A Novel Task for Modeling Negative Reinforcement with Temporal Precision

by

J.P. Gamboa

B.A., College of the Holy Cross, 2011

M.A., Universität Tübingen, 2015

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

John Paul Gamboa

It was defended on

January 5, 2022

and approved by

Susanne Ahmari M.D., Ph.D., Associate Professor, Department of Psychiatry

Caroline Runyan, Ph.D., Assistant Professor, Department of Neuroscience

Yan Dong, Ph.D., Professor, Department of Neuroscience

Thesis Advisor/Dissertation Director: Susanne Ahmari M.D., Ph.D., Associate Professor, Department of Psychiatry

A Novel Task for Modeling Negative Reinforcement with Temporal Precision

John Paul Gamboa, M.S.

University of Pittsburgh, 2022

Obsessive-compulsive disorder (OCD) is a burdensome psychiatric illness that affects 1%-3% of the world's population (Kessler, Chiu, Demler, Merikangas, & Walters, 2005). The disorder is characterized by recurrent thoughts, images, and impulses that are disturbing (i.e., obsessions) and repetitive behaviors or mental acts (i.e., compulsions). Currently, the pathophysiology of OCD is not well-understood. One theory in the field is that compulsions are ingrained by negative reinforcement. Specifically, compulsive behaviors are thought to be negatively reinforced because they temporarily relieve anxiety. If negative reinforcement is indeed a key component of OCD's pathophysiology, neural circuits mediating this reinforcement step are a valuable target of investigation for developing novel treatments of OCD. Neural mechanisms of negative reinforcement have typically been studied in the laboratory using active and passive avoidance learning paradigms. Although current active avoidance paradigms have allowed us to identify some neural circuits critical for avoidance acquisition, expression, and extinction, they have a serious limitation. Specifically, it is challenging to determine exactly when the avoidance behavior initiates, and this in turn makes it difficult to time-lock neural signals to the time of avoidance. Here, we present a solution to this issue by developing a novel task in which mice avoid by pressing a lever. This kind of task provides a discrete time point for the behavioral response, defining clear periods before, during, and after the conditioned response to time-lock acquisition of neural signals or deliver interventions. We also describe how several variations of this task affect acquisition and behavioral performance, show that animals continue to engage in the task at a wide range of

reinforcement probabilities (including 0%), and demonstrate that the conditioned response is extinguishable through an animal analog of exposure with response prevention (ERP) therapy.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Obsessive-compulsive disorder (OCD) is a burdensome psychiatric illness that affects 1%-3% of the world's population (Kessler et al., 2005). The disorder is characterized by recurrent thoughts, images, and impulses that are disturbing (i.e., obsessions) and repetitive behaviors or mental acts (i.e., compulsions). Currently, the pathophysiology of OCD is not well-understood. However, it is clear that obsessions cause emotional distress, and clinical reports indicate that patients often perform compulsive behaviors with the conscious aim of relieving anxiety associated with their obsessions. Accordingly, one theory in the field is that compulsions are ingrained by negative reinforcement. Specifically, compulsive behaviors are thought to be negatively reinforced because they temporarily relieve anxiety. If negative reinforcement is indeed a key component of OCD's pathophysiology, neural circuits mediating this reinforcement step are a valuable target of investigation for developing novel treatments of OCD.

Unfortunately, preclinical studies of negative reinforcement and its relevance to OCD are sparse (Geramita, Yttri, & Ahmari, 2020). Neural mechanisms of negative reinforcement have typically been studied in the laboratory using both active and passive avoidance learning paradigms. Active avoidance tasks involve performing a behavior to avoid harm, whereas passive avoidance requires inhibiting a behavior. In typical preclinical active avoidance studies, rodents are trained to avoid shocks by moving to the opposite side of a shuttle box upon presentation of a conditioned stimulus (CS). This paradigm has revealed some of the neural circuitry necessary for acquisition and expression of avoidance. There is an emerging consensus that avoidance learning occurs in stages mediated by different neural circuits (Cain, 2019; LeDoux, Moscarello, Sears, & Campese, 2017; Manning, Bradfield, & Iordanova, 2021). For example, lesioning the basolateral amygdala (BLA) prior to training impairs acquisition of active avoidance, but it has no effect on expression in over-trained animals (Lázaro-Muñoz, LeDoux, & Cain, 2010). An alternative active avoidance task developed by the lab of Greg Quirk involves training animals to avoid shocks by mounting a platform (Diehl, Bravo-Rivera, & Quirk, 2019). This task has been particularly useful for investigating neural mechanisms of persistent avoidance and avoidance extinction (Bravo-Rivera, Roman-Ortiz, Brignoni-Perez, Sotres-Bayon, & Quirk, 2014; Bravo-Rivera, Roman-Ortiz, Montesinos-Cartagena, & Quirk, 2015; Martínez-Rivera, Bravo-Rivera, Velázquez-Díaz, Montesinos-Cartagena, & Quirk, 2019; Martínez-Rivera et al., 2020; Rosas-Vidal et al., 2018).

Although current active avoidance paradigms have allowed us to identify some neural circuits critical for avoidance acquisition, expression, and extinction, they have a serious limitation. Specifically, it is challenging to determine exactly when the avoidance behavior initiates, and this in turn makes it difficult to time-lock neural signals to the time of avoidance. For example, does avoidance occur when the animal initiates movement, reaches safety, or somewhere in-between? While this ambiguity is unproblematic for deciding whether avoidance occurs on any given trial, it is a problem for applying temporally precise experimental techniques (e.g., determining if particular neuronal populations are active at initiation of avoidance).

Here, we present a solution to this issue by developing a novel task in which mice avoid by pressing a lever. This kind of task provides a discrete time point for the behavioral response, defining clear periods before, during, and after the conditioned response to time-lock acquisition of neural signals or deliver interventions. Previously, Fernando, Mar, Urcelay, Dickinson, and Robbins (2015) successfully trained rats on a free-operant lever-press avoidance task. However, a task developed for mice would also be valuable given the broader range of available transgenic mouse strains (Lee, Lavoie, Liu, Chen, & Liu, 2020) and technical challenges associated with delivering light to the larger rat brain in optogenetic experiments (Igarashi et al., 2018). In addition to presenting a solution to the time-locking problem in active avoidance experiments, we describe how several variations of this task affect acquisition and behavioral performance, show that animals continue to engage in the task at a wide range of reinforcement probabilities (including 0%), and demonstrate that the conditioned response is extinguishable through an animal analog of exposure with response prevention (ERP) therapy.

2.0 Methods and Materials

2.1 Animals

Male 8- to 12-week-old C57BL/6 mice from Jackson Laboratories were housed four per cage. All animals were maintained on a reverse 12/12-hour light-dark cycle (lights off at 7:00 AM; on at 7:00PM) with *ad libitum* access to food and water except during the food restriction condition described below. All experiments were approved by the Institutional Animal Use and Care Committee at the University of Pittsburgh in compliance with National Institutes of Health guidelines for the care and use of laboratory animals.

2.2 Apparatus

All animals were trained and tested in operant chambers from Med Associates. Animals were habituated to the chambers for 2 days prior to training. For the first experiment comparing a reward group and no reward group, animals had access to the full operant chamber. In all other experiments, a barrier was inserted to reduce the chamber size by half to encourage engagement with the lever.

2.3 Positive Reinforcement Training

Animals in the reward group were food-restricted to 85%-90% of their free-feeding bodyweight during positive reinforcement training. No other animals were food restricted. The reward group was first trained to receive chocolate pellets on a variable time 60 second schedule (VT 60) for 3 days. This was followed by 2 days of training to lever-press for chocolate pellets on a fixed ratio 1 schedule (FR 1). The mice were then returned to *ad libitum* feeding and given a day to recover their initial bodyweight, followed by one last day of FR1 training. The next day, the reward group began negative reinforcement training. For each day the reward group underwent positive reinforcement training, animals in the no reward group spent equivalent time in the operant chambers with no access to reward or lever. Both groups began negative reinforcement training on the same day.

