloxone-trained outcomes. Reinstatement was assessed on both levers in both contexts. If rats were able to distinguish between the outcomes in the naloxone-trained context then they should show reinstatement selectively on the lever paired with the specific delivered outcome.

The results of this experiment are presented in Fig. 3. A two-way repeated measures ANOVA using within-subjects factors of drug-trained context (naloxone vs. vehicle) and reinstatement (separating performance on the lever previously paired with the reinstating outcome from that on the other lever), revealed a significant main effect of reinstatement ($F_{1,15}=5.99$, $P=0.03$) but, importantly neither a main effect of drug-trained context ($F_{1,15}=0.12$, $P=0.74$), nor a context×reinstatement interaction ($F_{1,15}=0.008$, $P=0.93$). These data indicate that in both the vehicle-trained and naloxone-trained context, rats show selective reinstatement on the lever on which they were trained to press for the pellet that was used to induce the reinstatement. More importantly, these data indicate that when in the naloxone-paired context rats were indeed capable of discriminating between the sucrose and grain outcomes, ruling out a deficit here as an explanation of the current results.

Taken together, therefore, the results of Experiment 2 add to those of Experiment 1 to suggest that blockade of opioid receptors during learning results in instrumental actions that are insensitive to devaluation of the specific outcome earned during training under naloxone. Although rats were unable to update performance based on the change in outcome value in the context in which they learned the task under opioid receptor blockade, the same rats were able to maintain goal-directed control of their actions in the context in which they were trained after vehicle injections. These data demonstrate, therefore, that blockade of opioid receptors during training does not result in an inability to discriminate between the two training contexts. Furthermore, acute opioid receptor blockade on test did not alter the sensitivity of performance to outcome devaluation in the vehicle-trained context, nor did it reverse the insensitivity to outcome devaluation found in the naloxone-trained context, confirming that the effect of naloxone was not state-dependent. Hence, only chronic naloxone treatment prior to training in one context influenced action sensitivity to outcome devaluation. These results are, therefore, consistent with the conclusions from Experiment 1, that opioid-blockade during instrumental training increases stimulus control of performance, reducing goal-directed control and enhancing the habitual control of actions.

DISCUSSION

The results of these experiments demonstrate that when the activity of endogenous opioid peptides at opioid receptors is blocked during instrumental learning the performance of actions rapidly becomes insensitive to subsequent changes in the value of the outcome indicative of habitual, stimulus–response control of behavior (Dickinson, 1985; Balleine, 2005). Indeed, in Experiment 1 we found that the effect of naloxone on instrumental performance was similar to that induced by over-training, which has been previously shown to produce habits in rodents (Coutureau and Killcross, 2003). Interestingly, the data from Experiment 2 indicate that chronic, rather than acute, opioid receptor blockade during the course of learning is necessary to produce inflexible, habitual behavior, suggesting that opioid receptor activation contributes to normal goal-directed reward learning, but not to the expression of this learning in performance. Furthermore, this experiment demonstrated that the effect of naloxone was context specific, increasing the rate at which antecedent stimuli, such as the specific naloxone-trained context, exerted control over lever press performance at the expense, or possibly as a result, of a lack of control by outcome value. This experiment also revealed that the effect of naloxone on goal-directed reward learning was not due to state-dependent control of reward value.

Although it has long been thought that the endogenous opioid system is important for hedonic tone (Narayanan et al., 2004; Skoubis et al., 2005), here we show that opioid processes are critical for the acquisition of normal goal-directed control of actions. Indeed, blockade of opioid receptors acutely has previously been shown to negatively impact sucrose (Cleary et al., 1996) and self-stimulation (Trujillo et al., 1989) reward in progressive ratio tests. Similarly, mu opioid receptor (Papaleo et al., 2007), beta-endorphin- and enkephalin-knockout mice also show reduced responding for a food reward on a progressive ratio schedule (Hayward et al., 2002). Although they suggest that the mu opioid receptor and its two main opioid peptide ligands positively contribute to the incentive motivation used to drive reward-related actions, these studies did not specifically address what aspects of these actions were controlled by opioid receptor-related processes. Here, using outcome devaluation procedures, we were able to

provide a direct assessment of the specific aspects of action control influenced by the endogenous opioid system and found that compromising this system by opioid-receptor blockade resulted in accelerated habitual control of actions.

