[1] Table S1

	AP frequency (Hz)	AP amplitude (mV)	AP duration (ms)	AP threshold (mV)	AP rise-time (ms)
Unlesioned	2.1 ± 0.3	89 ± 2	1.7 ± 0.1	-56 ± 2	0.27 ± 0.02
Lesioned	0.6 ± 0.2*	86 ± 3	1.8 ± 0.1	-55 ± 2	0.26 ± 0.01
	AHP amplitude (mV)	Ih (pA)	Inward rectification ratio	Input resistance (MOhm)	Time constant (ms)
Unlesioned	16 ± 2	342 ± 53	7 ± 2	216 ± 30	25 ± 4
Lesioned	18 ± 2	268 ± 46	9 ± 2	256 ± 41	35 ± 6

Related to Figure 2: Cellular electrophysiological properties of CINs in the pDMS from non-lesioned (n = 7) and lesioned (n = 9) cerebral hemispheres. AP frequency is spontaneous action potential firing frequency in cell-attached configuration, AP amplitude is amplitude of action potential above threshold, AP duration is width of action potential at threshold, AP rise-time is time taken between 10-90% for the rise of action potential, AHP is amplitude of after hyperpolarization measured from threshold, Ih is current measurement difference (at -120 mV hyperpolarization step) between the beginning of "sag" and the end of pulse protocol with a duration of 400 ms, inward rectification ratio (measured at the end of pulse protocol) is ratio of current amplitude at -110 to -120 mV step over the amplitude at -70 to -80 mV step, input resistance is measured from voltage change between -70 to -80 mV with a current pulse injection into the cell under current-clamp, time-constant or tau (measured from trace about -80 mV) is the time taken to charge up the cell membrane to 62%. Unlesioned and lesioned groups were statistically compared using unpaired t-test and significance was $p < 0.05$. Only the test of AP frequency was reliable: * denotes $p = 0.0002$.