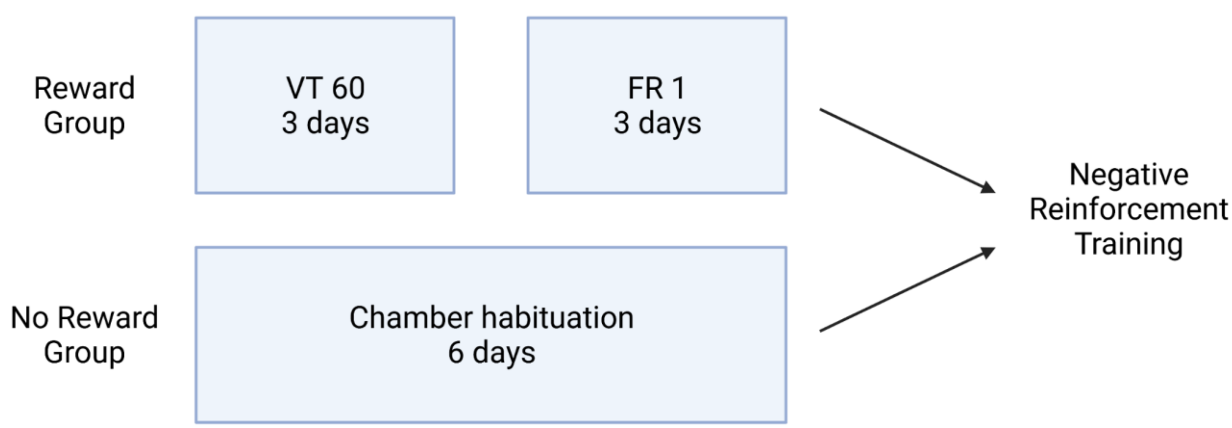


Figure 1 – Positive reinforcement training.

2.4 Negative Reinforcement Training

Each training session began with a 180 second acclimation period during which only the house light and fan were on. Sessions consisted of 5 trials signaled by the extension of a single lever (counterbalanced between animals to extend on the left or right side of the chamber) and simultaneous onset of a 30s light cue above the lever. For animals with prior positive reinforcement training, the lever was the same one used during the FR1. There were 3 possible outcomes on each trial (Figure 2). If an animal pressed the lever at any point within 30 seconds, the trial immediately ended, signaled by the offset of the light CS and retraction of the lever. Trials were separated by a pseudorandom intertrial interval (ITI) of 30s +/- 10s. If an animal did not press the lever during the avoid period, the light CS turned off and a series of 20 footshocks began. Each shock was set to 0.3mA, lasted 2s, and was separated by a 15s inter-stimulus interval (ISI). For animals conditioned with a tone CS in addition to the light CS (double CS condition), a 3 kHz tone played for the duration of this escape period. If an animal pressed the lever at any point within the escape period, the trial ended. Shocks were immediately terminated, the lever retracted, and the tone CS turned off if applicable. If an animal failed to press the lever during both the avoid and escape periods, it experienced the full series of 20 footshocks before moving on to the next trial or session end. All animals were trained for 15 consecutive days, and acquisition was defined as receiving ≤25 of the 100 possible shocks in a session by day 15.

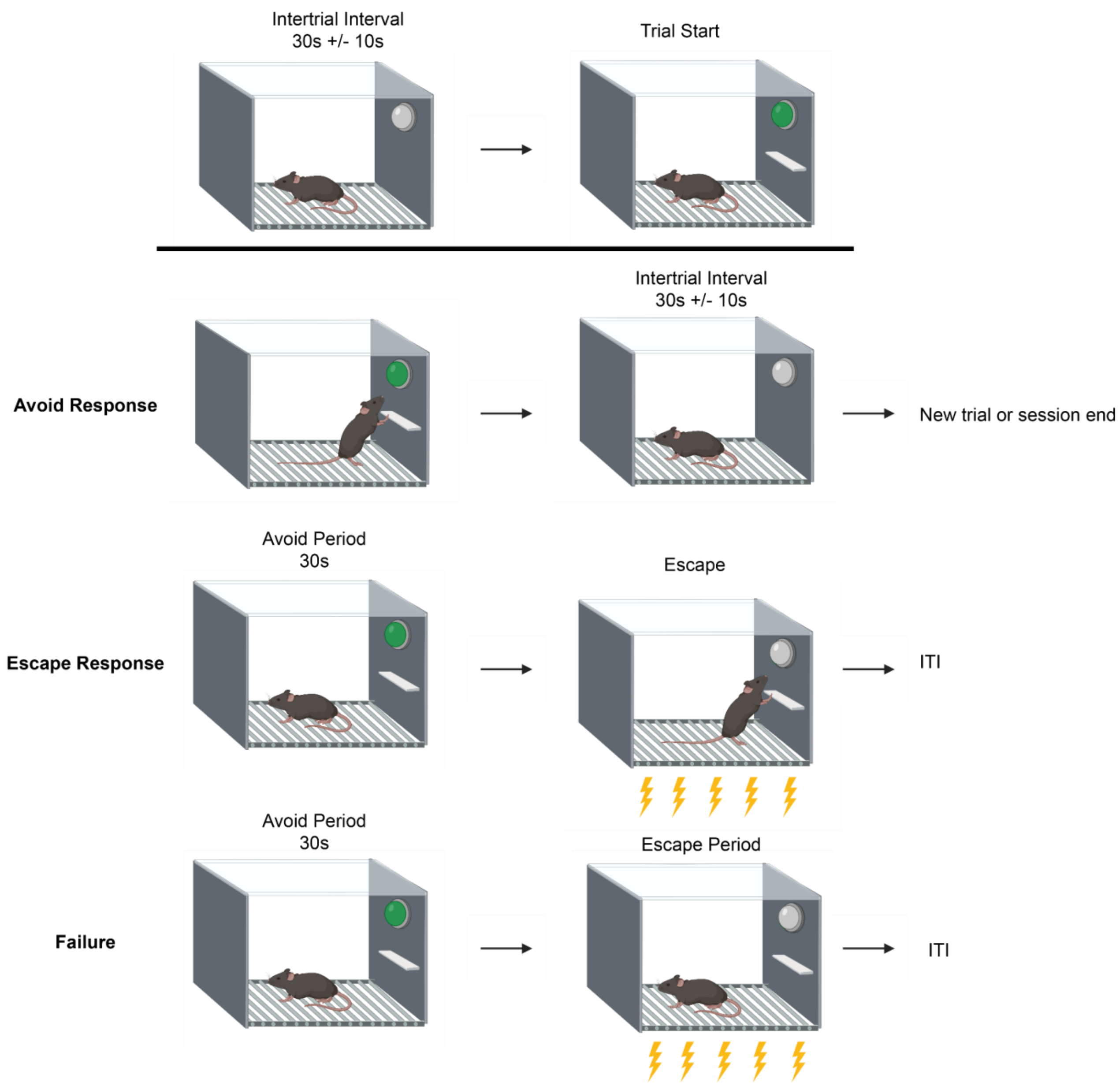


Figure 2 – Task structure. Trial start followed by three possible outcomes on each trial.

2.5 Probabilistic Reinforcement Testing

Sessions consisted of 20 trials with 5 possible shocks each. The acclimation period, conditioned stimuli, shock intensity, shock duration, ISI, and ITI were all the same as during negative reinforcement training. On reinforcing trials, a single lever-press immediately ended the trial. On non-reinforcing trials, the lever remained extended for the full duration of the trial regardless of how many times it was pressed. Lever-presses during the 30s warning signal were counted as "avoid attempts" and presses during the shock series were counted as "escape attempts". The first session for all animals was set to 100% reinforcement probability, meaning all 20 trials were reinforcing. Each subsequent day, the reinforcement probability was lowered by a fixed amount. For the reward and no reward groups in the first experiment this was 5%, and for all other groups in subsequent experiments it was 10%.

For testing at 5%, trial 1 was always reinforcing, and the rest were non-reinforcing. For all other tested probabilities, the first two trials of every session were always reinforcing, and which of the remaining 18 trials were reinforcing or non-reinforcing was randomized each day for each animal according to that day's probability. For example, on the 50% reinforcement probability day, the first two trials and 8 other random trials in the session were reinforcing, while the remaining 10 were non-reinforcing.

2.6 Retraining and CS-US Extinction

Animals from the single and double CS groups were retrained for 3 days at which point all of them met task acquisition criterion (<25 shocks received). They were retrained under the single

CS condition (light CS during the avoid period). One group then repeated their previous testing: probabilistic reinforcement testing as described above followed by 7 days at 0% reinforcement probability. The other group was immediately retested for 7 days at 0% after retraining. Following the seventh day at 0%, both groups underwent 3 days of CS-US extinction training in which they were given 20 presentations of the light CS (30s each) in the absence of the lever and US. The ITI between CS presentations was the same as before (30s +/- 10s). After 3 days of extinction training, animals were given an extinction test. The test consisted of 20 CS presentations again in the absence of the US, but the lever extended for the duration of the CS and retracted during each ITI.