Compromised reward function appears to potentiate the acquisition of habits. For example lesion studies have found that damage to the reward system can increase the habitual control of behavior (Balleine et al., 2003; Corbit et al., 2003; Killcross and Coutureau, 2003; Corbit and Balleine, 2005). Therefore, our data can be interpreted as suggesting that chronic naloxone treatment during training disrupts goal-directed learning, allowing the habitual system to rapidly take over action control. Moreover, these experiments established that disruption of the endogenous opioid system does not alter goal-directed actions generally, but only affected actions trained under naloxone. Additionally, acutely administered naloxone, which may induce an aversive state in rodents (Skoubis et al., 2005), and that should have acted to recall the naloxone-induced training state, did not alter the expression of goal-directed behavior when rats were responding in the context in which they were trained under vehicle. The ability of naloxone during training to reduce the sensitivity of actions to outcome devaluation was shown not to be state-dependent; responding in the naloxone-trained context was habitual whether tested off drug, or in the naloxone-induced state. This effect was also not due to an effect of naloxone on the rats' ability to discriminate the two outcomes used in these studies; outcome-specific reinstatement in the naloxone-paired context could not have emerged if outcome-discriminability were impaired.

These data suggest, therefore, that the blockade of opioid receptors during learning enhanced the ability of contextual stimuli to control the performance of actions, an effect that is characteristic of stimulus–response habit learning (Dickinson, 1985). Given the recently described role of the endogenous opioid system in encoding reward value (Wassum et al., 2009) and in reward learning (Shippenberg et al., 1987; Hayward et al., 2002; Papaleo et al., 2007) these data appear to indicate that the effect of naloxone during training was to block the reward processing necessary for appropriate goal-directed learning, thereby forcing control by the habit learning system. However, it remains a possibility that naloxone also acted directly to enhance habit learning, perhaps by releasing an inhibition of habit formation, or by increasing stimulus control of actions by enhancing context processing. Given the wide distribution of opioid receptors (Mansour et al., 1994) and their involvement in reward hedonia (Skoubis et al., 2005; Pecina et al., 2006), we cannot at present rule out these alternatives and future studies are planned to investigate these possibilities more directly.

Nevertheless, the effect of naloxone during instrumental learning on subsequent sensitivity to outcome devaluation is similar to that induced by specific disruption of the cortico-striatal network, involved in action–outcome learning, or the basolateral amygdala, involved in reward-value learning (Corbit et al., 2001; Balleine et al., 2003; Ostlund and Balleine, 2005; Yin et al., 2005). Importantly, these goal-directed learning circuits are densely populated with mu, delta and kappa opioid receptors (Mansour et al., 1994; Ding et al., 1996; Daunais et al., 2001). Activation of these $G_{i/o}$-protein-coupled receptors induces hyperpolarization and inhibition of neurotransmitter release via activation of potassium conductances and inhibition of calcium channel activity, respectively, and could, therefore, be involved in regulating the neuronal output of these regions to control goal-directed actions. Indeed, both mu and kappa opioid receptor activation has been shown to attenuate inhibitory post-synaptic conductances in striatal medium spiny neurons (Hjelmstad and Fields, 2003; Miura et al., 2007).

The activity of dopamine, specifically, within the ventral striatum may be involved in aspects of goal-directed action performance (Everitt and Robbins, 2005; Wickens et al., 2007; Palmiter, 2008). Interestingly, there is evidence to suggest that mu and delta opioid receptor agonists increase dopamine release in the nucleus accumbens (Spanagel et al., 1990), that dopamine release is decreased in this region in mu and delta opioid receptor deficient mice (Chefer et al., 2003), and that tonic activation of mu opioid receptors is necessary for the maintenance of accumbens dopamine release, likely via disinhibition of dopaminergic neurons in the ventral tegmental area (Spanagel et al., 1992). Mu opioid receptors are also involved in the regulation of dopamine release and metabolism in the dorsal striatum (Piepponen et al., 1999), a region which is involved in both goal-directed and habit learning (Yin et al., 2004). Moreover, nigrostriatal dopamine has been shown to play a role in habit formation (Faure et al., 2005). Therefore, naloxone could directly affect these learning processes. Indeed, a previous study reported that altered dopaminergic activity produced by sensitization to amphetamine induced rapid habit acquisition (Nelson and Killcross, 2006). More directly, we have recently found that naloxone-induced blockade of opioid receptors in the basolateral amygdala blocked the ability of rats to encode changes in the value instrumental outcome (Wassum et al., 2009), consistent with the claim that endogenous opioid blockade negatively affects the reward learning circuitry necessary for goal-directed learning.

