Title: Goal-directed actions transiently depend on dorsal hippocampus

Author list:

Laura A. Bradfield[1,2*#], Beatrice K. Leung[3*], Susan Boldt[4], Sophia Liang[3] and Bernard W. Balleine[3#]

Affiliations:

[1] Centre for Neuroscience and Regenerative Medicine, University of Technology Sydney, Sydney.

[2] St Vincent's Centre for Applied Medical Research, St Vincent's Health Network, Sydney.

[3] Decision Neuroscience Lab, School of Psychology, University of New South Wales, Sydney.

[4] School of Psychology, University of Sydney.

* These authors contributed equally

Corresponding authors

Correspondence to:

Bernard Balleine
Decision Neuroscience Lab
School of Psychology, UNSW Sydney
Randwick, NSW 2052
Australia
Bernard.balleine@unsw.edu.au

Laura Bradfield
Centre for Neuroscience and Regenerative Medicine,
School of Life Sciences,
University of Technology Sydney (St. Vincent's Campus)
405 Liverpool St
Darlinghurst, NSW 2010.
laura.bradfield@uts.edu.au

Main Text: 2000 words

Abstract

The role of the hippocampus in goal-directed action is currently unclear; studies investigating this issue have produced contradictory results. Here we reconcile these contradictions by demonstrating that, in rats, goal-directed action relies on dorsal hippocampus but only transiently, immediately after initial acquisition. Furthermore, we found that goal-directed action also depends transiently on physical context, suggesting a psychological basis for the hippocampal regulation of goal-directed action control.

One Sentence Summary: Goal-directed actions are hippocampally- and context-dependent immediately after initial encoding but not after a delay.

Introduction

Evidence regarding the role of the dorsal hippocampus in goal-directed action is mixed. Neuroimaging studies conducted in humans have produced several findings suggesting a central role for hippocampus in regulating non-navigational goal-directed decision-making[1,2]. Likewise, electrophysiological studies in rodents have implied a role for dorsal hippocampus in both navigational[3] and non-navigational goal-directed tasks[4]. Rodent lesion studies, on the other hand, have found non-navigational goal-directed actions to be intact despite comprehensive lesions of dorsal hippocampus[5,6].

One potentially crucial difference between these studies is the types and amount of training involved. Studies that have detected a relationship between hippocampal activity and goal-directed action have involved protocols in which participants were either trained and tested in a single day[1] or trained over several days on tasks that involve multiple contingency switches that could each be encoded as a novel event[2,4]. By contrast, studies that found no role for hippocampus in goal-directed action trained animals on stable contingencies over multiple days[5,6]. We sought, therefore, to investigate the hypothesis that non-navigational goal-directed actions rely on dorsal hippocampus, but only during initial learning.

We used outcome devaluation tests[7] to determine whether actions were goal-directed (see methods). Rats were trained to press two levers each delivering a unique outcome; either pellets or sucrose. After training, the value of one of these outcomes was reduced using sensory specific satiety[7] after which the rats were given a choice between the levers in an extinction test (i.e. in the absence of pellet or sucrose delivery). In such tests, rats typically respond more on the lever associated with the still-valued (non-prefed) outcome relative to the devalued lever, demonstrating control by both the current value of the outcome and the action-outcome association in accord with definitions of goal-directed action[7,8]. Our first series of experiments (Experiments 1-3) investigated whether dorsal hippocampal

involvement in goal-directed action depends on the amount of training. To achieve this, we inactivated dorsal hippocampus using either local infusions of γ-aminobutyric acid-A ($GABA_A$) receptor antagonist muscimol or chemogenetics.

Results

Goal-directed action is transiently affected by dorsal hippocampal inactivation

We first assessed the effect of intra-dorsal hippocampal infusions of muscimol given either prior to training (Experiment 1a) or test (Experiment 1b, see methods; **Extended Data Fig. 1a-c**). For each experiment, rats were first trained to press both left (A1) and right (A2) levers for polycose (O3) on days 1-5. On day 6, the left and right levers earned unique outcomes, either pellets or sucrose (i.e., A1→O1 and A2→O2; counterbalanced, design in **Fig. 1a**). Groups did not differ in lever press acquisition either for polycose or pellets and sucrose in either experiment (largest $F = 1.99$, $p = .176$, **Extended Data Fig. 2a-f**). On test, devaluation was intact (Valued > Devalued) for rats that received saline infusions on Day 6, but was impaired in muscimol-infused rats (Valued = Devalued, Experiment 1a; **Fig. 1b**), demonstrated by a group x devaluation interaction, $F(1,17) = 6.56$, $p = .02$, and a significant simple effect for group SALINE, $F(1,17) = 13.46$, $p = .002$, but not group MUSCIMOL, $F < 1$. Hippocampal inactivation on test (Experiment 1b) again disrupted devaluation relative to saline controls (**Fig. 1c**). There was a significant group x devaluation interaction, $F(1,14) = 6.09$, $p = .027$, and a simple effect in group SALINE, $F(1,14) = 5.86$, $p = .03$, but not in group MUSCIMOL, $F < 1$.

Experiment 2 (design in **Fig. 1d**) replicated and extended Experiment 1b, except we omitted polycose pretraining and used inhibitory hM4Di DREADDs (based on procedures validated previously[9,10]) to inactivate dorsal hippocampus, avoiding the need for multiple infusions. Experiment 2 also sought to establish the regional specificity of the effect by comparing

111 animals that received transfection directed towards the CA1 region of the dorsal hippocampus
112 with a group in which expression was confined to CA2 (**Fig. 1e**,**Extended Data Fig. 1d-e**).
113 Half of the rats in each control group (hM4Di+Veh and mCherry+CNO) received viral
114 transfection in CA1 and half in CA2.

115 For this experiment, rats were trained to a criterion of a minimum of 20 outcomes on
116 each lever over 1-2 days, then tested for devaluation performance the following day (see
117 methods). Vehicle and CNO injections were administered prior to test. Initial lever press
118 acquisition and number of outcomes earned on each lever was similar for all groups (all Fs <
119 1, **Extended Data Fig. 2g-h**). On test, however, whereas devaluation was intact in controls,
120 i.e., groups hM4Di+Veh, mCherry+CNO, and hM4Di CA2+CNO (Valued > Devalued), it
121 was impaired in group hM4Di CA1+CNO (**Fig. 1f**). There was a group (hM4Di CA1+CNO
122 vs. the others) x devaluation interaction, $F(1,31) = 5.19$, $p = .03$, supported by significant
123 simple effects for groups hM4Di+Veh, $F(1,31) = 7.03$, $p = .012$, mCherry+CNO, $F(1,31) =$
124 17.98, $p = .00$, and CA2+CNO, $F(1,31) = 7.32$, $p = .011$, but not group hM4Di CA1+CNO, F
125 < 1. Following this test, all groups were trained for a further 4-5 days and tested again
126 (**Extended Data Fig. 2i**). In contrast to the initial test, all groups now showed intact
127 devaluation (**Fig. 1g**). There was a main effect of devaluation, $F(1,31) = 77.57$, $p = .00$,
128 which didn't interact with any group differences, largest $F(1,31) = 2.2$, $p = .148$.

129 Next we investigated whether goal-directed actions also become hippocampally-
130 independent with the passage of time (Experiments 3a and 3b - design in **Fig. 1h**). For this
131 experiment, animals again received 1-2 days of lever press training to criterion, followed by
132 outcome devaluation testing the next day (immediate - Experiment 3a), or one week later
133 (delayed - Experiment 3b). Performance for dorsal hippocampus inactivated (hM4Di+CNO)
134 animals was compared against mixed (hM4Di+Veh and mCherry+CNO) controls (**Extended**
135 **Data Fig. 1f**). For both experiments, half of the animals in each group received CNO/Vehicle

injections prior to training and half prior to test.

Groups tested immediately did not differ in their lever pressing during acquisition, $F(1,34) = 1.091$, $p = .304$, or with regards to the number of outcomes earned ($F < 1$, **Extended Data Fig. 2j-k**). On test, however, devaluation performance was intact in control rats, but impaired in hippocampally-inactivated rats whether inactivation occurred prior to training or prior to test (**Fig. 1i**). There was a group (hM4Di+CNO vs. Controls) x devaluation interaction $F(1,34) = 4.622$, $p = .039$. Simple effects analysis found intact devaluation (Valued > Devalued) in Controls that received training injections, $F(1,34) = 6.606$, $p = .015$ and those that received test injections, $F(1,34) = 6.568$, $p = .015$, but impaired devaluation in the hM4Di+CNO groups, $Fs < 1$. Rats tested after a delay again did not show any differences in lever press acquisition ($Fs < 1$) or outcomes received, $F(1,33) = 2.388$, $p = .132$, (**Extended Data Fig. 2j, 2l**) but did show intact devaluation performance in all groups (**Fig. 1j**). There was a main effect of devaluation, $F(1,33) = 21.81$, $p = .00$, but no two-way interaction ($F < 1$). Collectively, these results support the claim that goal-directed actions transiently depend on dorsal hippocampus.