2.7 Simulated Exposure with Response Prevention (ERP)

Animals put through simulated ERP were first given negative reinforcement training as described above. One group then did probabilistic reinforcement testing until they reached 10% reinforcement probability, while the other group was consistently tested at 100% reinforcement over the same 10-day period. Both groups then did 4 or 2 days of ERP. During ERP, access to the lever was blocked by a transparent plexiglass barrier inserted into each operant chamber. In each trial, the light CS was illuminated for 30s as usual, and the ITI was the same as before. The lever extended at the beginning of each trial and retracted during the ITI. For the 4 days of ERP condition, animals were given 100 trials per session, totaling 400 exposures. For the 2 days of ERP condition, animals were given 20 trials per session, totaling 40 exposures. For the probe test following ERP, the barrier was removed, and all groups were given 40 trials of CS and lever exposure with the usual ITI.

2.8 Data Analysis

Statistical analyses of behavioral data were performed using two-way repeated measures ANOVA with Tukey's or Šídák post-hoc tests. For probabilistic reinforcement testing, data are presented as attempts per minute rather than total attempts to control for differences in attempt periods between sessions. For example, at 90% reinforcement, only 2 of 20 trials are non-reinforcing, giving animals a total of 60s to perform avoid attempts (30s per trial), whereas at 50% reinforcement, 10 of 20 trials are non-reinforcing, giving animals 300s. For analysis of the reward vs. no reward group at 0% reinforcement probability (Fig. 6a, b), a mixed-effects model was used because one animal was removed from further testing after day 5 due to reaching learned helplessness (i.e., no lever-presses during the session). For CS-US extinction training (Fig. 7a), both groups underwent 7 days of testing at 0% reinforcement probability prior to extinction training. Analysis before and after CS-US extinction compared avoid attempts on the last day at 0% to avoid attempts on the extinction test day (Fig. 7a). For analysis of avoid attempts before and after ERP (Fig. 7b, c), paired t tests of the probabilistic reinforcement groups were used. To analyze the data, we compared avoid attempts in the 18 unreinforced trials the day before ERP began (10% reinforcement probability day; 18/20 unreinforced trials = .9) and the first 18 trials of the probe. For the probe test after ERP, animals were tested for 40 trials. All analyses were run using Prism Graphpad 9.

3.0 Results

3.1 A Lever-Press Negative Reinforcement Task for Mice

It remains an open question whether naïve mice can learn to lever-press to avoid aversive stimuli. To test this possibility, we gave naïve mice negative reinforcement training and compared their performance with mice that had previous positive reinforcement training ("reward group"). Both groups were conditioned with a light CS during the avoid period and a tone CS during the escape period. There was no significant Day x Group interaction or main effect of Group over the full course of training (Fig. 3a). However, post-hoc testing revealed that the reward group avoided significantly more trials on day 1 of training (Fig. 3a; Šídák correction, $p < 0.05$). All animals in the reward group met the task acquisition criterion (8/8 animals) compared to only half in the no reward group (4/8 animals), indicating that the 'no reward' task was more challenging (Supplementary Table 1). Further analysis of only animals that acquired the task revealed that the two groups differed in performance over time with the no reward group improving at a faster rate and avoiding significantly more by the end of training (Fig. 3b; two-way repeated measures ANOVA, Day x Group interaction: $p < 0.0001$). These results suggest that prior positive reinforcement training helps to acquire the task, but it also worsens performance on it.

Given the superior performance of the acquired animals from the no reward group, all subsequent experiments were conducted without any positive reinforcement training and in operant chambers with barriers reducing their size by half (see Methods section). We next aimed to further streamline the task by testing the influence of the second CS during the escape period. We trained one group of mice with both conditioned stimuli as before (double CS group) and another group with only the light CS during the avoid period (single CS group). Most animals in both groups met our acquisition criterion (Supplementary Table 1; 9/12 in double CS group, 10/12 in the single CS group). There were no significant differences between the groups in avoidances or shocks received (Fig. 4a, b). Note that as long as avoidance is imperfect, it is possible for animals to perform similarly with respect to avoidances but differ in shocks received. Altogether, these results show that mice without any prior training can learn to lever-press to avoid ≥90% of shocks within 15 days, and a single CS during the avoid period is sufficient

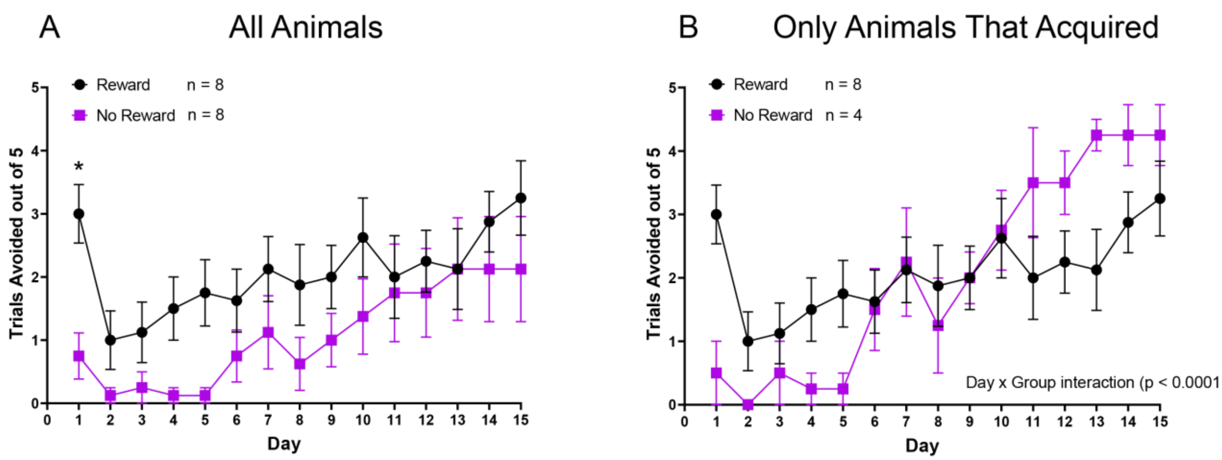


Figure 3 – Positive reinforcement training helps acquisition but not performance on the task. (A) There were no significant differences between the reward and no reward groups (two-way repeated measures ANOVA, Day x Group interaction $p = 0.55$). However, animals in the reward group avoid significantly more trials on day 1 (post-hoc test Šídák's correction, $p < 0.05$). Reward n = 8, No Reward n = 8. (B) Limiting analysis to animals that met the task acquisition criterion reveals significant differences in performance over the course of training (two-way repeated measures ANOVA, Day x Group interaction: $p < 0.0001$). Reward n = 8, No Reward n = 4.

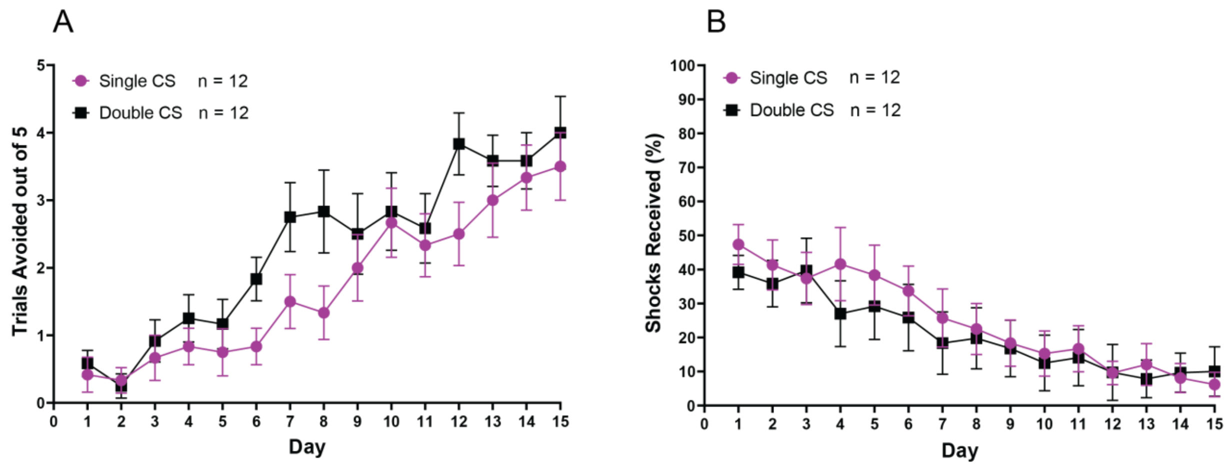


Figure 4 – Removing the second CS during the escape period does not significantly change performance. (A) Single and Double CS groups showed no significant differences in trials avoided during training (two-way repeated measures ANOVA, Day x Group interaction p = 0.34). (B) Groups did not significantly differ in shocks received (two-way repeated measures ANOVA, Day x Group interaction p = 0.94); Single CS n = 12, Double CS n = 12.