[2] Figure S1

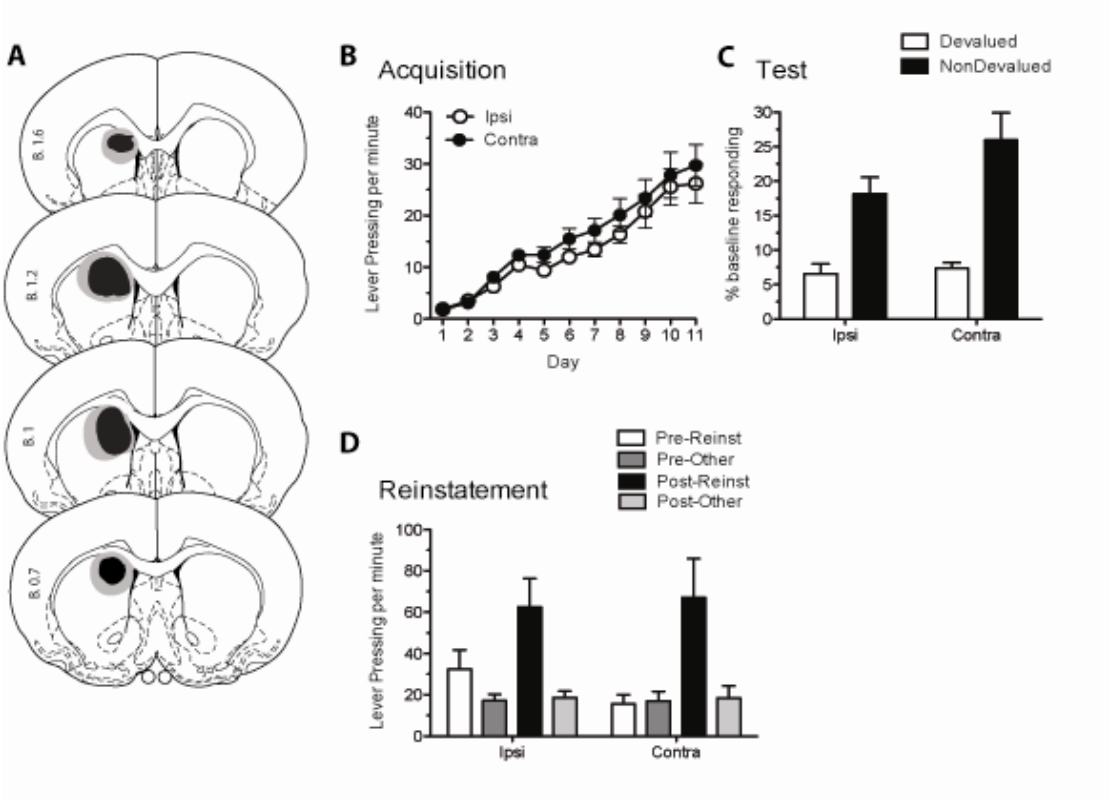

Figure S1. Related to Figure 4 J-K: The parafascicular – anterior dorsomedial striatum pathway does not regulate either goal-directed learning or changes in action-outcome contingencies.

(A) Schematic illustration of the minimal (black) and maximal (grey) unilateral lesions of anterior DMS. Lesions are represented on the left in this schematic, but were counterbalanced. (B-D) Mean or % baseline rate of lever pressing (± 1 SEM) during: (B) initial acquisition of action-outcome contingencies averaged over levers: The groups did not differ on acquisition. There was no main effect of group, $F(1, 13) = 1.02$, $p = .331$, there was a linear effect of acquisition, $F(1.13) = 74.89$, $p = .00$, and no acquisition x group interaction, $F < 1$. (C) Outcome devaluation testing: Both groups showed evidence of having learned the contingencies (i.e. nondevalued > devalued) and did not differ on test. There was a main effect of devaluation, $F(1, 13) = 35.39$, $p = .001$, but no main effect of group, $F(1, 13) = 2.96$, $p = .11$, and no group x devaluation interaction, $F(1, 13) = 1.88$, $p = .2$. (D) Reinstatement of the reversed contingencies, both groups showed evidence of having learned the reversed contingencies on reinstatement testing, and did not differ, there was a main effect of pre vs. post, $F(1, 13) = 12.25$, $p = .004$, but no group x post-outcome delivery interaction, $F < 1$, suggesting that all rats reinstated responding equally after outcome delivery, regardless of

group. Similarly, this was selective for the reinstated lever as there was an effect of lever, $F(1, 13) = 16.14$, $p = .001$, and this selective reinstatement occurred equally for both groups because there was no group x lever interaction, $F < 1$. Together with the results shown in main Figure 4J-K, these results show that the Pf-aDMS pathway is not involved in either the acquisition of initial or their integration with new action-outcome contingencies.

[3] Supplemental Text

We have argued that intact cholinergic interneuron (CIN) function in the posterior dorsomedial striatum (pDMS) is critical for the successful exploitation of new actions and the resistence to interference from prior learning. It must be considered how CINs compute this integration of new and old learning, and one possibility is that they encode a state prediction error that lead to reductions in contingency knowledge. The idea of state prediction differs from that of reward prediction. Specifically, predictions about states are based on prior experiences with state-to-state transitions that ultimately lead to the acquisition of an outcome. For example, upon placement in the conditioning chamber an animal is considered to be in its original state and pressing the lever whilst in that state causes the animal to transition to the next state when an outcome is delivered to the magazine. Within that next state the animal must enter the magazine to retrieve the outcome and, upon doing so, is transitioned to the terminal state. After reaching this state the animal returns to its original state and can then conduct a forward search based on these experiences to assess the probability that taking a particular action within a particular state will lead to reward.

State prediction errors are calculated when an animal is surprisingly transitioned into a state given their current estimate of state-to-state transition probabilities (Glascher et. al., 2010). In the example outlined above, the animal experiences a state prediction error (as well as reward prediction error) when pressing the lever for the first time surprisingly leads to the 'outcome delivered' state. Thus the finding that compromising pDMS CIN function leaves initial contingency learning intact suggests that CIN function does not affect the increases in learning that result from surprising transitions (i.e. state prediction error). Current results do, however, support the interpretation that intact CIN function is critical when competition between old and new learning creates a state prediction error resulting in a reduction in contingency knowledge. During contingency degradation, this competition is manifested when an animal that has previously learned to press a lever to transition to the ‘outcome delivered’ state is surprisingly transitioned to this state when an outcome is delivered in the absence of lever press. This surprising transition creates new learning (favouring context-outcome associations) and a concurrent reduction in the estimate that pressing the lever will cause a transition to the 'outcome delivered' state. Current results show that rats with intact pDMS CIN function reduced responding on the degraded lever, consistent with an ability to compute this state prediction error, whereas rats with

compromised CIN function appeared to be unable to compute this error as they continued to respond equally on both levers.

