

Thalamic Control of Dorsomedial Striatum Regulates Internal State to Guide Goal-Directed Action Selection

Laura A. Bradfield^{1,2} and  Bernard W. Balleine^{1,2}

¹School of Psychology, University of New South Wales, New South Wales 2052, Australia, and ²Brain and Mind Centre, University of Sydney, Camperdown, New South Wales 2050, Australia

We (Bradfield et al., 2013) have demonstrated previously that parafascicular thalamic nucleus (PF)-controlled neurons in the posterior dorsomedial striatum (pDMS) are critical for interlacing new and existing action–outcome contingencies to control goal-directed action. Based on these findings, it was suggested that animals with a dysfunctional PF–pDMS pathway might suffer a deficit in creating or retrieving internal contexts or “states” on which such information could become conditional. To assess this hypothesis more directly, rats were given a disconnection treatment using contralateral cytotoxic lesions of the PF and pDMS (Group CONTRA) or ipsilateral control lesions (Group IPSI) and trained to press a right and left lever for sucrose and pellet outcomes, after which these contingencies were reversed. The rats were then given an outcome devaluation test (all experiments) and a test of outcome-specific reinstatement (Experiments 1 and 3). We found that devaluation performance was intact for both groups after training of initial contingencies, but impaired for Group CONTRA after reversal. However, performance was restored by additional reversal training. Furthermore, when tested a second time after reversal training, rats in both groups demonstrated responding in accordance with the original contingencies, providing direct evidence of modulation of action selection by state. Finally, we found that external context could substitute for internal state and so could rescue responding in Group CONTRA, but only in the reinstatement test. Together, these findings suggest that animals use internal state information to guide action selection and that this information is modulated by the PF–pDMS pathway.

Key words: disconnection; goal-directed action; outcome devaluation; parafascicular thalamus; posterior dorsomedial striatum; reversal

Significance Statement

Individuals with Parkinson’s disease dementia often suffer a characteristic deficit in “cognitive flexibility.” It has been suggested that neurodegeneration in the pathway between the centromedian/parafascicular thalamic nucleus (PF) and striatum might underlie such deficits (Smith et al., 2014). In rats, we have similarly observed that a functional disconnection of the PF–posterior dorsomedial striatal pathway produces a specific impairment in the ability to alter goal-directed actions (Bradfield et al., 2013). It was suggested that this impairment could be a result of a deficit in state modulation. Here, we present four experiments that provide evidence for this hypothesis and suggest several ways (e.g., extended practice, providing external cues) in which this state modulation can be rescued.

Introduction

Recent evidence suggests that the ability of animals to select the appropriate action in the face of conflicting action–outcome contingencies depends on the parafascicular thalamus (PF) and its

projection to the posterior dorsomedial striatum (pDMS) (Bradfield et al., 2013). It has been hypothesized that this is achieved through the PF’s involvement in the control of the internal states on which specific action–outcome contingencies can become conditional and that thus provide a basis for retrieving and selecting specific actions as circumstances change (Deschênes et al., 1996; Gershman and Niv, 2012; Bradfield et al., 2013; Schoenbaum et al., 2013). Specifically, PF inputs to pDMS are thought to allow for conflicting contingencies to be partitioned into separate “states,” which in this instance can be thought of as something akin to an internal context (or latent cause) that enables actions appropriate to that state to be selected.

Consistent with this claim, rats have been found to show a deficit in action selection after PF–pDMS disconnection, but only after a shift in contingency. Although they were able to retain

Received Dec. 19, 2016; revised Feb. 2, 2017; accepted Feb. 13, 2017.

Author contributions: L.A.B. and B.W.B. designed research; L.A.B. performed research; L.A.B. analyzed data; L.A.B. and B.W.B. wrote the paper.

This work was supported by the National Health and Medical Research Council of Australia (Grant GNT1087689 to B.W.B. and L.A.B. and Senior Principal Research Fellowship GNT1079561 to B.W.B.). We thank Fred Westbrook and Andy Delamater for helpful conversations regarding these data.

The authors declare no competing financial interests.

Correspondence should be addressed to Bernard Balleine, Decision Neuroscience Laboratory, School of Psychology, University of NSW, Kensington, NSW 2052, Australia. E-mail: bernard.balleine@unsw.edu.au.

DOI:10.1523/JNEUROSCI.3860-16.2017

Copyright © 2017 the authors 0270-6474/17/373721-13\$15.00/0

action–outcome contingencies if they remained stable, subsequent changes produced by either contingency degradation or reversal produced a deficit in selecting the appropriate action. Furthermore, it has been reported recently that changes in state alter the activity of striatal cholinergic interneurons (CINs), the major target of PF inputs to the pDMS and the source of the effects that we observed previously, in a manner consistent with state modulation; in this case, action selection reflected the accuracy of state decoding by these neurons (Stalnaker et al., 2016).

Despite these findings, without direct evidence of state modulation questions remain regarding the nature of the deficit induced by the PF–pDMS disconnection. Here, we tested three predictions from the hypothesis that a deficit in encoding internal state is the source of the deficit in action selection induced by PF–pDMS disconnection after contingency reversal. Experiment 1 assessed whether the deficit in action selection could be rescued by additional training on the reversed contingency. If PF–pDMS disconnection induced a loss of internal state, then additional training should be successful in rescuing the deficit. This is because an inability to partition information into states should not in and of itself prevent the learning of new contingencies and should only attenuate such learning to the extent that the initial contingencies interfere with subsequent learning (Gershman and Niv, 2012). If, in contrast, the impairment caused by disconnection of the PF–pDMS pathway is a more general learning deficit that occurs whenever contingencies are altered, then it should not be possible for animals in the group receiving contralateral cytotoxic lesions of the PF and pDMS (Group CONTRA) to learn the reversed contingencies. Experiments 2A and 2B examined the consequences of inserting an interval between reversal training and test. Based on the results of several Pavlovian studies (Bouton and Peck, 1992; Romero et al., 2003), increasing the training-to-test interval might be expected to reduce any interference from the most recently acquired contingencies to favor the initial learning and might therefore prompt a return in responding in accordance with the initial contingencies. Such a result would provide strong evidence of state modulation in intact animals, suggesting that the original contingencies were retained despite animals also learning the reversed contingencies. Finally, Experiment 3 assessed whether the PF–pDMS disconnection produces a deficit in using state information of any kind by investigating whether, if internal cues are lost, animals with PF–pDMS disconnection can use external cues to guide action selection.

Materials and Methods

Subjects

For all experiments, subjects were experimentally naive male outbred hooded Wistar rats (350–450 g). They were housed in plastic boxes (2–4 rats per box) located in a climate-controlled colony room and maintained on a 12 h light/dark cycle. Rats were allowed 3–5 d to recover from surgery, during which time they were handled and weighed daily. Three days before the behavioral procedures, the rats were put on a food deprivation schedule to maintain them at ~85% of their *ad libitum* feeding weight. All experimental procedures were approved by the Animal Ethics Committee at the University of Sydney.

Behavioral apparatus

The behavioral procedures were performed in 16 identical MED Associates operant chambers enclosed in sound- and light-attenuating shells. Each chamber was equipped with a pump that was fitted with a syringe that delivered 20% sucrose solution (0.1 ml) and a pellet dispenser that delivered grain pellets (45 mg; Bioserv Biotechnologies) into a recessed magazine when activated. The chambers contained two retractable levers that could be inserted to the left and the right of the magazine. An infra-

red photobeam crossed the magazine opening, allowing for the detection of head entries. A house light (3 W, 24 V) provided constant illumination of the operant chamber and an electric fan fixed in the shell enclosure provided constant background noise (~70 dB). The house light was turned off briefly (for 2 s) each time sucrose was delivered to the magazine to indicate its delivery to the animal. The walls were clear Plexiglas and the floor was a stainless steel grid. A set of two microcomputers running MED Associates proprietary software (Med-PC) controlled all experimental events and recorded lever presses and magazine entries. Two distinct contexts were used in Experiment 3 only. One of these contexts constituted the bare, unadorned chamber with a paper towel placed in the bedding that had 0.5 ml of 10% peppermint essence added. For the other context, laminated sheets of black-and-white vertical stripes were positioned on the transparent walls of the chambers, smooth Plexiglas sheets were placed on the floor and paper towel with 0.5 ml of 10% vanilla essence was placed in the bedding. Therefore, these contexts differed along visual (transparent vs striped walls), tactile (grid vs smooth floor), and olfactory (peppermint vs vanilla) dimensions.