3.2 Negative Reinforcement under Uncertainty

We next asked at what reinforcement probability, if any, would animals stop engaging in the task. We further tested animals that had successfully passed training by gradually lowering the probability of reinforcement each session. The single and double CS groups showed no significant differences in reinforcing trials avoided or percent of avoidable shocks received (Fig. 5a, b). Interestingly, they also showed no main effect of reinforcement probability, revealing that engagement in the task was consistent throughout testing (Fig. 5a, b). The lack of significant differences and changes across different reinforcement probabilities were also found for the reward and no reward groups (Supplementary Fig. 1). During non-reinforcing trials, we found no group differences between the single and double CS groups with respect to avoid attempts per minute

(Fig. 5c). However, there was a main effect of reinforcement probability (Fig. 5c, two-way repeated measures ANOVA, main effect $p < 0.01$) with a significant decrease in rate of avoid attempts at 80% and 70% reinforcement probability compared to 90% (Fig. 5c, Tukey's test, $p < 0.05$). Both the main effect and initial decrease in rate of avoid attempts were replicated by probabilistically tested groups in the ERP experiments described below (Supplementary Fig. 2). Altogether, these data show that animals trained on the task continue lever-pressing at a wide range of reinforcement probabilities, and do not on average display learned helplessness.

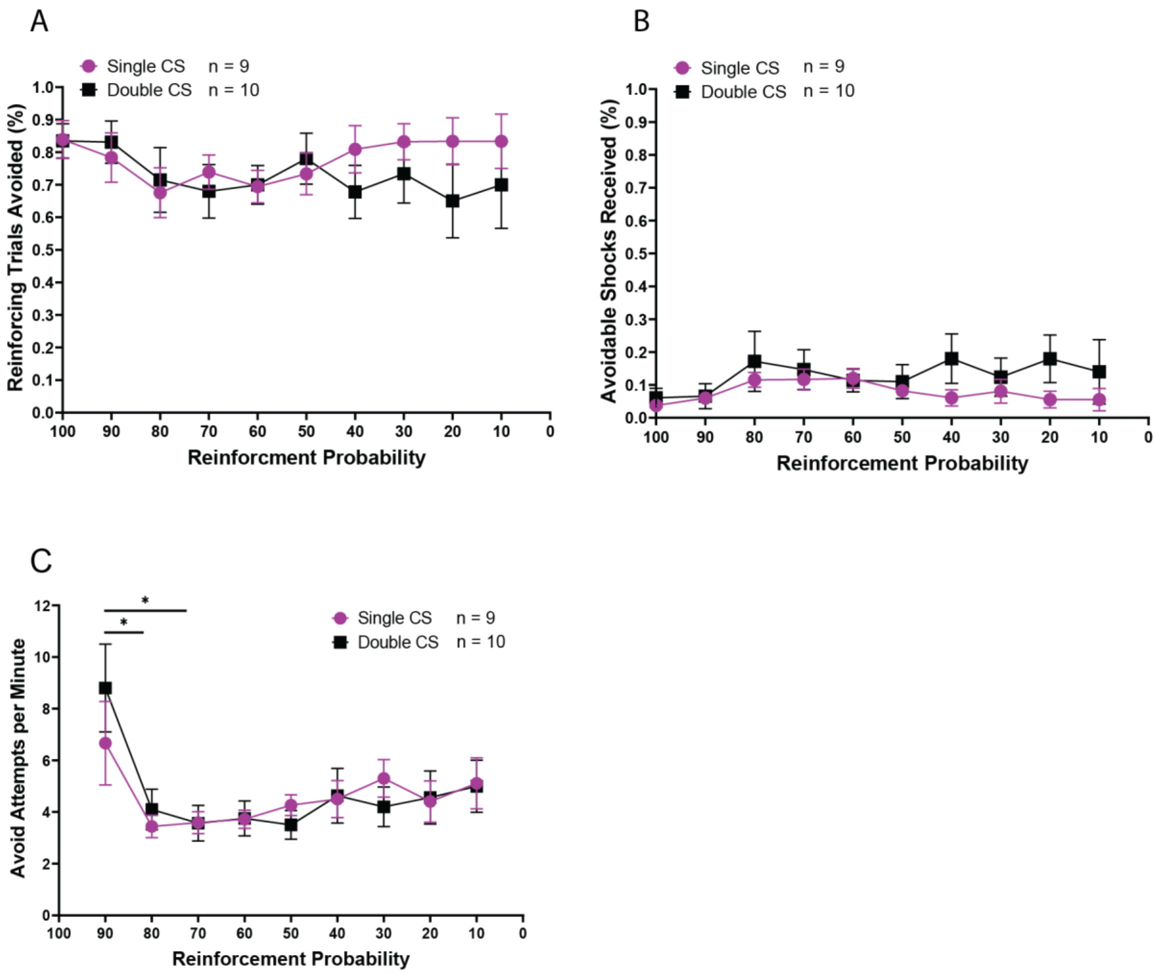


Figure 5 – Avoidance is resistant to decreasing reinforcement probability. (A) Single and Double CS groups continue avoiding on most trials as reinforcement probability decreases with no significant differences between them (two-way repeated measures ANOVA, Day x Group interaction $p = 0.58$). (B) The groups also did not differ in

percent of avoidable shocks received across all reinforcement probabilities (two-way repeated measures ANOVA, Day x Group interaction $p = 0.78$). (C) Both Single and Double CS groups significantly reduced avoid attempts per minute on non-reinforcing trials at 80% and 70% reinforcement probability compared to 90% (post-hoc Tukey's test, $p < 0.05$); Single CS $n = 9$, Double CS $n = 10$.

3.3 Engagement at 0% Reinforcement Probability

Because animals continued avoiding and attempting to avoid throughout all positive reinforcement probabilities, we next tested them at 0% (i.e., all 20 trials in the session were non-reinforcing). All groups were tested for 7 days. We pre-determined that animals would be removed from further testing if they failed to perform at least one lever-press during a session. Only one animal met this criterion, and it was an animal from the reward group on day 5. The no reward group maintained a higher average rate of avoid attempts per minute which decreased over the 7 days unlike the reward group's rate (Fig. 6a; mixed-effects analysis, Day x Group interaction $p < 0.005$), but the groups did not differ with respect to escape attempts per minute (Fig. 6b). By contrast, the single and double CS groups were significantly different in both rate of avoid attempts (Fig. 6c; two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$) and escape attempts (Fig. 6d; two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$). Specifically, the single CS group showed a small but significant increase in avoid attempts/minute, and fewer escape attempts/min. There was no main effect of day in either of the two comparisons. These results show that animals remain engaged in the task at 0% reinforcement probability for at least a week.

