Probability and rate of reinforcement in negative prediction error learning

David J. Sanderson, Joseph M. Austen, Anthony McGregor and Jasmin A. Strickland

Department of Psychology, Durham University, South Road, DH1 3LE, UK

Correspondence to David J. Sanderson, david.sanderson@durham.ac.uk

Probability and rate of reinforcement in negative prediction error learning

David J. Sanderson, Joseph M. Austen, Anthony McGregor and Jasmin A. Strickland

Department of Psychology, Durham University, South Road, DH1 3LE, UK

Correspondence to David J. Sanderson, david.sanderson@durham.ac.uk

Trial based theories of associative learning propose that learning is sensitive to the probability of reinforcement signalled by a conditioned stimulus (CS). Learning, however, is often sensitive to reinforcement rate rather than probability of reinforcement per trial, suggesting that temporal properties of cues may be more important than trial-based properties. In four experiments, the role of probability of reinforcement per trial was examined in appetitive Pavlovian conditioning in mice under conditions in which reinforcement rate was controlled. Experiments 1 and 2 examined the loss of conditioned responding caused by overexpectation of reinforcement. Probability of reinforcement per trial failed to affect acquisition and summation of conditioned responding and failed to affect overexpectation. It also failed to affect extinction of conditioned responding in Experiments 3 and 4. Experiments 2-4 contained non-reinforced trials in which responding at the offset of the conditioned stimulus could be measured. These probe trials did reveal an effect of probability of reinforcement per trial. Cues associated with 100% reinforcement elicited greater post-CS responding than cues associated with 50% reinforcement. The effect was also evident in summation trials (in Experiment 2) in which two 100% or 50% reinforced cues were presented in compound. The results show that mice learn about rate and probability information, but reinforcement rate determines anticipatory responding during the CS. Probability of reinforcement determines responding at the expected time of reinforcement. Thus, learning occurs continuously over the duration of experience and per episode of experience independent of duration.

Keywords: Learning, conditioned responding, trial, time, mice

A consequence of describing changes in associative strength between a conditioned stimulus (CS) and an unconditioned stimulus (US) in a trial-by-trial manner is that associative strength is predicted to be proportional to the probability of reinforcement signalled by a cue. An alternative approach is to describe changes in learning that happen iteratively over some constant unit of time such that the process is continuous rather than occurring in discrete events of potentially differing durations. This results in learning being updated not only during the periods of reinforcement but also in the periods of nonreinforcement during a CS prior to the occurrence of the US. Therefore, associative strength will be influenced by the reinforcement rate of the CS rather than the probability of reinforcement per trial. Indeed, cues may differ in their predicted probability of reinforcement per trial, but if they are matched for overall rate of reinforcement then learning should proceed similarly (Harris et al., 2015).

Research in appetitive Pavlovian magazine approach conditioning has shown that rats learn about both the reinforcement rate of cues over cumulative exposure, independent of how the exposure is structured in terms of trials, and the probability of reinforcement per trial, independent of reinforcement rate. Evidence for learning being determined by reinforcement rate rather than probability of reinforcement per trial has been shown in the acquisition of conditioned responding. Harris et al. (2015) found that, over a wide range of parameters, there was no effect of partial reinforcement on acquisition of conditioned responding when cues were matched for their overall reinforcement rate by manipulation of trial durations. For example, a cue that was on average 10 s long and reinforced on a third of its trials elicited levels of conditioned responding over acquisition similar to a cue that was on average 30 s long and reinforced on every trial. Furthermore, when probability of reinforcement per trial is held constant but reinforcement rate is

manipulated by differences in the average trial duration, the strength of conditioned responding increases as a function of reinforcement rate (Harris & Carpenter, 2011). This pattern of results has also been found in magazine approach behaviour in mice (Austen et al., 2021; Austen & Sanderson, 2019; Strickland et al., 2024).

Evidence for a role of probability of reinforcement per trial comes from research on the partial reinforcement extinction effect (PREE) in which conditioned responding reduces more slowly during extinction learning for a previously partially reinforced cue than a previously continuously reinforced cue (Amsel, 1962; Capaldi, 1966). Chan and Harris (2019) found the PREE occurred in Pavlovian conditioning for cues that were matched for their previous reinforcement rate. Therefore, slower extinction for partially reinforced cues could not be attributed to rate based explanations of the effect (see Gallistel & Gibbon, 2000), but reflected solely the effect of probability of reinforcement per trial.

Although probability of reinforcement per trial does not affect acquisition of conditioned responding during a CS, it does affect responding in the moments immediately after the termination of a CS. Harris and Kwok (2018) examined acquisition with two partially reinforced cues that differed in their probability of reinforcement per trial. Post-CS responding was examined on the nonreinforced trials in the absence of reinforcement. Rats showed greater immediate post-trial responding to a cue that had a higher probability of reinforcement compared to a cue with a lower probability. Furthermore, when cues differed in rate of reinforcement but were matched in probability of reinforcement per trial, post-CS responding was similar.

The dissociation between reinforcement rate and probability of reinforcement per trial suggests that trial-based learning determines responding when the US is omitted (Harris, 2019). During acquisition, this effect is limited to post-CS responding. Thus, it does

not affect the rate at which animals respond during the CS. However, in extinction learning

when the overall contingency between the CS and US is altered, probability of

reinforcement per trial determines the rate at which responding extinguishes. Therefore,

probability of reinforcement may determine the extent of learning as a consequence of

negative prediction error. If probability of reinforcement per trial determines negative

prediction error generally then it should affect losses of conditioned responding when it

occurs not only in the absence of reinforcement, as in extinction learning, but also in the

presence of reinforcement, as in overexpectation of reinforcement (e.g., Kremer, 1978).

Thus, negative prediction error will occur when the US is expected but the CS is

nonreinforced and also when a CS is reinforced but the amount of reinforcement is less than

expected (i.e., overexpectation).

The purpose of the present experiments was to test the circumstances under which

probability of reinforcement affects learning. Learning was assessed in Pavlovian appetitive

magazine approach conditioning in female mice. In all of the experiments, learning was

tested with cues that differed in probability of reinforcement but were matched for rate of

reinforcement. Experiments 1 and 2 examined the effect of probability of reinforcement on

overexpectation in which negative prediction occurs in the presence of reinforcement.

Experiments 3 and 4 examined extinction learning in the absence of reinforcement.

Experiment 1

Experiment 1 tested the role of probability of reinforcement per trial in overexpectation. Overexpectation occurs when two previously reinforced CSs (A+ and X+) are presented in compound and reinforced to the same extent as the previous individual CS presentations (AX+). At test, the compound conditioning results in a loss of conditioned responding to the individual cues. The Rescorla-Wagner (1972) model accounts for the loss of conditioned responding by proposing that the associative strength of the individual cues is summed when first presented in compound. Because the summed associative strength of the compound is greater than the associative strength supported by the level of reinforcement presented, the associative strength of the individual cues is reduced over the course of compound training. In other words, when initially presented with the compound, animals expect a level of reinforcement that is greater than is presented.

The design of Experiment 1 is shown in Table 1. Overexpectation was tested in a within-subjects design. The manipulation of probability of reinforcement per trial, however, was between-subjects. Mice were trained with four cues, A, X, B and Y. Cues A and X were used to test overexpectation. For mice in group 100%, cues A and X were on average 20 s in duration and reinforced on every trial. For group 50%, cues A and X were on average 10 s in duration and reinforced on 50% of trials. Therefore, although the groups differed in probability of reinforcement per trial, because of the difference in average duration, the cumulative rate of reinforcement of A and X was matched across groups. In stage 2, mice received trials in which A and X were presented in compound, reinforced on every trial and on average 20 s in duration. If overexpectation reflects an expectation of the summed

amount of reinforcement per trial, then during stage 2 there should be an overexpectation effect for group 100% but this effect will be attenuated or abolished for group 50%. If associative strength corresponds linearly to probability of reinforcement then because A and X were both associated with reinforcement on every trial, group 100% should expect twice as much reinforcement as was presented. In contrast, for group 50%, because A and X were previously associated with reinforcement on 50% of trials (and thus on average 50% reinforcement per trial), when presented with AX, they should expect one amount of reinforcement per trial, and consequently, overexpectation will not occur. If, however, overexpectation reflects summation of reinforcement rates, because the rates of reinforcement were matched across stage 1, both groups should show a similar overexpectation effect.

It is important to note that the Rescorla-Wagner (1972) model does not assume that associative strength corresponds linearly to probability of reinforcement because it assumes that the learning rate in the presence of reinforcement is greater than in the absence of reinforcement. This assumption is necessary for the model to account for findings such as the relative validity effect (Wagner et al., 1968). In contrast, however, for the model to be able to account for the correspondence between reinforcement rate and response rate in compound conditioning (Harris et al., 2012) it has to be assumed that the learning rates are equal in the presence and absence of reinforcement. Given these issues, the hypothesis is that the manipulation of probability should at the least attenuate overexpectation if not fully abolish overexpectation.

Cues B and Y served as control cues for the overexpectation effect. Their treatment was the same for both groups. In stage 1, cue B was on average 20 s long and reinforced on 100% of trials. Cue Y was also on average 20 s long but nonreinforced. In stage 2, in addition

to AX compound trials, mice were presented with trials in which B and Y were presented in compound, reinforced on every trial and on average, 20 s in duration. Because B was associated with reinforcement on every trial and Y was associated with non-reinforcement, when reinforced in compound on 100% of trials, there will be no overexpectation. The crucial comparison for overexpectation was between cues A and B. Whereas responding should reduce for A as a consequence of the compound training in stage 2, responding for B should not.

Various control methods for overexpectation have been used that range from methods similar to those used here to procedures such as continued reinforced training of a cue not presented in compound during stage 2 or no training at all during stage 2 (Arico & McNally, 2014; Garfield & McNally, 2009; Haney et al., 2010; Holland, 2016; Iordanova et al., 2016; Kehoe & White, 2004; Khallad & Moore, 1996; Kremer, 1978; Lattal & Nakajima, 1998; Lay et al., 2023a, 2023b; Lay et al., 2019; Lucantonio et al., 2015; McNally et al., 2004; Rescorla, 1970, 1999, 2006, 2007; Ruprecht et al., 2014; Sissons & Miller, 2009; Terao et al., 2022; Witnauer & Miller, 2009). The advantage of the use of cue B and Y as a within-subjects control procedure for overexpectation is that it allows various explanations for the reduced responding to A be ruled out. Because the procedure is within-subjects, differences between responding to A and B cannot be attributed to general effects on response levels, which may occur in between-subjects designs. Furthermore, certainly for group 100%, A and B are matched for manner of presentation and reinforcement contingency and pairings across stage 1 and 2. Therefore, differences in responding to A and B during the test phase cannot be attributed to differences in exposure in the previous stages that may, subsequently, result in differences in generalisation decrement. Each aspect of the

procedure designed to result in overexpectation for cue A is matched for cue B except for the overexpectation of reinforcement.

Cue durations were variable over trials. Variable durations reduce the influence of reinforcement timing on response rates within trials (Harris & Carpenter, 2011). Harris et al. (2015) showed that response rates are sensitive to reinforcement rates, independent of trial durations and probability of reinforcement, when animals are unable to time the occurrence of reinforcement. In addition to these factors, variable durations cues were chosen for the current study to reduce any generalisation decrement that may occur when the average duration of cues is changed between stage 1 and 2 for group 50%.

Methods

Subjects: Thirty-two naïve female C57BL/6J mice (Charles River UK Ltd), approximately 10 weeks old at the start of testing, with a mean free-feeding weight of 18.6g (range: 16.4-20.3g) were used. Mice were caged in groups of eight in a temperature-controlled housing room on a 12-hour light-dark cycle (lights on at 8:00 am). Prior to the start of the experiment, the weights of the mice were reduced by being placed on a restricted diet. Mice were then maintained at 85% of their free-feeding weights throughout the experiment. Mice had ad libitum access to water in their home cages. All procedures were in accordance with the United Kingdom Animals Scientific Procedures Act (1986); under project license number P8B95F992.

Apparatus: A set of eight identical operant chambers (interior dimensions: 15.9 x 14.0 x 12.7 cm; ENV-307A, Med Associates, Inc., Fairfax, VT, USA), enclosed in sound-attenuating cubicles (ENV-022V) were used. The operant chambers were controlled by Med-PC IV

software (SOF-735). The side walls were made from aluminium, and the front and back walls and the ceiling were made from clear Perspex. The chamber floors each comprised a grid of stainless-steel rods (0.32-cm diameter), spaced 0.79 cm apart, and running perpendicular to the front of the chamber (ENV-307A-GFW). A food magazine (2.9 x 2.5 x 1.9 cm; ENV-303M) was situated in the centre of one of the sidewalls of the chamber, into which sucrose pellets (14 mg, TestDiet) could be delivered from a pellet dispenser (ENV-203-14P). An infrared beam (ENV-303HDA) across the entrance of the magazine was used to record head entries at a resolution of 0.1 s. A fan (ENV-025F) was located within each of the sound-attenuating cubicles and was turned on during sessions, providing a background sound pressure level of approximately 65 dB. Auditory stimuli were provided by a white noise generator (ENV-325SM) outputting a flat frequency response from 10 to 25,000 Hz at 75 dB and a clicker (ENV-335M) operating at a frequency of 4 Hz at 75 dB. Visual stimuli were a 2.8 W house light (ENV-315M) which could illuminate the entire chamber, and two LEDs (ENV-321M) positioned to the left and right of the food magazine, which provided more localised illumination.

Procedure: Mice were randomly assigned to one of two groups (N = 16 per group) that differed in their training during stage 1. Group 50% received 12 sessions of training with two short duration cues, A and X, that were on average, across trials, 10 s (range: 2-18 s) in duration, and two other cues, B and Y, that were on average, across trials, 20 s (range: 2–38 s) in duration. The durations of the cues varied across trials according to a flat distribution with the constraint that the mean duration for A and X, and B and Y, was 10 s and 20 s respectively, within a session. Cues A and X were both reinforced on a random 50% of trials. Cue B was reinforced on 100% of trials and cue Y was nonreinforced. Cues were reinforced

by the presentation of a sucrose pellet at the termination of the cue presentation. Group 100% received similar training except that cues A and X were variable duration 20 s cues rather than 10 s and were reinforced on 100% of trials.

Within each session, Group 100% received three trials of each trial type and group 50% received six trials of A and X each and three trials of B and Y each. Trial types were presented in a random order with the constraint that for group 50%, each block of six trials consisted of two A trials, two X trials, one B trial and one Y trial, and for group 100%, each block of four trials consisted of one trial of each of the four different trial types. The difference in trial numbers for the two groups resulted in both groups receiving the same cumulative exposure to cues A and X and the same number of pairings with reinforcement. In order to equate the session lengths, the fixed inter-trial interval for group 50% was 120 s (from CS offset to CS onset) and for group 100%, it was 180 s. Sessions lasted 40 minutes.

Within both groups, for half of the mice, cues A and B were auditory (white noise and clicker) and X and Y were visual (house light and alternating left and right flashing LEDs, 0.5s on, 0.5 s off). Within the subgroups allocated auditory A and B and visual X and Y cues, the allocation of individual stimuli to cue types was fully counterbalanced in a two-by-two design. For the other half of mice, A and B were visual and X and Y were auditory and the allocation of stimuli to cue types was also fully counterbalanced in a two-by-two design.

In stage 2, the groups received identical treatment. Mice received three sessions in which cues were now presented in the compounds AX and BY. Each compound was presented for a variable 20 s duration and was reinforced at the offset of the compound. Each compound was presented for nine trials per session. Trial types were presented in a random order with the constraint that each block of six trials contained three trials of each trial type. The inter-trial interval was a fixed duration of 150 s. Sessions lasted 51 minutes.