Surgery

Rats were anesthetized using isoflurane (Laser Animal Health) mixed with oxygen and placed in the stereotaxic frame (Stoelting). An incision was made into the scalp to expose the skull surface and the incisor bar was adjusted to place bregma and lambda in the same horizontal plane. A small hole was then drilled into the skull above the pDMS (all coordinates in millimeters relative to bregma: anteroposterior, -0.1 , mediolateral, ± 2.3 , dorsoventral, -4.5) in either the left or the right hemisphere. Another small hole was drilled above the PF in either the same hemisphere (Group IPSI) or the opposite hemisphere (Group CONTRA) to the pDMS hole (all coordinates in millimeters relative to bregma: anteroposterior, -4.2 , mediolateral, ± 1.3 , dorsoventral, -6.2). Excitotoxic pDMS and PF lesions were made by infusing $0.6 \mu\text{l}$ (pDMS) or $0.4 \mu\text{l}$ (PF) of NMDA (10 mg/ml) into each structure; the needle was then left in place for 4 min before removal to allow for diffusion. Half of the rats in each experiment were assigned to Group IPSI and received PF and pDMS lesions in the same (ipsilateral) hemisphere. The remaining rats were assigned to Group CONTRA and received PF and pDMS lesions in opposite (contralateral) hemispheres. The amount of damage that occurred in the right and left hemispheres was counterbalanced between rats. That is, half of the rats in Group IPSI received lesions in the left hemisphere, half in the right. In Group CONTRA, half the rats received PF lesions in the left and pDMS lesions in the right hemisphere and the other half received the opposite arrangement.

Histology

Subsequent to behavioral testing, subjects received a lethal dose of sodium pentobarbital. The brains were removed and sectioned coronally at $40 \mu\text{m}$ through the IC and NAc core. Every third section was collected on a slide and stained with cresyl violet. The location and size of lesions was determined under a microscope by a trained observer unaware of the subjects' group designations using the boundaries defined by the atlas of Paxinos and Watson (2013). Subjects for whom the lesions were too large or small or the placements inaccurate were excluded from the statistical analysis.

Data analyses

For this and all remaining experiments, data were analyzed using a mixed-model ANOVA followed by simple-effects analyses to establish the source of any interactions. Statistical significance was set at $p \leq 0.05$. Data are presented as mean \pm SEM and averaged across counterbalanced conditions.

For the sake of simplicity, all test results are described in accordance with the contingencies most recently trained. That is, if the rats most recently experienced training on the initial contingencies, then the “devalued” and “nondevalued” levers are described according to those contingencies. If, however, rats most recently experienced training on the reversed contingencies, then devalued and nondevalued levers are described according to the reversed contingencies. This practice was followed regardless of the temporal delay between training and test.

Experiment 1: Effect of extended reversal training after PF–pDMS disconnection

The design of Experiment 1 is presented in Figure 1A. In Experiment 1, rats were first trained to press two levers (left and right) for two distinct outcomes (pellets and sucrose, counterbalanced). Rats were then tested for their performance in an outcome devaluation test (see below). After testing, all animals were trained on the reversed action–outcome contingencies for 4 d. For example, if the left lever previously earned pellets it now earned sucrose and if the right lever previously earned sucrose, it now earned pellets (and vice versa for the counterbalanced condition). Rats were then given a second outcome devaluation test followed by an outcome-selective reinstatement test. Rats were then trained for a further 5 d on the reversed contingencies and tested again for devaluation and reinstatement performance.

Subjects

Twenty-six experimentally naive male outbred hooded Wistar rats (350–450 g) served as subjects. The housing conditions were as described above.

Behavioral procedures

The behavioral procedures closely replicated those of Bradfield et al. (2013) except for some additional training and testing.

Magazine training. After 3 d of food deprivation, rats were given 2 sessions of magazine training for ~20 min. In each session of each experiment, the house light was turned on at the start of each session and turned off when the session was terminated. No levers were extended during magazine training. Twenty pellet and 20 sucrose outcomes were delivered into the magazine, each according to independent random time 60 s schedules.

Lever training. For the following 10 d, rats were trained to lever press on random ratio schedules of reinforcement. Each session lasted for 50 min or until 40 outcomes were earned (maximum 20 grain pellets and 20 sucrose deliveries within the session). This consisted of two 10 min sessions on each lever (i.e., 20 min on left lever and 20 min on right lever in total or until 20 of each outcome was earned) separated by a 2.5 min timeout period in which the levers were retracted and the house light was turned off. The order of presentation of each lever was pseudorandom. For half of the animals in each group, the left lever earned pellets and the right lever earned a sucrose solution. The remaining animals were trained on the opposite action–outcome contingencies. For the first 2 d, lever pressing was continuously reinforced. Rats were shifted to a random ratio (RR5) schedule for the next 3 d (i.e., each action delivered an outcome with a probability of 0.2), then to an RR10 schedule (or a probability of 0.1) for 3 d, then to an RR20 schedule (or a probability of 0.05) for the final 2 d.

Devaluation extinction tests. After the final day of RR20 training, rats were given access *ad libitum* to either the pellets (25 g placed in a bowl in the devaluation cage) or the sucrose solution (100 ml in a drinking bottle fixed to the top of the devaluation cage) for 1 h. The aim of this prefeeding procedure was to satiate the animal specifically on the prefed outcome, thereby reducing its value relative to the nonprefed outcome (cf. Balleine and Dickinson, 1998). Rats were then placed in the operant chamber for a 10 min choice extinction test. During this test, both levers were extended and lever presses recorded, but no outcomes were delivered. The next day, a second devaluation test was administered with the opposite outcome. That is, if rats were previously prefed pellets, they now received sucrose and if rats were previously prefed sucrose, they now received pellets. Rats were then placed back into the operant chambers for a second 10 min choice extinction test.

Contingency reversal training. Subsequent to devaluation testing, rats were trained to lever press on an RR20 schedule with previously trained contingencies reversed. That is, the lever that previously earned pellets now earned sucrose and the lever that previously earned sucrose now earned pellets. All other procedures were unchanged. Contingency reversal training continued for 4 d.

Devaluation extinction tests. Devaluation extinction tests were conducted exactly as described above.

Reinstatement testing. Subsequent to the 2 devaluation extinction tests, rats were retrained on the reversed contingencies on an RR20 schedule

for 1 d. The next day, an outcome-selective instrumental reinstatement test was conducted. The test session began with a 13 min period of extinction to lower the rats' rate of responding on both levers. They then received 4 reinstatement trials separated by 7 min each. Each reinstatement trial consisted of a single delivery of either the sucrose solution or the grain pellet. All rats received the same trial order: sucrose, pellet, pellet, sucrose. Responding was measured during the 2 min periods immediately before (pre) and after (post) each delivery.

Extended contingency reversal training and testing. Subsequent to reinstatement testing, rats were trained for a further 5 d on the reversed contingencies and then tested for devaluation and reinstatement performance exactly as described above.

Experiment 2A: Effect of increasing the reversal to test interval after PF–pDMS disconnection

The design of Experiment 2A is presented in Figure 2A. Rats were trained and tested on the original contingencies in the manner described for Experiment 1. Rats then received 4 d of training on the reversed contingencies, followed by 10 d in their home cages (remaining on food deprivation) before being tested for outcome devaluation performance. Rats received a further test at 3 weeks after reversal.

Subjects

Twenty experimentally naive male outbred hooded Wistar rats (350–450 g) served as subjects.

Behavioral procedures

Magazine training, lever training, devaluation extinction tests, and contingency reversal training were conducted as described above. Rats were then kept on food deprivation in their home cages for a further 10 d without additional training. They were then given the outcome devaluation and outcome-selective reinstatement tests according to the procedures described above. Rats were then kept on food deprivation in their home cages for a further 7 d without additional training and were then given a further outcome devaluation test exactly 3 weeks after contingency reversal training in the manner described above.

Experiment 2B: Effect of testing at 3 weeks after reversal without prior testing

The design of Experiment 2B is presented in Figure 3A. This experiment was conducted identically to Experiment 2A except that the rats were not tested for their knowledge of the original contingencies so as not to extinguish responding and were tested for their knowledge of the reversed contingencies 3 weeks (rather than 10 d) after reversal training. Rats were returned to the home cage for a further 2 weeks and then given a further outcome devaluation test.

Subjects

Twenty experimentally naive male outbred hooded Wistar rats (350–450 g) served as subjects.

Behavioral procedures

Magazine training and lever training were conducted as described above. However, rats in this experiment were not given an outcome devaluation test after the first stage of training. Rats were then trained on the reversed contingencies for 4 d in the manner described above. They were then kept on food deprivation in their home cages for 3 weeks without any additional training after which they were given an outcome devaluation test in the manner described above. Rats were not tested for reinstatement so as to minimize extinction learning, but were kept in their home cages on food deprivation for a further 2 weeks without additional training, after which they were given a second devaluation test as described above.

Experiment 3: Effect of altering the physical context during reversal on tests in rats with PF–pDMS disconnection

The design of Experiment 3 is presented in Figure 4A. Similar to each of the prior experiments, rats in Experiment 3 were initially trained on two distinct action–outcome contingencies except that this training took place in one of two contexts (Context A) that differed along visual, tactile,

and olfactory dimensions. All rats were then trained and tested (devaluation and reinstatement) on the reversed action–outcome contingencies in the second context: Context B.

Subjects

Sixteen experimentally naive male outbred hood Wistar rats (350 g–450 g) served as subjects.