However, contingency degradation confounds state and reward prediction error such that this is not the only interpretation of these results. By contrast, when animals learn new action-outcome contingencies in place of existing contingencies reward prediction error is negligible but state prediction error is large. For example, when placed in conditioning chamber in this phase of the experiment, the animals' expectation in its original state is that pressing the left lever will cause a transition to the next state in which pellets are delivered (for example). When sucrose is delivered instead, the animal experiences a surprising transition that cannot be captured by reward prediction error models because the outcomes (pellets and sucrose) are similarly rewarding, The observed deficit in rats with a compromised pDMS CIN function cannot, therefore, be due to a deficit in computing reward prediction error. Rather, the animal experiences competition between learning about new contingencies (e.g. "pressing the left lever now causes transitions to the 'sucrose delivered' state") and reducing existing contingency knowledge (e.g. "pressing the left lever no longer causes transitions to the 'pellet delivered' state") that is captured by a state prediction error.

Current results show that intact animals successfully integrate both processes to learn the new contingencies as evidenced by outcome devaluation and reinstatement performance. Rats with impaired CIN function, however, show a specific deficit consistent with an inability to decrease contingency knowledge because they respond non-selectively on both levers at test (devalued = non devalued). Had these rats been unable to learn new contingencies altogether, they should have responded in accordance with prior contingencies on test (devalued > nondevalued). Thus, taken together, the behavioural impairments we observed in rats with compromised pDMS CIN function are consistent with a role in computing state prediction errors when new and existing learning compete requiring a reduction in contingency knowledge.

Mechanistically, the state prediction error may be computed in parafascicular thalamus (Matsumoto et al, 2004; Kimura,et al, 2004) and, in contrast to reward prediction errors, may not result in distinct patterns of firing but rather a distinct pattern of circuit activity resulting in an altered pattern of CIN modulation of plasticity at distinct MSNs or their dendrites. Although speculative, this interpretation provides a unique but critical modulatory function to CINs and warrants further investigation.

[4] Supplemental Methods.

Experimental Procedures.

Subjects

For the behavioral studies, male Long-Evans rats, weighing between 300 – 380g at the beginning of the experiment, were used as subjects. For electrophysiology experiments male Long-Evans rats between 5 and 6 weeks old were used, weighing between 120-150 g. Adult rats were housed in pairs and young rats in groups of four in transparent yellow-tinted plastic tubs (.5m^3) located in a temperature- and humidity- controlled vivarium. Rats that experienced behavioral training and testing were maintained at ~ 85% of their free-feeding body weight by restricting their food intake to between 8 and 12g of their maintenance diet per day. All procedures were approved by the University of Sydney Ethic Committee.

Surgery

Approximately 10 min before surgery rats were injected intraperitoneally (i.p.) with 1.3 ml/kg of the anaesthetic ketamine (CenVet Australia Pty Ltd, Sydney, Australia) at a concentration of 100 mg/mL and with the muscle relaxant xylazine (0.3mL/kg; Rompun; Bayer, Sydney, Australia) at a concentration of 20 mg/mL. Each rat was then placed in a stereotaxic frame (Stoelting, Wood Dale, IL, USA). 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.