In stage 3, mice received two sessions that consisted of reinforced presentations of the compounds AX and BY and nonreinforced probe presentations of A and B. All trial types were a fixed duration of 20 s. Sessions consisted of six presentations of A and B each and three presentations of AX and BY each. Each block of six trials consisted of two trials of A and B each and one trial of AX and BY each. The order of trial types within a six-trial block was random with the constraint that within groups the order of A and B was counterbalanced with respect to the allocation of stimuli to trial types. The inter-trial interval was a fixed duration of 150 s. Sessions lasted 51 minutes.

Data and Statistical Analysis: The frequency of head entries into the food magazine was recorded per-second during the CS presentations and the for the 10 s pre-CS period. Response rates are expressed as responses per minute (RPM). Across the different stages, response rates were analysed using multifactorial ANOVAs that included the factor of group and cue identity. Interactions were analysed with simple main effects analysis using the pooled error term from the original ANOVA or separate ANOVAs for repeated measures with more than two levels. Where sphericity of within-subjects variables could not be assumed, a Greenhouse-Geisser correction was applied.

In instances in which statistical tests of key predictions failed to be significant, the strength of evidence for the null hypothesis was tested with Bayesian analyses carried out in JASP (Team, 2024) using the default priors. For Bayesian ANOVAs, Bayes factors for the main effects or interactions on which the evidence for the hypothesis depends reflect comparison of models that include the main effect or interaction with similar models that exclude the main effect or interaction.

In order to simplify the analysis of pre-CS response rates, the analyses are collapsed across trial types because for all experiments, the order of trial types was either randomised within particular constraints or was counterbalanced across animals.

Transparency and Openness: The study was not pre-registered. The data are available upon request. Sample sizes were approximately N=16 per between-subjects condition. This number was chosen to ensure sufficient numbers for counterbalancing of multiple factors and to achieve 80% power for effects size approximately equal to η_p^2 = .20.

Results and discussion

Stage 1: Acquisition of responding is shown in Figure 1. Mice acquired conditioned responding to cues A, B and X. Responses to Y initially increased but then decreased over training. Response rates were similar between the two groups. A 4 (cue: A, B, X and Y) by 12 (session) by 2 (group: 50%, 100%) ANOVA was conducted. There were significant effects of cue, $F(3,90)$ = 14.48, $p < 0.001$, η_p^2 = .33, 90% CI [.18, .42], and session, $F(11,330)$ = 28.50, $p < 0.001$, η_p^2 = .49, 90% CI [.41, .52], and a significant interaction of factors, $F(33,990)$ = 9.02, $p < 0.001$, η_p^2 = .23, 90% CI [.17, .24]. The effect of group was not significant, $F < 1$, and there were no other significant interactions of factors, F-values < 1. Post-hoc t-tests (Holm-Bonferroni corrected) that explored the effect of cue showed there was no significant difference between cues A, B and X (p-values > 0.5) but each of those cues elicited significantly greater responding than the nonreinforced cue, Y (p-values < 0.001). Simple main effects analysis of the cue by session interaction revealed that there was a significant effect of cue from session 3 onwards (p-values < 0.04), but not before (p-values > 0.20).

A Bayesian 2 (group) by 3 (cue: A, B and Y) by 12 (session) ANOVA was conducted to test evidence for the null hypothesis that there was no difference in the acquisition of responding to partially and continuously reinforced cues. For the cue by group interaction, BF_{incl} = 0.37. For the cue by group by session interaction, BF_{incl} = 1x10^{-5}. Therefore, the results provided at least 2.7 times evidence for the null hypothesis over the alternative hypothesis.

Additional analyses were conducted comparing response rates between cues under matched conditions (see Supplemental Material). The analyses compared responding across similar durations within the first trial of each trial type across sessions in the latter half of training (see Supplemental Figure 3). There was no significant difference between partially and continuously reinforced cues and Bayesian analyses found evidence in favour of the null hypothesis.

Pre-CS responding, collapsed across trial types, reduced over the course of training for both groups (see Figure 1). There was a significant effect of session, $F(11,330)$ = 26.19, p < 0.001, η_p^2 = .47, 90% CI [.39, .50], no significant effect of group, $F(1,30)$ = 1.80, p = 0.18 and no significant interaction of group and session, $F(11,330)$ = 1.33, p = 0.21.

Stage 2: Performance during the compound conditioning phase is shown in Figure 2. Mice initially showed greater responding to the compound AX than to BY but by the third session response rates for AX reduced to a level that was less than for BY. A 2 (compound: AX, BY) by 2 (group: 50%, 100%) by 3 (session) ANOVA revealed that there was no significant effect of compound, F < 1, but there was a significant compound by session interaction, $F(2,60)$ = 17.66, p < 0.001, η_p^2 = .37, 90% CI [.20, .49]. Simple main effects analysis showed that there was a significant effect of session for AX, $F(2,60)$ = 11,59, p < 0.001, but not for BY, $F(2,60)$ = 2.5, p = 0.09. There was a significant effect of compound on session 1, $F(1,30)$ =

4.44, p = 0.04, but not on sessions 2 and 3 (largest $F(1,30) = 3.59$, $p = 0.07$). The effect of session was not significant, $F(2,60) = 1.12$, $p = 0.33$. There were no other significant main effects or interactions of factors, F-values < 1.

Pre-CS response rates increased over the course of Stage 2 training. There was a significant effect of session, $F(2,60) = 11.08$, $p < 0.01$, $\eta_p^2 = .27$, 90% CI [.11, .39]. There was no significant effect of group or interaction between group and session, F-values < 1.

Test: The results of the overexpectation test are shown in Figure 3. Both groups responded more to cue B than cue A and the extent of the difference in responding to cues was similar across groups. In a 2 (cue: A, B) x 2 (group) ANOVA, it was found that there was a significant effect of cue, $F(1, 30) = 19.56$, $p < 0.001$, $\eta_p^2 = .39$, 90% CI [.16, .55]. The effect of group and the cue by group interaction was not significant, F-values < 1.

In order to test the strength of evidence for the null hypothesis a Bayes factor was calculated in JASP using the default priors, testing the prediction that the overexpectation effect would be greater for group 100% than group 50%. A Bayesian between-subjects t-test was conducted comparing the difference between response rates for the two cues (response rate for cue B minus the response rate for A). $BF_{10} = 0.31$, thus providing 3.2 times evidence for the null hypothesis over the alternative hypothesis.

An additional analysis of responding to the reinforced compounds, AX and BY, and the pre-CS rates of responding (see Figure 4) showed no significant effect of group, F < 1, or interaction of factors, F < 1. Response rates significantly differed between the compounds and the pre-CS period, $F(2,60) = 9.77$, $p < 0.001$, $\eta_p^2 = .25$, 90% CI [.09, .37]. Post-hoc t-tests (Holm-Bonferroni corrected) showed that there was no significant difference between responses to the two reinforced compounds (p = 0.65), but pre-CS response rates were significantly lower than both compounds, p-values < 0.003.

Probability of reinforcement per trial failed to affect overexpectation. Therefore, the results did not support the hypothesis that negative prediction error reflects an expectation of reinforcement that is dependent on probability information. As anticipated, there was also no significant effect on acquisition of responding. There was also no significant effect on summation of conditioned responding, consistent with other findings showing that responding to compounds reflects the summed rates of reinforcement of the individual cues (Andrew & Harris, 2011; Harris et al., 2012).

Experiment 2a and 2b

The primary purpose of Experiments 2a and 2b was to rule out potential accounts of the failure to observe an effect of probability of reinforcement per trial in Experiment 1. In Experiment 1, manipulation of probability was achieved by group 50% switching from 50% reinforcement per trial in stage 1 to 100% in stage 2. It is possible that the reduced responding in group 50% to cue A compared to cue B was due to generalisation decrement caused by the manipulation of probability and/or the average trial duration between stages in 1 and 2 rather than overexpectation. Such an effect may have masked an effect of probability of reinforcement per trial on overexpectation. Therefore, in Experiments 2a and 2b, the procedure of Experiment 1 was replicated but instead of group 50% switching from partial reinforcement in stage 1 to continuous reinforcement in stage 2, probability of reinforcement was now manipulated by group 100% switching to 50% reinforcement in stage 2 (see Table 2). If overexpectation is observed when group 50% are maintained on 50% reinforcement per trial for both stage 1 and 2, then this would rule out the possibility that the low responding to A compared to B, in Experiment 1, was a consequence of generalisation decrement across the stages. Furthermore, if overexpectation is dependent

on the probability of reinforcement per trial, then shifting group 100% to 50% reinforcement in stage 2 should result in a larger overexpectation effect than for group 50%. On the basis that summation of learning reflects summation of expected amount of reinforcement and with the assumption that probability of reinforcement linearly relates to associative strength, while group 50% should expect one reinforcement per trial in stage 2 rather than half on average, group 100% should expect two lots of reinforcement rather than half.

Another motivation for Experiment 2a and 2b was to test the effect of probability of reinforcement per trial on overexpectation after fewer compound conditioning trials. It is possible that probability of reinforcement per trial has an effect on initial decrements in responding in overexpectation but the effect diminishes across compound conditioning. Such an effect has been found in extinction learning in which the probability of reinforcement per trial affects the number of nonreinforced trials required before a decrement in responding is observed (Chan & Harris, 2017, 2019; Harris & Andrew, 2017). Thus, cues associated with a low probability of reinforcement require more nonreinforced trials before extinction is observed than cues with a higher probability of reinforcement. Because the minimum number of trials necessary for overexpectation to occur was not known, we chose to present 24 trials of AX and BY each reinforced on 50% of trials before presenting probe trials of cues A and B intermixed with additional partially reinforced compound trials. The 12 reinforced compound trials, before the first probe trial, in which overexpectation may occur, were fewer than in Experiment 1 in which mice received 27 reinforced compound trials of each trial type before the first probe trial. The results show that 12 reinforced trials was sufficient for a reduction in summation to be observed over sessions.

Experiment 2b also allowed the inclusion of an additional control procedure for overexpectation. In Experiment 1, cue Y was presented in compound with cue B in stage 2. Cue Y had been nonreinforced in stage 1 and mice had learnt the discrimination between cue Y and reinforced cues A, B and X. Although cue Y had lower associative strength than the reinforced cues, it is possible that it may have undergone other processes that may have affected its ability to function as a control for cue X in as far as being associatively neutral. It is possible that during stage 1 it may have come to predict the absence of reinforcement such that it acquired inhibitory associative strength. Indeed, if the context gained excitatory associative strength then cue Y would have undergone feature-negative training, which typically, establishes a cue as conditioned inhibitor (Rescorla, 1969). Although we have found no evidence of this in similar discrimination training procedures (Austen et al., 2022), if this were the case, then subsequent reinforcement of the BY compound may have led to further increments in the associative strength of B because the combined associative strength of the compound BY was lower than for B alone. In other words, the inhibitory associative strength of cue Y allows B to undergo superconditioning (Williams & McDevitt, 2002). This would result in cue B gaining greater associative strength than cue A such that the difference between responding to the cues in the test phase would not reflect overexpectation alone. In order to avoid the possibility of cue Y being inhibitory in stage 2, in Experiment 2b, cue Y was a novel cue at the start of stage 2. During stage 1, instead of nonreinforced exposure to cue Y, mice received nonreinforced trials of cue Z, which was a different cue but from the same modality as cues X and Y.

A final purpose of Experiments 2a and 2b was to examine post-trial response rates. While a number of studies have demonstrated that acquisition of conditioned responding is sensitive to reinforcement rate and not probability of reinforcement per trial (Chan &

Harris, 2017, 2019; Harris, 2019; Harris & Andrew, 2017; Harris et al., 2015), it has been found that responding in the period immediately after the CS offset reflects probability of reinforcement (Harris & Kwok, 2018). Thus, when two cues differed in probability of reinforcement per trial but were matched for overall reinforcement rate, rats showed greater responding during the post-CS period for the cue with a high probability of reinforcement per trial compared to the other CS despite responding during the two CSs being similar. A reason why examining post-CS response rates would be useful given the results of Experiment 1 is that it may allow detection of an effect of probability of reinforcement per trial in the absence of an effect on overexpectation and summation during CS responding. Because both groups received partial reinforcement during stage 2 compound training, it was possible to compare post-CS response on the nonreinforced trials for compounds AX and BY (see Table 2) across both groups.

Methods

Experiment 2a

Subjects: Thirty-two naïve female C57BL/6J mice (Charles River UK Ltd), approximately 10 weeks old at the start of testing, with a mean free-feeding weight of 17.8 g (range: 15.5-19.6g) were used. All other details were the same as Experiment 1.

Apparatus: All other details were the same as Experiment 1.

Procedure: Mice were randomly assigned to one of two groups (N = 16 per group). In stage 1, group 50% received 12 sessions of training with four cues, A, X, B and Y that were, on average across trials, 10 s in duration. Cues A, B and X were reinforced on a random 50% of

trials. Cue Y was nonreinforced. Group 100% received similar training except that cues A and X were variable duration 20 s cues rather than 10 s. Trial types were presented in a random order with the following constraints. For group 50%, each block of eight trials consisted of two trials of each trial type (A, B, X and Y). For group 100%, each block of six trials consisted of one trial of A and X each and two of B and Y each. Within each session, group 100% received three trials of A and X each and six of B and Y each. Group 50% received six trials of each trial type. The difference in trial numbers for the two groups resulted in both groups receiving the same cumulative exposure to cues A and X and the same number of pairings with reinforcement. In order to equate the session lengths, the fixed inter-trial interval for group 50% was 90 s (from CS offset to CS onset) and for group 100%, it was 120 s. Sessions lasted 40 minutes. The allocation of stimuli to trial types across mice was counterbalanced in the same manner as used for Experiment 1.

On session 13-14 and 17-18, mice received 12 compound presentations of AX and BY each that were reinforced on a random 50% of trials. Compound presentations were a variable duration of 10 s. The order of trial types was random with the constraint that were equal numbers of each trial every four trials. The inter-trial interval was a fixed duration of 90 s. Sessions lasted 40 minutes.

On sessions 15-16 and 19-21, mice received 50% reinforced presentations of compounds AX and BY and also nonreinforced presentations of A and B. Compounds AX and BY were presented for a variable duration 10 s but cues A and B were presented for a fixed duration of 10 s. On Sessions 15 and 16, there were two trials of A and B each and 10 presentations of AX and BY each. On session 15, trials 1 and 2 consisted of one trial each of A and B. This was also true for trials 23 and 24. On session 16, trials 13 and 14 consisted of one trial each of A and B. This was also true for trials 23 and 24. On sessions 19 and 20,

there were four trials of A and B each and eight trials of AX and BY each. For both sessions 19 and 20, trials 5 and 6, 11 and 12, 17 and 18, and 23 and 24 consisted of one trial each of A and B. On session 21, there were eight trials of A and B each and four trials of AX and BY each. A and B trials occurred on trials 5-12 and 17-24, in a double alternating manner. For sessions 15, 16 and 19-21, the order of A and B across the trials was counterbalanced within groups with respect to the allocation of stimuli to trial types across mice. The order of AX and BY was random. Sessions lasted 40 minutes.

Experiment 2b

Subjects and apparatus: Thirty-two naïve female C57BL/6J mice (Charles River UK Ltd), approximately 10 weeks old at the start of testing, with a mean free-feeding weight of 18.3 g (range: 16.5-20.0g) were used. All other details were the same as Experiment 1 and 2a. The apparatus was the same as Experiments 1 and 2a except an additional cue was provide by a pure tone generator (ENV-323AM) that produced a 2,900 Hz tone at 80dB.

Procedure: The procedure was similar to Experiment 2a but with a number of exceptions. The counterbalancing of the allocation of stimuli to cues A, B, X and Y was the same as Experiment 2a except that the alternating flashing left and right LED stimulus was replaced with a flashing presentation of the right LED alone. In stage 1, in contrast to Experiment 2a, mice did not received trials of Y but instead received nonreinforced presentations of cue Z. For mice that were allocated an auditory stimulus for cue X (noise/clicker), cue Z was the pure tone. For mice that were allocated a visual stimulus for cue X (house light/flashing right LED), cue Z was a constant presentation of the left LED. The replacement of cue Y with cue Z in stage 1 resulted in cue Y being novel at the start of stage 2 training.