Behavioral procedures

Each rat received two magazine training sessions as described above, except that one occurred in Context A and the other occurred in Context B to familiarize each rat with both contexts. The identity of each context (i.e., clear walls, grid floor, peppermint odor, or striped walls, smooth floor, and vanilla odor) was counterbalanced.

Lever training proceeded as described above, except that it was conducted in Context A, the identity of which was counterbalanced.

Contingency reversal training, devaluation extinction testing, and reinstatement testing proceeded as described previously, except that it was conducted in Context B (identity counterbalanced).

Results

Experiment 1: Effect of extended reversal training after PF–pDMS disconnection

The initial results of this experiment replicated our prior findings (Bradfield et al., 2013): that functionally disconnecting the PF–pDMS pathway by way of contralateral cytotoxic lesions does not impair outcome devaluation performance after the initial acquisition of action–outcome contingencies, but does impair devaluation and outcome-selective reinstatement performance after reversal of those contingencies. However, rats were next subjected to 5 additional days of training on the reversed contingencies and tested again. After additional training, these rats were able to learn and express the reversed contingencies on both devaluation and reinstatement tests.

Histology

Figure 1, B and C, shows a representation of each overlapping placement of excitotoxic pDMS and PF lesions, respectively. Lesions of both structures produced substantial cell loss and shrinkage in their respective regions. Six rats were excluded because of incorrect lesion placement or size, yielding the following group sizes: Group IPSI, $n = 10$ and Group CONTRA, $n = 10$.

Behavioral results

Mean (\pm SEM) response rates for each group during acquisition of the initial contingencies are shown in Figure 1D averaged across levers. From this figure, it is clear that rats in both groups linearly increased their lever-pressing response over days and that group assignment did not affect responding during acquisition. Statistically, all rats linearly increased responding across days ($F_{(1,18)} = 152.445, p = 0$) and this did not interact with group ($F_{(1,18)} = 2.67, p = 0.12$). Figure 1E shows responding during the outcome devaluation test of the original contingencies and both Group IPSI and Group CONTRA demonstrated a devaluation effect (nondevalued > devalued). This is supported by statistical analysis: there was a main effect of devaluation ($F_{(1,18)} = 14.511, p = 0.001$) and no group \times devaluation interaction ($F < 1$).

Mean (\pm SEM) response rates during training on the reversed action–outcome contingencies are shown in Figure 1F. Again, it is clear that both groups linearly increased responding over days: statistically, there was a linear main effect ($F_{(1,18)} = 45.775, p = 0$) and this did not interact with group ($F < 1$). However, during the outcome devaluation test of the reversed contingencies (Fig. 1G), only Group IPSI showed a devaluation effect (nondevalued > devalued), whereas Group CONTRA responded equally on both levers (nondevalued = devalued). This is supported by the statis-

tical analysis: there was a group \times devaluation interaction ($F_{(1,18)} = 4.661, p = 0.045$). Follow-up simple-effects analysis revealed the nature of this interaction: Group IPSI showed a devaluation effect ($F_{(1,18)} = 9.4, p = 0.007$), whereas Group CONTRA did not ($F < 1$). Similarly, during the outcome-selective reinstatement test (Fig. 1H), although both groups demonstrated an elevation of responding after outcome delivery, this was only selective for the reinstated lever (reinstated > nonreinstated) in Group IPSI, whereas Group CONTRA reinstated responding similarly on both levers (reinstated = nonreinstated). Indeed, statistical analysis showed that there was a main effect of preoutcome versus postoutcome delivery responding (post > pre; $F_{(1,18)} = 19.709, p = 0$) and this did not interact with group ($F < 1$). However, there was a group \times reinstatement interaction ($F_{(1,18)} = 11.501, p = 0.003$) and follow-up simple-effects analyses showed a significant simple effect (reinstated > nonreinstated lever) for Group IPSI ($F_{(1,18)} = 37.712, p = 0$), but not for Group CONTRA (reinstated = nonreinstated; $F_{(1,18)} = 1.809, p = 0.195$).

Together with the test of the original contingencies, these findings replicate our prior findings (Bradfield et al., 2013) showing that, although animals with a functional disconnection of the PF–pDMS pathway were able to acquire and express initial goal-directed learning, they were unable to alter this learning when those contingencies were reversed. We and others (Schoenbaum et al., 2013) have argued that these results might indicate that the PF–pDMS pathway might be essential for partitioning conflicting contingencies into two separate states, preventing interference between them. One prediction that can be derived from this account is that, with additional training, even animals with PF–pDMS disconnection should eventually be able to learn and express the reversed contingencies as measured by the outcome devaluation and reinstatement tests.

Mean (\pm SEM) response rates during the 5 d of extended training on the reversed action–outcome contingencies are shown in Figure 1I. Both groups continued to linearly increase responding over days of extended training: there was a linear main effect ($F_{(1,18)} = 64.638, p = 0$) and this did not interact with group ($F < 1$). Remarkably, however, the rats in Group CONTRA now demonstrated intact outcome devaluation similar to that demonstrated by control rats in Group IPSI (Fig. 1J, nondevalued > devalued). This is supported by statistical analysis: there was a main effect of devaluation (nondevalued > devalued; $F_{(1,18)} = 24.052, p = 0$) and no group \times devaluation interaction ($F < 1$). Similarly, the previously observed deficit in outcome-selective reinstatement after 4 d of reversal training was not observed after extended training because there was a main effect of reinstatement (reinstated > nonreinstated lever; Fig. 1K; $F_{(1,18)} = 38.76, p = 0$) and no group \times reinstatement interaction ($F_{(1,18)} = 1.054, p = 0.318$). These findings show that, in support of a “state” account of the function of the PF–pDMS pathway, animals with a functional PF–pDMS disconnection can overcome their initial inability to learn and express reversed contingencies with extended training.

Experiment 2A: Effect of increasing the reversal to test interval after PF–pDMS disconnection

Experiment 1 showed that, although animals with a functional PF–pDMS disconnection demonstrate a deficit in outcome devaluation and reinstatement performance after reversal, this can be overcome with additional training. Nevertheless, conducting the additional training meant that the subsequent testing also differed from the tests conducted immediately after reversal with regard to temporal