Whatever the proximal cause of the acceleration in habit learning observed in this study, these data add to a growing body of evidence suggesting that the endogenous opioid system is important for motivational aspects of behavior. We show here that activation of endogenous opioid receptors is necessary for normal goal-directed learning. Notably, when this system was chronically compromised, instrumental actions rapidly became habitual, and inflexible to negative changes in the value of the outcome. Chronic administration of drugs of abuse, such as morphine, cocaine and ethanol, can, like naloxone, compromise the integrity of the endogenous opioid system in regions previously shown to be important for goal-directed learning (Turchan et al., 1999). The current results suggest, therefore, one means by which the chronic administration of abused substances results in enhanced habitual

control of actions through which addicted individuals seek access to those substances.

Acknowledgments—This research was supported by grant AA18014 from NIAAA to Bernard W. Balleine, grants DA09359 and DA05010 from NIDA to Nigel T. Maidment, and a Ruth L. Kirschstein National Research Service Award and Hatos scholarship to Kate M. Wassum.

REFERENCES

Balleine B (2001) Incentive processes in instrumental conditioning. In: Handbook of contemporary learning theories (Mowrer R, Klein S, eds), pp 307–366. Hillsdale, NJ: Erlbaum.

Balleine BW (2005) Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol Behav 86:717–730.

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419.

Balleine BW, Killcross AS, Dickinson A (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. J Neurosci 23:666–675.

Chefer VI, Kieffer BL, Shippenberg TS (2003) Basal and morphine-evoked dopaminergic neurotransmission in the nucleus accumbens of MOR- and DOR-knockout mice. Eur J Neurosci 18:1915–1922.

Cleary J, Weldon DT, O'Hare E, Billington C, Levine AS (1996) Naloxone effects on sucrose-motivated behavior. Psychopharmacology (Berl) 126:110–114.

Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. Behav Brain Res 146:145–157.

Corbit LH, Balleine BW (2005) Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. J Neurosci 25:962–970.

Corbit LH, Muir JL, Balleine BW (2001) The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. J Neurosci 21:3251–3260.

Corbit LH, Muir JL, Balleine BW (2003) Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. Eur J Neurosci 18:1286–1294.

Coutureau E, Killcross S (2003) Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behav Brain Res 146:167–174.

Daunais JB, Letchworth SR, Sim-Selley LJ, Smith HR, Childers SR, Porrino LJ (2001) Functional and anatomical localization of mu opioid receptors in the striatum, amygdala, and extended amygdala of the nonhuman primate. J Comp Neurol 433:471–485.

Dickinson A (1985) Actions and habits: the development of behavioural autonomy. Philos Trans R Soc Lond B 308:67–78.

Dickinson A, Balleine BW (2002) The role of learning in motivation. In: Steven's handbook of experimental psychology: learning, motivation, and emotion, Vol. 3 (Gallistel R, ed). New York: John Wiley & Sons.

Ding YQ, Kaneko T, Nomura S, Mizuno N (1996) Immunohistochemical localization of mu-opioid receptors in the central nervous system of the rat. J Comp Neurol 367:375–402.

Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci 8:1481–1489.

Faure A, Haberland U, Conde F, El Massioui N (2005) Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. J Neurosci 25:2771–2780.

Gallistel CR (2009) The importance of proving the null. Psychol Rev 116:439–453.

Hayward MD, Pintar JE, Low MJ (2002) Selective reward deficit in mice lacking beta-endorphin and enkephalin. J Neurosci 22:8251–8258.

Hjelmstad GO, Fields HL (2003) Kappa opioid receptor activation in the nucleus accumbens inhibits glutamate and GABA release through different mechanisms. J Neurophysiol 89:2389–2395.

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex 13:400–408.

Mansour A, Fox CA, Burke S, Meng F, Thompson RC, Akil H, Watson SJ (1994) Mu, delta, and kappa opioid receptor mRNA expression in the rat CNS: an in situ hybridization study. J Comp Neurol 350:412–438.