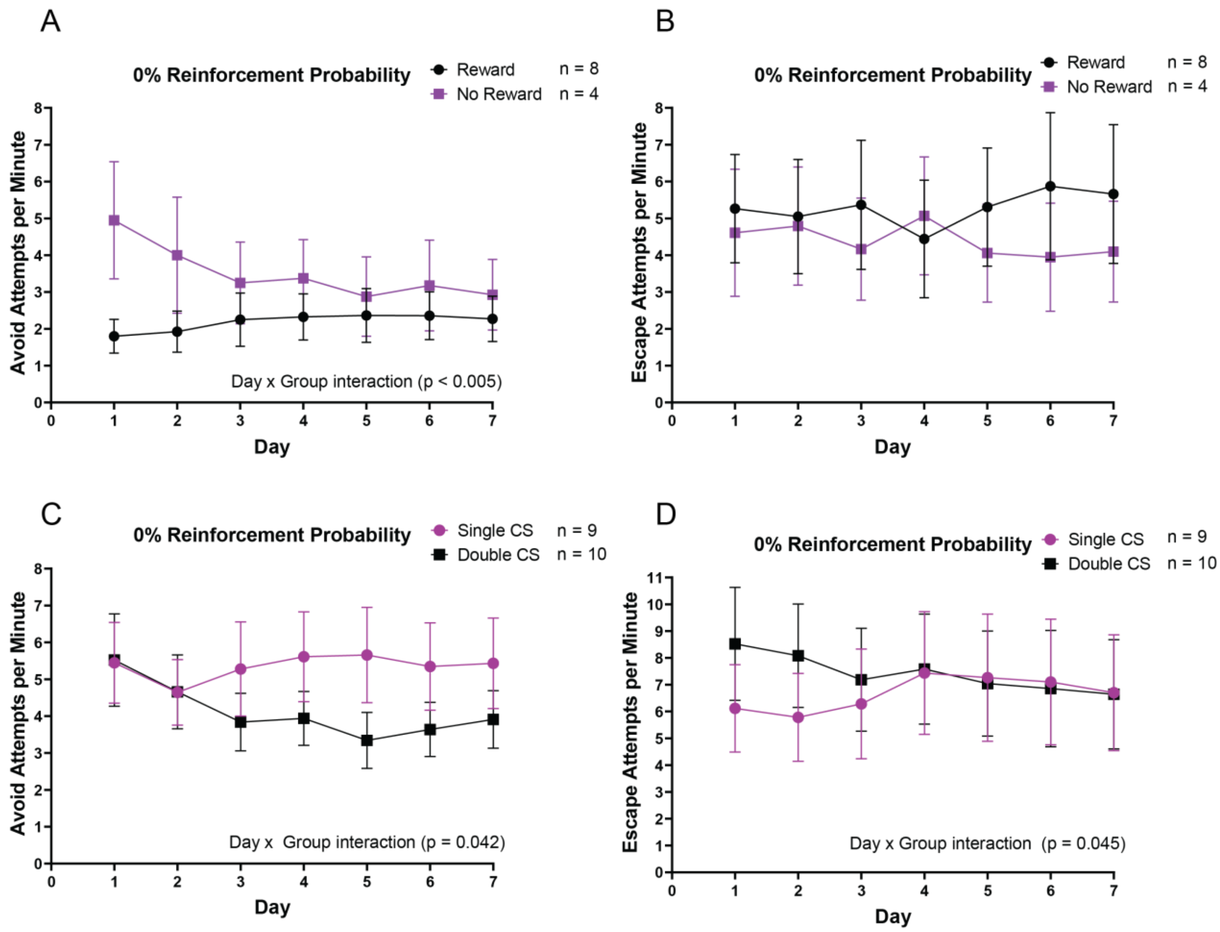


Figure 6 – Animals continue performing avoid and escape attempts over 7 days at 0% reinforcement probability. (A) The reward group's rate of avoid attempts remained stable over 7 days whereas the no reward group's rate decreased (mixed-effects analysis, interaction $p < 0.005$). (B) Both groups had higher rates of escape attempts compared to avoid attempts, but the groups did not significantly differ (two-way repeated measures ANOVA, Day x Group interaction $p = 0.76$). (C) Avoid attempts per minute were significantly different between the Single CS and Double CS groups (two-way repeated measures ANOVA, Day x Group interaction $p = 0.042$). (D) Both groups also differed in escape attempts per minute, though this only narrowly reached statistical significance (two-way repeated measures ANOVA, Day x Group interaction $p = 0.045$); Reward $n = 8$, No Reward $n = 4$, Single CS $n = 9$, Double CS $n = 10$.

3.4 Extinguishing Avoidance Behavior

Lastly, we tested various procedures for extinguishing conditioned behavior. We first retrained animals from the single and double CS groups. Since we previously found no significant differences between them, all animals were retrained under the single CS condition. After retraining to our task acquisition criterion, one group repeated their previous testing (probabilistic reinforcement over 10 days followed by 7 days at 0%) while the other was immediately retested at 0% reinforcement probability. After the seventh day at 0% reinforcement probability, all animals underwent 3 days of CS-US extinction training followed by an extinction test. We then compared avoid attempts the day before and after extinction training. The groups showed no significant differences, but there was a significant reduction in avoid attempts between the two days, indicating that animals successfully extinguished the avoidance response (Fig. 7a; two-way repeated measures ANOVA, Day main effect $p < 0.005$).

We next simulated exposure with response prevention therapy (ERP). It is currently considered the most effective treatment for obsessive-compulsive disorder (Ferrando & Selai, 2021). A new group of animals was trained for 15 days and then tested from 100% reinforcement probability down to 10% in 10% steps as in previous experiments. After ERP over 4 days (100 exposures per session), animals were given a probe test in which access to the lever was restored. There was a significant reduction in avoid attempts (Fig. 7b; paired t test, $p < 0.0005$). During the probe test, there were no significant differences in avoid attempts between trials (Supplementary Fig. 3a; two-way ANOVA, trial main effect $p = 0.2109$). Since 400 exposures with response prevention was relatively extensive, we repeated the experiment with a new group of animals but used 2 days of ERP with only 20 trials per session. This time, there was also a significant drop in avoid attempts (Fig. 7c; paired t test, $p < 0.05$). There was again no main effect of trial during the probe test (Supplementary Fig. 3b; two-way ANOVA , trial main effect $p = 0.2752$).

For both experiments in which animals were given probabilistic reinforcement testing prior to ERP, we concurrently tested a separate group at 100% reinforcement probability for the same number of days. We found that groups maintained at 100% reinforcement perform significantly better on the task in terms of reinforcing trials avoided (Supplementary Fig. 4a, two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$; Supp. Fig. 4b, Day x Group interaction $p < 0.01$) and percent of avoidable shocks received (Supplementary Fig. 4c, two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$; Supp. Fig. 4d, Day x Group interaction $p < 0.005$). These data suggest that although animals tested under probabilistic reinforcement continue to engage in the task, they lose some motivation as reflected by their impaired performance relative to animals tested consistently at 100% reinforcement.