Goal-directed action transiently depends on physical context

We next considered the psychological basis of hippocampal involvement in goal-directed action. One potential source of the above effects lies in the well documented role of the hippocampus in representing spatial context[11]. Although goal-directed actions are independent of context when stable contingencies are trained over multiple days[12], their dependency on context early in training has not been assessed so we sought to determine whether goal-directed actions also transiently depend on the physical context. To achieve this, rats again received 5 days of polycose pretraining before a single day of training during which the levers earned either pellets or sucrose (Experiment 4a – design in **Fig. 2a**). In parallel, another cohort of rats were trained such that the levers earned pellets and sucrose

across all 6 days of training (Experiment 4b - design in **Fig. 2a)**. After day 6, rats in both experiments were given an outcome devaluation test in either the training context (group SAME) or in a different but familiar context (group DIFFERENT). Based on the hippocampal results (Experiment 1a), we predicted that only rats given minimal training would show transient control of goal-directed action by context. As shown in Figure 2, this was the observed result. There were no group differences during lever press acquisition in either experiment (all Fs < 1, **Extended Data Fig. 3a-d**). On test, however, animals given minimal training (Experiment 4a) showed intact devaluation (Valued > Devalued) if tested in the same context, but impaired devaluation (Valued = Devalued) if tested in the different context (**Fig. 2b**); there was a significant group x devaluation interaction, $F(1,23) = 4.55$, $p = .044$, and simple effect for group SAME, $F(1,23) = 7.076$, $p = .014$, but not group DIFFERENT, $F < 1$. By contrast, rats given additional training (Experiment 4b) showed intact devaluation regardless of test context; there was a significant main effect of devaluation, $F(1,19) = 11.78$, $p = .003$, but no interaction, $F < 1$ (**Fig. 2c**).

In a final study (Experiment 5 - design in **Fig. 2d**) we sought to determine whether goal-directed actions also become independent of their physical context with the passage of time. Rats were again trained to lever press to criterion over 1-2 days (without pretraining) after which they were tested in the same or different context either immediately (1 day later, groups SAME-IMM and DIFF-IMM), or 1 week later (groups SAME-DELAY and DIFF-DELAY). Again, lever pressing did not differ during acquisition (all Fs < 1), nor did delivery of pellets/sucrose, $F(1,22) = 2.411$, $p = .135$, **Extended Data Fig. 3e-f**). On test, goal-directed action was context-dependent only when testing was conducted immediately after training, and was context-independent after a one-week delay (**Fig. 2e**). There was a significant group x lever interaction $F(1,22) = 4.584$, $p = .044$, supported by significant simple effects for groups SAME-IMM, $F(1,22) = 5.489$, $p = .029$, SAME-DELAY, $F(1,22) =$

186 10.385, p = .004, and DIFF-DELAY F(1,22) = 6.018, p = .023, but not in group DIFF-IMM,
187 F < 1.

188 **Discussion**

189 Together, our results support our central claim that goal-directed actions transiently
190 depend on dorsal hippocampus and imply that a hippocampal representation of the physical
191 context is the source of its role in this effect. Our findings support this specific claim rather
192 than a role in the contextual regulation of goal-directed action generally because Experiment
193 5 essentially involved two contextual alterations: physical (i.e. the change in physical
194 context) and temporal (i.e. testing immediately vs. after one-week). If goal-directed actions
195 were broadly context-dependent after initial training we would expect the delay alone to
196 impair devaluation even in the absence of an altered physical environment. In contrast to this
197 prediction, however, our results found devaluation to be intact in both the same *and* different
198 contexts after a delay (**Fig. 2e**). Furthermore, these results cannot be explained by
199 hippocampal or context manipulations causing generalized rather than specific satiety
200 following devaluation, even transiently, because that would be expected to reduce responding
201 selectively on the valued lever. Such an effect was only statistically supported in Experiment
202 5 (**Extended Data Fig. 4i**), where specific satiety was induced *before* the context
203 manipulation was applied. Thus, these results are more consistent with an effect on the
204 capacity for goal-directed action, although whether this reflects a deficit in action-outcome
205 memory or in decision-making itself remains to be tested.

206 The effects reported here fit well with several of the known functions of
207 hippocampus, such as systems consolidation[13,14] and episodic memory [15,16], and, for the first
208 time, link these functions directly with decision-making involving choice. The systems
209 consolidation account[13,14] suggests that the dorsal hippocampus) regulates short term memory
210 formation and recall, which becomes hippocampally independent, possibly migrating to

frontal cortical structures over the course of a week. The vast majority of the evidence for this theory has come from studies of conditioned reflexes in Pavlovian conditioning, particularly Pavlovian fear conditioning[17,18]. In light of the current findings, however, a link between systems consolidation and goal-directed action control warrants further consideration. Relatedly, our findings are also consistent with the view that goal-directed actions initially rely on (contextually and hippocampally-dependent) episodic memories which become context-free, extra-hippocampal semantic memories with the passage of time[19,20].

Acknowledgments: We thank Genevra Hart for helpful discussions and Fred Westbrook for his feedback on the manuscript. **Funding:** This work was supported by grants GNT1087689 and GNT1148244 from the National Health and Medical Research Council in Australia to L.A.B. and B.W.B.

Author contributions: L.A.B, B.K.L., and B.W.B designed the experiments, L.A.B., B.K.L., S.L. and S.B. performed the experiments, L.A.B, B.K.L. and B.W.B wrote the manuscript.

Competing Interests Statement: Authors declare no competing interests.

References

1. Vikbladh, O. M. *et al.* Hippocampal contributions to model-based planning and spatial memory. *Neuron* **102,** 683-693.e4 (2019).

2. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science.* **364,** eaaw5181 (2019).

3. Wikenheiser, A. M. & Redish, A. D. Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* **18,** 289–294 (2015).

4. Zhou, J. *et al.* Complementary Task Structure Representations in Hippocampus and Orbitofrontal Cortex during an Odor Sequence Task. *Curr. Biol.* **29,** 3402-3409.e3 (2019).

5. Corbit, L. H. & Balleine, B. W. The Role of the Hippocampus in Instrumental Conditioning. *J. Neurosci.* **20,** 4233 LP–4239 (2000).

6. Corbit, L. H., Ostlund, S. B. & Balleine, B. W. Sensitivity to Instrumental Contingency Degradation Is Mediated by the Entorhinal Cortex and Its Efferents via the Dorsal Hippocampus. *J. Neurosci.* **22,** 10976 LP–10984 (2002).

7. Dickinson, A. & Balleine, B. Motivational Control of Instrumental Action. *Curr. Dir. Psychol. Sci.* **4**, 162–167 (1995).

8. Dickinson, A. Instrumental conditioning. in *Animal learning and cognition.* 45–79 (Academic Press, 1994). doi:10.1016/B978-0-08-057169-0.50009-7.

9. Laurent, V., Chieng, B. & Balleine, B. W. Extinction Generates Outcome-Specific Conditioned Inhibition. *Curr. Biol.* **26,** 3169–3175 (2016).

10. Bradfield, L. A., Dezfouli, A., Van Holstein, M., Chieng, B. & Balleine, B. W. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations.

Neuron **88,** 1268–1280 (2015).

11. Fanselow, M. S. & Dong, H. W. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* **65,** 7–19 (2010).

12. Thrailkill, E. A. & Bouton, M. E. Contextual control of instrumental actions and habits. *J. Exp. Psychol. Anim. Learn. Cogn.* **41,** 69–80 (2015).

13. Russell, W. R. & Nathan, P. W. Traumatic amnesia. *Brain* **69,** 280–300 (1946).

14. Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20,** 11 LP–21 (1957).

15. Vargha-Khadem, F. *et al.* Differential effects of early hippocampal pathology on episodic and semantic memory. *Science.* **277,** 376 LP–380 (1997).

16. Tulving, E. & Markowitsch, H. J. Episodic and declarative memory: Role of the hippocampus. *Hippocampus* **8,** 198–204 (1998).

17. Kim, J. J. & Fanselow, M. S. Modality-specific retrograde amnesia of fear. *Science.* **256,** 675–677 (1992).

18. Schafe, G. E., Nader, K., Blair, H. T. & LeDoux, J. E. Memory consolidation of Pavlovian fear conditioning: A cellular and molecular perspective. *Trends Neurosci.* **24,** 540–546 (2001).

19. Gershman, S. J. & Daw, N. D. Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annu. Rev. Psychol.* **68,** 101–128 (2016).

20. Verschure, P. F. M. J., Pennartz, C. M. A. & Pezzulo, G. The why, what, where, when and how of goal-directed choice: neuronal and computational principles. *Philos. Trans. R. Soc. B Biol. Sci.* **369,** 20130483 (2014).