Miura M, Saino-Saito S, Masuda M, Kobayashi K, Aosaki T (2007) Compartment-specific modulation of GABAergic synaptic transmission by mu-opioid receptor in the mouse striatum with green fluorescent protein-expressing dopamine islands. J Neurosci 27: 9721–9728.

Narayanan S, Lam H, Christian L, Levine MS, Grandy D, Rubinstein M, Maidment NT (2004) Endogenous opioids mediate basal hedonic tone independent of dopamine D-1 or D-2 receptor activation. Neuroscience 124:241–246.

Nelson A, Killcross S (2006) Amphetamine exposure enhances habit formation. J Neurosci 26:3805–3812.

Ostlund SB, Balleine BW (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J Neurosci 25:7763–7770.

Ostlund SB, Balleine BW (2008) Differential involvement of the basolateral amygdala and mediodorsal thalamus in instrumental action selection. J Neurosci 28:4398–4405.

Palmiter RD (2008) Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. Ann N Y Acad Sci 1129:35–46.

Papaleo F, Kieffer BL, Tabarin A, Contarino A (2007) Decreased motivation to eat in mu-opioid receptor-deficient mice. Eur J Neurosci 25:3398–3405.

Pecina S, Smith KS, Berridge KC (2006) Hedonic hot spots in the brain. Neuroscientist 12:500–511.

Piepponen TP, Honkanen A, Kivastik T, Zharkovsky A, Turtia A, Mikkola JA, Ahtee L (1999) Involvement of opioid mu1-receptors in opioid-induced acceleration of striatal and limbic dopaminergic transmission. Pharmacol Biochem Behav 63:245–252.

Poulin JF, Chevalier B, Laforest S, Drolet G (2006) Enkephalinergic afferents of the centromedial amygdala in the rat. J Comp Neurol 496:859–876.

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. Psychon Bull Rev 16:225–237.

Shippenberg TS, Bals-Kubik R, Herz A (1987) Motivational properties of opioids: evidence that an activation of delta-receptors mediates reinforcement processes. Brain Res 436:234–239.

Skoubis PD, Lam HA, Shoblock J, Narayanan S, Maidment NT (2005) Endogenous enkephalins, not endorphins, modulate basal hedonic state in mice. Eur J Neurosci 21:1379–1384.

Skoubis PD, Matthes HW, Walwyn WM, Kieffer BL, Maidment NT (2001) Naloxone fails to produce conditioned place aversion in mu-opioid receptor knock-out mice. Neuroscience 106:757–763.

Spanagel R, Herz A, Shippenberg TS (1990) The effects of opioid peptides on dopamine release in the nucleus accumbens: an in vivo microdialysis study. J Neurochem 55:1734–1740.

Spanagel R, Herz A, Shippenberg TS (1992) Opposing tonically active endogenous opioid systems modulate the mesolimbic dopaminergic pathway. Proc Natl Acad Sci U S A 89:2046–2050.

Trujillo KA, Belluzzi JD, Stein L (1989) Opiate antagonists and self-stimulation: extinction-like response patterns suggest selective reward deficit. Brain Res 492:15–28.

Turchan J, Przewlocka B, Toth G, Lason W, Borsodi A, Przewlocki R (1999) The effect of repeated administration of morphine, cocaine and ethanol on mu and delta opioid receptor density in the nucleus accumbens and striatum of the rat. Neuroscience 91:971–977.

Wang SH, Ostlund SB, Nader K, Balleine BW (2005) Consolidation and reconsolidation of incentive learning in the amygdala. J Neurosci 25:830–835.

Wassum KM, Ostlund SB, Maidment NT, Balleine BW (2009) Distinct opioid circuits determine the palatability and the desirability of rewarding events. Proc Natl Acad Sci U S A 106:12512–12517.

Wellman LL, Gale K, Malkova L (2005) GABAA-mediated inhibition of basolateral amygdala blocks reward devaluation in macaques. J Neurosci 25:4577–4586.

Wickens JR, Horvitz JC, Costa RM, Killcross S (2007) Dopaminergic mechanisms in actions and habits. J Neurosci 27:8181–8183.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513–523.

APPENDIX

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroscience.2009.06.071.

(Accepted 30 June 2009)
(Available online 18 July 2009)