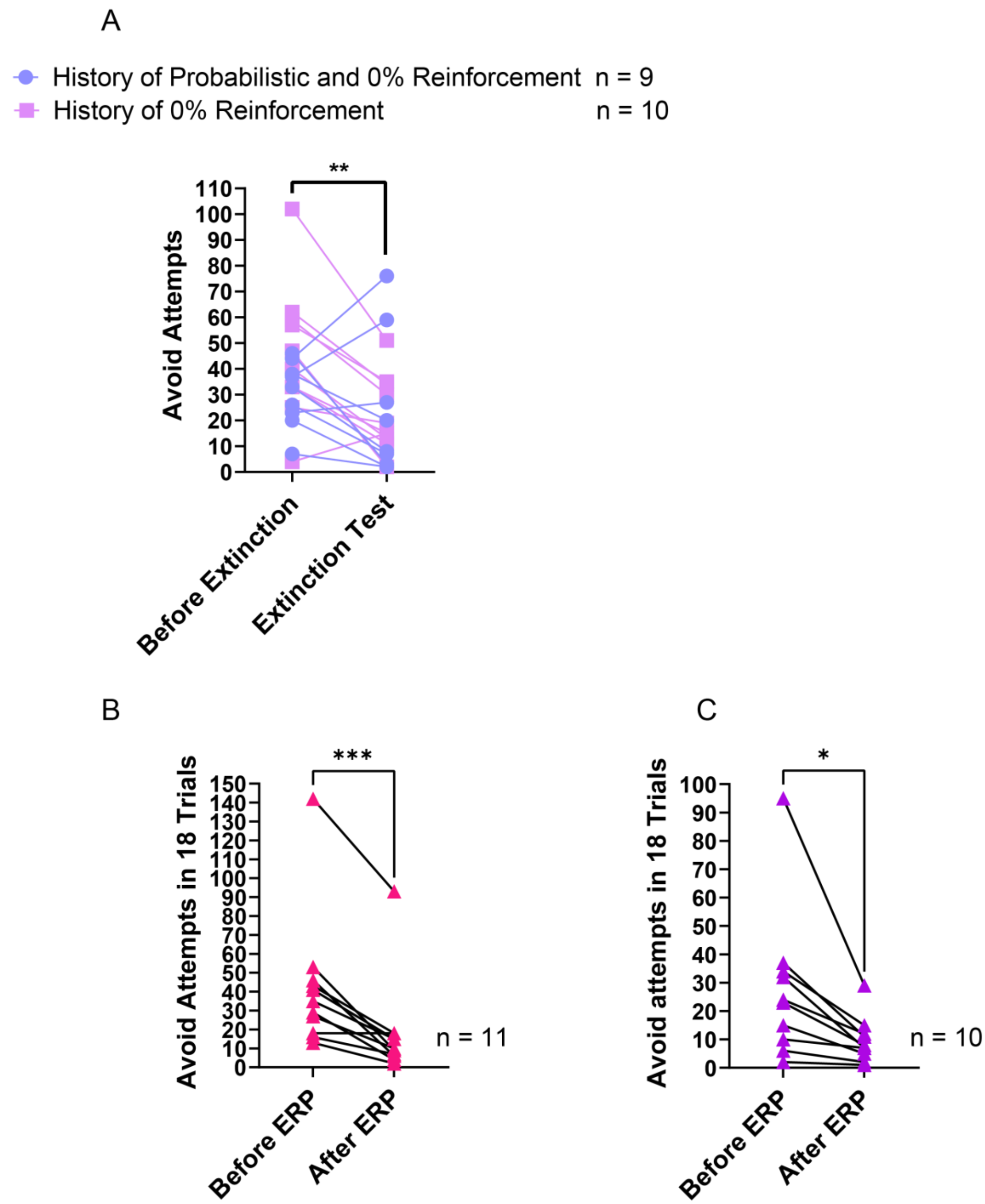


Figure 7 – Avoid attempts extinguish after extinction training and ERP. (A) Retrained animals successfully extinguish after 3 days of CS-US extinction training (two-way repeated measures ANOVA, Day main effect $p < 0.005$). (B) Extensive ERP (100 exposures per day over 4 days) significantly reduces avoid attempts in animals previously given probabilistic negative reinforcement (paired t test, $p < 0.0005$). (C) Two days of ERP (20 exposures per session) is also sufficient to extinguish avoid attempts (paired t test, $p < 0.05$).

4.0 Discussion

Negatively reinforced behavior is a promising target for treatment of OCD. Here we developed a novel active avoidance task for modeling negative reinforcement in mice. Our initial experiments aimed at streamlining the task. We first found that naïve animals avoid on significantly more trials than animals previously trained to lever-press for reward. Although prior reward training helps with acquisition of the negative reinforcement task, most animals successfully acquire without it when the chamber size is reduced with a barrier (~80% across all our experiments; Supplementary Fig. 1). Moreover, including the positive reinforcement stage complicates interpretation of behavior on the task by associating multiple reinforcers with the same behavior. We next showed that training avoidance with a second CS during the escape period does not help acquisition or behavioral performance. Removing this CS is optimal as it only adds unnecessary perceptual processing to the task.

Modeling extinction of negative reinforcement is another important goal for preclinical studies. Most active avoidance studies have focused on acquisition and expression (Bravo-Rivera et al., 2015). However, extinguishing negatively reinforced behavior faces unique challenges. As Manning et al. (2021) remark, well-trained animals rarely encounter the shock reinforcer due to high levels of avoidance, and this can prevent them from detecting the change in contingency when the CS is presented in extinction. This helps explain why many experimenters have observed since the 1950s that conditioned avoidance is especially resistant to extinction (Baum, 1970). Continued performance of the conditioned response is actually detrimental to extinction learning. Interestingly, the opposite is true in the case of behaviors conditioned through positive reinforcement. In rats well-trained to lever-press for reward, extinction training through presentation of the CS with an opportunity to lever-press is more effective at extinguishing the conditioned response than CS presentation in the absence of the lever (Bouton, Trask, & Carranza-Jasso, 2016). Removing the opportunity to perform the behavior is what prevents animals from learning that it is no longer reinforced.

Given the above considerations, we decided to attempt extinction of avoidance without allowing animals to perform the conditioned response. There are two broad options in a lever-press task. The first option is to present the CS in the absence of the lever and US. We called this CS-US extinction training because animals learn that the two are no longer associated. Our results show that this is an effective way of extinguishing avoid attempts in animals with a history of probabilistic negative reinforcement testing. However, one might think that this procedure is in fact a kind of ERP because animals are prevented from responding in the presence of the CS. We would argue that this is a semantic issue, and one could reasonably conceptualize the procedure as CS-US extinction training or an analog of ERP. A more substantive issue is whether animals associate the lever itself with shock. If so, keeping the lever retracted during extinction training does not present the full CS. This raises the second option for extinction: simulating ERP by blocking access to the visual cue and lever with a clear barrier. Given the possibility that animals associate the lever with shock, this is the preferable option, and we found that both 2 and 4 days of this procedure lead to significant reductions in avoid attempts.

In typical active avoidance paradigms, performance of the conditioned response within the avoid or escape period is always reinforcing. In platform-mediated avoidance, for example, every avoid attempt is reinforcing because movement onto the platform precludes delivering footshocks. This feature compromises the ecological validity of such experiments because threats in the real world are uncertain, and behaviors aimed at avoiding them are not guaranteed to succeed. After successful training on our task, the probability of reinforcement can be easily modified to better model negative reinforcement under uncertainty. One option is to program each lever-press to reinforce with a certain probability. However, we reasoned that in such a scenario animals could avoid on most trials even at low reinforcement probabilities by simply pressing repeatedly. For some studies, assigning a probability of reinforcement to each press may be desirable. Alternatively, experimenters can assign a probability to each trial by fixing the proportion of reinforcing to non-reinforcing trials and randomizing them in the session as we have done here.

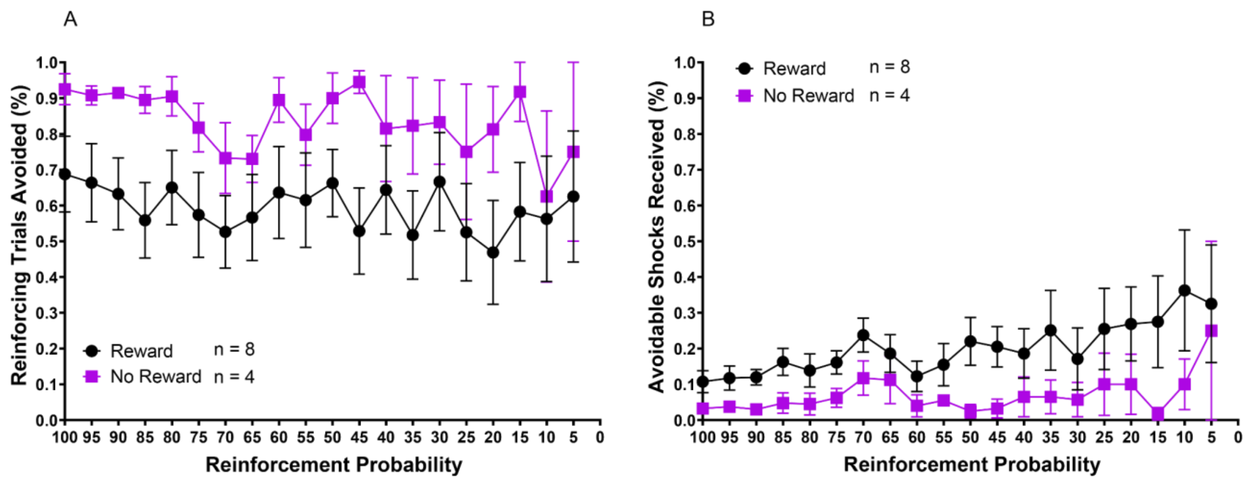
One of our goals in developing this task was to facilitate future investigations that rely on various experimental techniques. In this respect, the main advantage of our task over typical active avoidance paradigms is that it defines clear periods before, during, and after avoidance to time-lock acquisition of neural signals or deliver interventions. Avoidance in the task also requires little movement, making it easier to test animals with cable implants and reducing movement artifacts. The task could also be modified for experiments requiring head fixation. Although it is widely acknowledged in clinical practice that negative reinforcement of compulsive behaviors is likely a significant contributor to the pathophysiology of OCD, there are not enough preclinical studies of these neural mechanisms. The task presented in this paper opens many possibilities for experimenters to investigate these circuits in ways not possible with typical active avoidance tasks.