Figure Legends

Fig. 1. Inactivation of dorsal hippocampus transiently impaired outcome devaluation performance when testing was immediate, but not after additional training or a delay. a) Top - Experiment 1a design. Bottom - Experiment 1b design, **b**) Mean lever presses on test for Experiment 1a (n = 19 rats), **c**) Mean lever presses on test for Experiment 1b (n = 16 rats), **d**) Experiment 2 design. **e**) Representative micrographs of dorsal hippocampal neurons transfected with AAV8-hSyn-hM4D(Gi)-mCherry DREADDs virus (M4; red) in CA1 region (top row; n = 19 rats showed similar infection in Experiment 2) and lateral CA2 region (bottom row; n = 16 rats showed similar infection in Experiment 2). **f**) Mean lever presses during the initial test (days 3-4; n = 35 rats), **g**) Mean presses during the extended test (days 9-10; n = 35 rats), **h**) Top – Design of Experiment 3a. Bottom – Design of Experiment 3b, **i**) Mean lever presses on test for Experiment 3a (n = 38 rats), **j**) Mean lever presses on test for Experiment 3b (n = 37 rats). Data are shown as individual dot plots and mean ± SEM. A = action, O = outcome.

Fig 2. Outcome devaluation performance is impaired by a context switch immediately after limited training, but not after additional training or after a delay. a) Top - Design of Experiment 4a. Bottom - Design of Experiment 4b. **b**) Mean lever presses during test for rats that had limited training (Experiment 4a; n = 25 rats), **c**) Mean lever presses on test for rats that had extended training (Experiment 4b; n = 21 rats). **d**) Design of Experiment 5. **e**) Mean presses during devaluation test (n = 26 rats). Data are shown as individual dot plots and mean ± SEM. A = action, O = outcome.

Methods

All procedures were approved by the University of New South Wales Ethics Committee and/or or the University of Sydney Ethics Committee.

Subjects and Exclusions

Long-Evans Rats, between 12-24 weeks of age at the beginning of the experiment, were housed in transparent amber plastic boxes (0.5 m^3; 3-4 rats per box) located in a temperature- and humidity-controlled vivarium and were maintained on a 12 h light/dark cycle (lights on between 7:00 A.M. and 7:00 P.M.). Experiments 1a, 1b, 4a and 4b were conducted with male rats. All other experiments were conducted using approx. 50% male and 50% female rats. Males weighed between 350-500g and females weighed between 200-300g at the beginning of each experiment.

Experiment 1a

A total of 24 male rats were used as subjects. Four rats were excluded from the final analysis due to cannula misplacement and 1 rat was excluded due to infection. Thus the final experiment numbers were 9 rats in group SALINE, and 10 rats in group MUSCIMOL.

Experiment 1b

A total of 19 male rats were used as subjects. Three rats were excluded from the final analysis due to cannula misplacement. Thus the final experiment numbers were 9 rats in group SALINE, and 7 rats in group MUSCIMOL.

Experiment 2

A total of 55 rats (28 males and 27 females) were used as subjects. Four animals were excluded for misplaced DREADDs expression. Sixteen animals were excluded from the final test for not reaching criterion (> 20 outcomes per lever).

Experimental numbers for the first round of devaluation testing (after 1-2 days of training) were: 7 rats in group hM4Di+Veh, 16 rats in mCherry+CNO, 6 rats in hM4Di CA1+CNO, and 6 rats in hM4Di CA2+CNO.

For hM4Di-transfected animals, half of those that received CNO during the first round of testing were switched to Vehicle (Veh) for the second round of testing (after 6 days of training), and vice versa. The remaining half received consistent injections prior to both tests. Experimental numbers for the second round of devaluation testing (after 6 days of training) were 8 rats in group hM4Di+Veh, 16 rats in mCherry+CNO, 5 rats in hM4Di CA1+CNO, and 6 rats in hM4Di CA2+CNO.

Experiment 3a

A total of 52 rats (25 males and 27 females) were used as subjects. Six animals were excluded for misplaced DREADDs expression or an infection at an injection site. Eight animals were excluded from the final test for not reaching criterion (> 20 outcomes per lever).
Experimental numbers on test were: 7 rats in group Controls-TRAINING, 9 M4-TRAINING, 12 rats in Controls-TEST and 10 rats in M4-TEST.

Experiment 3b

A total of 55 rats (28 males and 27 females) were used as subjects.

Six animals were excluded for misplaced DREADDs expression or an infection at an injection site. Twelve animals were excluded from the final test for not reaching criterion (> 20 outcomes per lever).

Experimental numbers on test were: 7 rats in group Controls-TRAINING, 9 M4-TRAINING, 10 rats in Controls-TEST and 11 rats in M4-TEST.

Experiment 4a

A total of 25 male rats were used as subjects, 13 in group SAME, and 12 in group DIFFERENT.

Experiment 4b

A total of 23 male rats were used as subjects. Two animals were excluded for failing to lever press, thus final experimental numbers were 11 rats in group SAME, and 10 rats in group DIFFERENT.

Experiment 5

A total of 46 rats (22 males and 24 females) were used as subjects. Twenty animals were excluded from the reported test for not reaching criterion (> 20 outcomes per lever).

Experimental numbers on test were 6 rats in group SAME-IMM, 6 rats in group DIFF-IMM, 8 rats in group SAME-DELAY, and 6 rats in group DIFF-DELAY.

Apparatus

For all behavioral experiments, training was conducted in 16 MED Associates operant chambers enclosed in sound- and light-attenuating cabinets. Each chamber was fitted with a pellet dispenser capable of delivering a 45 mg grain food pellet (F0165, BioServ Biotechnologies), to a recessed magazine inside the chamber, as well as two pumps fitted with syringes outside the chamber, capable of delivering 0.2 mL of either 20% sucrose solution (white sugar, Coles) diluted in water or 20% maltodextrin solution (Poly-Joule, Nutrica) diluted in water, each delivered to separate compartments of the recessed magazine inside the chamber. The chambers also contained two retractable levers that could be inserted individually on the left and right sides of the magazine. Head entries into the magazine were detected via an infrared photobeam. Unless otherwise stated, the operant chambers were fully illuminated during all experimental stages, illumination was provided by a 3W, 24V house light located on the upper edge of the wall opposite to the magazine. All training sessions were pre-programmed on two computers located in a separate room through the MED Associates software (Med-PC), these computers also recorded the experimental data from each session.

Surgery

Rats were anaesthetized with 3% inhalant isoflurane gas mixed with oxygen, delivered at a rate of 0.5 L/min throughout surgery. Anaesthetized rats were placed in a stereotaxic frame (Kopf Instruments). 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 Experiments 1a and 1b, a small hole was then drilled into the skull above the hippocampus (all co-ordinates in millimeters relative to bregma: anteroposterior, -3.8, mediolateral, ±3.2, dorsoventral, -2.5 at an angle of 15 degrees) in each hemisphere. A 26 gauge guide cannula (Plastics One) was then implanted into the hole, the tip of which was

aimed towards the CA1 region of the hippocampus. Guide cannulas were maintained in position with dental cement, and dummy cannulas were kept in each guide at all times except during microinfusions. The wound was subsequently stapled and cleaned, after which rats were injected with a prophylactic (0.4 mL) dose of 300 mg/kg procaine penicillin interperitoneally (i.p), and 0.1mL of the analgesic Rymadil subcutaneously (s.c.). Rats were given one week of recovery following surgery during which they were weighed and monitored each day.

For Experiments 2, 3a and 3b, following the scalp incision a small hole was drilled into the skull above the hippocampus, either above the CA1 region (Half of the animals in Experiment 2, and all animals in Experiment 3a and 3b, anteroposterior, -3.8, mediolateral, ±2.5, dorsoventral, -3.5 for males, and anteroposterior, -3.6, mediolateral, ±2.5, dorsoventral, -3.5 for females) or above the CA2 region (Half of the animals in Experiment 2 only, anteroposterior, -3.8, mediolateral, ±2.5, dorsoventral, -3 for males, anteroposterior, -3.6, mediolateral, ±2.5, dorsoventral, -3 for females). A 1.0 µL glass syringe (Hamilton Company) connected to an infusion pump (Pump 11 Elite Nanomite, Harvard Apparatus) was lowered into the brain, and rats received 0.75µL infusions of either AAV8-hSyn-hM4D(Gi)-mCherry or AAV8-hSyn-mCherry, at a rate of 0.15µL per minute. All DREADD viruses and control fluorophores were obtained from UNC Vector Core or Addgene based on plasmids gifted by Bryan Roth and Karl Deisseroth. Following infusions, the needle was left in place for a further 2 min for diffusion before being retracted. The wound was subsequently closed with staples (Stoelting) and cleaned, after which rats were injected with a prophylactic (0.4 mL) dose of 300 mg/kg procaine penicillin (i.p), and 0.1mL of the analgesic Carprofen (s.c.). Rats were given a minimum of 3 weeks of recovery time following surgery to allow sufficient viral expression.

