Neuroscience 154 (2008) 1195–1204

GUIDANCE OF INSTRUMENTAL BEHAVIOR UNDER REVERSAL CONDITIONS REQUIRES DOPAMINE D1 AND D2 RECEPTOR ACTIVATION IN THE ORBITOFRONTAL CORTEX

C. CALAMINUS AND W. HAUBER*

Abteilung Tierphysiologie, Biologisches Institut, Universität Stuttgart, Pfaffenwaldring 57, D-70550, Stuttgart, Germany

Abstract—The orbitofrontal cortex (OFC) plays a critical role in learning a reversal of stimulus–reward contingencies. Dopamine (DA) neurons probably support reversal learning by emitting prediction error signals that indicate the discrepancy between the actually received reward and its prediction. However, the role of DA receptor-mediated signaling in the OFC to adapt behavior to changing stimulus–reward contingencies is largely unknown. Here we examined the effects of a selective D1 or D2 receptor blockade in the OFC on learning a reversal of previously acquired stimulus–reward magnitude contingencies. Rats were trained on a reaction time (RT) task demanding conditioned lever release with discriminative visual stimuli signaling in advance the upcoming reward magnitude (one or five food pellets). After acquisition, RTs were guided by stimulus-associated reward magnitudes, i.e. RTs of responses were significantly shorter for expected high versus low reward. Thereafter, stimulus–reward magnitude contingencies were reversed and learning was tested under reversal conditions for three blocks after pre-trial infusions of the selective D1 or D2 receptor antagonists R(+)-7-chloro-8-hydroxy-3-methyl-1-phenyl-2,3,4,5-tetrahydro-1H-3-benzazepinhydrochloride (SCH23390), eticlopride, or vehicle. For comparisons, we included intra-OFC infusions of the selective *N*-methyl-D-aspartate receptor antagonist AP5. Results revealed that in animals subjected to intra-OFC infusions of SCH23390 or eticlopride learning a reversal of previously acquired stimulus reward-magnitude contingencies was impaired. Thus, in a visual discrimination task as used here, D1 and D2 receptor-mediated signaling in the OFC seems to be necessary to update the reward-predictive significance of stimuli.

Key words: prefrontal, glutamate, SCH23390, eticlopride, AP5.

Behavioral studies in humans and animals demonstrated that different subregions of the prefrontal cortex (PFC) mediate distinct types of behavioral flexibility. For instance, lesion studies in rats suggest that the medial prefrontal cortex (mPFC) appears necessary for the acquisition of novel strategies or rules and the inhibition of a previously learned strategy (e.g. Birrell and Brown, 2000; Ghods-Sharifi et al., 2008; Ragozzino et al., 1999). By contrast, subjects with lesions of the orbital region of the PFC (OFC) had difficulties in adapting behavioral responding when previously established contingencies between stimuli and outcomes were reversed (Ferry et al., 2000; Izquierdo et al., 2004; Mishkin, 1964; Schoenbaum et al., 2002, 2003). Therefore, the OFC appears to be fundamental in adapting behavior to changing stimulus–reward contingencies (see Brown and Bowman, 2002; Murray et al., 2007 for review). Correspondingly, electrophysiological recordings revealed that OFC neurons were activated in anticipation of outcomes after sampling of predictive stimuli and changed their stimulus-selective firing during a reversal of stimulus-outcome contingencies (Thorpe et al., 1983; Schoenbaum et al., 1999).

While a fundamental role of the OFC for using stimulus-outcome learning to guide action selection is well established, little is known about neurochemical substrates in the OFC mediating reversal learning. Behavioral studies in rodents implicated *N*-methyl-D-aspartate (NMDA) receptor activity in the mPFC in associative learning and cognitive flexibility (Baldwin et al., 2002; Stefani and Moghaddam, 2003), however, the involvement of NMDA receptors in reversal learning is still poorly defined. A recent study suggests that NMDA receptor-mediated signaling in the OFC seems to be critical in guiding instrumental behavior under reversal conditions (Bohn et al., 2003a). Also, the role of OFC DA receptor activity in reversal learning is largely unexplored. This fact is surprising for several reasons. First, the OFC receives a prominent dopamine (DA) input from mesocorticolimbic fibers (Berger et al., 1991). Second, models of temporal difference learning (Schultz, 2006; Sutton and Barto, 1990) suggest that learning of predictions under reversal conditions requires a prediction error signal that indicates the discrepancy between actually received reward and its prediction. Accordingly, empirical studies show that DA neurons emit prediction error signals (Schultz, 2006; Roesch et al., 2007) and indicate the presence of such signals in target areas of DA neurons such as the OFC (McClure et al., 2003; O'Doherty, 2003). These findings imply that DA signals in the OFC could be necessary to update the predictive significance of stimuli in response to changing stimulus–reward contingencies. However, contrasting with this account, DA depletion of the OFC did not affect serial discrimination reversal learning (Clarke et al., 2007).

The aim of the current study was to further analyze the role of OFC DA in behavioral flexibility and to explore

*Corresponding author. Tel: +49-711-685-65003; fax: +49-711-685-65090.
E-mail address: hauber@bio.uni-stuttgart.de (W. Hauber).
Abbreviations: ANOVA, analysis of variance; AP5, DL-2-amino-5-phosphonovaleric acid; BLA, basolateral amygdala; DA, dopamine; mPFC, medial prefrontal cortex; MT, movement time; NMDA, *N*-methyl-D-aspartate; OFC, orbitofrontal cortex; PFC, prefrontal cortex; RT, reaction time; SCH23390, R(+)-7-chloro-8-hydroxy-3-methyl-1-phenyl-2,3,4,5-tetrahydro-1H-3-benzazepinhydrochloride.

0306-4522/08$32.00+0.00 © 2008 IBRO. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuroscience.2008.04.046

whether D1 and D2 receptor-mediated signals in the OFC are essential for adapting instrumental behavior to changing stimulus–reward magnitude contingencies. To this end, rats were trained in a reaction time (RT) task demanding conditioned lever release with a predictive stimulus signaling in advance the upcoming reward magnitude (five or one pellet) (Calaminus and Hauber, 2006; Giertler et al., 2005). After acquisition of the task, RTs of responses to expected high versus low reward were significantly shorter indicating that instrumental responding was guided by reward-predictive stimuli. Thereafter, we reversed stimulus–reward contingencies and rats received intra-OFC infusions of selective D1 or D2 receptor antagonists or of vehicle. For comparisons, we included intra-OFC infusions of the selective NMDA receptor antagonist AP5, a drug that has been shown to impair guidance of instrumental behavior under reversal conditions in this task (Bohn et al., 2003a). If D1 or D2 receptors in the OFC transmit a prediction error signal that shapes learning a reversal of stimulus–reward contingencies, we expected an impairment after a DA receptor blockade.

EXPERIMENTAL PROCEDURES

Animals

Eighty-four Lister-Hooded rats (Harlan-Winkelmann, Borchen, Germany) were housed in transparent plastic cages (55×39× 27 cm, Ferplast, Nürnberg, Germany). Temperature (20±2 °C) and humidity (50–60%) in the animal house were kept constant and a 12-h light/dark schedule was used with lights on between 7:00 h and 19:00 h. Rats were given *ad libitum* access to water; food was restricted to 15 g per animal and day. On days without behavioral testing, rats received 15 g standard laboratory maintenance chow (Altromin, Lage, Germany). On days with behavioral tests, rats received in the testing apparatus 5.4 g food pellets as reward (45 mg pellets, Bioserv, Frenchtown, NJ, USA). On these days, the amount of standard laboratory chow given was reduced to 9.6 g per animal. Rats weighed 200–210 g on arrival and 250–270 g at the time of surgery. All animal experiments were conducted according to the European Communities Council Directive of 24 November 1986 (86/609/EEC) as well as the German Law on Animal Protection and were approved by the proper authorities in Stuttgart, Germany. All efforts were made to minimize the number of animals used and their suffering.