5.0 Supplementary Figures

Training parameters **Proportion acquired**

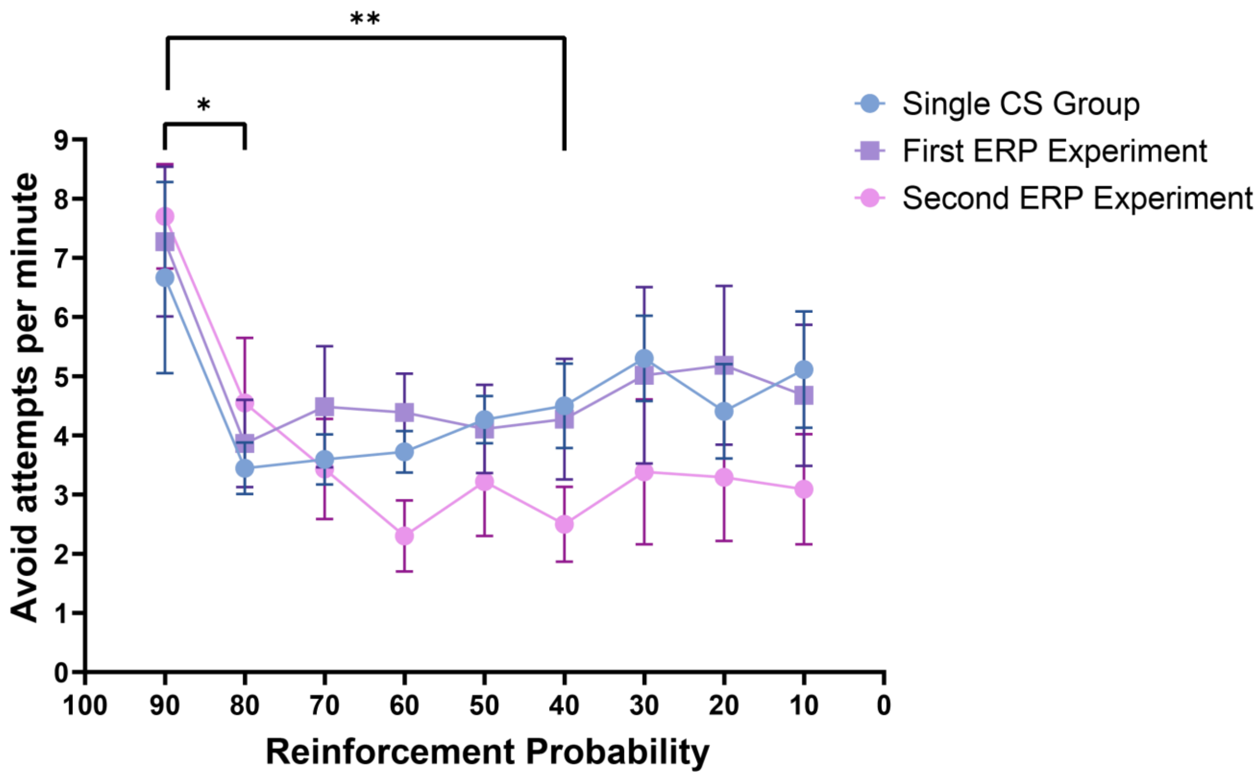
Supplementary Table 1 – Proportion of animals that met acquisition criterion for each task variation.

• Full chamber • Double CS condition	Reward Group	8/8
	No Reward Group	4/8
• Half chamber • No Reward condition	Double CS Group	9/12
	Single CS group	10/12
	Single CS condition	61/72



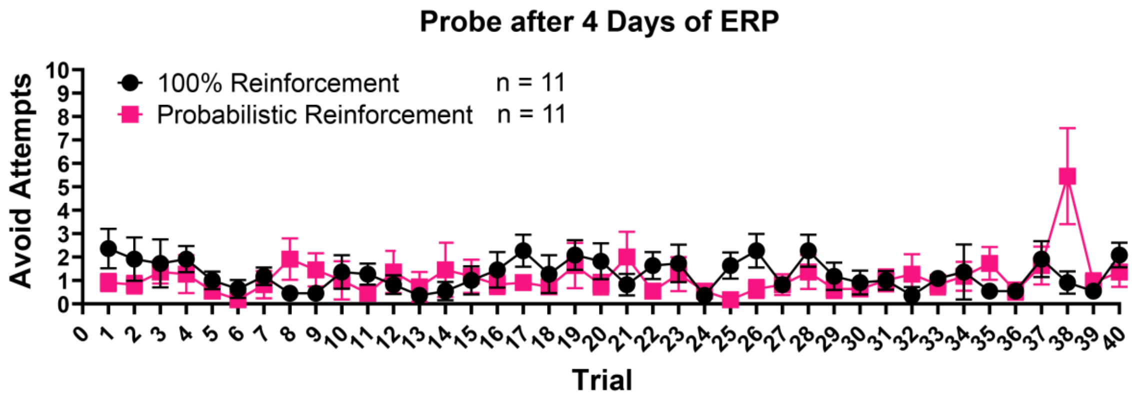
Supplementary Figure 1 – Avoidance is resistant to decreasing reinforcement probability. (A) The reward and no reward groups were tested from 100%-5% reinforcement probability increments. Both groups continued avoiding on most trials with no significant differences between them (two-way repeated measures ANOVA, Day x Group interaction p = 0.99) and no main effect of reinforcement probability (main effect p = 0.6049). (B) Both groups showed no significant differences in percent of avoidable shocks received across the tested reinforcement probabilities (two-way repeated measures ANOVA, Day x Group interaction p = 0.99) and no main effect of reinforcement probability (main effect p = 0.3454). Reward n = 8, No Reward n = 4.

A

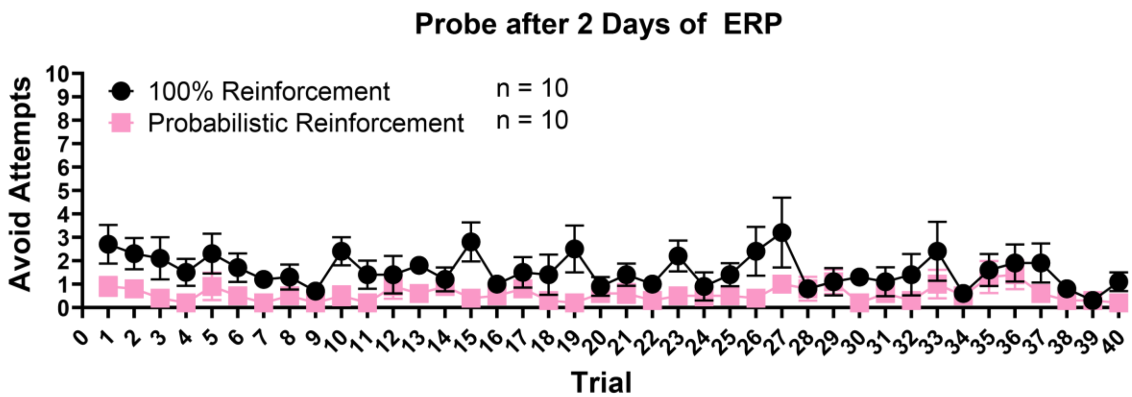


Supplementary Figure 2 – Three separate cohorts given negative reinforcement testing show an initial decrease in rate of avoid attempts. (A) Analysis of the three groups showed no significant differences between them (two-way repeated measures ANOVA, Day x Group interaction $p = 0.62$). Post-hoc test showed a significant reduction in rate of avoid at 80% (Tukey's test, $p < 0.05$) reinforcement probability and 70%, 60%, 50%, and 40% (Tukey's test, $p < 0.01$) compared to 90%. Single CS Group n = 9, First ERP Experiment n = 11, Second ERP Experiment n = 10.