Drug Infusions

In Experiments 1a and 1b, rats received bilateral intra-hippocampal infusions of 5-aminomethyl-3-hydroxyisoxazole (muscimol, M1523, Sigma). A total of 0.5 µL of muscimol (0.5mg/mL) was infused at 0.32 µL/min. Control rats received a saline infusion at the same rate. Microinfusions were conducted using a 33 gauge infusion cannulas (Plastics One) that extended 1 mm beyond the guide cannula. Infusion cannulas were inserted into the guide cannulas and connected to 25 µL glass syringes (Hamilton Company) fitted on an infusion pump (PHD ULTRA 4400, Harvard apparatus). The infusion cannulas were left in place for 1 min following infusions to allow diffusion of the drug. Rats were placed back in their home cage for 20 min prior to behavioral training/testing to permit the drug to take effect.

Drugs for i.p. injection

In Experiments 2, 3a and 3b, Clozapine-N-Oxide (CNO; RTI international) was dissolved in 0.8% HCl in water to a concentration of 7 mg/mL (pH 3-4). A solution of 0.1% HCl in water of the same pH was used as vehicle. Drug or vehicle was injected i.p. 1 h prior to the onset of instrumental training session or testing session, at a volume of 1 mL/kg, hence the dosage was 7 mg/kg. We have previously demonstrated the viability of this procedure, and its efficacy in reducing firing in hM4Di DREADDs-transfected cells in dorsal hippocampus using electrophysiology[1].

Contexts

Experiments 4a, 4b, and 5 employed two distinct contexts. 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 a 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.

For these experiments, animals received one magazine training session in each context during which both pellet and sucrose outcomes were delivered, one on each day (order counterbalanced), and two more pre-exposures to the 'different' context after lever press training sessions. This served to familiarize the animals to the different context and reduce neophobia. Pre-exposure sessions lasted 40 min during which no levers were extended and no food was delivered.

Food restriction

Rats underwent 3 d of food restriction prior to the onset of magazine training and this continued throughout the duration of the experiment. During this time, they received 5 g (for females) or 8 g (for males) of home chow daily for the first two days, and 7-12 g (females) and 10-15 g (males) from the third day until the end of the experiment. Their weight was monitored daily to ensure it remained above 85% of their pre-surgery body weight at all times.

Behavioral Procedures

Please note that we employed random ratio schedules during training[7] as well as choice tests[21] in all experiments to promote goal-directed behavior and prevent the transition to habitual responding, even after multiple days of training.

Magazine training

Rats in Experiments 1a, 1b, 2, 3a and 3b were given one session of magazine training and rats in Experiments 4a, 4b and 5 were given two sessions of magazine training (one in each context). For these sessions, the house light was turned on at the start of the session and turned off when the session was terminated. No levers were extended. For rats that received polycose pretraining first (Experiments 1a, 1b, and 4a), 30 deliveries of 20% polycose solution were delivered into the magazine on a random time 60 s schedule (RT60). For rats that only had lever press training for pellets and sucrose solution (Experiments 2, 3a, 3b, 4b and 5), 20 deliveries of pellets and 20 deliveries of 20% sucrose solution were delivered on independent RT60 schedules.

Polycose Pretraining

Rats that had polycose pretraining were trained to press two levers that earned the same outcome polycose, prior to receiving lever press training for pellet and sucrose solution. Each session lasted for a maximum of 50 min and consisted of four periods where a single lever was inserted into the chamber (i.e. two periods for each lever) separated by a 2.5 min time out period in which the lever was retracted and the house light was turned off. Each period ended after 20 outcomes had been earned or 10 min elapsed. The order of presentation of each lever was pseudorandom.

On day 1, lever presses on each lever were continuously reinforced with a polycose solution. On days 2 and 3, the schedule of reinforcement shifted to a random ratio (RR) 5 schedule such that the probability of polycose delivery was 0.2 for each action. On days 4 and 5, the schedule of reinforcement shifted to a RR10 schedule such that the probability of a delivery of the outcome was 0.1 for each action.

Lever press training for pellets and sucrose solution

Each session lasted for a maximum of 50 min and consisted of four periods where a single lever was inserted into the chamber (i.e. 2 periods for each lever) separated by a 2.5 min time out period in which the lever was retracted and the house light was turned off. Each period ended after 20 outcomes had been earned or 10 min elapsed. The order of presentation of each lever was pseudorandom.

Rats that had previously received polycose pretraining on days 1-5 (Experiments 1a, 1b, and 4a), received a single contingency training session on day 6 where lever presses earned sucrose solution and pellets on a RR10 schedule. For half of the animals in each group, the left lever earned pellets and the right lever earned a sucrose solution, and the remaining animals were trained on the opposite action-outcome contingencies (counterbalanced).

For rats that received 6 days of lever press training (Experiment 4b) lever presses were continually reinforced with pellets or sucrose solution on day 1. After, the schedule of reinforcement shifted to a RR5 schedule on days 2-3 and a RR10 schedule on days 4-6. Action-outcome contingencies were counterbalanced for each group.

For rats that received 1-2 days of lever press training (Experiments 2, 3a, 3b, and 5), lever presses were initially continually reinforced with pellets or sucrose solution in the first period each lever was extended (i.e. first 10 min on each lever). If rats earned more than 10 outcomes on both levers in the first 25 min, they were moved to a RR5 schedule. Rats that did not earn more than 10 outcomes on both levers remained on a continually reinforced schedule during the entire 50 min session. Animals that earned more than 20 outcomes per lever by the end of day 1 (i.e. animals that reached criterion) were not trained on day 2. Action-outcome contingencies were counterbalanced for each group.

Outcome Devaluation Tests

During devaluation tests, rats were first placed in the devaluation chambers (which were in a separate room from the operant chambers) and provided with ad libitum access to one of the previously earned outcomes (pellets or sucrose solution) for 1 h. After prefeeding, animals were placed in the operant chambers and given a choice test with both levers available for 5 min but no outcomes were delivered. The following day, rats were given another devaluation test for which they were prefed the opposite outcome. That is, if previously prefed on pellets they were now prefed on sucrose solution, and vice versa.
All rats received one 1 h pre-exposure session to the devaluation chambers following the final lever press training session, in which they were fed a little bit of their daily chow. This served to habituate animals to the devaluation chambers to reduce neophobia.

Each individual experiment employed a variant on the same procedures, as described below.

Randomization

Rodents were randomly allocated to groups. During lever press training, the order in which each (left and right) lever was presented on each day was chosen randomly by the experimenter, and the same order was never presented more than 3 days consecutively.

Experiments 1a

Rats first received polycose pretraining for 5 days. On day 6, rats received an intra-dorsal hippocampal infusion of saline or muscimol as described above. Rats were then placed back into their home cages for 20 min for the drug to take effect. After 20 min, rats were placed into the operant chambers for a single contingency training session for pellets and sucrose solution. After, rats received outcome devaluation testing across two days (day 7-8) as described above.

Experiment 1b

Rats first received polycose training for 5 days followed by 1 day of lever press training for pellets and sucrose solution. After, rats received outcome devaluation testing as described above (day 7-8). Following prefeeding on each devaluation day, rats were immediately administered intra-dorsal hippocampal infusions of saline or muscimol. Rats were then placed back into their home cages for 20 min for the drug to take effect before being tested in the operant chambers.

Experiment 2

Rats received lever press training for pellets and sucrose solution for 1-2 days. After, rats received outcome devaluation testing as described above (day 3-4). Rats then received additional 4 daily lever press training sessions (day 5-8) for pellets and sucrose solution. Following this extended training, rats received a second set of devaluation tests (day 9-10).

Prior to each devaluation day, animals were subject to i.p. Veh or CNO injections immediately before being placed into the devaluation chambers. Once in these chambers, prefeeding procedures took place as described before. After prefeeding, rats were immediately placed into the operant chambers and given a 5 min choice test between levers.

Experiment 3a

Rats received lever press training for pellets and sucrose solution for 1-2 days. The following day, rats received outcome devaluation testing as describe above (day 3-4).

Half of the animals were subject to i.p. Veh or CNO injections (depending on their group assignment) 1 h prior to the lever press training session and the other half were subject to i.p. Veh or CNO injections (depending on their group assignment) 1 h prior to the 5 min choice test (i.e. immediately before being placed into the devaluation chambers).

Experiment 3b

Rats received lever press training for pellets and sucrose solution for 1-2 days. After a 7 day delay, rats received outcome devaluation testing as describe above (day 9-10).

Half of the animals were subject to i.p. Veh or CNO injections (depending on their group assignment) 1 h prior to the lever press training session and the other half were subject to i.p. Veh or CNO injections (depending on their group assignment) 1 h prior to the 5 min choice test (i.e. immediately before being placed into the devaluation chambers).