Surgery

For stereotaxic surgery, animals were anesthetized with ketamine (120 mg/kg i.m.) (Bela-Pharm GmbH, Vechta, Germany) and xylazine (4 mg/kg i.m.) (Bayer AG, Leverkusen, Germany) and secured in a Kopf stereotaxic apparatus (Kopf Instruments, Tujunga, USA). Bilateral 15 mm stainless steel guide cannulae with an outer diameter of 0.7 mm were aimed at the OFC and implanted using standard stereotaxic procedures. The coordinates were 3.2 mm anterior bregma, 2.4 mm lateral to midline, and 4.0 mm ventral from the skull with the tooth bar −3.3 mm under the interaural line. Coordinates were determined from the atlas of Paxinos and Watson (1997). The guide cannulae were occluded by stainless steel stylets. Each rat was given at least 7 day to recover from surgery before behavioral testing was started.

Drug injection

Animals received bilateral intra-OFC injections of the selective D1 receptor antagonist R(+)-7-chloro-8-hydroxy-3-methyl-1-phenyl-2,3,4,5-tetrahydro-1H-3-benzazepinhydrochloride (SCH23390, Research Biochemicals, Natick, USA) (1 μg in 0.5 μl 0.9% sterile saline), the D2 receptor antagonist eticlopride (Sigma-Aldrich Chemie GmbH, Steinheim, Germany) (1 μg in 0.5 μl 0.9% sterile saline), the NMDA receptor antagonist DL-2-amino-5-phosphonovaleric acid (AP5) (5 μg in 0.5 μl 0.9% sterile saline) or vehicle (0.5 μl 0.9% sterile saline). The doses of AP5, SCH23390 and eticlopride were based on previous studies (e.g. Bohn et al., 2003; Capriles et al., 2003; Schweimer and Hauber, 2006). On injection days, stainless steel stylets were removed and injection cannulae (outer diameter: 0.45 mm, length: 17 mm) were lowered to the final site of infusion and attached via polyvinylchloride tubing to microliter syringes controlled by a syringe pump (Med Associates, St. Albans, VT, USA). Drugs were delivered over a 1-min interval and injection cannulae were left in position for a further minute to allow for diffusion. After injection, each rat remained in its home cage for an additional 10 min before being placed in the test chamber.

Apparatus

Six experimental chambers (24×21×30 cm) (Med Associates) were used. Each chamber was supplied with a retractable lever, two stimulus lights (one above the retractable lever, the other above the food receptacle) and encased in a sound attenuating cubicle with a fan providing a constant low level of background noise. Each food receptacle was equipped with an infrared head entry detector. The experiments were controlled online by a Windows 98™–based computer system equipped with SmartControl®-Interfaces and the MedPC™-Software (Med Associates).

RT task

A simple RT task as used in previous studies (e.g. Giertler et al., 2004; Calaminus and Hauber, 2006) was employed. A simplified scheme on the order of trial events is given in Fig. 1. First, an instructive stimulus above the food receptacle was turned on at one of two brightness levels indicating the upcoming reward magnitudes (one or five pellets, 45 mg pellets, Bioserv), 3 s later the lever was inserted. Thereafter, a trained rat pressed the inserted lever spontaneously. After a foreperiod of 0.3 s, an imperative stimulus provided by a stimulus light above the lever signaled the animal to release the lever. Responses with RT <2 s were considered as being correct and were rewarded. Early responses initiated before the onset of the imperative stimulus or late responses (RT ≥2 s) caused the trial to be repeated. After an inter-trial interval of 3 s, the instructive stimulus was turned on again thereby signaling the beginning of a new trial. The reward magnitude (low/high) for each trial was pseudo-randomly determined in advance. Brightness levels of instructive stimuli were balanced, i.e. for 50% of the rats a bright stimulus was associated with low reward and a dim stimulus was associated with high reward. For the other 50% of the rats, the opposite pattern was used. A daily session demanded 40 correct trials, i.e. 20 correct trials for low and high reward.

RT was defined as latency from the onset of the imperative stimulus to lever release, movement time (MT) was defined as latency from lever release to photobeam disruption in the food receptacle. Both measures were recorded with an accuracy of <10 ms and calculations on RT and MT values were conducted with data from correct trials (RT <2 s). Furthermore, the overall number of trials (early+correct+late responses) to reach the criterion of 40 correct responses was counted and used as an index of the accuracy of performance. These measures allow a detailed analysis of the guidance of instrumental behavior by reward-predictive stimuli. Response latencies measured by RTs become shorter for expected high versus low reward. Therefore, the RT differences of responses for expected low and high reward are a sensitive index of discriminative learning of stimulus-associated

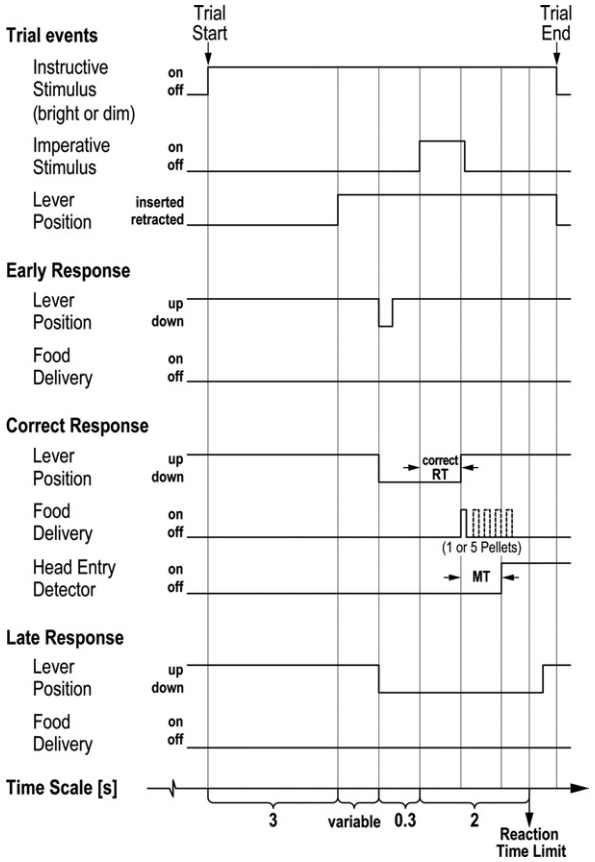

Fig. 1. Schematic representation of the order of trial events. First, an instructive stimulus was turned on at one of two brightness levels indicating the upcoming reward magnitudes (one or five pellets). Thereafter, the rat pressed the inserted lever spontaneously. After a foreperiod of 0.3 s, the imperative stimulus signaled the animal to release the lever. Responses with RT$<$2 s were considered as being correct and were rewarded. Early responses initiated before the onset of the imperative stimulus or late responses (RT$\geq$2 s) caused the trial to be repeated.

reward values. The speed of response execution was measured by MTs. MTs became also generally shorter during task acquisition, however, unlike RTs MTs are influenced by a number of parameters such as body position relative to the manipulandum. Therefore, MT is a less appropriate parameter for reward expectancy (Giertler et al., 2003; Hollerman et al., 1998) and was used here as a sensitive parameter to control for gross drug-induced motor effects (Giertler et al., 2003). Response accuracy measured by the number of responses necessary to reach a fixed criterion of correct responses was used as an index of learning the instrumental contingency. Incorrect, i.e. early and late responses, permitted further assessment of drug effects on response preparation.