Statistical Analyses: The results of both Experiments 2a and 2b were combined and analysed with experiment as included as a factor. All other details were the same as Experiment 1.

Results and discussion

Stage 1 acquisition: Acquisition of responding is shown in Figure 5. Mice acquired conditioned responding to cues A, B and X over the course of training. Responding to the nonreinforced cue (Y in Experiment 2a and Z in Experiment 2b) initially increased at the start of training but then subsequently decreased. A 4 (cue: A, B, X and Y) by 12 (session) by 2 (group: 50%, 100%) by 2 (experiment) ANOVA showed a significant effect of cue, $F(3,180) = 29.19$, $p < 0.001$, $\eta_p^2 = .24$, 90% CI [.15, .32], and session, $F(11,660) = 52.92$, $p < 0.001$, $\eta_p^2 = .47$, 90% CI [.42, .50], and cue by session interaction, $F(33,1980) = 16.17$, $p < 0.001$, $\eta_p^2 = .21$, 90% CI [.17, .22]. Post-hoc tests (Holm-Bonferroni corrected) showed that mice responded more to cues A, B and X compared to the nonreinforced cue (Y/Z), p values < 0.001. There were no significant differences in responding to the reinforced cues (smallest non-corrected p-value = 0.036, corrected alpha = 0.0167). Simple main effects analysis of the interaction showed that there was no significant effect of cue for the first three sessions, F-values < 1, but there was for the remaining sessions, smallest $F(3,58) = 4.98$, $p = 0.004$. There was no significant effect of group, $F < 1$, or experiment, $F(1,60) = 3.09$, $p = 0.08$. There was a significant session by experiment interaction, $F(11,660) = 15.26$, $p < 0.001$, $\eta_p^2 = .20$, 90% CI [.15, .23]. The three-way cue by session by experiment interaction failed to reach significance $(33,660) = 1.96$, $p = 0.051$. No other interactions of factors were significant, p-values > 0.2. Simple main effects analysis of the session by experiment interaction showed that mice from Experiment 2a responded at a significantly lower level than Experiment 2b in

sessions 2-4, but at a significantly higher level that Experiment 2b on sessions 7-12, smallest $F(1,60) = 6.14$, $= 0.02$. There was no significant difference between experiments on the remaining sessions, p-values > 0.2.

A Bayesian ANOVA was conducted to test evidence for the null hypothesis that there was no difference in the acquisition of responding to partially and continuously reinforced cues. For the cue by group interaction, $BF_{incl} = 0.019$. For the cue by group by session interaction, $BF_{incl} = 2\times10^{-15}$. Therefore, the results provided at least 52 times greater evidence for the null hypothesis over the alternative hypothesis.

Additional analyses were conducted comparing response rates between cues under matched conditions (see Supplemental Material). The analyses compared responding across similar durations within the first trial of each trial type across sessions in the latter half of training (see Supplemental Figure 4). There was no significant difference between partially and continuously reinforced cues and Bayesian analyses found evidence in favour of the null hypothesis.

Pre-CS response rates initially increased and then declined over the course of training. A 2 (group) by 2 (experiment) by session ANOVA revealed a significant effect of session, $F(11,660) = 31.35$, $p < 0.001$, $\eta_p^2 = .34$, 90% CI [.29, .38], and experiment, $F(1, 60) = 10.20$, $p < 0.001$, $\eta_p^2 = .14$, 90% CI [.03, .28], but no significant effect of group, $F < 1$. There was a significant experiment by session interaction, $F(11,660) = 19.25$, $p < 0.001$, $\eta_p^2 = .33$, 90% CI [.19, .28], reflecting that mice in Experiment 2a responded at a lower rate than mice in Experiment 2b early on in training but came to respond more than Experiment 2b later on in training. Simple main effects analysis showed a significant effect of experiment on sessions 2-3 and 5-12, smallest $F(1,60) = 7.17$, $p = 0.01$. There was no significant effect on

sessions 1 and 4, F-values < 1. There were no other significant interactions of factors, p-values > 0.06.

Stage 2: Performance during the compound training phase is shown in Figure 6. In session 13, the first session of compound training, mice responded more to AX than BY. This effect was not evident in session 14. The results of the first block of compound training (session 13 and 14) were analysed using a 2 (compound: AX, BY) by 2 (session) by 2 (group: 50%, 10)%) by 2 (experiment) ANOVA. There were significant effects of compound, $F(1,60) = 7.91$, $p < 0.007$, $\eta_p^2 = .12$, 90% CI [.02, .25], and session, $F(1,60) = 7.60$, $p < 0.008$, $\eta_p^2 = .11$, 90% CI [.02, .24], and a significant compound by session interaction, $F(1,60) = 30.57$, $p < 0.001$, $\eta_p^2 = .34$, 90% CI [.18, .47]. Simple main effects analysis of the interaction revealed that mice responded more to AX than BY on session 13, $F(1,60) = 21.74$, $p < 0.001$, but not on the session 14, $F < 1$. Responding to BY significant increased over sessions, $F(1,60) = 27.29$, $p < 0.001$, but responding to AX did not significantly differ across sessions, $F < 1$. There was a significant effect of experiment, $F(1,60) = 7.05$, $p = 0.01$, $\eta_p^2 = .11$, 90% CI [.01, .23]. The effect of group was not significant, $F < 1$. There were no other significant interactions of factors, p-values > 0.10, except for a significant four-way interaction, $F(1,60) = 4.57$, $p = 0.04$, $\eta_p^2 = .07$, 90% CI [.002, .19]. Inspection of the results for each experiment suggested that the summation effect on the first session was weaker in the 50%-50% group compared to 100%-50% in Exp 2a. Separate analyses of each experiment revealed that for both experiments, the three-way interaction was not significant, although approaching significance in Exp 2a (Exp 2a: $F(1,30) = 4.10$, $p = 0.052$, Exp 2b: $F < 1$).

Responding during the 10 s period immediately after the presentation of the compounds on nonreinforced trials is shown in Figure 7. In first session of compound

training, session 13, mice responded more in the period immediately after AX than the period after BY and this effect was greater for group 100% than group 50%. These effects were less apparent in the subsequent sessions. The data from the first block of training, sessions 13 and 14, was analsyed using a 2 (compound: AX, BY) by 2 (session) by 2 (group: 50%, 100%) by 2 (experiment) ANOVA. There was a significant effect of compound, $F(1,60) = 71.69$, $p < 0.001$, $\eta_p^2 = .54$, 90% CI [.40, .64], that significant interacted with group, $F(1,60) = 10.13$, $p < 0.002$, $\eta_p^2 = .14$, 90% CI [.03, .28], and there was also a significant cue by group by session interaction, $F(1,60) = 4.31$, $p = 0.042$, $\eta_p^2 = .12$, 90% CI [.07, .18]. Simple main effects analysis of the cue by group interaction revealed that group 100% responded significantly more than group 50% in the post-CS period for AX, $F(1,60) = 7.14$, $p = 0.01$, but there was no significant difference between groups in the post-CS period for BY, $F < 1$. Both groups responded more in the post-CS period for AX than BY, smallest $F(1,60 = 13.96$, $p < 0.001)$. There was a significant effect of session, $F(1,60) = 10.07$, $p < 0.002$, $\eta_p^2 = .14$, 90% CI [.03, .28], a cue by session interaction, $F(1,60) = 4.50$, $p = 0.04$, $\eta_p^2 = .07$, 90% CI [.002, 19] and a three-way cue by session by experiment interaction, $F(1,60) = 10.12$, $p = 0.002$, $\eta_p^2 = .14$, 90% CI [.03, .28]. Separate analyses of each experiment revealed a significant cue by session experiment for Exp 2a ($F(1,60) = 13.74$, $p < 0.001$, but not Exp 2b, $F < 1$. There were no other significant interactions of factors, F-values < 1.

Analysis of pre-CS rates of responding revealed that mice in Exp 2a responded more than mice in Exp 2b, $F(1,60) = 17.05$, $p < 0.001$, $\eta_p^2 = .22$, 90% CI [.08, .36]. There were no other significant effects of group, experiment or session, or interactions of factors, (p-values > 0.06)

The second block of compound training, after the initial test sessions, was analysed separately, in a similar manner as the first block. During the compounds, responding

significantly decreased over the two sessions, $F(1,60) = 19.59$, $p < 0.001$, $\eta_p^2 = .25$, 90% CI [.10, .38]. There was a significant cue by experiment interaction, $F(1,60) = 8.49$, $p = 0.005$, $\eta_p^2 = .12$, 90% CI [.02, .25]. Simple main effects analysis revealed that this was due to significantly lower responding to AX than BY in Exp 2b, $F(1,60) = 9.87$, $p = 0.003$, but not Exp 2a, $F < 1$. There were no other significant main effects or interactions, p-values > 0.06.

In the post-CS periods, responding reduced over the two sessions, $F(1,60) = 7.60$, $p = 0.008$, $\eta_p^2 = .11$, 90% CI [.02, .24]. There were no other significant main effects or interactions, p-values > 0.08.

Pre-CS levels of responding were similar across groups. There were no significant effects or interactions of factors, p-values > 0.09.

Test: The response rates collapsed across test trials for each experiment are shown in Figure 8. Overall response rates were lower in Experiment 2b than 2a, but all groups showed lower rates of responding to A, the overexpectation cue, than the control cue B (see Supplemental Figure 1 for the results broken down by trial). Mice responded more to cue B than cue A. A 2 (group: 50%, 100%) by 2 (cue: A, B) by 2 (experiment) ANOVA showed a significant effect of cue, $F(1,60) = 7.16$, $p = 0.01$, $\eta_p^2 = .11$, 90% CI [.02, .23]. The cue by group interaction was not significant, $F(1,60) = 1.07$, $p = 0.31$. Mice in Experiment 2a responded at a higher rate than mice in Experiment 2b, $F(1,60) = 4.79$, $p = 0.03$, $\eta_p^2 = .07$, 90% CI [.003, .19]. There were no other significant main effects or interactions of factors (smallest p-value = 0.14).

Although the cue by group interaction was not significant, the overexpectation effect was numerically smaller in group 50% than group 100%. In order to test the strength of evidence for the null hypothesis a Bayes factor was calculated in JASP using the default priors, testing the prediction that the overexpectation effect would be greater for group

100% than group 50%. A Bayesian between-subjects t-test was conducted comparing the difference between response rates for the two cues (response rate for cue B minus the response rate for A). BF_{10} = 0.67, thus providing 1.5 times evidence for the null hypothesis over the alternative hypothesis.

Responding during the reinforced compounds and the pre-CS periods (collapsed across all trial types) is shown in Figure 9. Response rates were analysed using a 3 (period: pre-CS, AX, BY) by 2 (experiment) by 2 (group: 50%, 100%) ANOVA. There was a significant effect of period, $F(2,120)$ = 81.17, $p < 0.001$, η_p^2 = .58, 90% CI [.47, .64], a significant effect of experiment, $F(1,60)$ = 5.48, $p = 0.023$, η_p^2 = .08, 90% CI [.006, .21], and significant period by experiment interaction, $F(2,120)$ = 5.04, $p = 0.02$, η_p^2 = .08, 90% CI [.01, .15]. There were no other significant main effects or interactions, F-values < 1. Post-hoc analysis of the effect of period confirmed that mice responded significantly more during AX and BY compared to the pre-CS period, p-values < 0.001, but there was no significant difference between AX and BY, $p = 0.14$. Simple main effects analysis of the experiment by period interaction revealed that mice in Experiment 2b responded at a significantly lower level than mice in Experiment 2a for AX, $F(1,30)$ = 9.62, $p = 0.003$, but not the pre-CS period or during BY, p-values > 0.1.

Similar to the results of Experiment 1, probability of reinforcement failed to affect the strength of overexpectation. Therefore, regardless of whether group 50% switched to 100% reinforcement for stage 2 compound training, as in Experiment 1, or group 100% switched to 50% compound training, as in Experiment 2a and 2b, the manipulation of probability of reinforcement per trial failed to affect the extent of overexpectation. Furthermore, the extent of overexpectation was similar across the experiments suggesting that whether cue Y was novel or not made little difference to the pattern of results.

Although there was no significant cue by group interaction, the overexpectation effect was numerically smaller in the 50% groups compared to the 100% groups. Bayesian analysis, however, found evidence in favour of the null hypothesis that there was no difference between groups. While the cue by group by experiment three-way interaction was not significant, it is possible that the ability to detect a cue by group interaction depended on the different treatments in Experiment 2a and 2b. Separate analyses of each experiment (see Supplemental Material) failed to reveal evidence of a cue by group interaction for either experiment (F-values < 1). It is possible that the study was not sensitive enough to detect a cue by group interaction, however, there was no evidence for a cue by group interaction from the statistical analysis and, furthermore, the lack of effect of group in Experiment 2 is consistent with the lack of effect in Experiment 1.

One concern with Experiment 1 was that if probability of reinforcement per trial affects the rate at which overexpectation occurs but not the overall extent of overexpectation then overtraining in the compound training phase may have masked a possible effect. This is unlikely to be the case in Experiment 2 because, although variable over trials, there was a trend, albeit nonsignificant, for the overexpectation effect to emerge over training trials (see Supplemental Figure 1). Therefore, the test of overexpectation likely occurred at the earliest measurable point.

The lack of effect of probability of reinforcement per trial on overexpectation is in contrast to the effects observed in post-CS responding. During the summation test, probability of reinforcement failed to affect summation during the CS, but in the post-CS periods, group 100% responded more to AX than group 50%. Therefore, the post-CS summation effect was greater for group 100% than group 50% despite no significant difference in extent of summation during presentation of the compound. Rather than being

an effect on summation per se, this likely reflects that for group 100%, cues A and X individually signalled 100% probability of reinforcement, whereas for group 50% the cues signalled 50% reinforcement.

Experiment 2b was conducted because of the possibility that the effect in Experiment 1, rather than reflecting overexpectation, may instead be the result of increased associative strength of cue B rather than a reduction in associative strength of cue A. This could have occurred because, during stage 1, cue Y may have acquired inhibitory associative strength through signalling a reduction in the occurrence of the US compared to the context (Miller et al., 1991). In stage 2, when the compound BY was reinforced, the inhibitory associative strength of cue Y would have led to positive prediction error that would allow cue B to increase in associative strength. To rule out this possibility, cue Y was a novel cue in Experiment 2b and, therefore, was associatively neutral. It was found that the extent of overexpectation, as indicated by the difference in responding between cues A and B, did not differ between Experiment 2b and 2a in which cue Y was a nonreinforced cue during stage 1. It is possible, however, that a novel cue may result in external inhibition such that there is an increase in the prediction error at the start of training on reinforced trials with BY compared to the previous training with B alone. For this to result in greater responding to cue B than to cue A in the test phase, it would have to be assumed that the effect of external inhibition is greater for cue Y than cue X. But if external inhibition is caused by generalisation decrement that results in a failure of learning to transfer completely between training with the element B and the compound BY then the potential for external inhibition would be matched between the cues Y and X. Therefore, this account would not explain a difference in responding between cues A and B. Furthermore, even if cue Y was able to elicit greater external inhibition than cue X then any new learning about cue B as a consequence

While the results of Experiment 2 are consistent with Experiment 1 and other demonstrations of an overexpectation effect, it is possible that it not a pure test of overexpectation. Experiment 2 differs from Experiment 1 in that in stage 2, compound training, the compound was partially reinforced rather continuously reinforced. If extinction is a trial-dependent process then on nonreinforced trials the amount of extinction that occurs for AX would be expected to be greater than for BY. This is because the discrepancy between the expectation of the US and nonreinforcement on AX compound trials is greater than for BY. According to error-correction models of learning (e.g., Rescorla & Wagner, 1972), this leads to a greater reduction in associative strength. Therefore, AX may have undergone greater extinction on nonreinforced trials than BY in addition to the reduction caused by overexpectation on reinforced trials. This depends on extinction occurring in a trial-by-trial manner rather than continuously over cumulative nonreinforced exposure. The fact that Experiment 1, in which compounds AX and BY were continuously reinforced, produced similar results to Experiment 2 results suggests that the potential for deepened

extinction did not qualitatively change the pattern of results in Experiment 2. Thus, the

manipulation of probability of reinforcement per trial did not affect reductions in

conditioned responding as a consequence of overexpectation alone in Experiment 1 or

overexpectation with the potential for extinction in Experiment 2.