Experiment 4 a

Rats first received polycose pretraining in one context (context alterations are described above). After, rats received a single session of lever press training on pellets and sucrose solution in that same context. The following day, rats received outcome devaluation tests as described above. For each choice test, animals were placed in operant chambers with the

same or a different context to that in which they had been trained. Context assignment for each test did not change between days, such that animals were either tested in the 'same' context on both test days, or the 'different' context on both test days (although context identity was counterbalanced, such that the 'same' context was the vanilla/stripy/smooth context for half of the rats, and the unadorned/grid/peppermint context for the rest, and likewise for the 'different' context).

Experiment 4b

Rats received lever training sessions for pellets and sucrose solution for 6 days. After, rats received outcome devaluation testing in the same or different context as described in Experiment 4a. Again, context assignment for test did not change between days, such that animals were either tested in the 'same' context on both test days (i.e. one with each outcome), or the 'different' context on both test days.

Experiment 5

Rats received lever press training for pellets and sucrose solution for 1-2 days. Outcome devaluation was conducted identically to that described for Experiments 4a and 4b, except that half of the animals received the devaluation tests one day after the final lever press session (day 3-4), whereas the other half received the devaluation tests after a 7-day delay (day 9-10). Again, context assignment did not change between days.

Histology

Rats were rapidly anaesthetized with sodium pentobarbital (300 mg/kg i.p., Virbac) and transcardially perfused with 400 mL of 4% paraformaldehyde in 0.1 M sodium phosphate buffer (PB; pH 7.4).

Brains collected for cannula placement only were post-fixed for 1 h in the same fixative and placed in 20% sucrose in phosphate buffered saline (PBS pH 7.2) overnight. 40 µm coronal sections were cut using a cryostat (CM1520, Leica Microsystems). Every third section was collected on a slide and stained with cresyl violet. Slides were examined for placement of the cannula tip and evidence of infusion.

Brains collected for further immunofluorescence analysis to determine the location of viral expression were post-fixed overnight at 4°C. Coronal sections (30 µm) were collected with a vibratome (VT1000, Leica Microsystems) and stored at -20°C in a solution containing 30% ethylene glycol, 30% glycerol in PB, until they were rinsed four times for 10 min in PBS, mounted on slides and cover-slipped in Vectashield mounting medium (VEH1400, Vector Laboratories). Images were obtained using an Olympus FV1000 confocal microscope. For each rat, sections were selected along the rostral-caudal axis of dorsal hippocampus.

The location of the cannula in Experiment 1a and 1b and the extent of the DREADDs expression Experiment 2, 3a and 3b was determined using the boundaries defined by Paxinos and Watson[22] (**Extended Data Fig. 1**). Animals were excluded when the placement of the guide cannulas were misplaced, when the DREADDs expression was not in the boundaries of the targeted region, or when expression was minimal or not observed or when there was an infection at the cannula or injection site.

Cannula placements and evaluation of virus transfection was conducted by 1-2 experimenters on an Olympus FV1000 confocal microscope, all of whom were blind to the experimental conditions.

Statistics

Data was collected automatically by Med-PC (versions 4 or 5) and was uploaded to MS excel automatically using MedPC2Excel software.

Sampling was random. Sample size was determined in accordance with both power calculations and prior experience.

According to G*power 3 if we control the per-family error rate at alpha = .05 and assume the default correlation of 0.5 among repeated measures and a minimum of 2 groups, a group size of n = 12 will provide 0.65 power (1-β) to detect a medium effect size (0.25 for this type of analysis according to Cohen's size of effect). For experiments with 4 groups, n = 9 will provide 0.65 power (1-β). We determined that this was sufficient power based on our prior experience, and the fact that we employed the slightly less conservative (but valid for orthogonal contrasts) per-contrast error rate. Our final group sizes were slightly smaller than this after exclusions.

All data was analyzed using complex orthogonal contrasts controlling the per-contrast error rate at α = .05 using the statistical procedure described by Hays[23]. The software used to for these analyses was PSY statistical software which is freely available for download here: https://www.psy.unsw.edu.au/research/research-tools/psy-statistical-program. For lever press acquisition data there was typically a between-subjects factor of group and a linear trend contrast to test the within-subjects acquisition of lever pressing. Some simpler analyses (e.g. Day 6 analyses in Experiment 1a) only had a between-subjects comparison of group. For test data there was a between-subjects factor of group and a within-subjects factor of lever (devalued vs. valued). If an interaction (or interactions) was detected, follow up simple effects analyses were calculated to determine the source of the interaction. All statistical analyses can be found in the relevant data files at the link: https://osf.io/vd4an/?view_only=b161002919a24ca196ce23f7b2df84ad. All of the statistical analyses used can be considered to be two-tailed. Data distribution was assumed to be normal, but this was not formally tested.

Finally, suppression ratios were calculated separately for the devalued and valued levers for each animal according to the following equation:

$$SR = \frac{Lever\ press\ rate\ on\ test}{(Lever\ press\ rate\ on\ test + Lever\ press\ rate\ during\ training)}$$

A ratio of below 0.5 indicates that responding was suppressed during training, whereas a ratio of 0.5 or above indicates no suppression. These calculations were performed to determine whether animals suppressed responding on the devalued and/or valued levers on test relative to presses on that same lever training. Results were analyzed using the same complex orthogonal contrasts that were applied to the raw test data, and if an interaction was detected, follow up simple effects were calculated for the valued and devalued levers (separately). Results of this analysis are shown in the caption for **Extended Data Fig. 4**.

Please refer to the **Life Sciences Reporting Summary** for further information.

Data Availability Statement

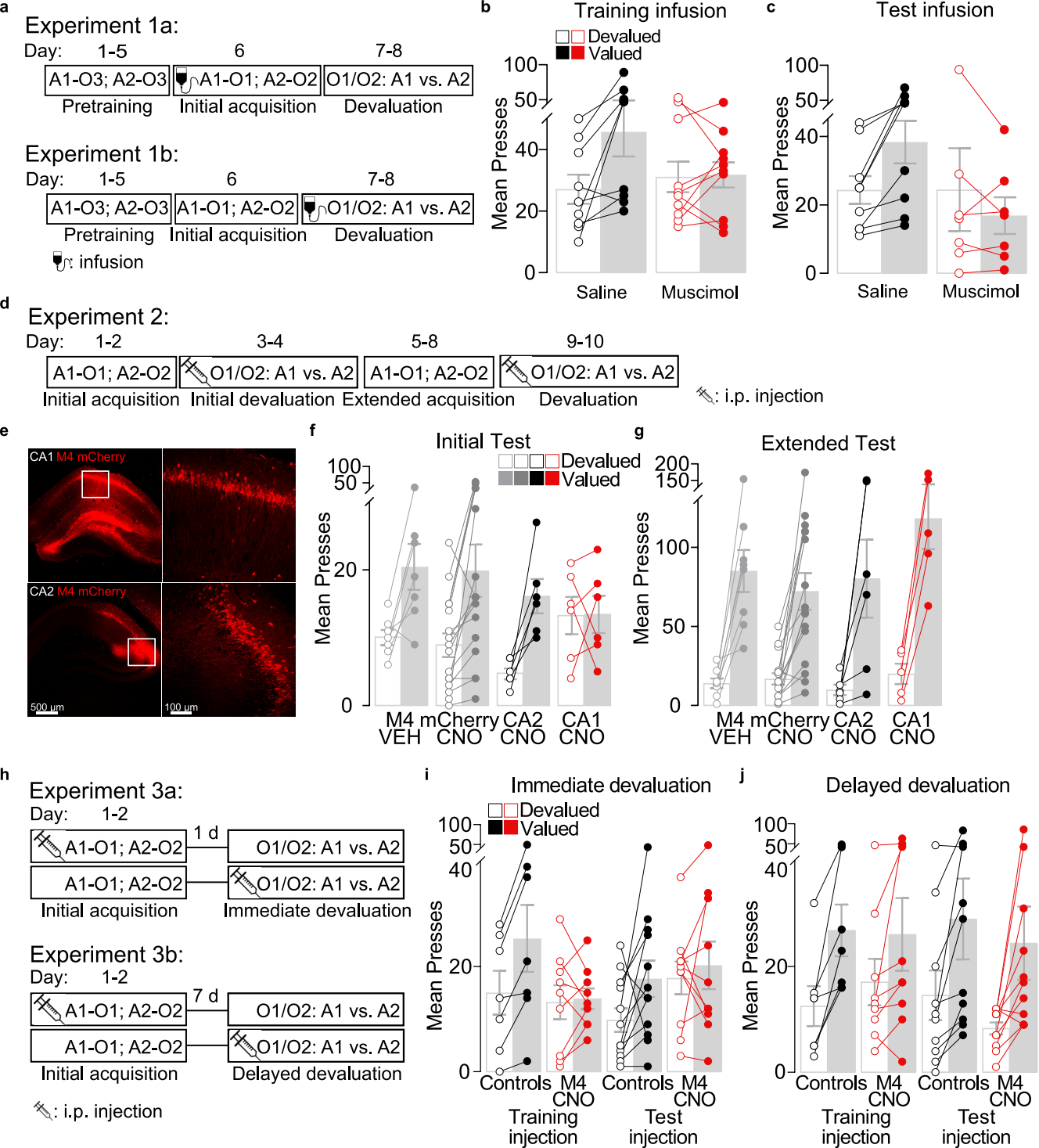
Research data for this article (**Figs. 1-2** and **Extended Data Figs. 2-4**) is available for download at the following link: https://osf.io/vd4an/?view_only=b161002919a24ca196ce23f7b2df84ad.