Experimental procedures

Experiment 1 examined the effects of an intra-OFC D1 receptor antagonism, experiment 2 the effects of an intra-OFC D2 receptor antagonism, experiment 3 the effects of an intra-OFC NMDA receptor antagonism on a reversal of discrimination learning. The design of all three experiments was identical. Different groups of rats were trained in the RT task described above. Acquisition was tested throughout 8 days with one daily session. Thereafter stimulus–reward magnitude contingencies were reversed and instrumental learning was analyzed throughout 6 days within one daily session. During reversal, different groups of animals received microinjections of either SCH23390, eticlopride or AP5; respective control groups received microinjections of vehicle.

(1) *Preoperative habituation.* In the first two sessions, subjects were habituated to the experimental chamber with access to food pellets placed into the food receptacle. In the following five sessions, a habituation program with a fixed ratio-1 schedule commenced until a criterion of 20 consecutive lever responses was attained. Afterward, rats were subjected to surgery.

(2) *Acquisition.* After postoperative recovery the experiment was started with one daily session; data from the initial session were not evaluated. During days 1–8, task acquisition was examined. On days 6, 7 and 8, animals received a sham injection procedure including handling procedure, insertion of injection cannulae dummies and operation of the injection pump (without running an injection) before the onset of individual behavioral testing.

(3) *Reversal.* Subsequently, learning of reversed stimulus–reward magnitude contingencies was tested on days 9–14, i.e. rats had to learn that the stimulus formerly predicting high reward was associated with low reward and vice versa. During reversal rats received a drug or vehicle microinjection (1 μg eticlopride, n=17/vehicle n=13; 1 μg SCH23390, n=14/vehicle, n=13; 5 μg AP5, n=14/vehicle, n=13) before the onset of behavioral testing.

Data analysis

In line with earlier studies (e.g. Bohn et al., 2003), subjects perceived brightness levels of instructive stimuli equally, i.e. for a given reward magnitude level mean accuracy and RT values obtained with a bright or a dim stimulus did not differ significantly (data not shown). Therefore, response measures for a given reward magnitude obtained with bright and dim instructive stimuli were collapsed.

Data are expressed as means from blocks of two sessions $\pm$standard error of the mean (S.E.M.). When averaging RT data, a geometric mean was calculated for each rat and session, as the geometric mean is less influenced by outlying data points than is the arithmetic mean. Overall, RT means of responses associated with the high and low reward magnitude represent the arithmetic average of the geometric means of individual rats (Brasted et al., 1997).

Data from all experiments were subjected to separate repeated measures analysis of variance (ANOVA). Numbers of correct and early responses of correct responses during acquisition (blocks 1–4) and reversal (blocks 5–7) were compared using an ANOVA with group (groups to be treated or treated with vehicle, eticlopride, SCH23390 or AP5) as between-subjects factor and reward magnitude and blocks as within-subjects (repeated measures) factors. RTs of correct responses for high and low reward during acquisition and reversal were subjected to a planned contrast analysis, i.e. for each block and group, mean $RT_{low\ reward}$ versus mean $RT_{high\ reward}$ were compared separately by linear contrasts. All statistical computations were carried out with Statistica™ (version 7.1, StatSoft®, Inc., Tulsa, OK, USA). The level of statistical significance (α-level) was set at $P<0.05$.

Histology

After completion of behavioral testing, animals were killed by an overdose of sodium pentobarbital (150 mg/kg, i.p.) (Sigma-Aldrich, Taufkirchen, Germany) to control for correct placement of cannulae. Brains were rapidly removed, fixed in 10% formalin for 2.5 h and stored in 30% glucose. Brain sections (30 μm) were cut with a cryostat (Reichert & Jung, Heidelberg, Germany), mounted on coated slides and stained with Cresyl Violet. The location of cannulae tips is shown in Fig. 2. No rat was excluded from analysis due to cannulae misplacements.

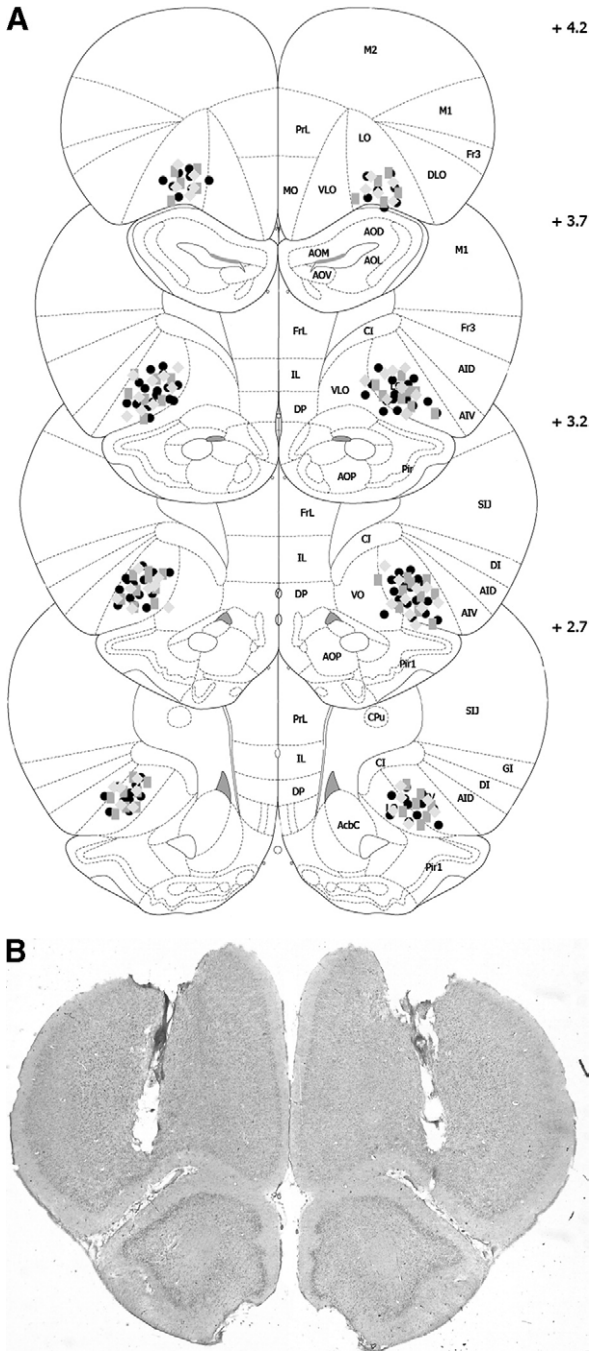

Fig. 2. (A) Location of microinjection cannulae tips in rats of experiments 1–3 (● experiment 1, ♦ experiment 2, ■ experiment 3). Plates are adaptations from the atlas of Paxinos and Watson (1997)s. Numbers beside each plate correspond to millimeters anterior to bregma. (B) Nissl stain of a coronal section showing the cannulae tracks.

RESULTS

According to their performance during acquisition, animals in experiments 1–3 were divided in two groups to be treated with vehicle or drug during reversal. The groups were chosen that no significant difference of RT was detectable during acquisition ($F<1$).