Experiment 3

The lack of effect of probability of reinforcement per trial on overexpectation in

Experiments 1 and 2 is in contrast its effect on extinction of magazine approach behaviour in

rats (Chan & Harris, 2019). While this may suggest a dissociation between overexpectation

and extinction, there are numerous differences between the procedures used in

Experiments 1 and 2, which failed to find an effect, and those used for extinction learning

other than the difference in species. The purpose of Experiment 3 was to test the effect of

probability of reinforcement per trial on extinction learning using procedures similar to

those used in Experiments 1 and 2 (see Table 3). Thus, the manipulation of probability of

reinforcement per trial was between subjects: one group was trained with a 50% reinforced

cue that was on average 10 s in duration (group 50%). The other group was trained with a

100% reinforced cue that was on average 20 s in duration (group 100%). The manipulation

of extinction was within-subjects: mice were initially trained with two cues (A and B) before

one cue was nonreinforced in the extinction phase (cue B) and other continued to be

reinforced (cue A). In order to control for any differences in generalisation between cues on

the basis of cue duration and probability of reinforcement per trial, for half of the mice

within each group, the reinforced cue was a variable duration 10 s cue that was reinforced

on 50% of trials (see Table 3). For the remaining mice within each group, the reinforced cue was a variable duration 20 s cue that was reinforced on 100% of trials. Therefore, within each group the similarity between cues A and B was balanced in terms of their trial durations and reinforcement contingencies.

Methods

Subjects and apparatus: Thirty-two naïve female C57BL/6J mice (Charles River UK Ltd), approximately 9 weeks old at the start of testing, with a mean free-feeding weight of 18.3 g (range: 16.2-20.8g) were used. One mouse died before the start of testing. All other details were the same as Experiment 1 except that only the auditory cues (noise and clicker) were used.

Procedure: Mice were randomly assigned to group 100% (N = 15) or group 50%. (N = 16). In stage 1, mice were trained with two cues, A and B. Cue A was a cue that had a constant reinforcement contingency throughout stage 1, the acquisition phase, and stage 2, the extinction phase. Cue B was the cue that received nonreinforced extinction training in stage 2. Within groups 100% and 50%, for half of the mice, cue A was on average, across trials, 20 s in duration and reinforced on 100% of trials. For the remaining mice, cue A was on average 10 s in duration and reinforced on a random 50% of trials. This counterbalancing resulted in four sub-groups (see Table 3). In stage 1, for group 100%, cue B was on average 20 s in duration and reinforced on 100% of trials. For group 50%, cue B was on average 10 s in duration and reinforced on a random 50% of trials. Cues that were 100% reinforced were presented six times per session. Cues that were 50% reinforced were presented 12 times per session. Thus, the number of CS-US pairings was matched between cues and groups.

The inter-trial interval differed between groups and subgroups in order that the overall session length was matched across groups. For mice trained with two 100% reinforced cues (A and B), the inter-trial interval was 240 s. For mice trained with one 50% reinforced cue and one 100% reinforced cue, the inter-trial interval was 160 s. For mice trained with two 50% reinforced cues, the inter-trial interval was 120 s. Sessions lasted 52 minutes. There were six sessions of stage 1. This resulted in the same number of CS-US pairing during acquisition as in Experiments 1 and 2. The allocation of stimuli (noise and clicker) to cues A and B was counterbalanced within groups.

In stage 2, for both groups, cue B was on average 10 s in duration and now nonreinforced. Mice received 12 trials with cue B per session. Cue A was presented and reinforced in the same manner as in stage 1. For mice for which cue A was 50% reinforced, the inter-trial interval was 120 s. For mice for which cue A was 100% reinforced, the inter-trial interval was 160 s. Sessions lasted 52 minutes. Stage 2 lasted for six sessions.

Data and statistical analyses: The reinforcement contingencies of cue A, either variable 10 s cue reinforced on 50% of trials or a variable 20 s cue reinforced on 100% of trials, did not significantly affect the outcome of the key factors of interest. Therefore, the results are presented collapsed across the factor of cue A and the factor is not included in the statistical analyses. The data (see Supplemental Figure 2) and additional analyses that include the factor cue of A are presented in the Supplemental Material.

Results and discussion

Acquisition: In stage 1, conditioned responding to cues A and B increased similarly over sessions between the two group (see Figure 10). A 2 (cue: A, B) x 6 (session) x 2 (group)

ANOVA showed a significant effect of session, $F(5,145) = 63.38$, $p < 0.001$, $\eta_p^2 = .69$, 90% CI [.61, .73]. There was no significant effect of cue, $F < 1$. No other main effects or interactions were significant, p-values > 0.1.

A Bayesian ANOVA was conducted to test evidence for the null hypothesis that there was no difference in the acquisition of responding to partially and continuously reinforced cues. For the cue by group by group interaction, $BF_{incl} = 0.78$. For the cue by group by session interaction, $BF_{incl} = 0.14$. Therefore, the results provided at least 1.28 times greater evidence for the null hypothesis over the alternative hypothesis.

Additional analyses were conducted comparing response rates between cues under matched conditions (see Supplemental Material). The analyses compared responding across similar durations within the first trial of each trial type in the first extinction session (see Supplemental Figure 5). There was no significant difference between partially and continuously reinforced cues and Bayesian analyses found evidence in favour of the null hypothesis.

Pre-CS response rates, collapsed across trial types (see Figure 10), initially increased with training, but then subsequently reduced. A 2 (group) by 12 (session) ANOVA showed a significant effect of session, $F(5,145) = 5.04$, $p < 0.001$, $\eta_p^2 = .15$, 90% CI [.05, .21], but no significant effect of group, or interaction between session and group, F-values < 1.

Extinction stage: Both groups reduced responding to cue B over the course of extinction training while maintaining responding to cue A (see Figure 10). A 2 (cue: A, B) x 6 (session) x 2 (group) ANOVA showed a significant effect of cue, $F(1,29) = 74.44$, $p < 0.001$, $\eta_p^2 = .72$, 90% CI [.55, .80], session, $F(5,145) = 9.68$, $p < 0.001$, $\eta_p^2 = .25$, 90% CI [.13, .32], and a significant cue by session interaction, $F(5,145) = 16.45$, $p < 0.001$, $\eta_p^2 = .36$, 90% CI [.24, .43]

The effect of group and remaining interactions of factors were not significant, p-values >

0.16.

A Bayesian ANOVA was conducted in order to assess the strength of evidence for the

null hypothesis compared to the alternative hypothesis. To simplify the analysis the effect of

cue (reinforced vs. nonreinforced) was reduced to the difference between response rates

such that a group by session ANOVA was conducted. For the effect of group, BF_{incl} = 0.35

and for the group by session interaction, BF_{incl} = 0.46. Therefore, evidence for the null

hypothesis was at least 2.18 times greater than for the alternative hypothesis.

Pre-CS response rates reduced over the extinction phase similarly for both groups.

The effect of session was significant, $F(5,145)$ = 3.01, p = 0.032, η_p^2 = .09, 90% CI [.01, .15],

but the effect group and group by session interaction were not significant (smallest p-value

= 0.2)

Response rates in the 10 s period immediately after the presentation of B, the

extinction cue, are shown in Figure 11. The 100% group showed greater responding than

group 50%. This effect reduced over the course of extinction as responding reduced overall

for both groups. A 2 (group) by 6 (session) ANOVA showed a significant effect of session,

$F(5,145)$ = 60.29, p < 0.001, η_p^2 = .68, 90% CI [.59, .72]. The effect of group failed to reach

significance, $F(1,29)$ = 3.91, p = 0.06, but the group by session interaction was significant,

$F(5,145)$ = 4.63, p = 0.015, η_p^2 = .14, 90% CI [.04, .20]. Simple main effects analysis of the

group by session interaction revealed that group 100% responded significantly more than

group 50% on the first session of extinction, $F(1,29)$ = 6.65, p = 0.015, but not thereafter, p-

values > 0.1.

In contrast to the findings in rats (Chan & Harris, 2019), probability of reinforcement

per trial failed to affect extinction learning. Similar to the results of Experiment 2, mice

responded more in the post-CS period after the previously 100% reinforced cue than after 50% reinforced cue. Therefore, mice had learnt information about the probability of reinforcement per trial but it did not affect the rate of extinction learning.

Experiment 4

Experiment 3 failed to find an effect of probability of reinforcement on extinction learning. Experiment 4 was similar to Experiment 3 but the manipulation of probability of reinforcement per trial was within-subjects rather than between-subjects (see Table 4). Although within-subjects procedures have the potential to increase sensitivity of measures, they introduce the potential that generalisation of learning between cues may diminish effects on responding that would otherwise occur. In order to reduce the effect of generalisation between cues, the 50% reinforced cue was of a different modality than the 100% reinforced cue. Furthermore, each animal was trained with two 50% and two 100% reinforced cues and one cue of each type was nonreinforced in the extinction phase while the other cues continued to be reinforced. This design allows the detection of any generalisation effect based on reinforcement contingency (i.e., 50% or 100%) that might reduce the detection of a partial reinforcement extinction effect.

Methods

Subjects and apparatus: Sixteen naïve female C57BL/6J mice (Charles River UK Ltd), approximately 14 weeks old at the start of testing, with a mean free-feeding weight of 22.1 g (range: 19.6-24.6g) were used. All other details were the same as Experiment 1.

Procedure: In stage 1, mice received 12 sessions of training with four different cues. Cues A and X were a variable duration of 10 s and were reinforced on a random 50% of trials. Cues B and Y were a variable duration of 20 s and reinforced on 100% of trials. There were six trials of A and X each per session and three of B and Y per session. The difference in the number of trials resulted in all cues having the same cumulative exposure per session and the same number of pairings with reinforcement. Trial types were presented in a random order with the constraint that for every block of six trials there were two trials of A and X each and one of B and Y each. The inter-trial interval was a fixed duration of 120 s. Sessions lasted 40 minutes. For half of the mice, cues A and X were auditory (noise and clicker) and B and Y were visual (house light and alternating left and right flashing LEDs, 0.5s on, 0.5 s off). For the remaining mice, A and X were visual cues and B and Y were auditory cues. Within each subgroup, the allocation of stimuli to cues was counterbalanced in 2 by 2 manner.

In stage 2, mice continued to receive presentations of cue A reinforced on 50% of trials, six trials per session and cue B reinforced on 100% of trials, 3 trials per session. Mice also received nonreinforced presentations of cues X and Y on separate trials. Half of the mice received six presentations of X and Y per session and each trial was a fixed duration of 10 s. The remaining half of mice received three presentations of X and Y and each trial was a fixed duration of 20 s. Trial types were presented in a random order with the constraint that for the mice for which X and Y were 10 s long, there was two presentations of A, X and Y, and one of B every block of seven trials. For the mice for which X and Y were 20 s long, each block of five trials consisted of one each of B, X and Y, and two of A. Mice received 14 sessions in Stage 2. The inter-trial interval was a fixed duration of 120s. This results in sessions lasting for 46 mins for mice that received six 10 s presentations of the

nonreinforced cues and 34 minutes for mice that received three 20 s presentations of the nonreinforced cues.

Data analysis: The design of the within-subjects manipulation resulted in counterbalancing of the modality of the 50% and 100% reinforced cues across mice. We have previously found that mice show greater magazine entries for auditory cues paired with food than visual cues (Sanderson et al., 2016). In order to assess the effect of the key manipulations independent of the variance caused by this counterbalancing factor, the counterbalancing of modality was included as a factor in the analysis of response rates but the main effect and any interactions with other factors were ignored.

For the results of the extinction phase, preliminary analysis of the effect of the number of extinction trials per session (i.e., three 20-s trials or six 10-s trials) showed that trial number had little effect (F-values < 1). For ease of exposition, the results are shown collapsed across this factor and the factor is not included in the reported analysis.

Results and discussion

Stage 1: Mice acquired responding to the 50% reinforced cues in a similar manner as the 100% reinforced cues (see Figure 12). The data were analysed, collapsing over the two cues that were 50% reinforced and the two cues that were 100% reinforced, using a 2 (probability: 50%, 100%) by 12 (session) by 2 (modality counterbalancing factor) ANOVA. There was a significant effect of session, $F(11,154) = 5.41$, $p = 0.002$, $\eta_p^2 = .28$, 90% CI [.14, .32]. The effect of partial reinforcement was not significant, $F(1,14), = 3.45$, $p = 0.08$, and did not significantly interact with session, $F(11,154) = 1.48$, $p = 0.14$.

A Bayesian ANOVA was conducted to test evidence for the null hypothesis that there was no difference in the acquisition of responding to partially and continuously reinforced

cues. For the effect of cue, BF_{incl} = 0.39. For the cue by session interaction, BF_{incl} = 0.13.

Therefore, the results provided at least 2.56 times greater evidence for the null hypothesis

over the alternative hypothesis.

Additional analyses were conducted comparing response rates between cues under

matched conditions (see Supplemental Material). The analyses compared responding across

similar durations within the first trial of each trial type across sessions in the latter half of

training (see Supplemental Figure 6). There was no significant differences between partially

and continuously reinforced cues.

Pre-CS responses reduced over training. There was a significant effect of session

($F(11,154)$ = 11.21, $p < 0.001$, η_p^2 = .44, 90% CI [.31, .49].

Extinction stage: The results of the extinction stage are shown in Figure 13. Results were

analysed with a 2 (extinction: nonreinforced/reinforced) by 2 (probability: 50%/100%) by 2

(modality counterbalancing factor) by 14 (session) ANOVA. Mice responded significantly less

to the nonreinforced cues than reinforced cues. $F(1,14)$ = 9.14 $p = 0.01$, η_p^2 = .40, 90% CI

[.07, .60]. There was a significant effect of session, $F(13,182)$ = 2.82, $p = 0.049$, η_p^2 = .17, 90%

CI [.04, .19], which significantly interacted with reinforcement, $F(13,182)$ = 5.63, $p < 0.001$,

η_p^2 = .29, 90% CI [.15, .32]. The effect of probability, the probability by reinforcement

interaction and the three-way reinforcement by session by probability interactions were not

significant, F-values < 1. The remaining main effects and interactions were not significant, p-

values > 0.3. Simple main effects analysis of the reinforcement by session interaction

showed that responding to nonreinforced cues (X and Y) was significantly lower than to

reinforced cues (A and B) on sessions 17 (fifth extinction session), and 21-26 (last six

extinction sessions), largest $p = 0.02$, but not on the remaining sessions, smallest $p = 0.06$.

A Bayesian ANOVA was conducted in order to assess the strength of evidence for the null hypothesis compared to the alternative hypothesis. To simplify the analysis, the effect of extinction was reduced to the difference between response rates to reinforced and non-reinforced cues such that a cue (probability: 50% vs 100%) by session ANOVA was conducted. For the effect of cue, BF_{incl} = 0.33 and for the cue by session interaction, BF_{incl} = 0.01. Therefore, evidence for the null hypothesis was at least 3 times greater than for the alternative hypothesis.

Pre-CS response rates were stable over extinction training. There was no significant effect of session, $F(13,182)$ = 1.64, p = 0.17.

Mice showed greater post-CS responding to the previously reinforced 100% cue than the 50% reinforced (see Figure 13, right panel). Responding was analysed with a 2 (cue: 50%, 100%) by 2 (modality counterbalancing factor) by 14 (session) ANOVA. There was a significant effect of cue, $F(1,14)$ = 7.46, p = 0.016, η_p^2 = .35, 90% CI [.04, .56], significant effect of session, $F(13,182)$ = 5.95, p < 0.001, η_p^2 = .30, 90% CI [.16, .34], but no significant cue by session interaction, F < 1.