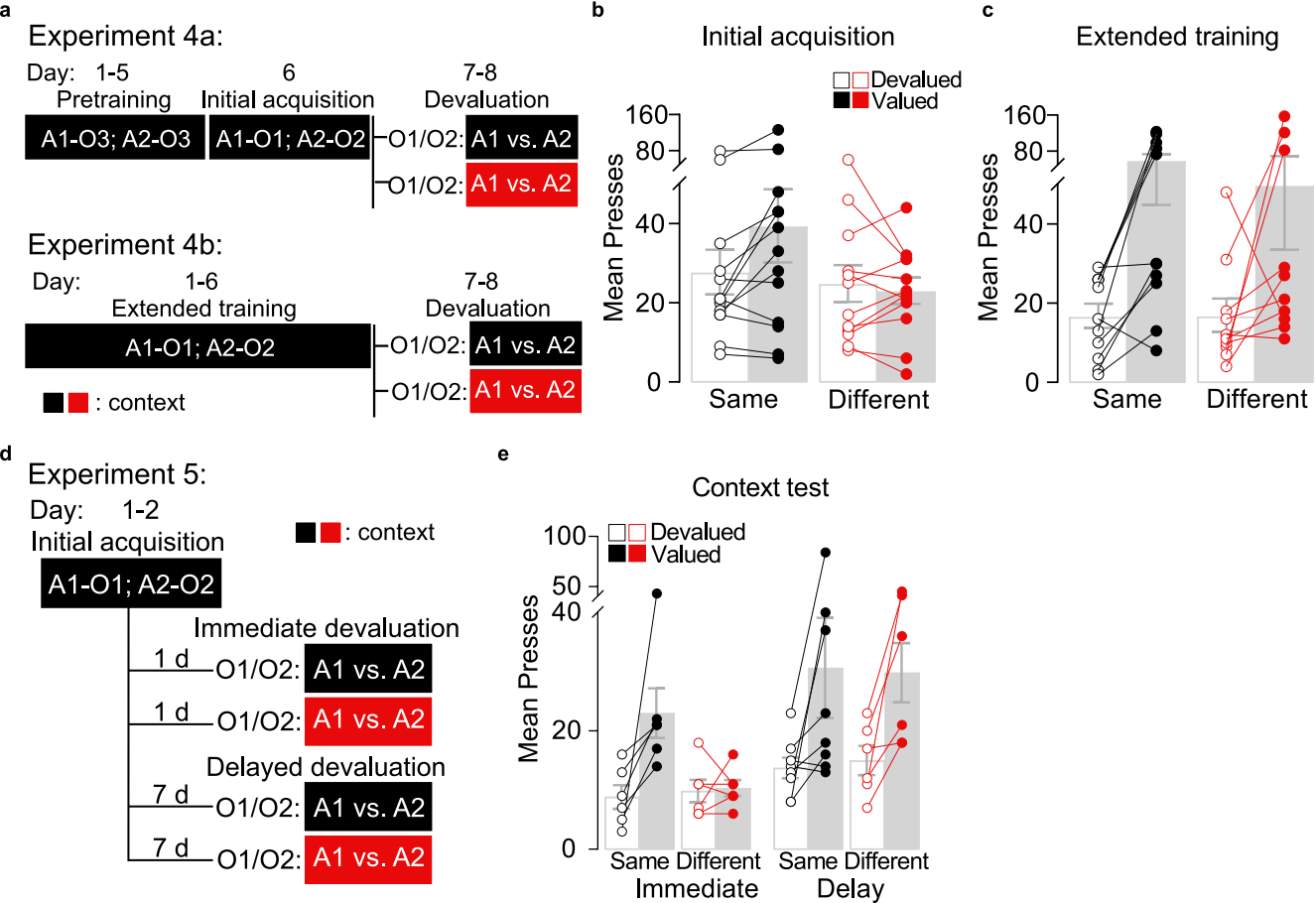
Methods-only References

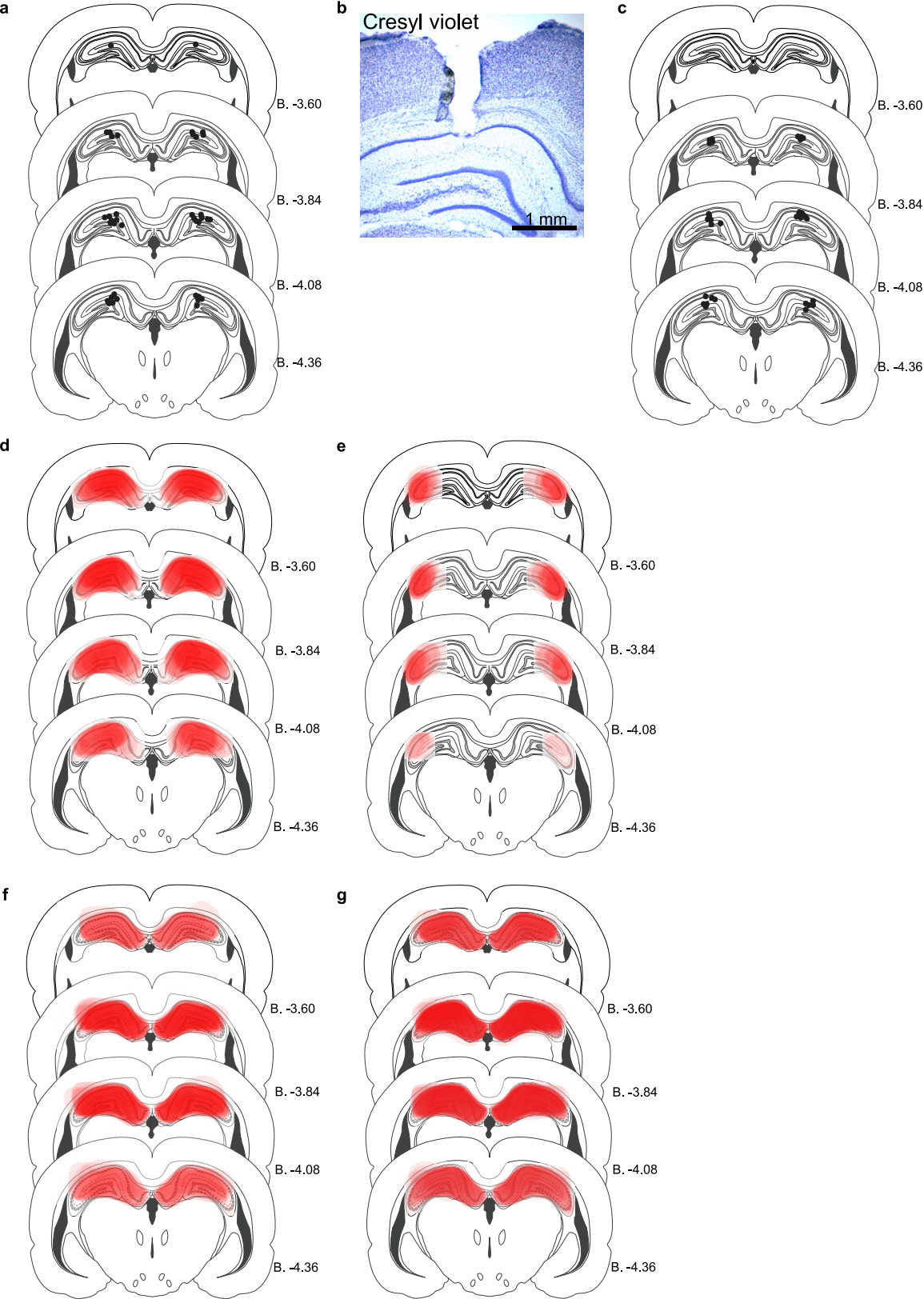
21. Kosaki, Y. & Dickinson, A. Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *Journal of Experimental Psychology: Animal Behavior Processes* **36,** 334–342 (2010).