For retrograde tracing a small hole was then drilled into the skull above the pDMS (all coordinates in millimetres relative to bregma: anteroposterior, -0.4, mediolateral, $\pm$ 2.2, dorsoventral, -4.5) in either the left or right hemisphere. FG was injected into the pDMS via a glass micropipette attached to the end of a nanoject (Drummond Scientific company, Broomali, PA, USA). Three 52nL quantities of FG were injected at 30 s intervals. The micropipette was left in place for a further 4 min to allow the FG to diffuse. For the electrophysiology experiment rats received a prior unilateral Pf lesion (in the right or left hemisphere, counterbalanced). A small hole was drilled above the PF. Due to the young age and therefore small size of the rats in this experiment only, excitotoxic PF lesions were made by infusing 0.2μl of NMDA (10mg/mL) in sterilized 0.1M phosphate buffered saline (PBS) pH 7.2 over 2 min. The relevant coordinates, in mm relative to bregma, are anteroposterior, -

4.1, mediolateral, ± 1.2, dorsoventral, -5.8). The needle was left in place for 4 min prior to removal to allow for diffusion. For bilateral Pf lesions the same procedures were used except that 0.4μl of NMDA was infused into the Pf bilaterally, and the relevant co-ordinates were, in mm relative to bregma, are anteroposterior, -4.2, mediolateral, ± 1.3, dorsoventral, -6). Sham-operated rats underwent the same procedures but no neurotoxin was infused. For the Pf-pDMS disconnection experiment rats received sham or excitotoxic, ipsilateral or contralateral lesions of PF and pDMS. Excitotoxic lesions were made by infusing 0.6μl (pDMS lesions: anteroposterior, -0.4, mediolateral, ± 2.2, dorsoventral, -4.5) or 0.4μl (PF lesions: anteroposterior, 4.2, mediolateral, ± 1.3, dorsoventral, -6) of NMDA (10mg/mL) in sterilized 0.1M phosphate buffered saline (PBS) pH 7.2 over 6 min (DMS lesions) or 4 min (PF lesions). For the oxotremorine experiment rats received a unilateral Pf lesion (as described), then a 26 gauge guide cannula (Plastics One) was then implanted into a small hole drilled above the pDMS. The tip of the guide cannula was aimed at the pDMS (coordinates relative to bregma: anteroposterior, -0.4, mediolateral, ± 2.2, dorsoventral, -3.5). The guide cannulas were maintained in position with dental cement, and dummy cannulas were kept in each guide at all times except during microinfusions. For all groups in all experiments, the amount of damage that occurred in each hemisphere was counterbalanced and needle insertions in the Sham group were similarly counterbalanced (i.e. half ipsilateral, half contralateral, and half of each in each hemisphere).

Retrograde Tracing

Four days after surgery, the two male Long-Evans rats were deeply anesthetized with sodium pentobarbital (100mg/kg i.p.) and perfused transcardially with 400mL of 4% paraformaldehyde in 0.1M phosphate buffer (PB), pH 7.4. Prior to perfusion a .25mL injection of a 1% sodium nitrite and heparin (5000 i.u./mL) solution (in the ratio 2:1, NaNitrite: heparin) was injected directly into the heart. Brains were postfixed for 1 h in the same fixative and placed in PBS (pH 7.2) 20% sucrose solution overnight. Free-floating 30μM sections were obtained in a -20 °C cryostat (Leica, etc) and stored in 0.1% sodium azide in 0.1 M PBS pH 7.2. Sections were washed for 3 x 10 min in 0.1 M PBS (pH 7.2), then mounted and cover slipped with fluoromount (Sigma Aldrich, St. Louis, MO, USA). The locations of the

injections sites and the retrogradely labelled neurons in the thalamus were imaged using an Olympus B61W1 microscope.

Brain slice preparation

Male Long-Evan rats (6-7 weeks old) were killed under deep anaesthesia by isoflurane inhalation (4% in air), decapitated and the brain removed. Coronal brain slices (300 µm thick) containing the posterior DMS were cut using a vibratome (Leica Microsystems VT1200S, Germany) in ice-cold oxygenated sucrose cutting solution containing (in mM): 241 sucrose, 28 $NaHCO_3$, 11 glucose, 1.4 NaH_2PO_4, 3.3 KCl, 0.2 $CaCl_2$, 7 $MgCl_2$. Slices were hemisected at midline and maintained at 33ºC in a submerged chamber containing physiological saline with composition (in mM): 126 NaCl, 2.5 KCl, 1.4 NaH_2PO_4, 1.2 $MgCl_2$, 2.4 $CaCl_2$, 11 glucose and 25 $NaHCO_3$, and equilibrated with 95% O_2 and 5% CO_2.