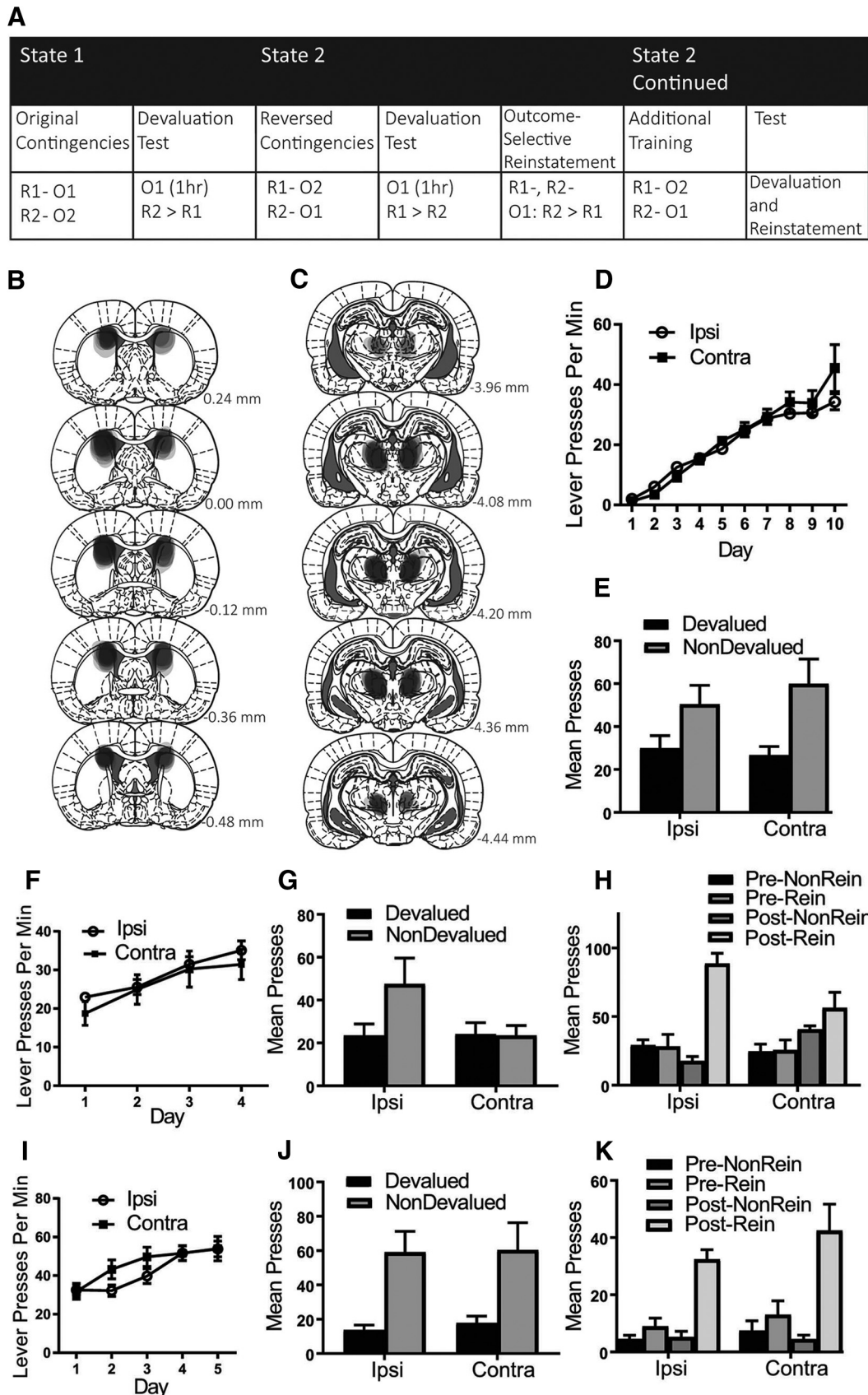


Figure 1. Experiment 1: Extended training rescues the reversal deficit in animals with a dysfunctional PF–pDMS connection. **A**, Rats were first trained to press two levers for two separate outcomes under the initial state, State 1, and then tested in devaluation. During the proposed switch to State 2, animals were trained on the reverse of these contingencies and then tested again in devaluation and reinstatement. Rats then received additional training on the reversed contingencies and were tested again. **B, C**, Representation of lesion placements showing each overlapping cytotoxic lesion for pDMS (**B**) and PF (**C**). **D**, Mean lever presses per minute (± 1 SEM) during acquisition of the lever press response. **E**, Mean total (*Figure legend continues*.)

context. That is, although the initial postreversal tests were conducted immediately after reversal training, the additional tests were conducted 10 d after this training (i.e., after 1 d of rest, 5 d of additional training, 2 d of devaluation testing, 1 d of retraining, and 1 d of reinstatement testing). Therefore, it is possible that the increased training-to-test interval, rather than the additional training itself, accounted for the recovery.

Experiment 2A tested for this possibility by training rats on the original and reversed contingencies as described above. Rats then remained in their home cages for a further 10 d without any additional training and were then given the outcome devaluation and reinstatement tests. We expected that animals in Group CONTRA would show an impairment in devaluation despite the increased training-to-test interval and that animals in Group IPSI might also show a somewhat attenuated devaluation effect on this delayed test due to enhanced interference (i.e., by increasing ambiguity and removing recency as a cue to which contingency might be currently active).

Finally, we wished to determine whether rats could return to responding in accordance with the original contingencies after learning the reversed contingencies (i.e., devalued > nondevalued). If so, this would provide evidence that animals retain these contingencies despite reversal, suggesting that the original/reversed contingencies are parsed into separate states. Some Pavlovian studies (Bouton and Peck, 1992) have suggested that animals will return to their original responding when tested at intervals longer than 10 d; therefore, in the current study, we tested after a 3 week (21 d) interval. We expected that rats with an intact PF–pDMS pathway (Group IPSI) would respond in accordance with the original contingencies (devalued > nondevalued). We speculated that animals with a disconnected pathway (Group CONTRA) might again demonstrate interference (devalued = nondevalued) or, like Group IPSI, might also respond in accordance with the original contingencies because they have consistently demonstrated intact learning of the original contingencies here (Experiment 1) and elsewhere (Bradfield et al., 2013).

Histology

Figure 2, B and C, shows a representation of each overlapping placement of excitotoxic pDMS and PF lesions, respectively. Five rats were excluded because of incorrect lesion placement or size, yielding the following group sizes: Group IPSI, $n = 8$ and Group CONTRA, $n = 7$.

Behavioral results

As before, rats in both the IPSI and CONTRA groups acquired lever pressing during training of the initial contingencies and subsequently demonstrated intact outcome devaluation (nondevalued > devalued) as shown in Figure 2, D and E, respectively (mean \pm SEM response rates). Statistically, during acquisition of lever pressing, there was a linear main effect across days ($F_{(1,13)} =$

151.691, $p = 0$) and this did not interact with group ($F_{(1,13)} = 1.078$, $p = 0.318$). During outcome devaluation testing, there was a main effect of devaluation ($F_{(1,13)} = 14.931$, $p = 0.002$) and no group \times devaluation interaction ($F < 1$).

During acquisition of the reversed contingencies, both groups again linearly increased responding and did not differ from each other in this acquisition (mean \pm SEM response rates shown in Fig. 2F). Statistically, there was a linear main effect over days ($F_{(1,13)} = 34.053$, $p = 0$) and this did not interact with group ($F < 1$). The pattern of responding during the outcome devaluation test of the reversed contingencies (mean \pm SEM response rates shown in Fig. 2G) was similar to that observed previously: whereas animals in Group IPSI showed intact devaluation (nondevalued > devalued), animals in Group CONTRA did not (devalued = nondevalued). This is supported by statistical analysis because there was a group \times devaluation interaction ($F_{(1,13)} = 6.763$, $p = 0.022$). Follow-up simple-effects analysis revealed the nature of this interaction: Group IPSI showed a devaluation effect ($F_{(1,13)} = 11.337$, $p = 0.005$), whereas Group CONTRA did not ($F < 1$). It should be noted that, during this test, overall responding was low in both groups. We suggest that this is likely the result of confusion stemming from enhanced ambiguity as a consequence of the increased training-test interval. Without the recency of the reversed contingencies to cue responding, animals may have confused the original, the conflicting, and the extinction contingencies, the latter being learned in the original devaluation tests. Alternatively, it is possible that instrumental responses are simply not subject to spontaneous recovery. Whatever the cause, despite low rates of responding, it is clear that animals in Group CONTRA did not learn or express the reversed contingencies simply as a result of the increased training-to-test interval, suggesting that the ability to do so in the extended test in Experiment 1 was indeed a result of the additional training.

Mean \pm SEM response rates during the 3 week postreversal outcome devaluation test are shown in Figure 2I. As before, responding during this test was low; however, it is clear from this figure that rats in both Group IPSI and Group CONTRA responded in accordance with the original contingencies (i.e., devalued > nondevalued, as represented according to the reversed contingencies). This is supported by statistical analysis, because there was a main effect of higher responding on the devalued lever relative to the nondevalued lever ($F_{(1,13)} = 4.992$, $p = 0.044$) and this did not interact with group ($F < 1$). It is worth noting that this pattern of responding was particularly evident in the first 5 min of devaluation Test 1 (Fig. 2H), before it was reduced by extinction. Indeed, in those first 5 min, statistically there was a main effect of devaluation (nondevalued > devalued; $F_{(1,13)} = 8.427$, $p = 0.012$) and this did not interact with group ($F < 1$). These results suggest that, in support of a state modulation account, rats in both Group IPSI and Group CONTRA retain the original contingencies despite having learned (Group IPSI) or suffered interference (Group CONTRA) from the reversed contingencies (as expressed 10 d after reversal).

Experiment 2B: Effect of testing at 3 weeks after reversal without prior testing

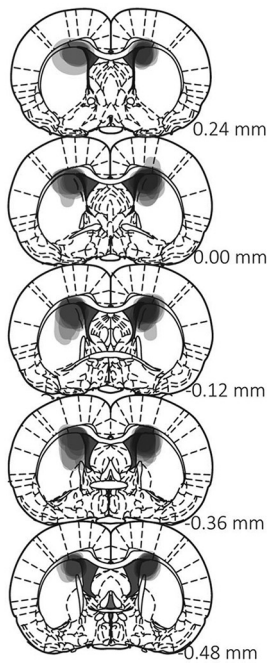
Experiment 2A demonstrated that altering the temporal context of devaluation testing by increasing the training-test interval by 10 d after reversal was not sufficient to recover performance in animals with PF–pDMS disconnection. Nevertheless, when the same rats were tested 3 weeks after reversal training, both groups responded in accordance with the original contingencies (devalued > nondevalued).

(Figure legend continued.) presses (± 1 SEM) during devaluation testing of the original contingencies. **F**, Mean lever presses per minute (± 1 SEM) during acquisition of the reversed contingencies. **G**, Mean total presses (± 1 SEM) during devaluation testing of the reversed contingencies. **H**, Mean total presses (± 1 SEM) during outcome-selective reinstatement testing of the reversed contingencies. **I**, Mean lever presses per minute (± 1 SEM) during additional training of the reversed contingencies. **J**, Mean total presses (± 1 SEM) during devaluation testing of the reversed contingencies after additional training. **K**, Mean total presses (± 1 SEM) during reinstatement testing of the reversed contingencies after additional training. R1, Response 1; R2, response 2; O1, outcome 1; O2, outcome 2. Group IPSI, $n = 10$ and Group CONTRA, $n = 10$.