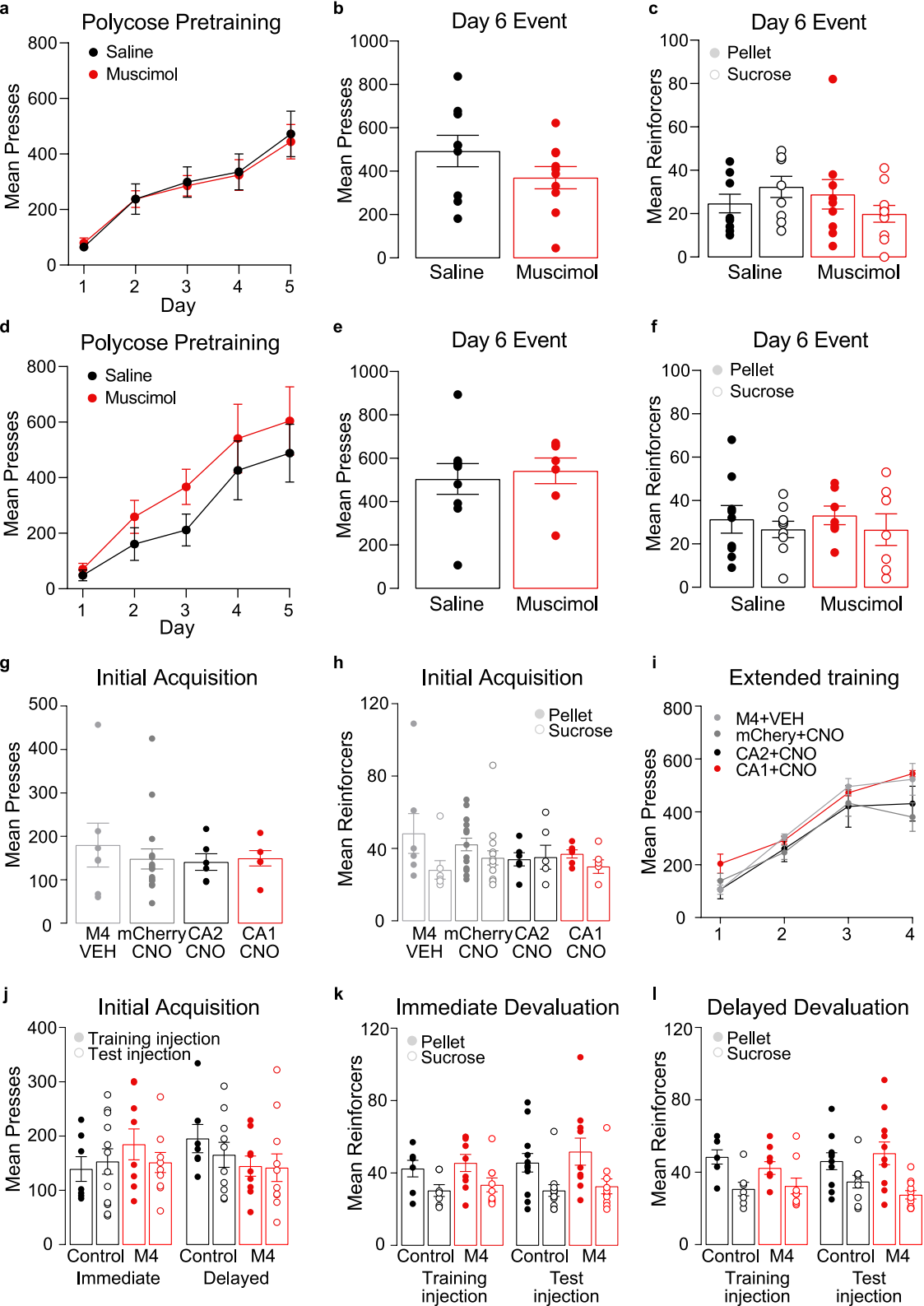
22. Paxinos, G. & Watson, C. *Paxinos and Watson's The Rat Brain in Stereotaxic Coordinates*. (Elsevier Academic Press, 2014).

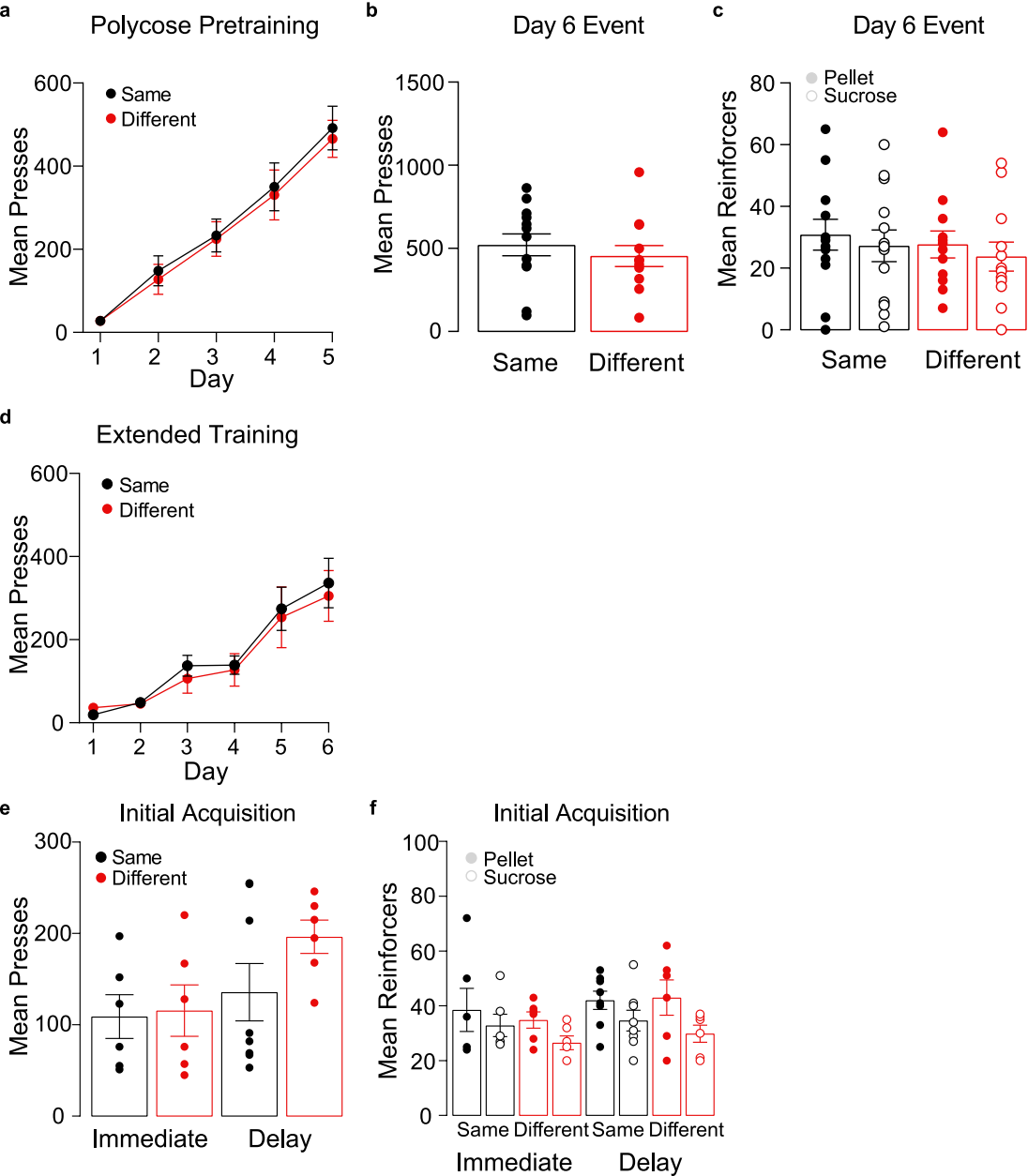
23. Hays, W. L. *Statistics for the Social Sciences*. (New York, Holt, Rinehart & Winston, 1973).