Experiment 1: Effects of an intra-OFC D1 receptor blockade

Accuracy of performance. As observed in previous studies (Giertler et al., 2005), rats needed approx. 50–60 trials to achieve the criterion of 40 correct responses (20 responses with RT <2 s for each reward magnitude). As incorrect trials largely reflected early errors, only this type of errors was given. Intra-OFC infusion of SCH23390 had no prominent effects on the number of early errors (Fig. 3) and the number of trials to reach criterion. Three-way ANOVAs on the number of early responses with group/treatment (vehicle, SCH23390) as between-subject factors and reward magnitude (low, high) and block as within-subjects factor revealed no main effects of group during acquisition (blocks 1–4) and reversal (blocks 5–7), but a main block effect during acquisition ($F_{(3,72)}=2.82$, $P=0.04$) as well as reversal ($F_{(2,48)}=3.21$, $P=0.05$). Similarly, an ANOVA on the trials to reach criterion during acquisition and reversal indicated no main effects of group, but a main block effect during acquisition ($F_{(3,72)}=2.82$, $P=0.04$) and reversal ($F_{(2,48)}=3.21$, $P=0.05$) (data not shown).

RTs. As shown in Fig. 4 during acquisition, RTs of animals to be treated with SCH22390 and vehicle significantly decreased over blocks ($F_{(3,69)}=62.71$, $P<0.0001$) and were guided by expected reward magnitude ($F_{(1,23)}=24.59$, $P<0.0001$). Planned contrast analysis revealed that in both groups RTs for expected high reward were shorter on blocks 2–4. For reversal, an ANOVA indicated a main effect of reward magnitude ($F_{(1,23)}=6.93$, $P<0.02$), but no main effect of treatment and treatment×reward magnitude interaction. Planned contrast analysis revealed that in vehicle-treated rats RT for expected low and high reward differed significantly on blocks 5 and 6, whereas on block 7 there was a trend for a significant difference ($P=0.07$). By contrast, in SCH23390-treated animals, no significant differences between RT for expected low and high reward were detected on blocks 5–7.

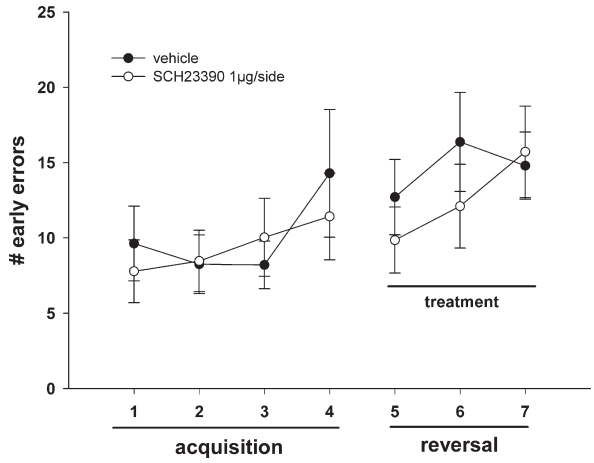

Fig. 3. Effects of intra-OFC infusion of SCH23390 on accuracy of responding. Mean number of early responses (±S.E.M.) in blocks of two sessions are given. SCH23390 at 1 μg ($n=14$) or vehicle ($n=13$) was given during blocks 5–7. Infusion of SCH23390 did not alter the number of early errors significantly.

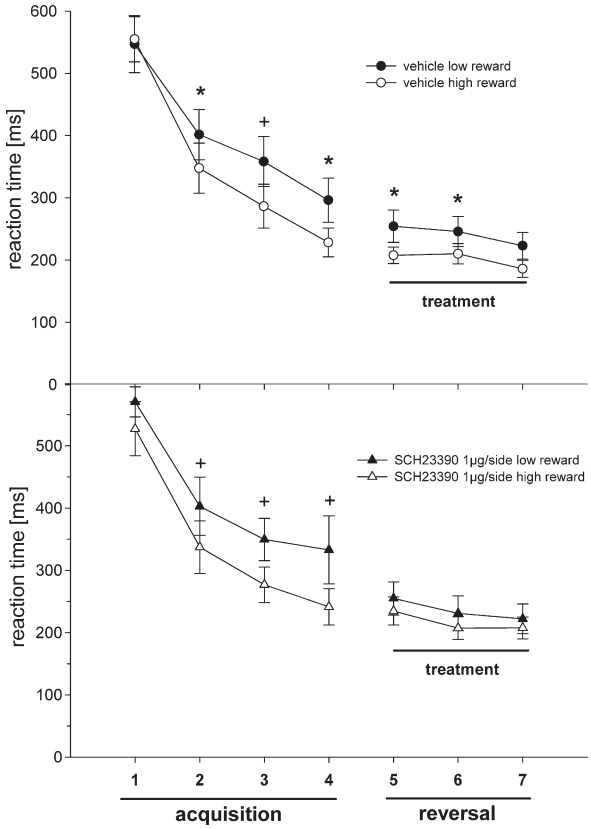

Fig. 4. Effects of intra-OFC infusion of SCH23390 on RT. Mean RTs (±S.E.M.) of correct responses in blocks of two sessions are given. SCH23390 at 1 µg (n=14) or vehicle (n=13) was given during blocks 5–7. * P<0.05, $^{+}$ P<0.01 (planned contrasts; $RT_{low\ reward}$ versus $RT_{high\ reward}$).

MTs. Intra-OFC infusion of SCH23390 had no effects on MTs (data not shown). During acquisition and reversal, MT for responses for expected high and low reward did not differ significantly as analyzed by planned contrasts. Separate ANOVAs on respective blocks revealed no significant main effects of reward magnitude or treatment and no significant reward magnitude×treatment interactions.

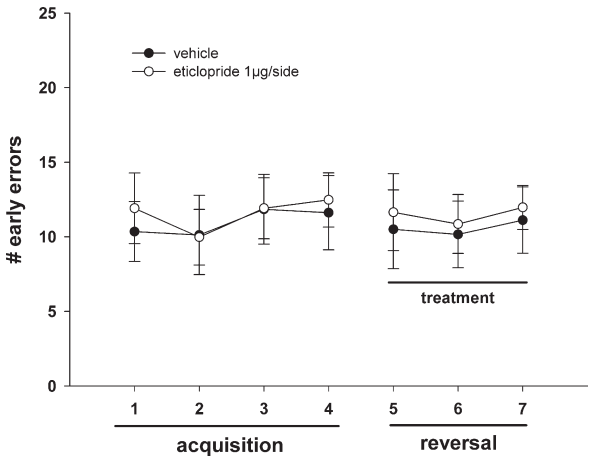

Fig. 5. Effects of intra-OFC infusion of eticlopride on accuracy of responding. Mean number of early responses (±S.E.M.) in blocks of two sessions are given. Eticlopride at 1 µg (n=17) or vehicle (n=13) was given during blocks 5–7. Infusion of eticlopride did not alter the number of early errors significantly.

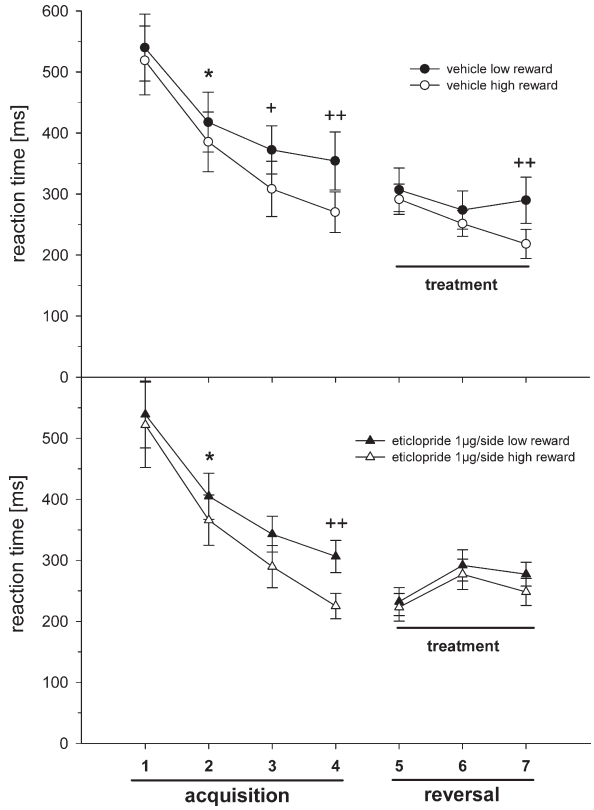

Fig. 6. Effects of intra-OFC infusion of eticlopride on RT. Mean RTs (±S.E.M.) of correct responses in blocks of two sessions are given. Eticlopride at 1 µg (n=17) or vehicle (n=13) was given during blocks 5–7. * P<0.05, $^{+}$ P<0.01, $^{++}$ P<0.001 (planned contrasts; $RT_{low\ reward}$ versus $RT_{high\ reward}$).