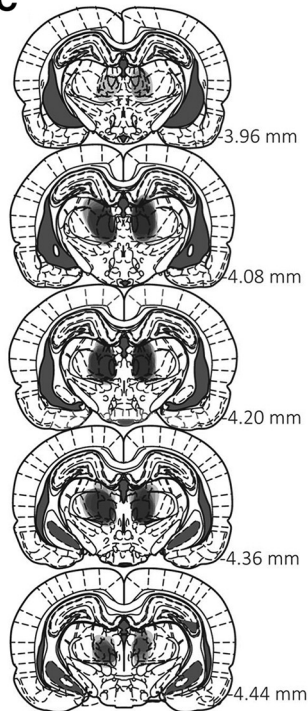
A

State 1		State 2		Return to State 1?		
Original Contingencies	Devaluation Test	Reversed Contingencies	Home Cage	Devaluation Test	Home Cage	Test
R1- O1 R2- O2	O1 (1hr) R2 > R1	R1- O2 R2- O1	10 days	O1 (1hr) R1 > R2	7 days	Devaluation R2 > R1 ?

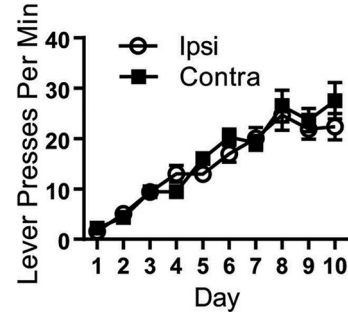
B



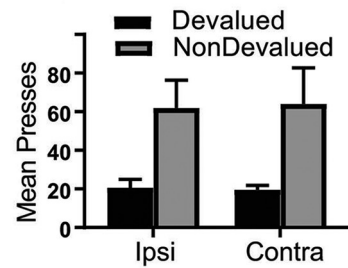
C



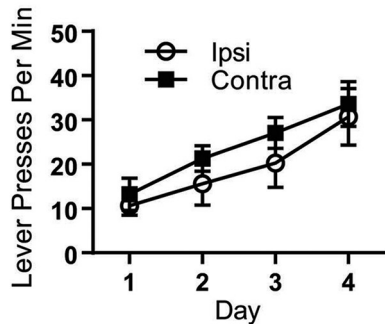
D



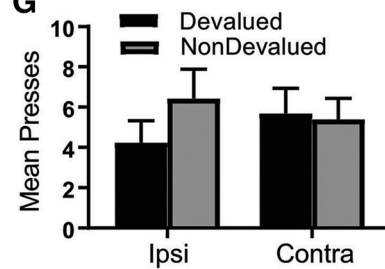
E



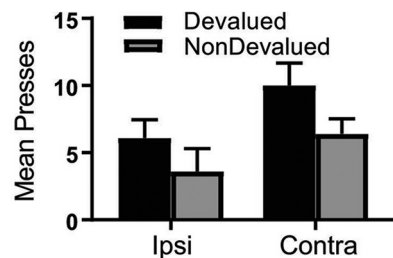
F



G



H



I

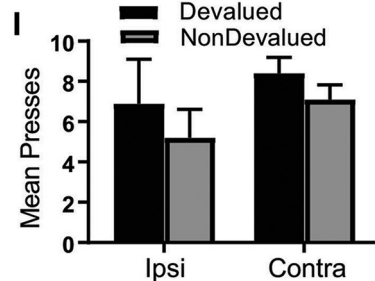


Figure 2. Experiment 2A: Increasing the training–test interval does not, but extinction does, encourage a return to responding in accordance with original contingencies rats with and without an intact PF–pDMS connection. **A**, Rats were trained to press two levers for two separate outcomes and then tested in devaluation. Rats were then trained on the reversed contingencies and then returned to their home cages for 10 d, followed by a devaluation test. Rats were then returned to their home cages for a further 7 d, followed by another devaluation test. **B**, **C**, Representation of lesion placements showing each overlapping cytotoxic lesion for pDMS (**B**) and PF (**C**). **D**, Mean lever presses per minute (± 1 SEM) during acquisition of the lever press response. (Figure legend continues.)

From Experiment 2A, however, it is unclear whether the return to the original contingencies was because of the 3 week interval or because prior devaluation testing at 10 d after reversal extinguished the reversed contingencies. Experiment 2B aimed to separate these possibilities. For this experiment, rats were again trained on the original followed by the reversed contingencies and then given an outcome devaluation test 3 weeks after reversal training. Unlike Experiments 1 and 2A, however, rats did not receive any testing in extinction before this 3-week test. Therefore, if it was the 3-week interval between training and test that promoted responding according to the original contingencies in Experiment 2A, then rats should again respond in this way in the current test (i.e., devalued > nondevalued in both groups). If, however, it was the prior extinction of the reversed contingencies that was crucial for this effect, then intact animals should respond according to the reversed contingencies (nondevalued > devalued) and PF–pDMS disconnected animals should respond indiscriminately (nondevalued = devalued, as is typically observed). If this latter result were obtained, then, upon further testing, all rats might be expected to respond in accordance with the original contingencies (i.e., devalued > nondevalued in both groups), as we observed during the second round of testing after reversal in Experiment 2A.

Histology

Figure 3, *B* and *C*, shows a representation of overlapping placement of excitotoxic pDMS and PF lesions, respectively. Five rats were excluded because of incorrect lesion placement or size and one rat became ill after surgery and was killed, yielding the following group sizes: Group IPSI, $n = 7$ and Group CONTRA, $n = 7$.

Behavioral results

As before, rats in both the IPSI and CONTRA groups acquired lever pressing during training of the initial contingencies and did not differ from each other, as shown in Figure 3*D* (mean \pm SEM response rates). Statistically, during acquisition of lever pressing, there was a linear main effect across days ($F_{(1,12)} = 42.713, p = 0$) and this did not interact with group ($F < 1$).

During acquisition of the reversed contingencies, both groups again linearly increased responding and did not differ from each other in this acquisition (mean \pm SEM response rates shown in Fig. 3*E*). Statistically, there was a linear main effect over days ($F_{(1,12)} = 20.113, p = 0.001$) and this did not interact with group ($F < 1$).

Mean \pm SEM response rates during outcome devaluation testing conducted 3 weeks after reversal training are shown in Figure 3*F*. It is clear from this figure that rats demonstrated the typical pattern of results after reversal: animals in Group IPSI showed an intact devaluation effect in accordance with the reversed contingencies (nondevalued > devalued) and animals in Group CONTRA did not (devalued = nondevalued). Statistically, there was a group \times devaluation interaction ($F_{(1,12)} = 7.461, p = 0.018$). Follow-up simple-effects analysis revealed that Group IPSI responded selectively on the nondevalued relative to

the devalued lever ($F_{(1,12)} = 9.564, p = 0.009$), whereas Group CONTRA responded equally on both ($F < 1$).

This result suggests that, in Experiment 2A, it was prior extinction of the reversed contingencies, rather than the 3-week training-to-test interval, that promoted a return to responding in accordance with the original contingencies. We therefore returned the rats in the current experiment to the home cage for a further 2 weeks before testing them again. This time, the rats in both groups did respond in accordance with the original contingencies (devalued > nondevalued, mean \pm SEM; response rates are shown in Fig. 3*G*). Statistically, there was a main effect of greater responding on the devalued lever than the nondevalued lever ($F_{(1,12)} = 5.003, p = 0.045$) and this did not interact with group ($F < 1$). Similar to the postreversal tests in Experiment 2A, responding was again low in this test. However, regardless of interval, this observation provides a replication of the result of Experiment 2A (i.e., devalued > nondevalued), suggesting that it is robust.

Experiment 3: Effect of altering the physical context during reversal on tests in rats with PF–pDMS disconnection

Together, Experiments 1, 2A, and 2B demonstrate that variations in the conditions of learning have mixed consequences for the ability of animals with a disconnected PF–pDMS pathway to learn and express reversed action–outcome contingencies. Specifically, they suggest that the PF–pDMS pathway is only recruited when there is interference between the original and reversed contingencies because, when this interference is removed (e.g., by additional training in Experiment 1 or extinction in Experiments 2A and 2B), animals in Group CONTRA act similarly to controls. For our final experiment, we investigated whether reducing interference by training the reversed contingencies in a different physical context might also restore accurate performance in animals with a disconnected PF–pDMS pathway.

For this experiment, rats were trained and tested on the original contingencies as described in Experiments 1 and 2, except that this training and testing took place in one of two contexts that differed along visual, tactile, and olfactory dimensions (i.e., Context A, the identity of which was counterbalanced). Training of the reversed contingencies took place in Context B. Rats were then given both an outcome devaluation and a reinstatement test in Context B. If the alteration in physical context during training/testing of the reversed contingencies was sufficient for animals with a disconnected PF–pDMS pathway to learn and express the reversed contingencies, then both animals in Group IPSI and Group CONTRA should demonstrate intact devaluation (nondevalued > devalued) and reinstatement (reinstated > nonreinstated) performance.

Histology

Figure 4, *B* and *C*, shows a representation of each overlapping placement of excitotoxic pDMS and PF lesions, respectively. Four rats were excluded because of incorrect lesion placement or size, yielding the following group sizes: Group IPSI, $n = 6$ and Group CONTRA, $n = 6$.