Electrophysiological recording

After equilibrating slices for 1 h, slices were then transferred to a recording chamber and visualised under an upright microscope (Olympus BX50WI) using differential interference contrast Dodt tube optics, and superfused continuously (1.5 ml min^{-1}) with oxygenated physiological saline at 33ºC. Cell-attached and whole-cell patch-clamp recordings were made using electrodes (2–5 MΩ) containing internal solution (in mM): 115 K gluconate, 20 NaCl, 1 $MgCl_2$, 10 HEPES, 11 EGTA, 5 Mg-ATP, and 0.33 Na-GTP, pH 7.3, osmolarity 285-290 mOsm l^{-1}. Biocytin (0.1%, Sigma-Aldrich, St. Louis, MO) was routinely added to the internal solution for marking the sampled neurons during whole-cell recording. Data acquisition was performed with an Axopatch 200A amplifier (Molecular Devices, Sunnyvale, CA), connected to a Macintosh computer and interface ITC-16 (Instrutech, Long Island, NY). In cell-attached mode, action potentials were sampled at 10 kHz (low pass filter 5 kHz) and whole-cell currents were sampled at 5 kHz (low pass filter 2 kHz, Axograph X, Molecular Devices, Sunnyvale, CA). Whole-cell recordings were established immediately following data collection in cell-attached mode. Data from cell-attached and whole cell recordings were only included in analyses if (1) the neurons appeared healthy under DIC on monitor screen, (2) cholinergic interneurons were spontaneously active during cell-attached recording, (3) action potential amplitudes were at least 70 mV after establishing whole-cell recording mode, and (4) neurons demonstrated physiological characteristics of cholinergic

interneurons such as presence of hyperpolarization-activated cation current Ih but no plateau low-threshold spiking (Kawaguchi et al., 1995), to ensure that only highly viable neurons were included.

For biocytin-filled neurons, immediately after whole-cell physiological recording, brain slices were fixed overnight in 4% paraformaldehyde/0.16 M phosphate buffer (PB) solution followed by placing them in 0.3% triton X-100/PB for 3 days to permeabilize cell membrane. Slices were rinsed in PB and then incubated in Cy3-conjugated ExtrAvidin/PB solution (1:500, Sigma-Aldrich, St. Louis, MO) for 2 h. Stained slices were rinsed, mounted onto glass slides, dried and coverslipped with Vectashield mounting medium (Vector Laboratories, Burlingame, CA).

Drugs

Working drug concentrations for incubations were: oxotremorine sesquifumarate (Oxo-S; 1 µM, Sigma-Aldrich, St. Louis, MO), scopolamine (Scop, 3 µM, Sigma-Aldrich, St. Louis, MO) and tetrodotoxin (TTX, 100 nM, Ascent Scientific, Bristol, UK). A cocktail of synaptic blockers was applied, combining picrotoxin (100 µM, Sigma-Aldrich, St. Louis, MO), CNQX disodium and DL-AP5 (10 µM and 100 µM, respectively, Ascent Scientific, Bristol, UK). During cell-attached and whole-cell recordings experiments, stock solutions of all drugs were diluted to working concentrations in the extracellular solution immediately before use and applied by continuous superfusion.

Tissue preparation and immunofluorescence

Rats were rapidly anaesthetized with sodium pentobarbital (300 mg/kg i.p., Virbac Pty. Ltd., Australia) and transcardially perfused with 400mL of 4% paraformaldehyde in 0.1 M sodium phosphate buffer (pH 7.4). Brains collected for lesion placement only were post-fixed for 1 h in the same fixative and placed in PBS pH 7.2 20% sucrose solution overnight. 40 µm coronal sections were cut using a cryostat. Every fourth section was collected on a slide and stained with cresyl violet. Slides were examined for placement and extent of the lesion; the latter was assessed by microscopically examining sections for areas of marked cell loss as well as general shrinkage of a region relative to controls. Brains collected for further immunofluoresent analysis were post-fixed overnight in 4% paraformaldehyde in 0.1 M sodium phosphate buffer (pH 7.4) and stored at 4°C. Coronal sections (30 µm) were cut with