Experiment 2: Effects of an intra-OFC D2 receptor blockade

Accuracy of performance. As with experiment 1, rats needed approx. 50–60 trials to achieve the criterion of 40 correct responses. Intra-OFC infusion of eticlopride did not significantly alter the number of early responses (Fig. 5) or the number of trials to reach criterion. Separate ANOVAs on the number of early responses during acquisition and reversal indicated no main effect of group/treatment and no effect of block. Likewise, an ANOVA on the trials to reach criterion during acquisition and reversal indicated no main effects of group/treatment (data not shown) and no effect of block.

RTs. RTs of animals to be treated with eticlopride and vehicle significantly decreased over acquisition blocks ($F_{(3,84)}$=24.86, P<0.0001) and were guided by expected reward magnitude ($F_{(1,28)}$=29.68, P<0.0001). Planned contrast analysis showed that in both groups RT for expected high reward became increasingly shorter on blocks 2–4 as shown in Fig. 6. However, on block 3, animals to be treated with eticlopride showed only a trend for a significant RT difference (P=0.051).

An ANOVA on reversal blocks showed a main effect of reward magnitude ($F_{(1,28)}=9.93$, $P<0.01$), blocks ($F_{(2,56)}=5.55$, $P<0.01$) and a significant block×magnitude×treatment interaction ($F_{(2,56)}=4.11$, $P<0.05$), but no main effect of treatment and no two-way interactions. Planned contrast analysis revealed that in vehicle-treated rats RT for expected low and high reward differed significantly on block 7, whereas in eticlopride-treated animals, no significant differences between RT for expected low and high reward were detected on blocks 5–7.

MTs. Intra-OFC infusion of eticlopride produced no effects on MTs (data not shown). During acquisition and reversal, MT for responses for expected high and low reward did not differ significantly as analyzed by planned contrasts. Separate ANOVAs on respective blocks revealed no significant main effects of reward magnitude or treatment and no significant reward magnitude×treatment interactions.

Experiment 3: Effects of an intra-OFC NMDA receptor blockade

Accuracy of performance. Intra-OFC infusion of AP5 increased the number of early errors (Fig. 7). An ANOVA on the number of early responses revealed an effect of treatment ($F_{(1,26)}=7.37$, $P=0.01$) and blocks ($F_{(2,52)}=4.39$, $P=0.02$) during reversal, but not during acquisition. Consequently, the number of trials to reach criterion was higher in animals that received microinfusion of AP5 and an ANOVA on the trials to reach criterion found main block ($F_{(2,52)}=4.39$, $P=0.02$) and treatment effects ($F_{(1,26)}=7.37$, $P=0.01$) during reversal, but no main block and treatment effects during acquisition (data not shown).

RTs. Animals to be treated with vehicle or AP5 showed a significant decrease in RT over acquisition blocks ($F_{(3,69)}=36.56$, $P<0.0001$) and a guidance by expected reward magnitude ($F_{(1,23)}=11.18$, $P<0.01$) (Fig. 8).

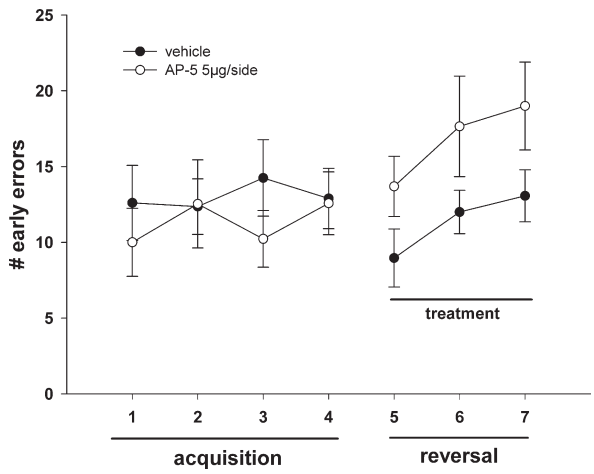

Fig. 7. Effects of intra-OFC infusion of AP5 on accuracy of responding. Mean number of early responses (±S.E.M.) in blocks of two sessions are given. AP5 at 5 µg ($n=14$) or vehicle ($n=13$) was given during blocks 5–7. Infusion of AP5 increased the number of early errors significantly.

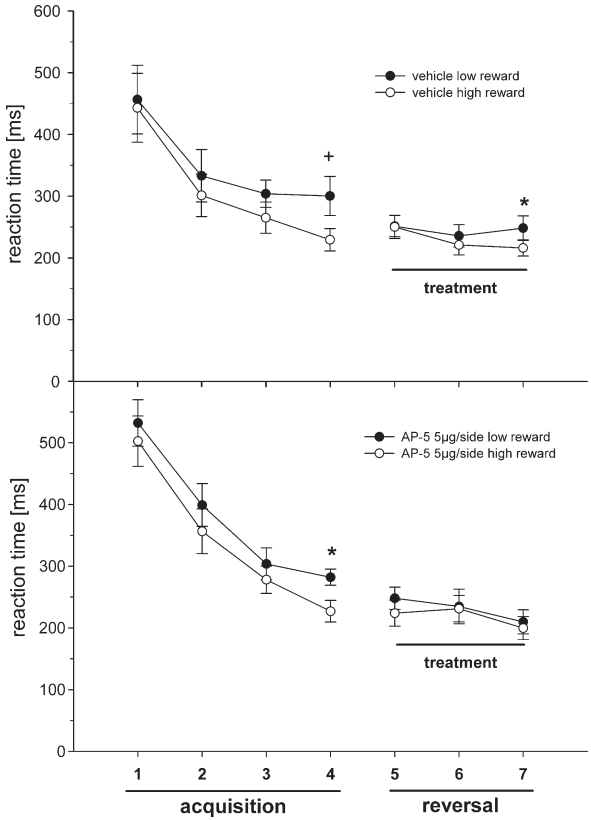

Fig. 8. Effects of intra-OFC infusion of AP5 on RT. Mean RTs (±S.E.M.) of correct responses in blocks of two sessions are given. AP5 at 5 µg ($n=14$) or vehicle ($n=13$) was given during blocks 5–7. * $P<0.05$, $^{+}$ $P<0.01$ (planned contrasts; $RT_{low\ reward}$ versus $RT_{high\ reward}$).

Planned contrast analysis revealed that in both groups RT for expected high reward were significantly shorter on block 4. Over reversal blocks, an ANOVA indicated a main effect of reward magnitude ($F_{(1,22)}=6.61$, $P<0.02$), but no main effect of treatment and no two-way interactions. Planned contrast analysis revealed that in vehicle-treated rats RT for expected low and high reward differed significantly on block 7, whereas in AP5-treated animals, no significant differences between RT for expected low and high reward were detected on blocks 5–7.