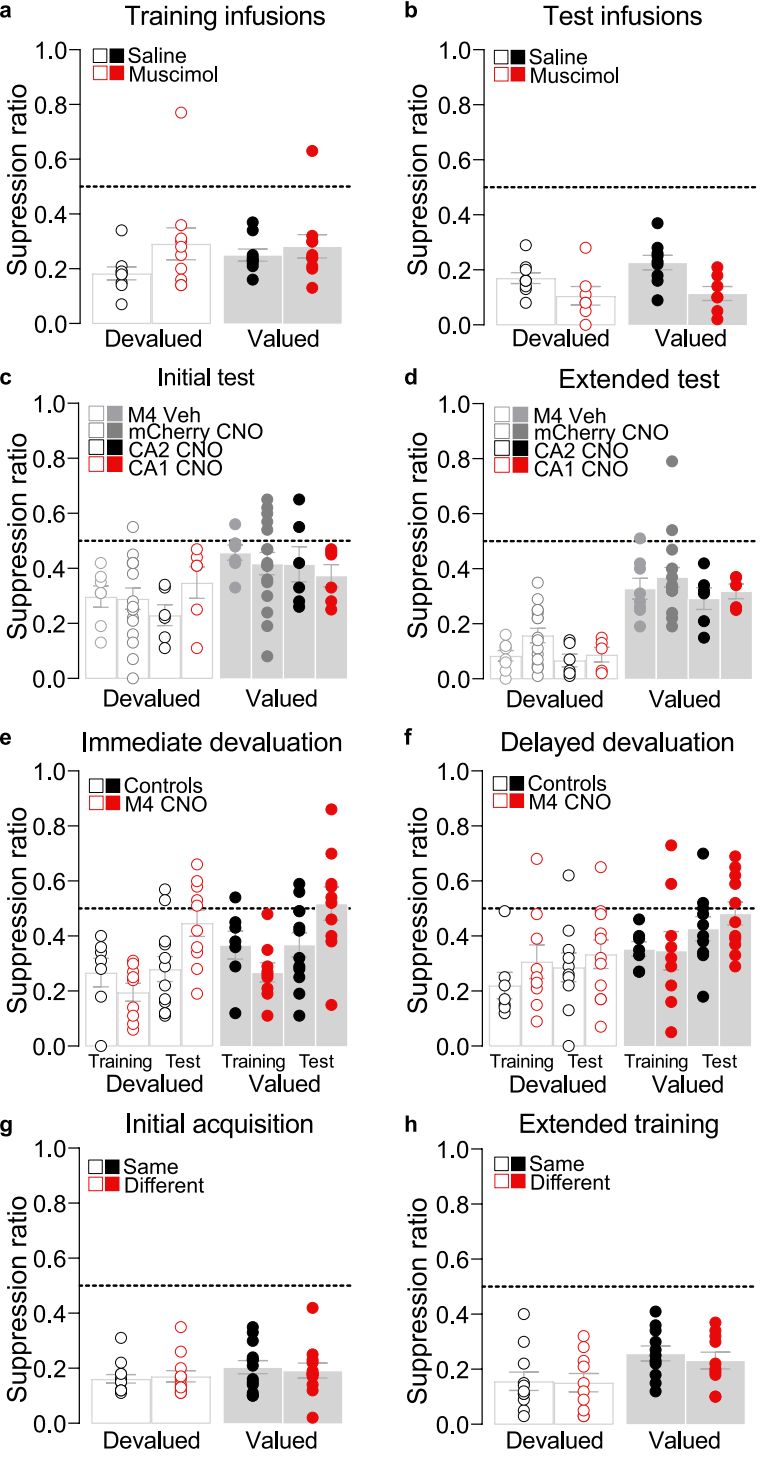

a
Experiment 1a:
Day: 1-5 6 7-8
A1-O3; A2-O3
A1-O1; A2-O2
O1/O2: A1 vs. A2
Pretraining
Initial acquisition
Devaluation
Experiment 1b:
Day: 1-5 6 7-8
A1-O3; A2-O3
A1-O1; A2-O2
O1/O2: A1 vs. A2
Pretraining
Initial acquisition
Devaluation
: infusion
b
Training infusion
Devalued
Valued
Mean Presses
100
50
40
20
0
Saline
Muscimol
c
Test infusion
Mean Presses
100
50
40
20
0
Saline
Muscimol
d
Experiment 2:
Day: 1-2 3-4 5-8 9-10
A1-O1; A2-O2
O1/O2: A1 vs. A2
A1-O1; A2-O2
O1/O2: A1 vs. A2
Initial acquisition
Initial devaluation
Extended acquisition
Devaluation
: i.p. injection
e
CA1 M4 mCherry
CA2 M4 mCherry
500 μm
100 μm
f
Initial Test
Devalued
Valued
Mean Presses
100
50
20
0
M4 mCherry VEH
CA2 CNO
CA1 CNO
g
Extended Test
Mean Presses
200
150
100
50
0
M4 mCherry VEH
CA2 CNO
CA1 CNO
h
Experiment 3a:
Day: 1-2
A1-O1; A2-O2
1 d
O1/O2: A1 vs. A2
A1-O1; A2-O2
O1/O2: A1 vs. A2
Initial acquisition
Immediate devaluation
Experiment 3b:
Day: 1-2
A1-O1; A2-O2
7 d
O1/O2: A1 vs. A2
A1-O1; A2-O2
O1/O2: A1 vs. A2
Initial acquisition
Delayed devaluation
: i.p. injection
i
Immediate devaluation
Devalued
Valued
Mean Presses
100
50
40
20
0
Controls
M4 CNO
Training injection
Test injection
j
Delayed devaluation
Mean Presses
100
50
40
20
0
Controls
M4 CNO
Training injection
Test injection


a
Experiment 4a:
Day: 1-5
Pretraining
6
Initial acquisition
7-8
Devaluation
A1-O3; A2-O3
A1-O1; A2-O2
O1/O2: A1 vs. A2
O1/O2: A1 vs. A2
Experiment 4b:
Day: 1-6
Extended training
7-8
Devaluation
A1-O1; A2-O2
O1/O2: A1 vs. A2
O1/O2: A1 vs. A2
: context
b
Initial acquisition
Devalued
Valued
Mean Presses
160
80
40
20
0
Same
Different
c
Extended training
Mean Presses
160
80
40
20
0
Same
Different
d
Experiment 5:
Day: 1-2
Initial acquisition
: context
A1-O1; A2-O2
Immediate devaluation
1 d
O1/O2: A1 vs. A2
1 d
O1/O2: A1 vs. A2
Delayed devaluation
7 d
O1/O2: A1 vs. A2
7 d
O1/O2: A1 vs. A2
e
Context test
Devalued
Valued
Mean Presses
100
50
40
20
0
Same
Different
Same
Different
Immediate
Delay


a
B. -3.60
B. -3.84
B. -4.08
B. -4.36
b
Cresyl violet
1 mm
c
B. -3.60
B. -3.84
B. -4.08
B. -4.36
d
B. -3.60
B. -3.84
B. -4.08
B. -4.36
e
B. -3.60
B. -3.84
B. -4.08
B. -4.36
f
B. -3.60
B. -3.84
B. -4.08
B. -4.36
g
B. -3.60
B. -3.84
B. -4.08
B. -4.36


a
Polycose Pretraining
Saline
Muscimol
Mean Presses
Day
b
Day 6 Event
Mean Presses
Saline
Muscimol
c
Day 6 Event
Pellet
Sucrose
Mean Reinforcers
Saline
Muscimol
d
Polycose Pretraining
Saline
Muscimol
Mean Presses
Day
e
Day 6 Event
Mean Presses
Saline
Muscimol
f
Day 6 Event
Pellet
Sucrose
Mean Reinforcers
Saline
Muscimol
g
Initial Acquisition
Mean Presses
M4 VEH
mCherry CNO
CA2 CNO
CA1 CNO
h
Initial Acquisition
Pellet
Sucrose
Mean Reinforcers
M4 VEH
mCherry CNO
CA2 CNO
CA1 CNO
i
Extended training
M4+VEH
mChery+CNO
CA2+CNO
CA1+CNO
Mean Presses
j
Initial Acquisition
Training injection
Test injection
Mean Presses
Control
M4
Immediate
Delayed
k
Immediate Devaluation
Pellet
Sucrose
Mean Reinforcers
Control
M4
Training injection
Test injection
l
Delayed Devaluation
Pellet
Sucrose
Mean Reinforcers
Control
M4
Training injection
Test injection


a
Polycose Pretraining
Same
Different
Mean Presses
Day
b
Day 6 Event
Mean Presses
Same
Different
c
Day 6 Event
Pellet
Sucrose
Mean Reinforcers
Same
Different
d
Extended Training
Same
Different
Mean Presses
Day
e
Initial Acquisition
Same
Different
Mean Presses
Immediate
Delay
f
Initial Acquisition
Pellet
Sucrose
Mean Reinforcers
Same Different Same Different
Immediate
Delay


a
Training infusions
Saline
Muscimol
Supression ratio
1.0
0.8
0.6
0.4
0.2
0.0
Devalued
Valued
b
Test infusions
Saline
Muscimol
Supression ratio
Devalued
Valued
c
Initial test
M4 Veh
mCherry CNO
CA2 CNO
CA1 CNO
Suppression ratio
Devalued
Valued
d
Extended test
M4 Veh
mCherry CNO
CA2 CNO
CA1 CNO
Suppression ratio
Devalued
Valued
e
Immediate devaluation
Controls
M4 CNO
Suppression ratio
Training
Test
Training
Test
Devalued
Valued
f
Delayed devaluation
Controls
M4 CNO
Suppression ratio
Training
Test
Training
Test
Devalued
Valued
g
Initial acquisition
Same
Different
Suppression ratio
Devalued
Valued
h
Extended training
Same
Different
Suppression ratio
Devalued
Valued
i
Context test
Same
Different
Suppression ratio
Immediate
Delay
Immediate
Delay
Devalued
Valued