a vibratome (Leica Microsystems VT1000, Germany) and stored at -20°C in a solution containing 30% ethylene glycol, 30% glycerol and 0.1 M sodium phosphate buffer, until they were processed for immunofluorescence. Individualized free-floating sections were rinsed in Tris-buffered saline with sodium fluoride (TBS-NaF; 0.25 M Tris, 0.5 M NaCl and 0.1mM NaF, pH 7.5), incubated for 5 min in TBS-NaF containing 3% H_2O_2 and 10% methanol, and then rinsed 10 min three times in TBS-NaF. After 20 min incubation in 0.2% Triton X-100 in TBS-NaF, sections were rinsed three times in TBS-NaF again. Choline acetyltransferase (ChAT) and the double phosphorylated form of S6 ribosomal protein (p-Ser$^{240-244}$-S6rp) were simultaneously detected through incubation with combined polyclonal goat anti-ChAT (1:500, #AB144P, Millipore, Billerica, MA) and polyclonal rabbit anti-p-Ser$^{240-244}$-S6rp (1:500, #2215, Cell Signaling Technology, Beverly, MA) primary antibodies diluted in TBS-NaF (4°C, overnight). Dopamine- and cAMP- regulated phosphoprotein (DARPP-32) and the double phosphorylated form of extracellular signal-regulated kinase 1/2 (p-Thr^{202}-Tyr^{204}-ERK1/2) were simultaneously detected through incubation with combined polyclonal mouse anti-DARPP-32 (1:300, #611520, BD Transduction Laboratories) and polyclonal rabbit anti-p-p44/42 MAPK (ERK1/2; Thr^{202}-Tyr^{204}) (1:300, #9101, Cell Signaling Technology, Beverly, MA) primary antibodies diluted in TBS-NaF (4°C, overnight). Sections were then rinsed 10 min in TBS-NaF three times and incubated 60 min with combined donkey anti-goat/mouse Alexa-594-coupled and donkey anti-rabbit Alexa-488-coupled secondary antibodies (Life Technologies, Carlsbad, CA) diluted 1:400 in TBS-NaF. Sections were rinsed four times 10 min in TBS before mounting in Vectashield fluorescence medium (Vector laboratories, Burlingame, CA). The same procedure was followed for simultaneous immunodetection of ChAT (see above) and muscarinic M2-receptors (M2Rs; primary antibody: polyclonal rabbit anti-m2, 1:1000, #M9558, Sigma-Aldrich, St. Louis, MO), although no NaF was added in this case.

Fluorescence analysis and cell counts

Images were obtained using sequential laser scanning confocal microscopy (Fluoview FV300, BX61WI microscope, Olympus, Shinjuku, Japan). All ChAT-immunoreactive neurons in dorso-medial and dorso-lateral striatal regions were detected in the microscope using a 60X objective (UPFL 60X oil) and centered in the acquisition area. Focal plane with optimal ChAT immunoreactivity was determined in channel 2 (Ch02; HeNe green laser). Sequential high-

resolution images (optical magnification: 60X; digital zoom: 4X; resolution: 1024x1024 px) were obtained for p-Ser$^{240-244}$-S6rp signal (Ch01) and corresponding ChAT signal (Ch02) with a Kaplan filter (5 averaging scans). Laser intensity, PMT voltage and offset were maintained constant in all acquisitions of the same experiment (Ar laser intensity [Ch01] usually: 25.5%; PMT: 740v; offset: 2%; HeNe green laser intensity [Ch02] usually: 35.0%; PMT: 720v; offset: 2%). Raw 16-bit images were then analyzed using ImageJ software (MacBiophotonics upgrade v. 1.43u, Wayne Rasband, National Institutes of Health, USA). In superimposed channels, somatic neuronal area (excluding the nucleus) was determined as ROI in Ch02 (ChAT signal) and mean fluorescence intensity (mean gray value, gray values ranging from 0 to 4096) was measured in corresponding Ch01 (p-Ser$^{240-244}$-S6rp signal). All clear ChAT-immunoreactive neurons were photographed (10-15 neurons per striatal region). For the quantification of p-Thr^{202}-Tyr^{204}-ERK1/2 immunoreactive neurons, high-resolution scanning confocal images of dorso-medial and dorso-lateral striatal regions were obtained per each individual (optical magnification: 20X; resolution: 2048x2048 px). Cell counts were performed using imageJ software (Cell counter plugin). Prior to all quantifications, all image files in each experiment were randomly renumbered using a MS Excel plug-in (Bio-excel2007 by Romain Bouju, France). A pseudo-color palette highlighting intensity of fluorescence (16-color LookUp Table, display range: 0 - 4096) was applied to representative neurons.