MTs. Intra-OFC infusion of AP5 had no effects on MTs (data not shown). During acquisition and reversal, MT for responses for expected high and low reward did not differ significantly as analyzed by planned contrasts. Separate ANOVAs on respective blocks revealed no significant main effects of reward magnitude or treatment and no significant reward magnitude×treatment interactions.

DISCUSSION

The present study demonstrates dissociable effects of a D1 and D2 receptor blockade versus an NMDA receptor blockade in the OFC on adapting instrumental behavior to changing stimulus–reward contingencies. A blockade of intra-OFC D1 or D2 receptors during reversal did not alter the number of early responses, but slowed learning to

discriminate the current stimulus–reward magnitude contingencies. As previously shown (Bohn et al., 2003a), blockade of intra-OFC NMDA receptor increased the number of early responses and impaired learning to discriminate the current stimulus–reward magnitude contingencies. These findings suggest that D1/D2 and NMDA receptor-mediated signaling in the OFC plays an important and partially overlapping role in adapting responding to changing stimulus–reward contingencies.

The possibility that drug diffusion from the OFC to adjacent regions contributes to behavioral effects cannot be discounted as little is known about the exact spread of each drug from the site of infusion. However, Granon et al. (2000) demonstrated that a substantial amount of radiolabeled SCH23390 injected into the mPFC in the same volume as used here remained in the vicinity of the injection site even 1 h after injection. Furthermore, a considerable number of studies showed that microinfusion of various drugs into adjacent prefrontal subregions such as OFC, prelimbic or infralimbic cortex produced dissociable behavioral effects (e.g. Capriles et al., 2003; De Bruin et al., 2000) suggesting that the functional spread of drugs was relatively limited. Therefore, behavioral effects seen here may primarily arise from drug actions within the OFC.

D1 and D2 receptor activity and learning under reversal conditions

Under reversal conditions, RTs of vehicle controls were guided by expected reward magnitude on block 3 or earlier. Notably, vehicle controls of experiment 1 adapted to changing stimulus–reward contingencies already on the first block. Such inter-group variability already observed in previous studies (Calaminus and Hauber, 2006) may reflect a sampling problem, i.e. some groups involve a high ratio of individuals rapidly responding to contingency changes, while most groups consist of about similar proportions of rapid and slow responders (data not shown). In addition, in an earlier study (Bohn et al., 2003a) adaption of RTs to reversed stimulus–reward magnitudes required more days as observed here and on initial reversal days RTs for low reward were even shorter than RTs for high reward. It is likely that faster reversal seen here is related to a shorter acquisition phase of 8 days (four blocks) in the present as compared with 15 days in our previous study (Bohn et al., 2003a).

Most importantly, results show that intra-OFC infusion of SCH23390 and eticlopride impaired rats' ability to adapt instrumental behavior to changing stimulus–reward magnitude contingencies. The doses of SCH23390 and eticlopride used here were based on pilot studies and data reported in the literature (Floresco et al., 2006; Ragozzino, 2002; Schweimer and Hauber, 2006; Seamans et al., 1998; Sun and Rebec, 2005). As the relative densities of prefrontal D1 and D2 receptors are different (Lidow et al., 1991; Goldman-Rakic et al., 1992; Gaspar et al., 1995; Sesack et al., 1995), we cannot rule out that, in behavioral terms, the doses of SCH23390 and eticlopride (1 μg, respectively) were not fully equipotent. However, this possibility is unlikely as this dose, but not a lower dose, of SCH23390 (Ragozzino, 2002) or eticlopride (Floresco et al., 2006) microinjected into the mPFC impaired set-shifting of rats. Moreover, motor execution was intact in SCH23390 and eticlopride-treated animals as shown by unaffected MTs suggesting that drug-induced nonspecific motor effects may not account for learning impairments seen here. Consistent with this view, OFC lesion studies provide no indications of nonspecific motor symptoms (Boulougouris et al., 2007; Ferry et al., 2000; Kim and Ragozzino, 2005; Mobini et al., 2002; Schoenbaum et al., 2002; Stalnaker et al., 2007). Furthermore, an impaired reward magnitude processing may not contribute to reversal deficits as OFC lesions did not affect the ability to discriminate large from small rewards (Bohn et al., 2003b). Notably, the number of early errors remained constant or even increased during the course of acquisition and reversal in vehicle controls and drug-treated animals. This observation corresponds with previous findings (e.g. Giertler et al., 2005) and largely reflects the fact that with progressive learning animals responded faster and thus made more erroneous early responses, in particular for expected high reward, suggesting that reward magnitude expectation interferes with response preparation (Giertler et al., 2005).

Our observation that a blockade of intra-OFC D1 and D2 receptors impaired the ability to modify behavior to changing stimulus–reward magnitude contingencies supports the general notion that the OFC is critical to facilitating rapid reversal learning (Bohn et al., 2003c; Brown and Bowman, 2002; Chudasama and Robbins, 2003; Dias et al., 1996; Schoenbaum et al., 2003; see Murray et al., 2007, for a recent overview). However, a recent primate study by Clarke et al. (2007) implicated 5-HT, not DA, neurotransmission in the OFC in reversal learning. They observed that OFC DA depletion did not increase the number of errors to criterion to discriminate a rewarded and a non-rewarded stimulus during serial reversals. Yet, it is important to note that reversal learning tasks used in this and our present study differ considerably and results of experimental manipulation of OFC function on reversal learning are therefore difficult to compare. For instance, Clarke et al. (2007) determined the error rate to criterion during reversal learning in a visual discrimination task using reward predictive and non-reward predictive stimuli. By contrast, in the task used here discrimination learning is not instrumental in that reinforcement is contingent on the subject's choosing between the stimuli and the task examined whether animals can discriminate stimuli predictive of different reward magnitudes. In addition, it is possible that discrimination of stimuli predictive of high versus low reward as used here is more difficult than discrimination of reward-predictive and nonreward-predictive stimuli. Preliminary data from our task provide support to this notion (Calaminus and Hauber, unpublished observations). Furthermore, subtle effects of OFC DA depletions on guidance of response vigor by reward stimuli may be easier to detect by using RT measures than error counts, i.e. RTs of responses to expected high versus low reward measured here may provide a more sensitive index of the motiva-

tional effects of predictive stimuli. Thus, by analyzing response vigor to discriminative stimuli, our results suggest that an OFC DA depletion impaired guidance of instrumental behavior under reversal conditions possibly by affecting discrimination of the incentive value of the stimuli.

Other task differences, i.e. the use of simple versus serial reversals, may be relevant as well, in particular because the role of prefrontal DA in simple and serial reversal learning seems to be distinct. Using *in vivo* microdialysis, van der Meulen et al. (2007) examined mPFC DA efflux in rats during serial reversals in an instrumental learning task. Importantly, they found that prefrontal DA activity was increased during the execution of an instrumental discrimination task and that this increase was only detected during the first, but not during later reversals. This observation indicates that prefrontal DA plays a critical role limited to initial reversal learning. However, our microinfusions were placed more laterally (about 0.5 mm) than microdialysis probes in the study by van der Meulen et al. (2007) and the extent to which they capture OFC DA may be relatively low (Rice et al., 1985; Nicholson, 2005). Yet, it is possible that DA not only in the mPFC but also in the OFC mediates initial adaptation to changing task contingencies, whereas serial reversal learning involves additional learning mechanisms sensitive to 5-HT depletion as shown by Clarke et al. (2007). Additional studies involving different types of reversal learning tasks will be required to provide a more detailed characterization of the role of OFC DA in this type of learning.