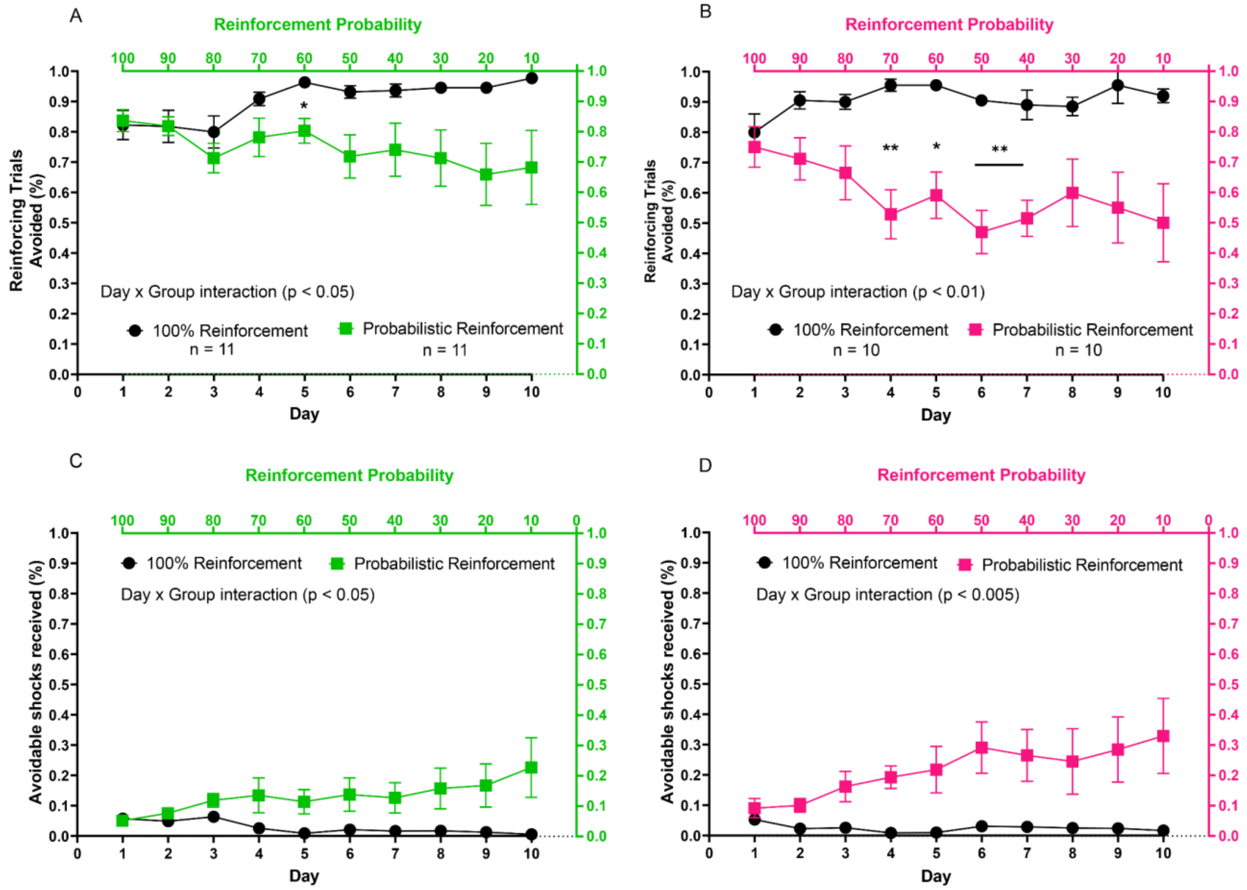
A



B



Supplementary Figure 3 – Avoid attempts are consistent across trials in a probe test after ERP. (A) After 4 days of ERP, the barrier was removed, and animals were given 40 trials of CS and lever presentations. There was no main effect of trial, suggesting that animals learned to inhibit avoid attempts during ERP as opposed to the early trials of the probe test (two-way repeated measures ANOVA, main effect p = 0.2109). (B) After 2 days of ERP in a separate cohort, there was again no main effect of trial (two-way ANOVA , trial main effect p = 0.2752).



Supplementary Figure 4 – Groups maintained at 100% reinforcement perform better than probabilistically tested groups. (A) 100% reinforcement and probabilistic reinforcement groups from the 4 day ERP experiment. The 100% group avoids on a higher percentage of its reinforcing trials each session (two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$). (B) Separate groups used for the 2 day ERP experiment replicate the significant difference in reinforcing trials avoided (Day x Group interaction $p < 0.01$). (C) 100% reinforcement and probabilistic reinforcement groups from the 4 day ERP experiment. The 100% group received significantly less avoidable shocks (two-way repeated measures ANOVA, Day x Group interaction $p < 0.05$). (D) Groups from the 2 day ERP experiment replicate the significant difference in avoidable shocks received (Day x Group interaction, $p < 0.005$).

Bibliography

Baum, M. (1970). Extinction of avoidance responding through response prevention (flooding). Psychological Bulletin, 74(4), 276-284. doi:10.1037/h0029789

Bouton, M. E., Trask, S., & Carranza-Jasso, R. (2016). Learning to inhibit the response during instrumental (operant) extinction. Journal of Experimental Psychology: Animal Learning and Cognition, 42(3), 246-258. doi:10.1037/xan0000102

Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F., & Quirk, G. J. (2014). Neural Structures Mediating Expression and Extinction of Platform-Mediated Avoidance. The Journal of Neuroscience, 34(29), 9736-9742. doi:10.1523/jneurosci.0191-14.2014

Bravo-Rivera, C., Roman-Ortiz, C., Montesinos-Cartagena, M., & Quirk, G. J. (2015). Persistent active avoidance correlates with activity in prelimbic cortex and ventral striatum. Frontiers in Behavioral Neuroscience, 9(184). doi:10.3389/fnbeh.2015.00184

Cain, C. K. (2019). Avoidance problems reconsidered. Current Opinion in Behavioral Sciences, 26, 9-17. doi:https://doi.org/10.1016/j.cobeha.2018.09.002

Diehl, M. M., Bravo-Rivera, C., & Quirk, G. J. (2019). The study of active avoidance: A platform for discussion. Neuroscience & Biobehavioral Reviews, 107, 229-237. doi:10.1016/j.neubiorev.2019.09.010

Fernando, A. B. P., Mar, A. C., Urcelay, G. P., Dickinson, A., & Robbins, T. W. (2015). Avoidance Behavior: A Free-Operant Lever-Press Avoidance Task for the Assessment of the Effects of Safety Signals. Current Protocols in Neuroscience, 70(1), 8.32.31-38.32.12. doi:https://doi.org/10.1002/0471142301.ns0832s70

Ferrando, C., & Selai, C. (2021). A systematic review and meta-analysis on the effectiveness of exposure and response prevention therapy in the treatment of Obsessive-Compulsive Disorder. Journal of Obsessive-Compulsive and Related Disorders, 31, 100684. doi:https://doi.org/10.1016/j.jocrd.2021.100684

Geramita, M. A., Yttri, E. A., & Ahmari, S. E. (2020). The two-step task, avoidance, and OCD. Journal of Neuroscience Research, 98(6), 1007-1019. doi:https://doi.org/10.1002/jnr.24594

Igarashi, H., Ikeda, K., Onimaru, H., Kaneko, R., Koizumi, K., Beppu, K., . . . Yawo, H. (2018). Targeted expression of step-function opsins in transgenic rats for optogenetic studies. Scientific Reports, 8(1), 5435. doi:10.1038/s41598-018-23810-8

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity

Survey Replication. Archives of General Psychiatry, 62(6), 617-627. doi:10.1001/archpsyc.62.6.617

Lázaro-Muñoz, G., LeDoux, J. E., & Cain, C. K. (2010). Sidman Instrumental Avoidance Initially Depends on Lateral and Basal Amygdala and Is Constrained by Central Amygdala-Mediated Pavlovian Processes. Biological Psychiatry, 67(12), 1120-1127. doi:10.1016/j.biopsych.2009.12.002

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: a reconceptualization of a troubled paradigm. Molecular Psychiatry, 22(1), 24-36. doi:10.1038/mp.2016.166

Lee, C., Lavoie, A., Liu, J., Chen, S. X., & Liu, B.-h. (2020). Light Up the Brain: The Application of Optogenetics in Cell-Type Specific Dissection of Mouse Brain Circuits. *Frontiers in Neural Circuits, 14*(18). doi:10.3389/fncir.2020.00018

Manning, E. E., Bradfield, L. A., & Iordanova, M. D. (2021). Adaptive behaviour under conflict: Deconstructing extinction, reversal, and active avoidance learning. Neuroscience & Biobehavioral Reviews, 120, 526-536. doi:https://doi.org/10.1016/j.neubiorev.2020.09.030

Martínez-Rivera, F. J., Bravo-Rivera, C., Velázquez-Díaz, C. D., Montesinos-Cartagena, M., & Quirk, G. J. (2019). Prefrontal circuits signaling active avoidance retrieval and extinction. Psychopharmacology, 236(1), 399-406. doi:10.1007/s00213-018-5012-7

Martínez-Rivera, F. J., Sánchez-Navarro, M. J., Huertas-Pérez, C. I., Greenberg, B. D., Rasmussen, S. A., & Quirk, G. J. (2020). Prolonged avoidance training exacerbates OCD-like behaviors in a rodent model. Translational Psychiatry, 10(1), 212. doi:10.1038/s41398-020-00892-5

Rosas-Vidal, L. E., Lozada-Miranda, V., Cantres-Rosario, Y., Vega-Medina, A., Melendez, L., & Quirk, G. J. (2018). Alteration of BDNF in the medial prefrontal cortex and the ventral hippocampus impairs extinction of avoidance. Neuropsychopharmacology, 43(13), 2636-2644. doi:10.1038/s41386-018-0176-8