The results were similar to those of Experiment 3. The within-subjects procedure failed to provide evidence of a partial reinforcement extinction effect, but post-CS responding demonstrated that mice were sensitive to the manipulation of probability of reinforcement per trial. In comparison to Experiment 3, extinction of responding took longer to emerge over the course of training. This may have been due to the increased generalisation between cues. Whereas mice learnt about only two cues in Experiment 3, mice learnt about four cues in Experiment 4 and in the extinction phase had to discriminate between stimuli within two different modalities. Despite the seemingly different rates of extinction between the two experiments, neither provided evidence of a partial

reinforcement extinction effect. Discussion of the potential differences between the results

of Experiment 3 and 4 and those in rats (Chan & Harris, 2019) are saved for the General

Discussion.

General discussion

Probability of reinforcement per trial failed to affect overexpectation and extinction. Furthermore, in each experiment, it failed to affect acquisition of conditioned responding and, in Experiment 1 and 2, summation of conditioned responding. The lack of effect is consistent with the finding that cumulative rate of reinforcement rather than probability of reinforcement per trial is the key determinant of conditioned responding (Harris et al., 2015). In contrast to the lack of effect on conditioned responding during the presentation of CSs, probability of reinforcement did affect post-CS responding on nonreinforced trials. Mice showed greater responding at the termination of cues that were previously 100% reinforced than for cues that were previously 50% reinforced. The dissociation between CS and post-CS responding suggest that mice learn about probability of reinforcement per trial as well rate of reinforcement, but it does not affect responding to a CS prior to the expected time of the US.

The results of Experiments 1 and 2 lead to several conclusions about the cause of overexpectation. The first concerns the role of summation. The Rescorla-Wagner (1972) model proposes that overexpectation is the result of summation of the associative strength of cues when presented in compound. In Experiments 1 and 2, the extent of summation was similar for cues that differed in probability of reinforcement per trial but were matched for reinforcement rate. Therefore, the learning that was summed likely reflects learning of the reinforcement rate of the individual cues and not the probability of reinforcement. This builds on work by Andrew and Harris (2011) that showed that in a summation test, responding to the test compound was similar to that for a cue or compound that had a reinforcement rate that was sum of the reinforcement rate of the individual cues in the test compound. This was true regardless of whether reinforcement rate was manipulated by CS

duration or probability of reinforcement per trial. Therefore, this may suggest that summation not only reflects learning about reinforcement rates but it reflects specifically summation of expected reinforcement rate. The second conclusion is that the learning process that determines the reduction of conditioned responding also reflects learning of reinforcement rate rather than probability of reinforcement per trial. Therefore, the cause of summation and the subsequent overexpectation reflects a common process.

The results are similar to a certain extent to findings in rats. Similar to Experiments 2-4, Harris and Kwok (2018) found that post-CS responding was a function of the probability of reinforcement per trial in cues matched for reinforcement rate. The difference occurred despite similar response rates during the CS presentation. Therefore, learning in mice and rats is sensitive to both probability of reinforcement per trial and reinforcement rate over cumulative exposure. The dissociation suggests that there are separate learning mechanisms that update either continuously over time, in order to learn about reinforcement rate, or episodically, to learn about probability of reinforcement per trial. The fact that animals are sensitive to trial-based information suggests that animals separate experience into discrete events.

The results do, however, differ distinctly from findings in rats. Harris and colleagues have shown that extinction learning in rats is sensitive to probability of reinforcement per trial under conditions in which reinforcement rate is controlled (Chan & Harris, 2017, 2019; Harris, 2024; Harris, Kwok, et al., 2019; Harris, Seet, et al., 2019; Jiao & Harris, 2024; Norton & Harris, 2022). There was no evidence that this was the case for mice. Experiment 3 used a between-subjects design and despite differences in post-CS responding, responding during the CSs extinguished at a similar rate over sessions. Experiment 4 replicated these findings but used a within-subjects design. Not only is this a failure to observe an effect of

probability of reinforcement per trial on extinction under conditions in which reinforcement rate is matched between the cues, it is a failure to observe a partial reinforcement extinction effect generally. Failures to observe a partial reinforcement extinction effect in Pavlovian procedures (e.g., Pearce et al., 1997) have been attributed to potential generalisation between cues in within-subjects procedures such that a difference in extinction is not observed or to extinction of magazine approach behaviour independent of extinction of the CS which may have masked a partial reinforcement extinction effect. These factors are unlikely to contribute the lack of effect in the present experiments because the inclusion of a reinforced cue (Experiment 3) or cues (Experiment 4) during the extinction phase maintained reinforcement of the magazine and reduced generalisation between cues because mice learnt to discriminate between reinforced and nonreinforced, extinguished cues. Indeed, in Experiment 4, in which a within-subjects manipulation of partial reinforcement was used, one cue from each modality (visual, auditory) was extinguished such that discrimination between cues could not be achieved by modality alone and required mice to discriminate between individual cues on the basis of their reinforcement contingency. Furthermore, the differences in post-CS responding indicate that mice discriminated between cues that extinguished at similar rates. This also suggests that is unlikely that the overall change in the amount of reinforcement between the acquisition and extinction phases resulted in increased generalisation between the cues.

A partial reinforcement extinction effect has been observed in mice under conditions that have not controlled for reinforcement rate. For example, it has been observed in the extinction of the speed of traversing a runway for appetitive reinforcement (Lee et al., 1974; Shanab & Ralph, 1979; Wong et al., 1971) and in extinction of Pavlovian conditioned freezing to shock (Huh et al., 2009). Therefore, it is unlikely that the absence of the effect in

the Experiments 3 and 4 is simply due to an inability to observe an effect of partial

reinforcement on extinction in mice generally. A recent study by Mallea et al. (2024) used

an appetitive Pavlovian conditioning procedure in mice that was similar to those used in the

current study. From Experiment 2 of their study, Mallea et al. (2024) concluded that partial

reinforcement, when reinforcement rate is controlled, does result in a partial reinforcement

extinction effect. This conclusion, however, is potentially undermined by initial differences

in responding as a result of partial and continuous reinforcement prior to extinction. Thus,

mice respond more for a continuously reinforced 24 s cue than a 12 s cue reinforced on 50%

of trials. While responding extinguished from an initial higher rate for the continuously

reinforced group than the partially reinforced group, it did not come to be extinguished

below the rate of responding for the partially reinforced group. Therefore, it is hard to come

to firm conclusions about the rate of extinction when it starts from different initial levels of

responding. The study by Mallea et al. (2024) did find, however, that mice responded less

when reinforced on 50% of trials than on 100% of trials despite matching of overall

reinforcement rates. The effect emerged during training as a consequence of mice reducing

responding to the 50% reinforced cue compared to the 100% cue. This finding is at odds

with the absence of an effect that was observed in the acquisition data across Experiment 1-

4 and with other findings in mice (Austen et al., 2021) and rats (Harris et al., 2015). While

there are a number of procedural differences between the studies, it is not clear what the

cause of the discrepancy in findings is between the Mallea et al. (2024) study and the

present experiments.

       While the reasons for why a PREE was not observed in mice are not clear, the

dissociation between the sensitivity of post-CS responding and extinction learning

demonstrates that the expectation of the US that determines post-CS responding does not

determine subsequent extinction of conditioned responding during the CS. Harris (2019) suggested that trial based learning is responsible for extinction rate because it is only at the end of a trial that an animal can determine that the US has not occurred. Thus, learning about nonreinforcement reflects trial-based probability learning because the termination of trial is an event the determines the US has not and will not occur. This results in partially reinforced cues requiring more nonreinforced trials for extinction to occur than continuously reinforced cues. The present results demonstrate that mice are sensitive to trial-based information but it does not determine the rate of extinction. Furthermore, a similar dissociation was observed in Experiment 2, demonstrating that, despite an effect on post-CS responding, trial-based information does not determine the extent of overexpectation. Thus, although mice responded at a significantly higher rate after a compound of two individually 100% reinforced cues compared to two 50% reinforced cues, the extent of overexpectation was similar. While the dissociation in mice does not mean that the cause of the post-CS responding effect is not the cause of the PREE in rats, it does demonstrate that the cause of the post-CS responding effect is not sufficient for the PREE.

Harris and Kwok (2018) demonstrated that the effect of probability of reinforcement per trial on post-CS responding was not simply due to learning that the offset of the cue predicted the probability of reinforcement. The post-CS responding effect was still present when the US occurred at a random interval within the CS rather than at the termination of the CS. Therefore, the offset of the CS was not a cue for the US. This finding is important because it would be expected that with variable duration CSs the CS offset is more informative than the CS onset about the occurrence of reinforcement. Therefore, despite this consequence of trial duration variability, the effect nonetheless depended on the probability of reinforcement per trial signalled by the presence of the CS prior to the

termination of the trial. In the present experiments, the US always occurred at the termination of a CS. We cannot rule out the possibility that mice had learnt that the offset of the cue signalled the probability of reinforcement. Indeed, given this possibility, an account of the dissociation observed in mice can be made without appealing to independent trial-based and time-based learning mechanisms. If it is assumed that the immediate trace of the CS is paired with the US (e.g., the trace CS undergoes simultaneous or backwards conditioning) then the strength of post-CS responding may reflect the reinforcement rate of the trace CS. For cues that differ in probability of reinforcement but are matched for reinforcement rate by virtue of differences in CS duration, the reinforcement rate of the trace of a partially reinforced cue will be lower than the reinforcement rate of the trace of a continuously reinforced cue. For this account to be plausible it would have to be assumed that the trace of a long duration cue has a similar saliency and decays at a similar rate as the trace of a short duration cue. While this is possible, it may be unlikely given that it has been proposed that the reduction of conditioned responding to a long CS compared to a short CS is the consequence of the salience of a CS decreasing over the duration of its presentation (Harris & Bouton, 2020) as a consequence of short-term habituation (Brandon et al., 2003; Vogel et al., 2019; Wagner, 1981). Further work is needed to conclude that the post-CS responding in mice is specifically a function of probability of reinforcement independent of CS duration. However, the dissociations present in Experiments 2-4 show that mice had clearly learnt information about the cues that differed in probability of reinforcement but it affected post-CS responding only.

It is possible that the difference between CS responding and post-CS responding simply reflects that post-CS responding is a more sensitive measure of learning about the CS

than CS responding. This is unlikely to be the case for magazine approach behavior in rats, because the factors that affect CS and post-CS responding can be doubly dissociated (Harris & Kwok, 2018). Thus, conditioned stimuli that are matched for reinforcement rate but differ in probability of reinforcement per trial elicit similar levels of conditioned responding during the CS but different levels after the CS. Conditioned stimuli that are matched for probability of reinforcement per trial but differ in reinforcement rate elicit different levels of conditioned responding during the CS but similar levels after the CS. In mice, magazine approach behaviour during a CS is sensitive to reinforcement rate, independent of whether it is manipulated by CS duration or probability of reinforcement (Austen & Sanderson, 2019; Austen et al., 2022; Strickland et al., 2024), but as shown in the present experiments, when reinforcement rate is controlled, probability of reinforcement fails to affect performance. If post-CS responding is simply a more sensitive measure of learning about the CS than responding during the CS then this would make the results of Experiments 3 and 4 even more surprising because rather than simply a failure to observe the partial reinforcement extinction effect, it would suggest that the opposite effect was found, i.e., slower extinction for a continuously reinforced CS compared to a partially reinforced CS. This is unlikely given the observations of a partial reinforcement extinction effect in mice in other procedures (Huh et al., 2009; Lee et al., 1974; Shanab & Ralph, 1979; Wong et al., 1971). Therefore, the results suggest that CS responding and post-CS responding reflect different forms of learning.

Across the experiments, probability of reinforcement per trial was manipulated under conditions in which the reinforcement rate was controlled. The absence of an effect of probability of reinforcement on acquisition is consistent with other findings (Austen et al., 2021; Harris et al., 2015) and given the positive evidence of an effect of reinforcement

rate on acquisition (Austen et al., 2021; Harris et al., 2015) suggests that reinforcement rate is the primary determinant of rate of conditioned responding and not probability of reinforcement. The failure to find an effect of probability of reinforcement on overexpectation and extinction may also suggest that reinforcement rate is more important, however, the experiments, although they included control procedures for overexpectation and extinction, did not include manipulations of reinforcement rate that held probability of reinforcement constant. It seems reasonable to assume that in the absence of an effect of probability of reinforcement per trial, overexpectation and extinction must rely on reinforcement rate but a positive demonstration would make this conclusion compelling. At the least, evidence from other effects that assess manipulations of prediction error and/or summation in compound conditioning procedures have provided evidence that reinforcement rate and not probability of reinforcement is sufficient to result in learning. Harris et al. (2012) found that in a summation test in rats, summation of responding reflection summation of reinforcements rates regardless of whether reinforcement rate was manipulated by trial duration or probability of reinforcement per trial. Similarly, conditioned inhibition was acquired for cues that signalled a decrease in reinforcement rate regardless of whether rate was manipulated by trial duration or probability of reinforcement per trial (Harris et al., 2014). For overexpectation, it would be predicted that the extent of overexpectation would depend on the summation of reinforcement rates and the degree to which the summation of associative strength exceeded the maximum associative strength supported by the US (lambda). For example, for a cue that signals a high reinforcement rate then overexpectation would be greater when that cue was reinforced in compound with another high reinforcement rate cue than with a lower reinforcement rate cue.

In conclusion, it remains to be seen whether overexpectation and extinction are determined by similar factors in rats. In mice, however, the present results suggest that they share a common a process that may reflect learning about reinforcement rate but not probability of reinforcement per trial. This suggests that learning that determines anticipatory responding is continuously updated over time rather than in a trial-dependent manner.

Group	Stage 1	Stage 2	Test
50%	A – 10 s, 50% B – 20 s, 100% X – 10 s, 50% Y – 20 s, 0%	AX – 20 s, 100% BY – 20 s, 100%	A – 20 s, 0% B – 20 s, 0% AX – 20 s, 100% BY – 20 s, 100%
100%	A – 20 s, 100% B – 20 s, 100% X – 20 s, 100% Y – 20 s, 0%		

Table 1. Design of Experiment 1. The letters A, B, X and Y denote cues. The durations denote the average trial duration of the cue (see methods for details). The percentages indicate the probability of reinforcement per trial. Cues that were on average 10 s were presented twice as often as 20 s cues.

Experiment	Group	Stage 1	Stage 2	Test
Exp 2a	50%	A – 10 s, 50% B – 10 s, 50% X – 10 s, 50% Y – 10 s, 0%		
	100%	A – 20 s, 100% B – 10 s, 50% X – 20 s, 100% Y – 10 s, 0%	AX – 10 s, 50% BY – 10 s, 50%	A – 10 s, 0% B – 10 s, 0% AX – 10 s, 50% BY – 10 s, 50%
Exp 2b	50%	A – 10 s, 50% B – 10 s, 50% X – 10 s, 50% Z – 10 s, 0%		
	100%	A – 20 s, 100% B – 10 s, 50% X – 20 s, 100% Z – 10 s, 0%		

Table 2. Design of Experiment 2a and 2b. The letters A, B, X and Y denote cues. The durations denote the average trial duration of the cue (see methods for details). The percentages indicate the probability of reinforcement per trial. Cues that were on average 10 s were presented twice as often as 20 s cues. The main difference experiments was that cue Y was a novel cue at the beginning of stage 2 in Experiment 2b, but a familiar cue in Experiment 2a.

Group	Cue A treatment	Stage 1	Stage 2
Continuous	100%	A – 20 s, 100% B – 20 s, 100%	A – 20 s, 100% B – 10 s, 0%
Partial		A – 20 s, 100% B – 10 s, 50%	
Continuous	50%	A – 10 s, 50% B – 20 s, 100%	A – 10 s, 50% B – 10 s, 0%
Partial		A – 10 s, 50% B – 10 s, 50%	

Table 3. Design of Experiment 3. The letters A and X denote cues. The durations denote the average trial duration of the cue (see methods for details). The percentages indicate the probability of reinforcement per trial. Cues that were on average 10 s were presented twice as often as 20 s cues.