Electrophysiological studies demonstrate that OFC neurons modify their responses to reward predictive stimuli to reflect changes in their current incentive value (Critchley and Rolls, 1996; Schoenbaum et al., 1999; Tremblay and Schultz, 1999). Thus, through signaling the current value of the expected outcome, the OFC is critical to facilitating rapid reversal learning (Murray et al., 2007). Recent studies provide evidence that DA neurons emit prediction error signals necessary to update the predictive significance of stimuli (Schultz, 2006; Roesch et al., 2007) and that such signals are present in the OFC, a target area of DA neurons (McClure et al., 2003; O'Doherty, 2003). The current findings are consistent with these findings and further suggest that D1 and D2 receptor activity in the OFC could support discriminative guidance of reward-directed behavior. Several open questions may be addressed in future studies. First, it is possible that not only reversal, but acquisition of this task also requires OFC D1 and D2 receptor activity. Second, a behavioral analysis of a combined intra-OFC D1 and D2 receptor blockade during reversal may provide further insights, e.g. additive deleterious effects could indicate that OFC D1 and D2 receptors mediate distinct behavioral functions.

DA has been implicated in a number of other aspects of OFC-mediated reward processing, e.g. delay discounting and evaluation of reinforcer magnitudes (Kheramin et al., 2004; Winstanley et al., 2005). For instance, DA depletion of the OFC increased the sensitivity to the relative size of reinforcers (Kheramin et al., 2004), however, such an effect is unlikely to account for the behavioral results seen here. Furthermore, intra-OFC infusion of considerably higher doses of D1 and D2 antagonists as used here decreased breaking points on a progressive ratio schedule suggesting a role for OFC DA in translating motivation into action (Cetin et al., 2004). Our data provide no evidence consistent with this idea because the overall level of RT—which may be viewed as an index of general motivation—was not markedly increased after DA receptor blockade. Notably, after a D1 or D2 receptor blockade in the nucleus accumbens we obtained the opposite pattern of behavioral effects as seen here, i.e. during reversal the overall level of RTs was increased, but RTs were rapidly guided by the current reward magnitude (Calaminus and Hauber, 2006) indicating a specific role for mesoaccumbal DA systems in mediating general motivation.

NMDA receptor activity and learning under reversal conditions

Animals that received intra-OFC AP5 infusions not only had an increased number of early responses but responded with similar latencies to expected high and low reward indicating a failure to discriminate the current incentive values of reward-predictive stimuli. These results largely replicate earlier data (Bohn et al., 2003a) and suggest a role for OFC NMDA receptors in response inhibition and adapting instrumental responses under changing stimulus–reward contingencies. As in AP5-treated animals MTs were not altered relative to controls, it is unlikely that drug-induced nonspecific motor effects account for these deficits. Likewise, in our earlier study the same dose of AP5 had no nonspecific effects (Bohn et al., 2003a). Our observations that a blockade of intra-OFC NMDA receptors delayed adaptation to reversed contingencies are consistent with the view that the OFC is critical to facilitating rapid reversal learning (Murray et al., 2007). Notably, in OFC neurons encoding of predicted outcome depends on input from the basolateral amygdala (BLA) (Schoenbaum et al., 2003). Thus, it is possible that an NMDA receptor blockade inhibited a BLA-mediated update of the value of reward predictive stimuli represented in the OFC.

Our results further demonstrate an increased rate of early responses in AP5-treated animals. As AP5 did not reduce RTs in parallel, this effect may not be simply a manifestation of a general tendency to respond quickly. Instead, this increased rate of early errors may be tentatively interpreted to reflect impaired response inhibition. Whether reduced response inhibition accounts for impaired reversal learning is questionable. Chudasama et al. (2007) reported that OFC lesions do not always produce a failure to inhibit previously successful responses suggesting that an inability to inhibit prepotent responses may not provide a general explanation for reversal learning deficits associated with OFC dysfunction.

The role of OFC DA and glutamate in behavioral flexibility

There is consistent evidence from human, primate and rodent studies that central DA systems modulate reversal learning (Cools et al., 2002; Kruzich and Grandy, 2004;

Mehta et al., 2001; Ridley et al., 1981). Computational (Frank and Claus, 2006) and empirical studies (Taghzouti et al., 1985, but see Collins et al., 2000) implicated striatal DA in reversal learning. Our data suggest that OFC D1 and D2 receptor activity may contribute to guidance of instrumental behavior under reversal conditions, possibly by enhancing discrimination of the current incentive values of reward-predictive stimuli.

In line with our previous study (Bohn et al., 2003a), the present data further indicate that learning of reversed stimulus-outcome contingencies requires NMDA receptor activity in the OFC. Notably, only NMDA, not D1 and D2 receptor activity is critical in inhibition of early responses under reversal conditions. Together, these findings suggest that D1/D2 and NMDA receptor-mediated signaling in the OFC plays an important and partially overlapping role in behavioral flexibility as examined here.

Acknowledgments—This research was supported by the DFG (Ha2340/8-1).

REFERENCES

Baldwin AE, Sadeghian K, Kelley AE (2002) Appetitive instrumental learning requires coincident activation of NMDA and dopamine D1 receptors within the medial prefrontal cortex. J Neurosci 22: 1063–1071.

Berger B, Gaspar P, Verney C (1991) Dopaminergic innervation of the cerebral cortex: unexpected differences between rodents and primates. Trends Neurosci 14:21–27.

Birrell JM, Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. J Neurosci 20:4320–4324.

Bohn I, Giertler C, Hauber W (2003a) NMDA receptors in the rat orbital prefrontal cortex are involved in guidance of instrumental behaviour under reversal conditions. Cereb Cortex 13:968–976.

Bohn I, Giertler C, Hauber W (2003b) Orbital prefrontal cortex and guidance of instrumental behavior of rats by visuospatial stimuli predicting reward magnitude. Learn Mem 10:177–186.

Bohn I, Giertler C, Hauber W (2003c) Orbital prefrontal cortex and guidance of instrumental behaviour in rats under reversal conditions. Behav Brain Res 143:49–56.

Boulougouris V, Dalley JW, Robbins TW (2007) Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. Behav Brain Res 179:219–228.

Brasted PJ, Humby T, Dunnett SB, Robbins TW (1997) Unilateral lesions of the dorsal striatum in rats disrupt responding in egocentric space. J Neurosci 17:8919–8926.

Brown VJ, Bowman EM (2002) Rodent models of prefrontal cortical function. Trends Neurosci 25:340–343.

Calaminus C, Hauber W (2006) Intact discrimination reversal learning but slowed responding to reward-predictive cues after dopamine D1 and D2 receptor blockade in the nucleus accumbens of rats. Psychopharmacology (Berl) 191:551–566.

Capriles N, Rodaros D, Sorge RE, Stewart J (2003) A role for the prefrontal cortex in stress- and cocaine-induced reinstatement of cocaine seeking in rats. Psychopharmacology (Berl) 168:66–74.

Cetin T, Freudenberg F, Fuchtemeier M, Koch M (2004) Dopamine in the orbitofrontal cortex regulates operant responding under a progressive ratio of reinforcement in rats. Neurosci Lett 370:114–117.

Chudasama Y, Kralik JD, Murray EA (2007) Rhesus monkeys with orbital prefrontal cortex lesions can learn to inhibit prepotent responses in the reversed reward contingency task. Cereb Cortex 17:1154–1159.

Chudasama Y, Robbins TW (2003) Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. J Neurosci 23:8771–8780.

Clarke HF, Walker SC, Dalley JW, Robbins TW, Roberts AC (2007) Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. Cereb Cortex 17:18–27.