Behavioral results

Similar to previous experiments, rats in both the IPSI and CONTRA groups acquired lever pressing during training of the initial contingencies as shown in Figure 4*D* (mean \pm SEM response rates). From this figure, it appears that Group CONTRA responded slightly more than Group IPSI and, statistically, there was a significant main effect of group ($F_{(1,10)} = 6.085, p = 0.033$).

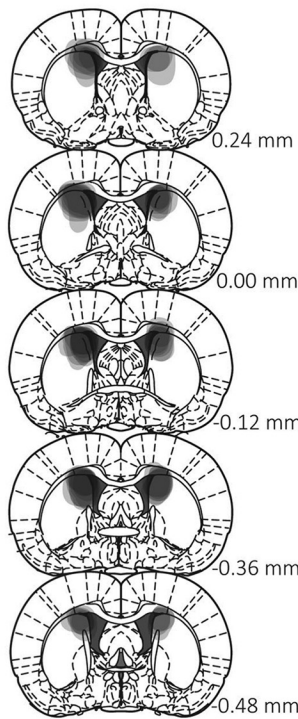
←

(Figure legend continued.) **E**, Mean total presses (± 1 SEM) during devaluation testing of the original contingencies. **F**, Mean lever presses per minute (± 1 SEM) during acquisition of the reversed contingencies. **G**, Mean total presses (± 1 SEM) during devaluation testing after 10 d in home cages. **H**, Mean total presses (± 1 SEM) during day 1 of devaluation testing after 7 additional days in the home cage. **I**, Mean total presses (± 1 SEM) averaged over both days of devaluation testing after 7 additional days in the home cage. R1, Response 1; R2, response 2; O1, outcome 1; O2, outcome 2. Group IPSI, $n = 8$ and Group CONTRA, $n = 7$.

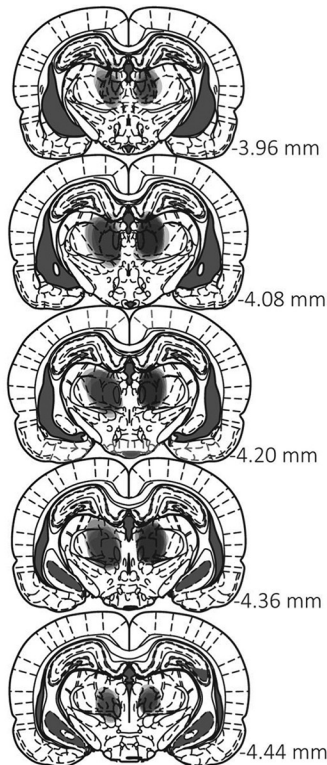
A

State 1		State 2		Return to State 1?	
Original Contingencies	Reversed Contingencies	Home Cage	Devaluation Test	Home Cage	Devaluation Test
R1- O1 R2- O2	R1- O2 R2- O1	3 weeks	O1 (1hr) R1 > R2	2 weeks	O1 (1hr) R2 >R1

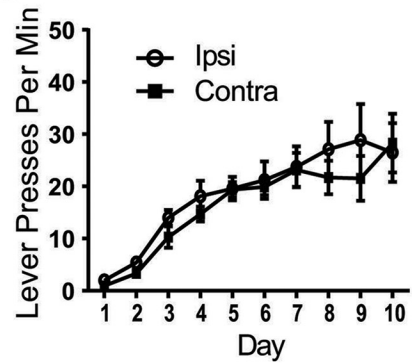
B



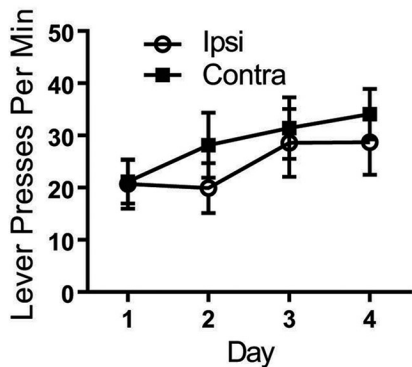
C



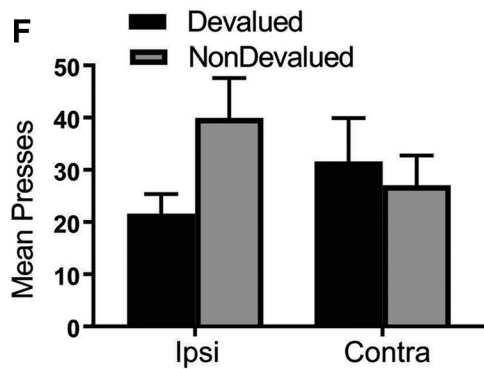
D



E



F



G

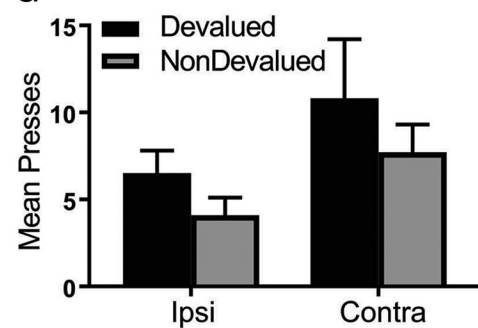
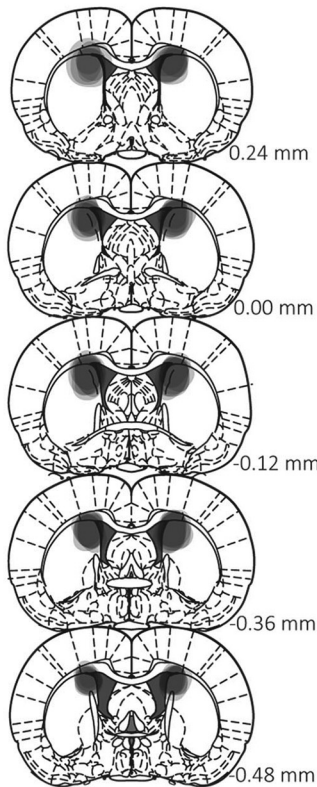


Figure 3. Experiment 2B: Extinction of reversed contingencies, rather than a specific training–test interval, is critical to return responding to the original contingencies. **A**, Rats were trained to press two levers for two separate outcomes, trained on the reversed contingencies, and then returned to their home cages for 3 weeks, followed by a devaluation test. Rats were then returned to their home cages for a further 2 weeks, followed by another devaluation test. **B, C**, Representation of lesion placements showing each overlapping cytotoxic lesion for pDMS (**B**) and PF (**C**). **D**, Mean lever presses per minute (± 1 SEM) during acquisition of the lever press response. **E**, Mean lever presses per minute (± 1 SEM) during acquisition of the reversed contingencies. **F**, Mean total presses (± 1 SEM) during devaluation testing after 3 weeks in the home cage. **G**, Mean total presses (± 1 SEM) during devaluation testing after an additional 2 weeks in the home cage. R1, Response 1; R2, response 2; O1, outcome 1; O2, outcome 2. Group IPSI, $n = 7$ and Group CONTRA, $n = 7$.

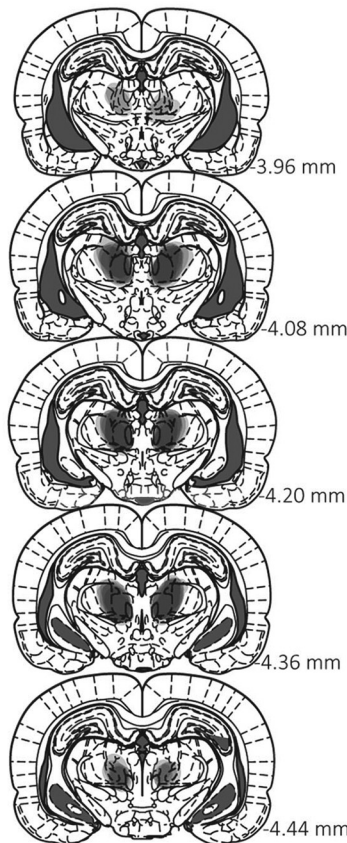
A

State 1 CONTEXT A		State 2 CONTEXT B		
Original Contingencies	Devaluation Test	Reversed Contingencies	Devaluation Test	Outcome-Selective Reinstatement
R1- O1 R2- O2	O1 (1hr) R2 > R1	R1- O2 R2- O1	O1 (1hr) R1 > R2	R1-, R2- O1: R2 > R1