Group	Stage 1	Stage 2
20 s x 3 trials	A – 10 s, 50% B – 20 s, 100% X – 10 s, 50% Y – 20 s, 100%	A – 10 s, 50% B – 20 s, 100% X – 10 s, 0% Y – 10 s, 0%
10 s x 6 trials		A – 10 s, 50% B – 20 s, 100% X – 20 s, 0% Y – 20 s, 0%

Table 4. Design of Experiment 4. The letters A, B, X and Y denote cues. The durations denote the average trial duration of the cue (see methods for details). The percentages indicate the probability of reinforcement per trial. Cues that were on average 10 s were presented twice as often as 20 s cues.

Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: some recent history and a theoretical extension. *Psychol Rev*, *69*, 306-328. https://doi.org/10.1037/h0046200

Andrew, B. J., & Harris, J. A. (2011). Summation of reinforcement rates when conditioned stimuli are presented in compound. *J Exp Psychol Anim Behav Process*, *37*(4), 385-393. https://doi.org/10.1037/a0024553

Arico, C., & McNally, G. P. (2014). Opioid receptors regulate blocking and overexpectation of fear learning in conditioned suppression. *Behav Neurosci*, *128*(2), 199-206. https://doi.org/10.1037/a0036133

Austen, J. M., Pickering, C., Sprengel, R., & Sanderson, D. J. (2021). Dissociating Representations of Time and Number in Reinforcement-Rate Learning by Deletion of the GluA1 AMPA Receptor Subunit in Mice. *Psychol Sci*, *32*(2), 204-217. https://doi.org/10.1177/0956797620960392

Austen, J. M., & Sanderson, D. J. (2019). Delay of reinforcement versus rate of reinforcement in Pavlovian conditioning. *J Exp Psychol Anim Learn Cogn*, *45*(2), 203-221. https://doi.org/10.1037/xan0000199

Austen, J. M., Sprengel, R., & Sanderson, D. J. (2022). Reinforcement rate and the balance between excitatory and inhibitory learning: Insights from deletion of the GluA1 AMPA receptor subunit. *J Exp Psychol Anim Learn Cogn*, *48*(4), 307-314. https://doi.org/10.1037/xan0000336

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behav Brain Res*, *110*(1-2), 67-72. https://doi.org/10.1016/s0166-4328(99)00185-0

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2003). Stimulus representation in SOP: I. Theoretical rationalization and some implications. *Behav Processes*, *62*(1-3), 5-25. https://doi.org/10.1016/s0376-6357(03)00016-0

Capaldi, E. J. (1966). Partial reinforcement: a hypothesis of sequential effects. *Psychol Rev*, *73*(5), 459-477. https://doi.org/10.1037/h0023684

Chan, C. K. J., & Harris, J. A. (2017). Extinction of Pavlovian conditioning: The influence of trial number and reinforcement history. *Behav Processes*, *141*(Pt 1), 19-25. https://doi.org/10.1016/j.beproc.2017.04.017

Chan, C. K. J., & Harris, J. A. (2019). The partial reinforcement extinction effect: The proportion of trials reinforced during conditioning predicts the number of trials to extinction. *J Exp Psychol Anim Learn Cogn*, *45*(1), 43-58. https://doi.org/10.1037/xan0000190

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychol Rev*, *107*(2), 289-344. https://doi.org/10.1037/0033-295x.107.2.289

Garfield, J. B., & McNally, G. P. (2009). The effects of FG7142 on overexpectation of Pavlovian fear conditioning. *Behav Neurosci*, *123*(1), 75-85. https://doi.org/10.1037/a0013814

Haney, R. Z., Calu, D. J., Takahashi, Y. K., Hughes, B. W., & Schoenbaum, G. (2010). Inactivation of the central but not the basolateral nucleus of the amygdala disrupts learning in response to overexpectation of reward. *J Neurosci*, *30*(8), 2911-2917. https://doi.org/10.1523/JNEUROSCI.0054-10.2010

Harris, J. A. (2019). The importance of trials. *J Exp Psychol Anim Learn Cogn*, *45*(4), 390-404. https://doi.org/10.1037/xan0000223

Harris, J. A. (2024). The sequencing of trials during partial reinforcement affects subsequent extinction. *J Exp Psychol Anim Learn Cogn*, *50*(1), 1-10. https://doi.org/10.1037/xan0000369

Harris, J. A., & Andrew, B. J. (2017). Time, trials, and extinction. *J Exp Psychol Anim Learn Cogn*, *43*(1), 15-29. https://doi.org/10.1037/xan0000125

Harris, J. A., Andrew, B. J., & Livesey, E. J. (2012). The content of compound conditioning. *J Exp Psychol Anim Behav Process*, *38*(2), 157-166. https://doi.org/10.1037/a0026605

Harris, J. A., & Bouton, M. E. (2020). Pavlovian conditioning under partial reinforcement: The effects of nonreinforced trials versus cumulative conditioned stimulus duration. *J Exp Psychol Anim Learn Cogn*, *46*(3), 256-272. https://doi.org/10.1037/xan0000242

Harris, J. A., & Carpenter, J. S. (2011). Response rate and reinforcement rate in Pavlovian conditioning. *J Exp Psychol Anim Behav Process*, *37*(4), 375-384. https://doi.org/10.1037/a0024554

Harris, J. A., Kwok, D. W., & Andrew, B. J. (2014). Conditioned inhibition and reinforcement rate. *J Exp Psychol Anim Learn Cogn*, *40*(3), 335-354. https://doi.org/10.1037/xan0000023

Harris, J. A., & Kwok, D. W. S. (2018). The probability of reinforcement per trial affects posttrial responding and subsequent extinction but not within-trial responding. *J Exp Psychol Anim Learn Cogn*, *44*(1), 23-35. https://doi.org/10.1037/xan0000158

Harris, J. A., Kwok, D. W. S., & Gottlieb, D. A. (2019). The partial reinforcement extinction effect depends on learning about nonreinforced trials rather than reinforcement rate. *J Exp Psychol Anim Learn Cogn*, *45*(4), 485-501. https://doi.org/10.1037/xan0000220

Harris, J. A., Patterson, A. E., & Gharaei, S. (2015). Pavlovian conditioning and cumulative reinforcement rate. *J Exp Psychol Anim Learn Cogn*, *41*(2), 137-151. https://doi.org/10.1037/xan0000054

Harris, J. A., Seet, M. S., & Kwok, D. W. S. (2019). The partial-reinforcement extinction effect does not result from reduced sensitivity to nonreinforcement. *J Exp Psychol Anim Learn Cogn*, *45*(2), 185-202. https://doi.org/10.1037/xan0000200

Holland, P. C. (2016). Effects of amygdala lesions on overexpectation phenomena in food cup approach and autoshaping procedures. *Behav Neurosci*, *130*(4), 357-375. https://doi.org/10.1037/bne0000149

Huh, K. H., Guzman, Y. F., Tronson, N. C., Guedea, A. L., Gao, C., & Radulovic, J. (2009). Hippocampal Erk mechanisms linking prediction error to fear extinction: roles of shock expectancy and contextual aversive valence. *Learn Mem*, *16*(4), 273-278. https://doi.org/10.1101/lm.1240109

Iordanova, M. D., Deroche, M. L., Esber, G. R., & Schoenbaum, G. (2016). Neural correlates of two different types of extinction learning in the amygdala central nucleus. *Nat Commun*, *7*, 12330. https://doi.org/10.1038/ncomms12330

Jiao, T., & Harris, J. A. (2024). Learning about trial sequences disrupts the partial reinforcement extinction effect in classical conditioning. *J Exp Psychol Anim Learn Cogn*, *50*(1), 11-24. https://doi.org/10.1037/xan0000370

Kehoe, E. J., & White, N. E. (2004). Overexpectation: response loss during sustained stimulus compounding in the rabbit nictitating membrane preparation. *Learn Mem*, *11*(4), 476-483. https://doi.org/10.1101/lm.77604

Khallad, Y., & Moore, J. (1996). Blocking, unblocking, and overexpectation in autoshaping with pigeons. *J Exp Anal Behav*, *65*(3), 575-591. https://doi.org/10.1901/jeab.1996.65-575

Kremer, E. F. (1978). The Rescorla-Wagner model: losses in associative strength in compound conditioned stimuli. *J Exp Psychol Anim Behav Process*, *4*(1), 22-36. https://doi.org/10.1037//0097-7403.4.1.22

Lattal, K. M., & Nakajima, S. (1998). Overexpectation in appetitive Pavlovian and instrumental conditioning. *Animal Learning & Behavior*, *26*, 351-360.

Lay, B. P. P., Choudhury, R., Esber, G. R., & Iordanova, M. D. (2023a). Correction to: Persistent disruption of overexpectation learning after inactivation of the lateral orbitofrontal cortex in male rats. *Psychopharmacology (Berl)*, *240*(3), 513. https://doi.org/10.1007/s00213-022-06215-4

Lay, B. P. P., Choudhury, R., Esber, G. R., & Iordanova, M. D. (2023b). Persistent disruption of overexpectation learning after inactivation of the lateral orbitofrontal cortex in male

rats. *Psychopharmacology (Berl)*, *240*(3), 501-511. https://doi.org/10.1007/s00213-022-06198-2

Lay, B. P. P., Nicolosi, M., Usypchuk, A. A., Esber, G. R., & Iordanova, M. D. (2019). Dissociation of Appetitive Overexpectation and Extinction in the Infralimbic Cortex. *Cereb Cortex*, *29*(9), 3687-3701. https://doi.org/10.1093/cercor/bhy248

Lee, C. T., Wong, P. T. P., & Chen, J. (1974). Durable Partial Reinforcement Effect and Social Dominance in 2 Inbred Mouse Strains. *Bulletin of the Psychonomic Society*, *4*(Nb4), 400-402. https://doi.org/Doi 10.3758/Bf03336732

Lucantonio, F., Kambhampati, S., Haney, R. Z., Atalayer, D., Rowland, N. E., Shaham, Y., & Schoenbaum, G. (2015). Effects of prior cocaine versus morphine or heroin self-administration on extinction learning driven by overexpectation versus omission of reward. *Biol Psychiatry*, *77*(10), 912-920. https://doi.org/10.1016/j.biopsych.2014.11.017

Mallea, J., Schulhof, A., Gallistel, C. R., & Balsam, P. D. (2024). Both probability and rate of reinforcement can affect the acquisition and maintenance of conditioned responses. *J Exp Psychol Anim Learn Cogn*, *50*(4), 254-266. https://doi.org/10.1037/xan0000386

McNally, G. P., Pigg, M., & Weidemann, G. (2004). Blocking, unblocking, and overexpectation of fear: a role for opioid receptors in the regulation of Pavlovian association formation. *Behav Neurosci*, *118*(1), 111-120. https://doi.org/10.1037/0735-7044.118.1.111

Miller, R. R., Hallam, S. C., Hong, J. Y., & Dufore, D. S. (1991). Associative structure of differential inhibition: implications for models of conditioned inhibition. *J Exp Psychol Anim Behav Process*, *17*(2), 141-150. https://doi.org/10.1037//0097-7403.17.2.141

Navarro, J. I., Hallam, S. C., Matzel, L. D., & Miller, R. R. (1989). Superconditioning and overshadowing. *Learning and Motivation*, *20*(2), 130-152.

Norton, S. E., & Harris, J. A. (2022). Delaying extinction weakens the partial reinforcement extinction effect. *J Exp Psychol Anim Learn Cogn*, *48*(4), 321-335. https://doi.org/10.1037/xan0000319

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Anim Learn Behav*, *30*(2), 73-95. https://doi.org/10.3758/bf03192911

Pearce, J. M., Redhead, E. S., & Aydin, A. (1997). Partial reinforcement in appetitive Pavlovian conditioning with rats. *Q J Exp Psychol B*, *50*(4), 273-294. https://doi.org/10.1080/713932660

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*, 77-94.

Rescorla, R. A. (1970). Reduction in the effectiveness of reinforcement after prior excitatory conditioning. *Learning and Motivation*, *1*(371-381).

Rescorla, R. A. (1999). Summation and overexpectation with qualitatively different outcomes. *Animal Learning & Behavior*, *27*, 50-62.

Rescorla, R. A. (2006). Spontaneous recovery from overexpectation. *Learn Behav*, *34*(1), 13-20. https://doi.org/10.3758/bf03192867

Rescorla, R. A. (2007). Renewal after overexpectation. *Learn Behav*, *35*(1), 19-26. https://doi.org/10.3758/bf03196070

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-99). Appleton-Century-Crofts.

Ruprecht, C., Izurieta, H. S., Wolf, J. E., & Leising, K. J. (2014). Overexpectation in the contect of reward timing. *Learning and Motivation*, *47*, 1-11.

Sanderson, D. J., Jones, W. S., & Austen, J. M. (2016). The effect of the amount of blocking cue training on blocking of appetitive conditioning in mice. *Behav Processes*, *122*, 36-42. https://doi.org/10.1016/j.beproc.2015.11.007

Shanab, M. E., & Ralph, L. (1979). Negative Contrast and Partial-Reinforcement Effects as a Function of Crowded Rearing Conditions in the Rat. *Journal of General Psychology*, *100*(1), 13-26. https://doi.org/Doi 10.1080/00221309.1979.9710521

Sissons, H. T., & Miller, R. R. (2009). Overexpectation and trial massing. *J Exp Psychol Anim Behav Process*, *35*(2), 186-196. https://doi.org/10.1037/a0013426

Strickland, J. A., Austen, J. M., Sprengel, R., & Sanderson, D. J. (2024). Knockout of NMDARs in CA1 and dentate gyrus fails to impair temporal control of conditioned behavior in mice. *Hippocampus*, *34*(3), 126-140. https://doi.org/10.1002/hipo.23593

Team, J. (2024). *JASP*. In (Version 0.19.3)

Terao, K., Matsumoto, Y., Alvarez, B., & Mizunami, M. (2022). Spontaneous recovery from overexpectation in an insect. *Sci Rep*, *12*(1), 9827. https://doi.org/10.1038/s41598-022-13800-2

Vogel, E. H., Ponce, F. P., & Wagner, A. R. (2019). The development and present status of the SOP model of associative learning. *Q J Exp Psychol (Hove)*, *72*(2), 346-374. https://doi.org/10.1177/1747021818777074

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Erlbaum.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *J Exp Psychol*, *76*(2), 171-180. https://doi.org/10.1037/h0025414

Williams, B. A., & McDevitt, M. A. (2002). Inhibition and superconditioning. *Psychol Sci*, *13*(5), 454-459. https://doi.org/10.1111/1467-9280.00480

Witnauer, J. E., & Miller, R. R. (2009). Contrasting the overexpectation and extinction effects. *Behav Processes*, *81*(2), 322-327. https://doi.org/10.1016/j.beproc.2009.01.010