Behavioral procedures:

Outcome Devaluation Procedure

Magazine training: On days 1 and 2 all rats were placed in operant chambers for ~20 min. In each session of each experiment the house light was illuminated at the start of the session and turned off when the session was terminated. No levers were extended during magazine training. 20 pellet and 20 sucrose outcomes were delivered to the magazine on an independent random time (RT) 60 s schedule.

Lever training: The animals were next trained to lever press on random ratio schedules of reinforcement. Each lever was trained separately, and for half of the animals in each group, the left lever earned pellets and the right lever earned the sucrose solution. The remaining animals were trained on the opposite action-outcome contingencies. The session was terminated when either 20 outcomes were earned, or when 30 min was reached, whichever came first. There were two training sessions per day, one with each action–

outcome pair. The animals had a break of at least 1 hr between sessions, and the order was alternated each day. For the first 2 days lever pressing was continuously reinforced. Rats were shifted to a random ratio (RR)-5 schedule for the next 3 days (i.e. each action delivered an outcome with a probability of .2), then to an RR-10 schedule (or a probability of .1) for 3 days, then to an RR-20 schedule (or a probability of .05) for the final 3 days.

Devaluation extinction tests: Following the final day of RR-20 training, rats were given access *ad libitum* to either the pellets (25g place in a bowl in the devaluation cage) or the sucrose solution (100mL in a drinking bottle fixed to the top of the devaluation cage) for 1 hr. The aim of this prefeeding procedure was to satiate the animal specifically on the prefed outcome, thereby reducing its value relative to the non-prefed 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 Degradation Procedure

Retraining: Subsequent to the devaluation extinction tests, rats received one day of retraining on an RR-20 schedule.

Contingency degradation training: During the 6 days of contingency degradation training rats continued to receive these same action-outcome pairings on an RR-20 schedule. In addition, one of the two outcomes (either pellets or sucrose) was delivered outside of the lever press-outcome contingency. That is, for each second that no lever pressing occurred, either sucrose or pellets were delivered with the same probability [p(outcome/no action) = .05] that a lever press earned that outcome. As a result, the probability of earning one of the two outcomes was the same whether the animal pressed the lever or not, which should degrade that specific action-outcome contingency. The alternative action-outcome contingency remained nondegraded because the rat was still required to press the lever to receive that outcome. For half of the animals the lever press–pellet contingency was degraded, and the lever press-sucrose contingency remained intact. The remaining animals received the opposite arrangement. Rats were given two 20 min training sessions each day,

one on each (i.e. right or left) lever with a break of 2 hr between sessions. The order of the sessions was alternated each day.

Contingency degradation extinction test: Following the final day of contingency training, rats in both groups received a 10 min choice extinction test. During this test both levers were extended and lever presses recorded, but no outcomes were delivered.

Contingency reversal procedure

Contingency reversal training: Subsequent to the contingency degradation extinction test rats were trained to lever press on an RR-20 schedule with the 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. Contingency reversal training continued for 4 days.

Devaluation Extinction Tests: Devaluation extinction tests took place as described for outcome devaluation.

Reinstatement Testing: Subsequent to the two devaluation extinction tests rats were retrained on the reversed contingencies on an RR-20 schedule for 1 day. The next day an outcome-selective instrumental reinstatement test was conducted. The test session began with a 15 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.