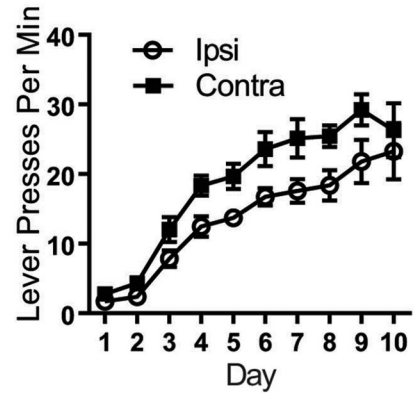
B



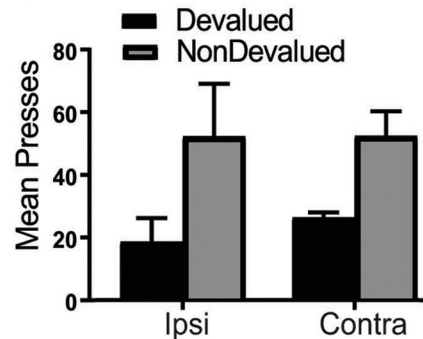
C



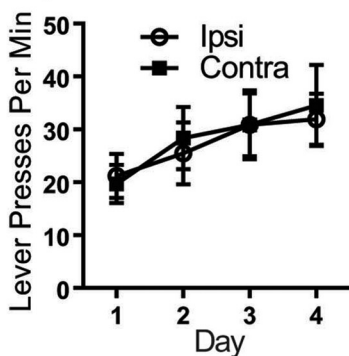
D



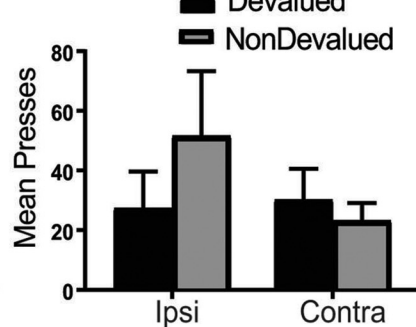
E



F



G



H

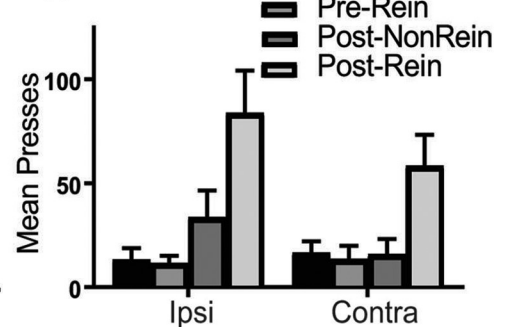


Figure 4. Experiment 3: Altering the reversal context rescues the deficit in PF–pDMS disconnected rats when tested in outcome-selective reinstatement but not devaluation. **A**, Rats were first trained to press two levers for two separate outcomes in a particular context (Context A) and then tested for devaluation performance in that same context. Rats were then trained on the reversed contingencies in a distinct context (Context B) and then tested for devaluation and reinstatement performance in that same new context. **B, C**, Representation of lesion placements showing each overlapping cytotoxic lesion for pDMS (**B**) and PF (**C**). **D**, Mean lever presses per minute (± 1 SEM) during acquisition of the lever press response. **E**, Mean total (Figure legend continues.)

Despite this difference, both groups linearly acquired lever pressing at the same rate because there was a linear main effect across days ($F_{(1,10)} = 151.515, p = 0$) and this did not interact with group ($F < 1$). Further, learning during lever press acquisition appears to have been similar between groups because rats in both groups demonstrated intact outcome devaluation performance, as shown in Figure 4E. Statistically, there was a main effect of devaluation ($F_{(1,10)} = 15.069, p = 0.003$) and this did not interact with group ($F < 1$).

During acquisition of the reversed contingencies, both groups again linearly increased responding and did not differ from each other in this acquisition (mean \pm SEM response rates shown in Fig. 4F). Statistically, there was a linear main effect over days ($F_{(1,10)} = 16.816, p = 0.002$) and this did not interact with group ($F_{(1,10)} = 2.218, p = 0.167$).

Mean \pm SEM response rates during postreversal outcome devaluation testing are shown in Figure 4G. From this figure, it is clear that, contrary to predictions, the alteration in physical context did not rescue performance in Group CONTRA. Rather, the typical pattern of results was observed: animals in Group IPSI showed an intact devaluation effect in accordance with the reversed contingencies (nondevalued > devalued) and animals in Group CONTRA did not (devalued = nondevalued). Statistically, there was a group \times devaluation interaction ($F_{(1,10)} = 5.723, p = 0.038$) and follow-up simple-effects analysis revealed that Group IPSI responded selectively on the nondevalued relative to the devalued lever ($F_{(1,10)} = 6.856, p = 0.026$), whereas Group CONTRA responded equally on both ($F < 1$). This result suggests an alteration in the physical context during reversal learning was not sufficient for Group CONTRA to learn/express the reversed contingencies, at least during outcome devaluation testing. We did, however, find that performance was restored in Group CONTRA during the outcome-selective reinstatement test (shown in Fig. 4H). Specifically, similar to rats in Group IPSI, rats in Group CONTRA also increased responding after outcome delivery and this increase was specific to the reinstated lever. This is supported by the statistical analysis: there was a main effect of responding postoutcome delivery ($F_{(1,10)} = 19.753, p = 0.001$) and this did not interact with group ($F_{(1,10)} = 2.476, p = 0.147$). Moreover, this increase was specific to the reinstated lever for both groups because there was a main effect of reinstatement ($F_{(1,10)} = 18.536, p = 0.002$) and this did not interact with group ($F < 1$).

This finding is in contrast to the reinstatement performance that we have observed both in Experiment 1 and previously (Bradfield et al., 2013). Importantly, it suggests that the alteration in physical context used here was sufficient for the animals in Group CONTRA to learn the reversed contingencies during training even though they were only able to express that learning during reinstatement testing. The implications of this finding are discussed below.

Discussion

The current results provide evidence supporting a role for the PF–pDMS pathway in controlling the internal state information used to

produce accurate goal-directed action selection. Experiment 1 replicated our previous finding (Bradfield et al., 2013) that animals with contralateral PF/pDMS lesions (Group CONTRA) can initially learn and express action–outcome contingencies, but are impaired relative to ipsilateral controls (Group IPSI) after those contingencies are altered; in this case reversed. Nevertheless, after additional reversal training, we found that the same animals demonstrated intact performance on both outcome devaluation and outcome-selective reinstatement tests. Experiment 2A demonstrated that this restoration of performance was not a consequence of the increased interval between reversal training and test; when animals were trained on reversed contingencies and then tested 10 d later without any additional training, Group CONTRA displayed an impaired devaluation effect relative to controls. When these same rats were retested 3 weeks after reversal training, both groups (IPSI and CONTRA) showed reliable devaluation in accordance with the original contingencies, confirming that the original contingencies survive reversal training consistent with state modulation of performance. Experiment 2B demonstrated that the rats' ability to respond according to the original contingencies was due to extinction of the reversed contingencies rather than testing after a 3 week interval per se. Specifically, rats in this experiment were tested in extinction for the first time 3 weeks after reversal training and yet did not retrieve the original contingencies until given a second extinction test. Our final experiment (Experiment 3) demonstrated that the disconnection-induced deficit was specific to a loss of the internal rather than external context. Rats were trained on the original contingencies in external Context A and then trained and tested on the reversed contingencies in external Context B. This alteration in context did not restore accurate performance in Group CONTRA in the outcome devaluation test, although it did restore outcome-specific reinstatement. Together, these results suggest that, consistent with a state modulation account, an intact PF–pDMS pathway acts to minimize interference between conflicting contingencies, but when that interference is reduced by other means (e.g., additional training, extinction, or alterations in context), action selection no longer relies on this pathway.

PF regulation of pDMS for state modulation

The current results provide direct evidence of state modulation by the PF–pDMS pathway. The most straightforward example is the return to the original contingencies after reversal observed in Experiments 2A and 2B. This finding is all the more remarkable when it is considered that the rats in Group IPSI completely altered their responding from initial postreversal testing in which they expressed the reversed contingencies, demonstrating that they must have simultaneously retained both sets of contingencies and that the selection of one or other was modulated by internal state. The finding that rats in Group CONTRA also reverted to responding in accordance with the original contingencies is particularly intriguing given our prior hypothesis (Bradfield et al., 2013) that new learning might permanently interfere with initial learning in these rats. One possibility, therefore, is that, rather than new learning overlaying initial learning, separate pools of neurons within the pDMS might encode the initial and reversed contingencies, as occurs in OFC (Schoenbaum et al., 2003; see also Sharpe and Schoenbaum, 2016) and PF-controlled CIN modulation determines which pool of neurons currently governs action selection. Furthermore, the pool of neurons most recently subject to plasticity (i.e., those that encoded the reversed contingencies), being the more active, might be disproportionately affected when any new learning occurs (such as extinction). In the current study, this could have led

←
(Figure legend continued.) presses (± 1 SEM) during devaluation testing of the initial contingencies. **F**, Mean lever presses per minute (± 1 SEM) during acquisition of the reversed contingencies. **G**, Mean total presses (± 1 SEM) during devaluation testing of the reversed contingencies. **H**, Mean total presses (± 1 SEM) during outcome-selective reinstatement testing of the reversed contingencies. R1, Response 1; R2, response 2; O1, outcome 1; O2, outcome 2. Group IPSI, $n = 6$ and Group CONTRA, $n = 6$.

action control to revert to the circuits that encoded the original contingencies, a hypothesis that we are currently investigating.