Wong, P. T. P., Lee, C. T., & Novier, F. H. (1971). Partial Reinforcement Effect (Pre) Sustained through Extinction and Continuous Reinforcement in 2 Strains of Inbred Mice. *Psychonomic Science*, *22*(3), 141-143. <Go to ISI>://WOS:A1971Y246000007
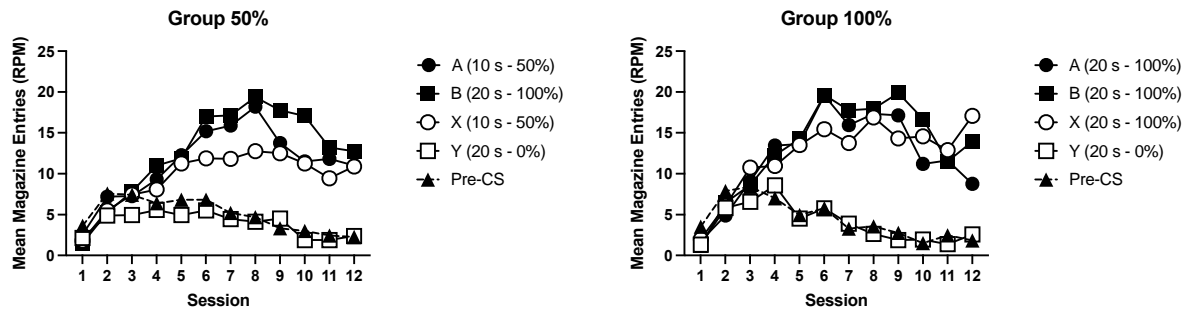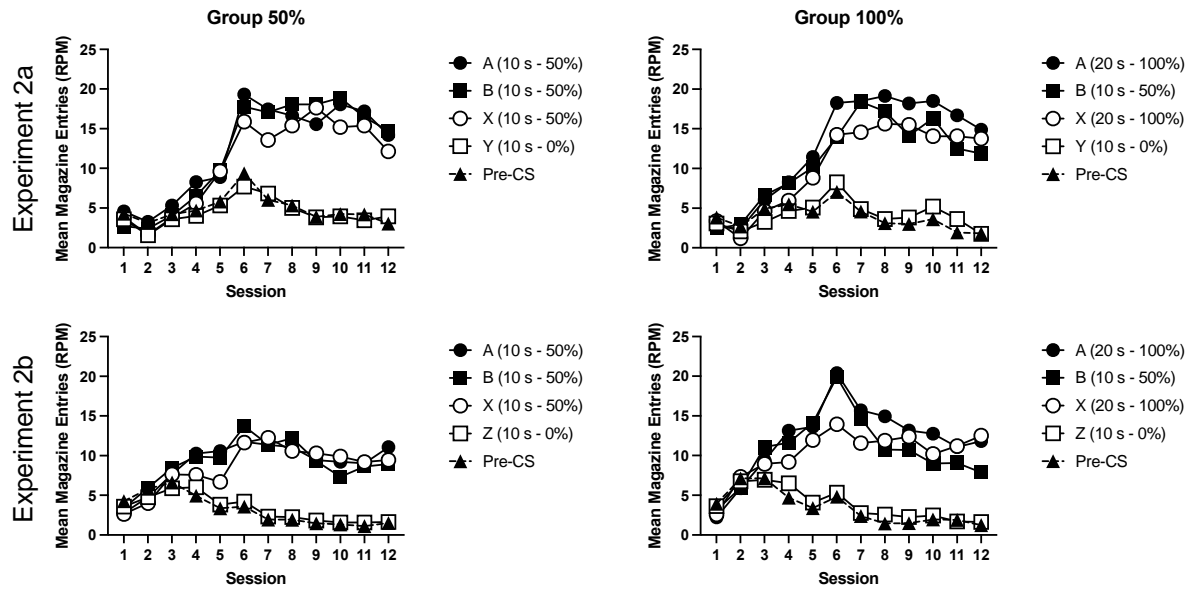
Figure 1. Acquisition of conditioned responding in stage 1, Experiment 1. Mean magazine entries are shown as rate per minute (RPM). Cues A, B and X were reinforced and cue Y was nonreinforced. For each cue, the durations and percentages indicate the mean trial duration and probability of reinforcement per trial.
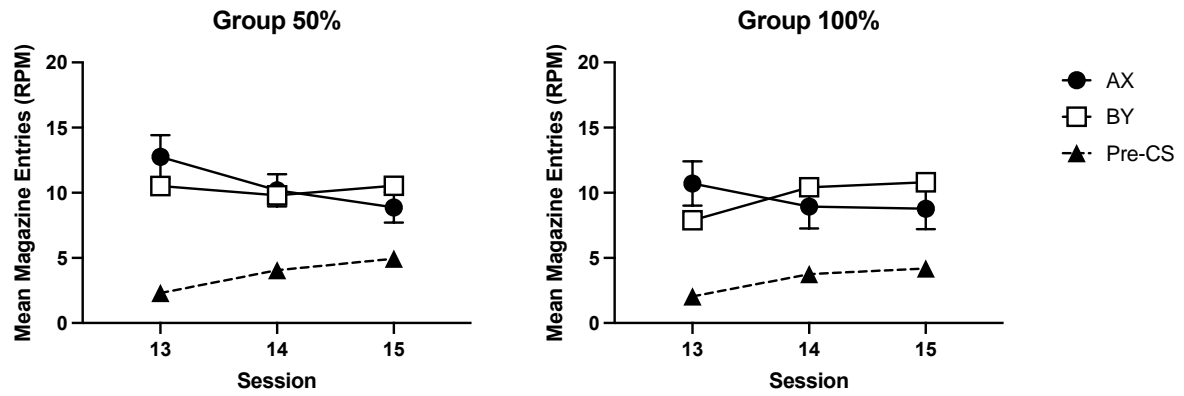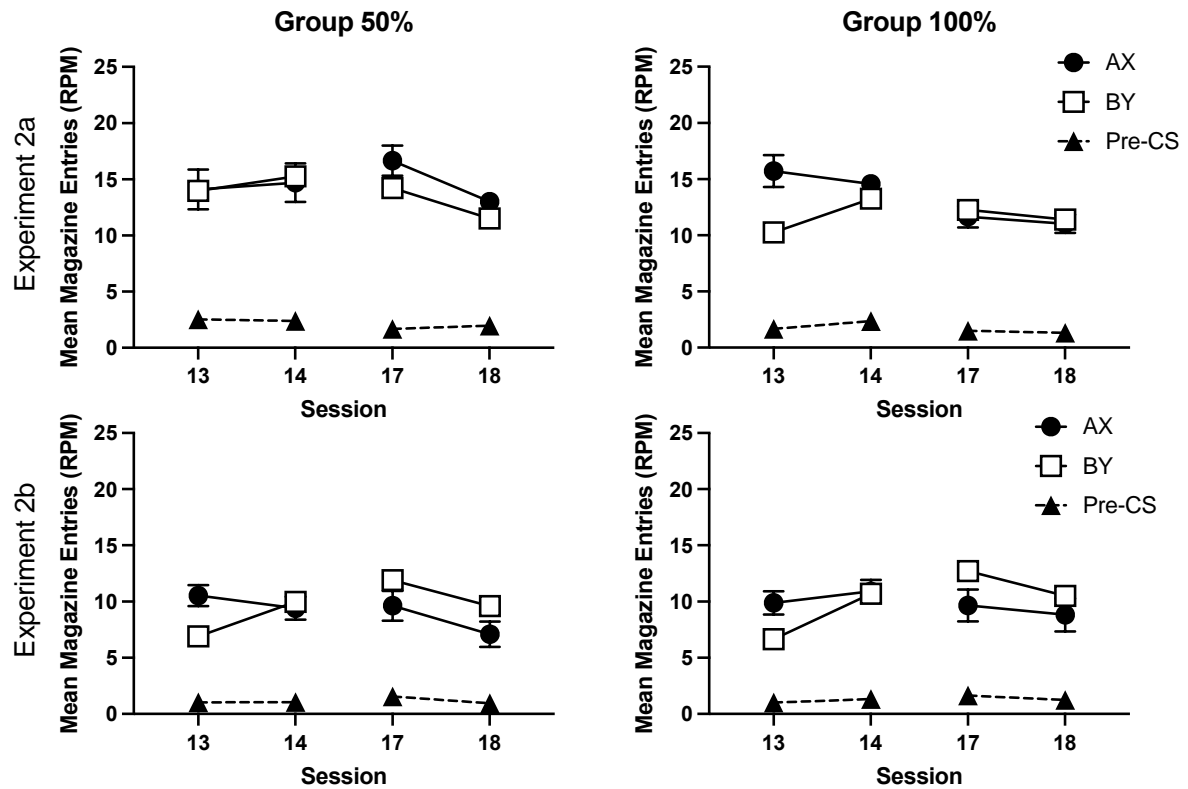
Figure 2. Compound conditioning in stage 2, Experiment 1. Mice initially showed summation of conditioned responding by responding more to AX than BY. Mean magazine entries are shown as rate per minute (RPM). The error bars for AX data indicate ±standard error of the mean of the within-subjects difference between AX and BY.
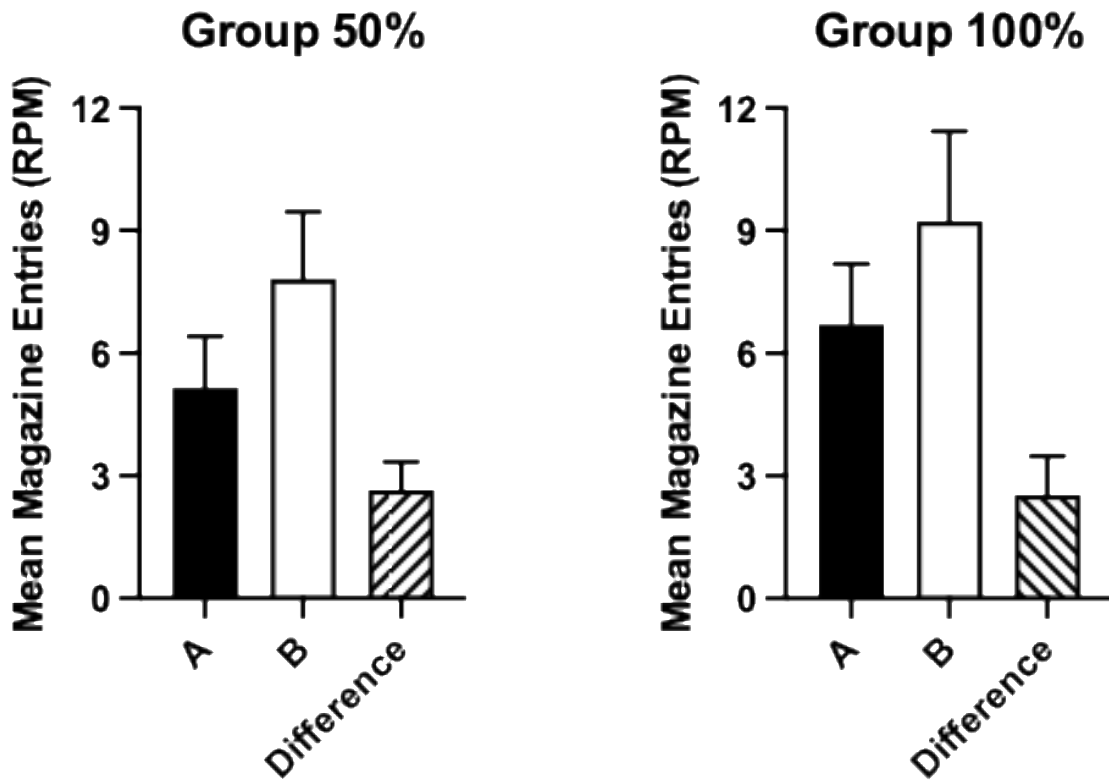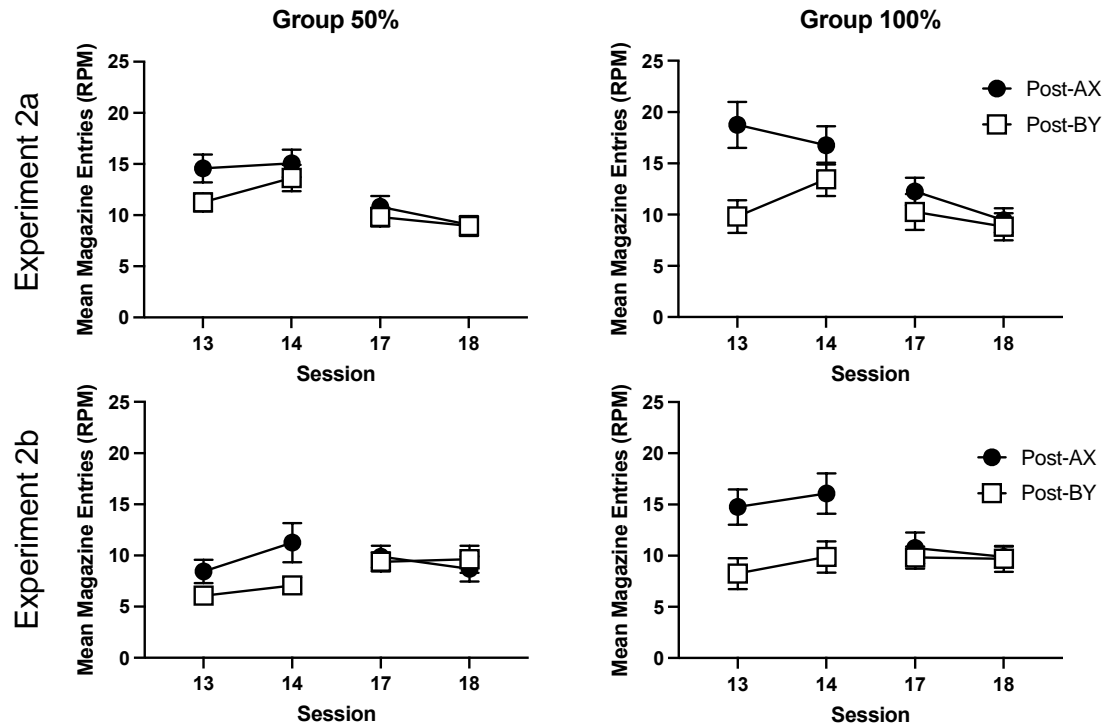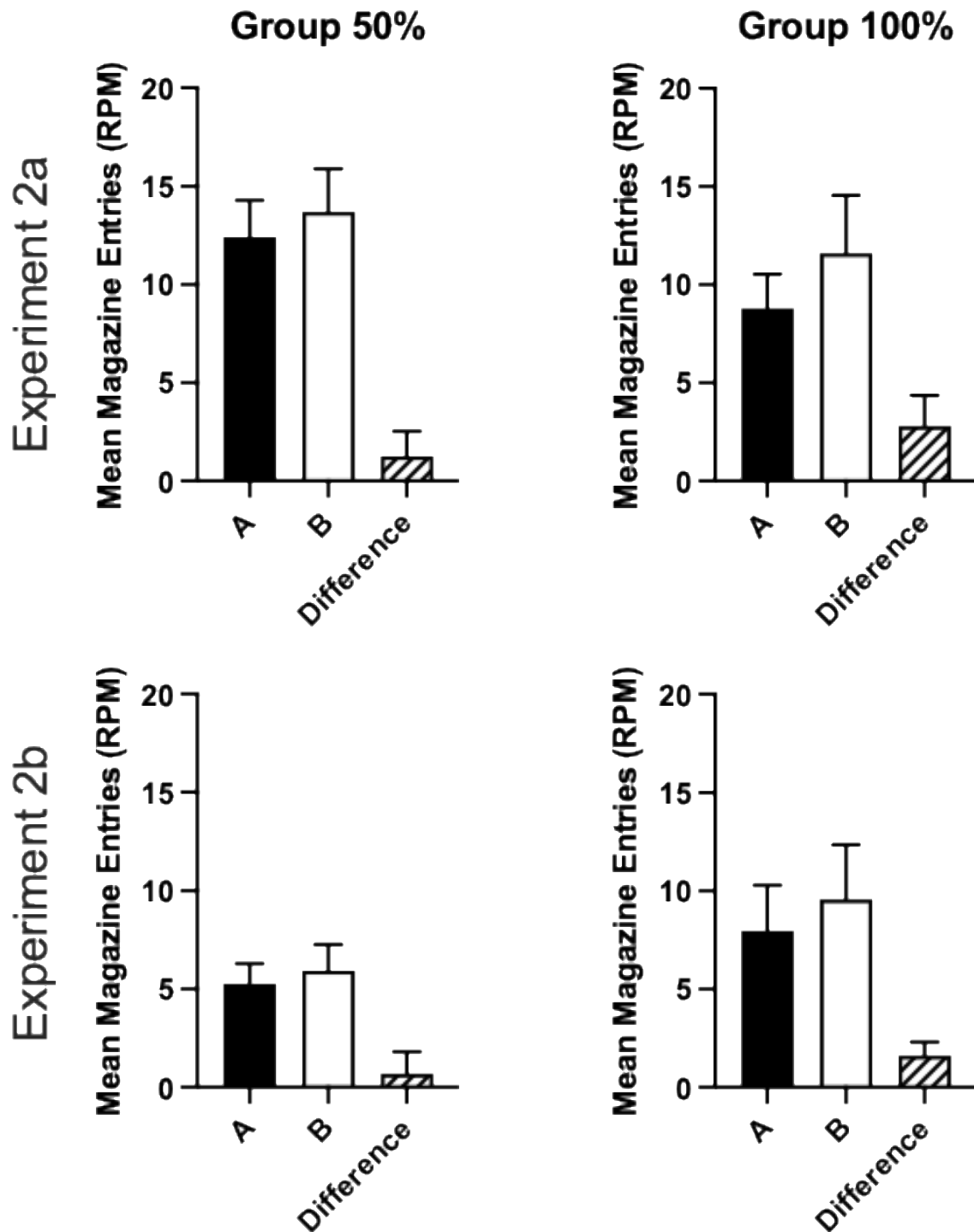
Figure 3. Overexpectation test, Experiment 1. Mice responded less to A, the overexpectation cue than B, the control cue. The striped bar indicates the mean difference in response rates between cue B and A. Mean magazine entries are shown as rate per minute (RPM). Error bars indicate the standard error of the mean.
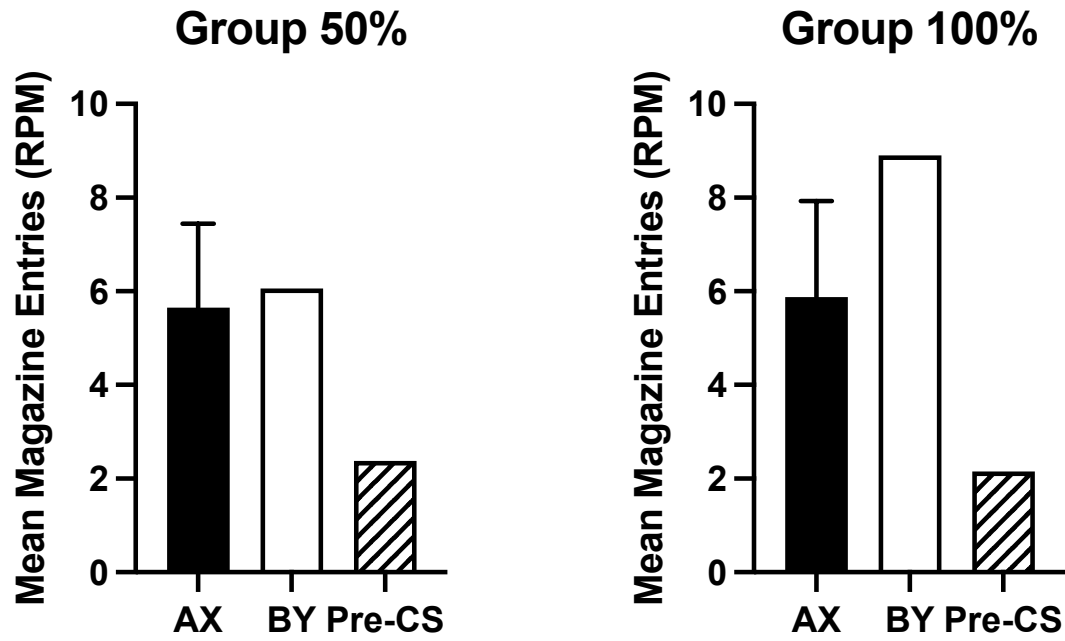
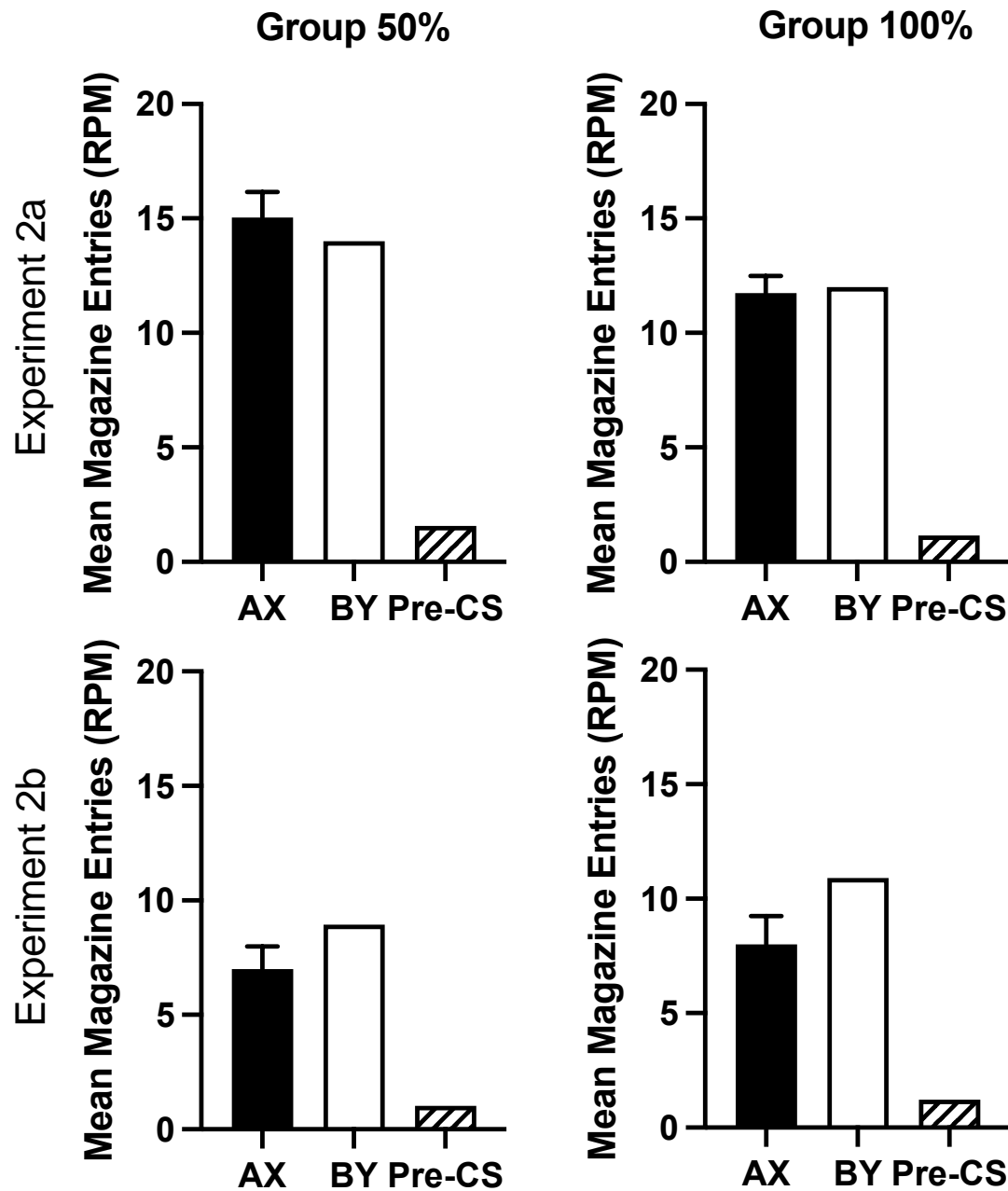*Figure 4. Responding to the reinforced compounds during the overexpectation test,*