Collins P, Wilkinson LS, Everitt BJ, Robbins TW, Roberts AC (2000) The effect of dopamine depletion from the caudate nucleus of the common marmoset (Callithrix jacchus) on tests of prefrontal cognitive function. Behav Neurosci 114:3–17.

Cools R, Clark L, Owen AM, Robbins TW (2002) Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. J Neurosci 22:4563–4567.

Critchley HD, Rolls ET (1996) Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. J Neurophysiol 75:1673–1686.

De Bruin JP, Feenstra MG, Broersen LM, Van Leeuwen M, Arens C, De Vries S, Joosten RN (2000) Role of the prefrontal cortex of the rat in learning and decision making: effects of transient inactivation. Prog Brain Res 126:103–113.

Dias R, Robbins TW, Roberts AC (1996) Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. Behav Neurosci 110:872–886.

Ferry AT, Lu XC, Price JL (2000) Effects of excitotoxic lesions in the ventral striatopallidal-thalamocortical pathway on odor reversal learning: inability to extinguish an incorrect response. Exp Brain Res 131:320–335.

Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MT (2006) Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. Neuropsychopharmacology 31:297–309.

Frank MJ, Claus ED (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol Rev 113:300–326.

Gaspar P, Bloch B, Le Moine C (1995) D1 and D2 receptor gene expression in the rat frontal cortex: cellular localization in different classes of efferent neurons. Eur J Neurosci 7:1050–1063.

Ghods-Sharifi S, Haluk DM, Floresco SB (2008) Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. Neurobiol Learn Mem 89:567–573.

Giertler C, Bohn I, Hauber W (2003) The rat nucleus accumbens is involved in guiding of instrumental responses by stimuli predicting reward magnitude. Eur J Neurosci 18:1993–1996.

Giertler C, Bohn I, Hauber W (2004) Transient inactivation of the rat nucleus accumbens does not impair guidance of instrumental behaviour by stimuli predicting reward magnitude. Behav Pharmacol 15:55–63.

Giertler C, Bohn I, Hauber W (2005) Involvement of NMDA and AMPA/KA receptors in the nucleus accumbens core in instrumental learning guided by reward-predictive cues. Eur J Neurosci 21:1689–1702.

Goldman-Rakic PS, Lidow MS, Smiley JF, Williams MS (1992) The anatomy of dopamine in monkey and human prefrontal cortex. J Neural Transm Suppl 36:163–177.

Granon S, Passetti F, Thomas KL, Dalley JW, Everitt BJ, Robbins TW (2000) Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. J Neurosci 20:1208–1215.

Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behaviour-related neuronal activity in primate striatum. J Neurophysiol 80:947–963.

Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci 24:7540–7548.

Kheramin S, Body S, Ho MY, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2004) Effects of orbital pre-

frontal cortex dopamine depletion on inter-temporal choice: a quantitative analysis. Psychopharmacology (Berl) 175:206–214.
Kim J, Ragozzino ME (2005) The involvement of the orbitofrontal cortex in learning under changing task contingencies. Neurobiol Learn Mem 83:125–133.
Kruzich PJ, Grandy DK (2004) Dopamine D2 receptors mediate two-odor discrimination and reversal learning in C57BL/6 mice. BMC Neurosci 5:12.
Lidow MS, Goldman-Rakic PS, Gallager DW, Rakic P (1991) Distribution of dopaminergic receptors in the primate cerebral cortex: quantitative autoradiographic analysis using [3H]raclopride, [3H]spiperone and [3H]SCH23390. Neuroscience 40:657–671.
McClure SM, Daw ND, Montague PR (2003) A computational substrate for incentive salience. Trends Neurosci 26:423–428.
Mehta MA, Swainson R, Ogilvie AD, Sahakian J, Robbins TW (2001) Improved short-term spatial memory but impaired reversal learning following the dopamine D(2) agonist bromocriptine in human volunteers. Psychopharmacology (Berl) 159:10–20.
Mishkin S (1964) The interdependence of clinical neurology and neurophysiology. An historical review of the vestibulo-ocular reflex. Mcgill Med J 33:80–97.
Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2002) Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology (Berl) 160:290–298.
Murray EA, O'Doherty JP, Schoenbaum G (2007) What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. J Neurosci 27:8166–8169.
Nicholson C (2005) Factors governing diffusing molecular signals in brain extracellular space. J Neural Transm 112:29–44.
O'Doherty J (2003) Can't learn without you: predictive value coding in orbitofrontal cortex requires the basolateral amygdala. Neuron 39:731–733.
Paxinos G, Watson C (1997) The rat brain in stereotaxic coordinates. New York: Academic Press.
Ragozzino ME, Wilcox C, Raso M, Kesner RP (1999) Involvement of rodent prefrontal cortex subregions in strategy switching. Behav Neurosci 113:32–41.
Ragozzino ME (2002) The effects of dopamine D(1) receptor blockade in the prelimbic-infralimbic areas on behavioral flexibility. Learn Mem 9:18–28.
Rice ME, Gerhardt GA, Hierl PM, Nagy G, Adams RN (1985) Diffusion coefficients of neurotransmitters and their metabolites in brain extracellular fluid space. Neuroscience 15:891–902.
Ridley RM, Haystead TA, Baker HF (1981) An analysis of visual object reversal learning in the marmoset after amphetamine and haloperidol. Pharmacol Biochem Behav 14:345–351.
Roesch MR, Calu DJ, Schoenbaum G (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat Neurosci 10:1615–1624.
Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci 19:1876–1884.
Schoenbaum G, Nugent SL, Saddoris MP, Setlow B (2002) Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport 13:885–890.
Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem 10:129–140.
Schultz W (2006) Behavioral theories and the neurophysiology of reward. Annu Rev Psychol 57:87–115.
Schweimer J, Hauber W (2006) Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. Learn Mem 13:777–782.
Seamans JK, Floresco SB, Phillips AG (1998) D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. J Neurosci 18:1613–1621.
Sesack SR, Snyder CL, Lewis DA (1995) Axon terminals immunolabeled for dopamine or tyrosine hydroxylase synapse on GABA-immunoreactive dendrites in rat and monkey cortex. J Comp Neurol 363:264–280.
Stalnaker TA, Franz TM, Singh T, Schoenbaum G (2007) Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. Neuron 54:51–58.
Stefani MR, Moghaddam B (2003) Distinct contributions of glutamate receptor subtypes to cognitive set-shifting abilities in the rat. Ann N Y Acad Sci 1003:464–467.
Sun W, Rebec GV (2005) The role of prefrontal cortex D1-like and D2-like receptors in cocaine-seeking behavior in rats. Psychopharmacology (Berl) 177:315–323.
Sutton RS, Barto AG (1990) Time derivative models of pavlovian reinforcement. In: Learning and computational neuroscience: foundations of adaptive networks (Gabriel M, Moore J, eds), pp 497–537. Cambridge, MA: MIT Press.
Taghzouti K, Louilot A, Herman JP, Le Moal M, Simon H (1985) Alternation behavior, spatial discrimination, and reversal disturbances following 6-hydroxydopamine lesions in the nucleus accumbens of the rat. Behav Neural Biol 44:354–363.
Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49: 93–115.
Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704–708.
van der Meulen JA, Joosten RN, de Bruin JP, Feenstra MG (2007) Dopamine and noradrenaline efflux in the medial prefrontal cortex during serial reversals and extinction of instrumental goal-directed behavior. Cereb Cortex 17:1444–1453.
Winstanley CA, Theobald DE, Dalley JW, Robbins TW (2005) Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. Neuropsychopharmacology 30:669–682.

(Accepted 23 April 2008)
(Available online 2 May 2008)