These findings complement and extend our previous study (Bradfield et al., 2013), based on which, we and others (Schoenbaum et al., 2013) speculated that PF-controlled CINs in the pDMS provide internal state information to modulate action selection. In that study, we presented several lines of evidence pointing to a specific role for pDMS CINs in our behavioral effects. Likewise, a more recent study demonstrated that, when recording from putative CINs in the pDMS during a task that involved responding to an unsignalled alteration in contingency, neural activity reflected internal state information (Stalnaker et al., 2016). We suggest, therefore, that the current effects of the PF–pDMS disconnection are also due to a modulation of pDMS CINs. Indeed, such effects have been replicated in mice in which the cholinergic interneurons were selectively ablated using a toxico-genetic approach (Matamales et al., 2016). It should be noted, however, that the current results do not address whether it is neurons in the PF or the pDMS CINs that create the state representations; indeed, the results of Stalnaker et al. (2016) suggest that state modulation by pDMS CINs relies on the OFC because OFC lesions abolished the state-related patterns of activity. The OFC and PF also have sparse but reciprocal connections (Reep et al., 1996; Hoover and Vertes, 2011), suggesting that a complex interplay between these structures could mediate this effect.

Internal versus external state control of goal-directed actions

Although the current results collectively speak to the role of the PF–pDMS pathway in regulating internal state modulation, they also relate to the contextual control of goal-directed actions more generally. There is a vast literature examining the role of contextual and temporal stimuli in various Pavlovian paradigms, such as extinction, latent inhibition, and counterconditioning (among others, for review, see Bouton, 1993), but a comparatively small literature looking the similar effects in instrumental conditioning. Nevertheless, Thrailkill and Bouton (2015) reported recently that contextual control is limited to habitual actions (i.e., actions controlled by a stimulus–response association), whereas the performance of goal-directed actions was found to be largely context independent. If true, this could explain why outcome devaluation but not reinstatement was impaired in Group CONTRA in Experiment 3 after the context alteration during reversal. The non-specific pattern of responding displayed by Group CONTRA during devaluation suggests that these animals suffered interference between the original and reversed contingencies, as might be expected if the original action–outcome contingencies survived the change of context. The same deficit would not be expected upon reinstatement testing, which has been shown to rely on outcome–response contingencies that are also learned during instrumental conditioning (Balleine and Ostlund, 2007). Because these contingencies constitute a form of stimulus–response association (with the outcome acting as the stimulus that cues responding), they are more likely to be context specific, with the outcome–response contingencies learned in Context A less potentially interfering with those learned in Context B. Therefore, there should be less involvement of internal state modulation and of the PF–pDMS pathway. A further implication of the results of Experiment 3, as well as several other recent findings (Im et al., 2016; Thrailkill et al., 2016), is that internal and external contexts (or states) can be at least partially dissociated by their ability to govern action selection, as predicted by certain reinforcement

learning models (Gershman et al., 2010; Gershman and Niv, 2012).

An alternative possibility is that learning the reversed contingencies in Group CONTRA was intact in all experiments and that the deficit actually occurs at the level of performance. That is, the PF–pDMS might govern states to determine the accurate performance of each action, rather than the underlying learning of contingencies. Therefore, in Experiment 3, state interference governing performance was minimized by the alteration in context, but only when all contextual stimuli were present during testing, including the outcome (which is present during reinstatement testing but not devaluation). However, our previous finding that disrupting the PF–pDMS pathway (by lesioning the PF and infusing M2/M4 agonist oxotremorine-S into the contralateral pDMS) during training, but not testing, of the reversed contingencies disrupted both devaluation and reinstatement performance suggests that, in the absence of a context alteration, disrupting this pathway disrupts the learning, not just the performance, of the reversed contingencies.

Clinical implications

The current results also have important clinical implications. Patients with Parkinson's and Parkinson's disease dementia (PDD) suffer from a characteristic deficit in cognitive flexibility, the nature of which could be analogous to the inability of rats with a disconnected PF–pDMS pathway to interlace new and existing action–outcome associations. Such patients suffer massive neuronal degeneration in the CM/PF (homologous to the PF in rats) and it has been suggested that this could underlie cognitive deficits in individuals with PDD (Smith et al., 2014). Similarly, dysfunction in striatal CINs has been associated with cognitive deficits in mice that could not be rescued by chronic treatment with L-DOPA, the conventional pharmacological treatment of Parkinson's (Tozzi et al., 2016). Furthermore, patients with PDD have displayed a consistent deficit in a set-shifting task (Richards et al., 1993), a task that is very likely to require intact state modulation. Therefore, we offer a number of behavioral methods, such as extinction and additional training, that could potentially be tested for their efficacy in overcoming deficits in cognitive inflexibility in such patients, in addition to any pharmacological treatments that might also be developed.

References

- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419. [CrossRef Medline](#)
- Balleine BW, Ostlund SB (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann N Y Acad Sci* 1104:147–171. [CrossRef Medline](#)
- Bouton ME (1993) Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychol Bull* 114:80–99. [CrossRef Medline](#)
- Bouton ME, Peck CA (1992) Spontaneous recovery in cross-motivational transfer (counterconditioning). *Learning & Behavior* 20:389–399.
- Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW (2013) The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. *Neuron* 79:153–166. [CrossRef Medline](#)
- Deschênes M, Bourassa J, Doan VD, Parent A (1996) A single-cell study of the axonal projections arising from the posterior intralaminar thalamic nuclei in the rat. *Eur J Neurosci* 8:329–343. [CrossRef Medline](#)
- Gershman SJ, Niv Y (2012) Exploring a latent cause theory of classical conditioning. *Learn Behav* 40:255–268. [CrossRef Medline](#)
- Gershman SJ, Blei DM, Niv Y (2010) Context, learning, and extinction. *Psychol Rev* 117:197–209. [CrossRef Medline](#)
- Hoover WB, Vertes RP (2011) Projections of the medial orbital and ventral orbital cortex in the rat. *J Comp Neurol* 519:3766–3801. [CrossRef Medline](#)

- Im HY, Bédard P, Song JH (2016) Long lasting attentional-context dependent visuomotor memory. *J Exp Psychol Hum Percept Perform* 42:1269–1274. [CrossRef Medline](#)
- Matamales M, Skrbis Z, Hatch RJ, Balleine BW, Götz J, Bertran-Gonzalez J (2016) Aging-related dysfunction of striatal cholinergic interneurons produces conflict in action selection. *Neuron* 90:362–373. [CrossRef Medline](#)
- Paxinos G, Watson C (2013) *The rat brain in stereotaxic coordinates*, Ed 7. San Diego: Academic.
- Reep RL, Corwin JV, King V (1996) Neuronal connections of orbital cortex in rats: topography of cortical and thalamic afferents. *Exp Brain Res* 111: 215–232. [Medline](#)
- Richards M, Cote LJ, Stern Y (1993) Executive function in Parkinson's disease: set-shifting or set-maintenance? *J Clin Exp Neuropsychol* 15:266–279. [CrossRef Medline](#)
- Romero MA, Vila JN, Rosas JM (2003) Time and context effects after discrimination reversal in human beings. *Psicologica* 24:169–184.
- Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2003) Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* 39: 855–867. [CrossRef Medline](#)
- Schoenbaum G, Stalnaker TA, Niv Y (2013) How did the chicken cross the road? With her striatal cholinergic interneurons, of course. *Neuron* 79: 3–6. [CrossRef Medline](#)
- Sharpe MJ, Schoenbaum G (2016) Back to basics: making predictions in the orbitofrontal-amygdala circuit. *Neurobiol Learn Mem* 131:201–206. [CrossRef Medline](#)
- Smith Y, Galvan A, Ellender TJ, Doig N, Villalba RM, Huerta-Ocampo I, Wichman T, Bolam JP (2014) The thalamostriatal system in normal and diseased states. *Front Syst Neurosci* 8:5. [CrossRef Medline](#)
- Stalnaker TA, Berg B, Aujla N, Schoenbaum G (2016) Cholinergic interneurons use orbitofrontal input to track beliefs about current state. *J Neurosci* 36:6242–6257. [CrossRef Medline](#)
- Thrailkill EA, Bouton ME (2015) Contextual control of instrumental actions and habits. *J Exp Psychol Anim Learn Cogn* 41:69–80. [CrossRef Medline](#)
- Thrailkill EA, Trott JM, Zerr CL, Bouton ME. Contextual control of chained instrumental behaviors. *J Exp Psychol Anim Learn Cogn* 42: 401–414.
- Tozzi A et al. (2016) Alpha-synuclein produces early behavioural alterations via striatal cholinergic synaptic dysfunction by interacting with GluN2D N-Methyl-D-Aspartate receptor subunit. *Biol Psychiatry* 79:402–414. [CrossRef Medline](#)