Experiment 1. Mean magazine entries are shown as rate per minute (RPM). The error bars

for AX indicate the standard error of the mean of the within-subjects difference between AX

and BY.

Figure 5. Acquisition of conditioned responding in stage 1, Experiment 2a and 2b. Mean magazine entries are shown as rate per minute (RPM). Cues A, B and X were reinforced. In Experiment 2a, cue Y was nonreinforced. In Experiment 2b, cue Z was nonreinforced. For each cue, the durations and percentages indicate the mean trial duration and probability of reinforcement per trial.

Figure 6. Compound conditioning in stage 2, Experiment 2a and 2b. Mice initially showed summation of conditioned responding by responding more to AX than BY. Mean magazine entries are shown as rate per minute (RPM). The error bars for AX data indicate ±standard error of the mean of the within-subjects difference between AX and BY.

Figure 7. Post-CS responding in the compound conditioning stage of Experiment 2a and 2b. Group 100% initially showed greater responding to AX than group 50%. Mean magazine entries are shown as rate per minute (RPM). Note that in contrast to previous figures, the error bars for post-AX and post-BY indicate ±standard error of the mean.

Figure 8. Overexpectation test, Experiment 2a and 2b. Mice responded less to A, the overexpectation cue than B, the control cue. The striped bar indicates the mean difference in response rates between cue B and A. Mean magazine entries are shown as rate per minute (RPM). Error bars indicate the standard error of the mean.

Figure 9. Responding to the reinforced compounds during the overexpectation test,

Experiment 2a and 2b. Mean magazine entries are shown as rate per minute (RPM). The

error bars for AX indicate the standard error of the mean of the within-subjects difference

between AX and BY.
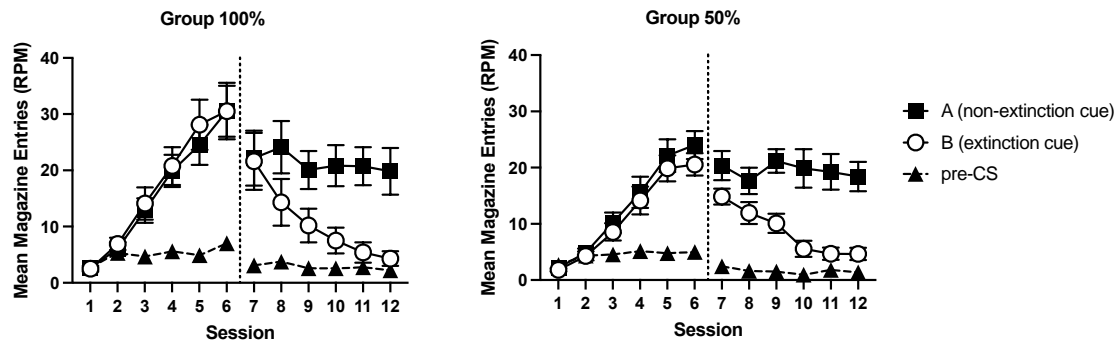
Figure 10. Stage 1 acquisition (sessions 1-6) and stage 2 extinction training (sessions 7-12_ in Experiment 3. Mice acquired responding to cues A and B during stage 1 and selectively extinguished responding to cue B during stage 2. Note that, similar to Figure 7, error bars indicate ±standard error of the mean.
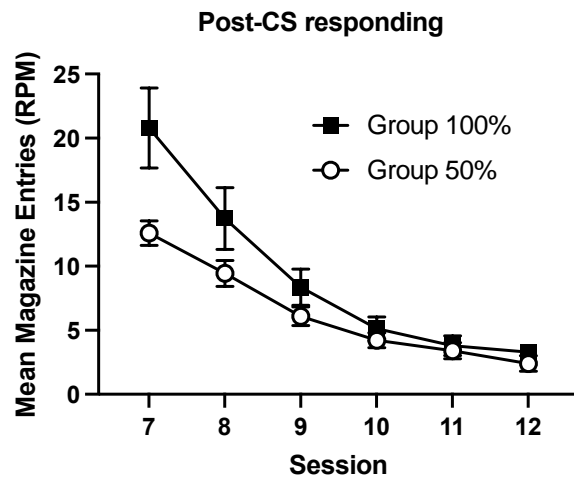
Figure 11. Post-CS responding to the nonreinforced cue B during the extinction phase of

Experiment 3. Group 100% initially responded at a higher rate than group 50%. Error bars

indicate ±standard error of the mean.
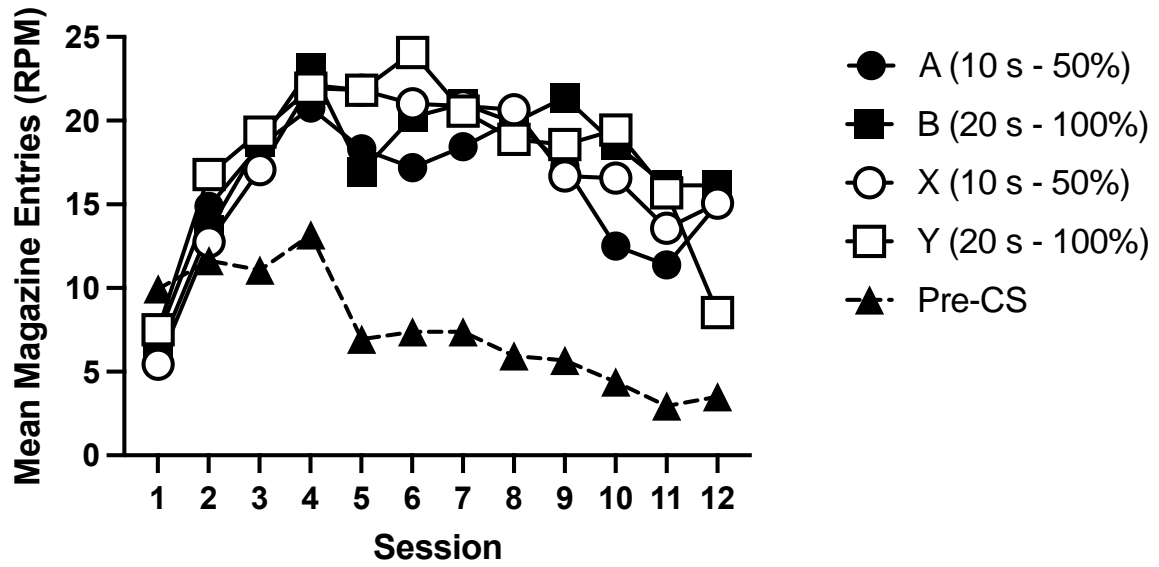
Figure 12. Acquisition of conditioned responding in stage 1, Experiment 4. For each cue, the durations and percentages indicate the mean trial duration and probability of reinforcement per trial.
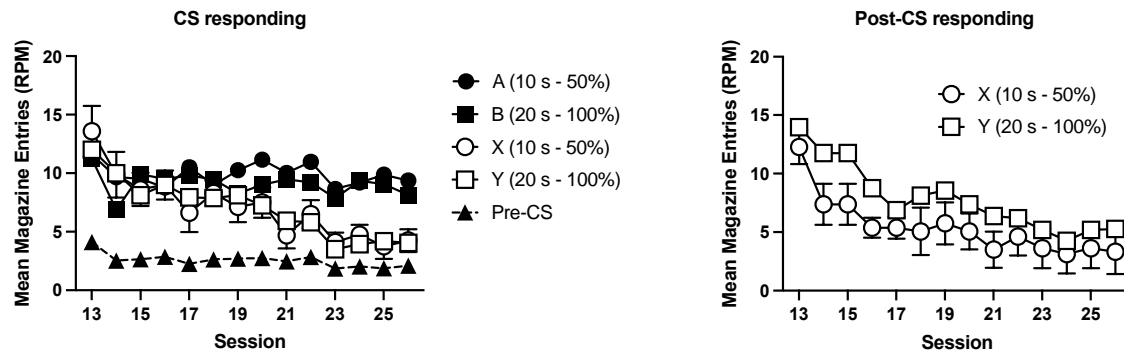
Figure 13. Extinction phase of Experiment 4. The left panel shows the selective extinction of conditioned responding to cues X and Y over the course of training. The right panel shows the post-CS responding to the nonreinforced cues, X and Y. Mice showed greater responding to the previously 100% reinforced cue than the 50% cue. For both panels, the error bars indicate ±standard error of the mean within-subject difference between cues X and Y.

Supplemental Material

Click here to access/download

Supplemental Material

Supplemental Material 2025-02-10.docx