

2019

# Contextual Control Of Instrumental Actions And Habits Following Retroactive Interference

Michael Steinfeld  
*University of Vermont*

Follow this and additional works at: <https://scholarworks.uvm.edu/graddis>



Part of the [Psychology Commons](#)

---

## Recommended Citation

Steinfeld, Michael, "Contextual Control Of Instrumental Actions And Habits Following Retroactive Interference" (2019). *Graduate College Dissertations and Theses*. 1041.  
<https://scholarworks.uvm.edu/graddis/1041>

This Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact [donna.omalley@uvm.edu](mailto:donna.omalley@uvm.edu).

CONTEXTUAL CONTROL OF INSTRUMENTAL ACTIONS AND HABITS  
FOLLOWING RETROACTIVE INTERFERENCE

A Thesis Presented

by

Michael Steinfeld

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements  
For the Degree of Master of Arts  
Specializing in Psychology

May, 2019

Defense Date: December 4, 2018  
Thesis Examination Committee:

Mark E. Bouton, Ph.D., Advisor  
Scott Mackey, Ph.D., Chairperson  
John T. Green, Ph.D.  
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

## **Abstract**

It is commonly accepted that instrumental responses that have been extinguished can return. For example, in a phenomenon known as the renewal effect, extinguished behaviors return upon removal from the extinction context. Another well-accepted notion is that instrumental behaviors can be thought of as goal-directed actions, which form over the course of moderate amounts of practice or training, and habits, which form after extended practice. Despite years of research on both topics, what happens to actions and habits following extinction is poorly understood. The present experiments examined the renewal of actions and habits following retroactive interference paradigms such as extinction and additional training. Experiment 1 examined renewal of an action following its extinction in a separate context, and demonstrated that the extinguished behavior renewed as an action upon return to the acquisition context. Experiment 2 asked the same question about habits, and found that the behavior renewed as a habit after extinction upon return to the acquisition context. Experiment 3 examined renewal of goal-directed responding in one context following extensive training and conversion into habit in another context. It demonstrated that a single response could manifest as a habit in one context, and renew as an action in the original training context. Experiment 4 asked if this effect depends on returning to the acquisition context, or simply removal from the habit training context. The results suggest that mere removal from the habit training context is sufficient to renew the goal-directed properties of a behavior. Together, the results suggest that actions and habits can be inhibited in a context-specific manner by extinction, and that instrumental behaviors can have both action and habit properties that can each renew under the proper circumstances. The results also expand on the notion that habits are especially context specific, while actions can transfer across contexts.

## **Acknowledgements**

I would like to acknowledge my advisor, Mark Bouton, for his help and encouragement throughout this process. I would also like to thank Eric Thrailkill for his tutelage and assistance in designing and conducting these, and other experiments. I would also like to thank my committee, the Biobehavioral Cluster, and the Department of Psychological Science as a whole. I would also like to acknowledge my funding from the R01 DA 033123 grant from the National Institute of Drug Abuse to Mark Bouton.

## Table of Contents

	Page
Introduction.....	1
Experiment 1: Renewal of Actions.....	13
Method.....	1
Subjects.....	14
Apparatus.....	15
Magazine Training.....	16
Instrumental Training.....	16
Reinforcer Devaluation.....	17
Extinction.....	18
Test.....	18
Consumption Test.....	18
Reacquisition Test.....	19
Results.....	19
Acquisition.....	19
Devaluation.....	19
Extinction.....	20
Test.....	20
Consumption Test.....	21
Reacquisition Test.....	22
Discussion.....	22
Experiment 2: Renewal of Habits.....	23
Method.....	24
Subjects and Apparatus.....	24
Procedure.....	24
Results.....	24
Acquisition.....	24
Devaluation.....	24
Extinction.....	25
Test.....	26
Consumption Test.....	26
Reacquisition Test.....	27
Discussion.....	27
Experiment 3: Renewal of Action after Extensive Training.....	28
Method.....	29
Subjects and Apparatus.....	29
Procedure.....	30
Results.....	30
Acquisition and Devaluation.....	30
Test.....	31

Consumption Test.....	32
Reacquisition Test.....	32
Discussion.....	33
Experiment 4: Return to the Action Acquisition Context Compared to a Neutral Context.....	33
Method.....	35
Subjects .....	35
Apparatus.....	35
Experiment 4a.....	35
Experiment 4b.....	36
Procedure.....	36
Acquisition.....	37
Devaluation.....	37
Test.....	37
Consumption and Reacquisition Tests.....	38
Results.....	38
Acquisition.....	38
Devaluation.....	38
Test.....	39
Consumption Test.....	40
Reacquisition Test.....	41
Discussion.....	42
General Discussion.....	42
Figures.....	51
References.....	63

## List of Tables

	Page
Table 1: Experiment 1 Design.....	14
Table 2: Experiment 2 Design .....	24
Table 3: Experiment 3 Design .....	29
Table 4: Experiment 4 Design.....	35

## List of Figures

	Page
Figure 1a: Experiment 1 acquisition data.....	56
Figure 1b: Experiment 1 extinction data.....	56
Figure 2: Experiment 1 devaluation data.....	56
Figure 3: Experiment 1 test data.....	57
Figure 4: Experiment 1 consumption data.....	57
Figure 5: Experiment 1 reacquisition data.....	58
Figure 6a: Experiment 2 acquisition data.....	59
Figure 6b: Experiment 2 extinction data.....	59
Figure 7: Experiment 2 devaluation data.....	59
Figure 8: Experiment 2 test data.....	60
Figure 9: Experiment 2 consumption data.....	60
Figure 10: Experiment 2 reacquisition data.....	61
Figure 11: Experiment 3 acquisition data.....	62
Figure 12: Experiment 3 devaluation data.....	62
Figure 13: Experiment 3 test data.....	63
Figure 14: Experiment 3 consumption data.....	63
Figure 15: Experiment 3 reacquisition data.....	64
Figure 16: Experiment 4 acquisition data.....	65
Figure 17: Experiment 4 devaluation data.....	65
Figure 18: Experiment 4 test data.....	66
Figure 19: Experiment 4 consumption data.....	66
Figure 20: Experiment 4 reacquisition data.....	67



## **Introduction**

Instrumental learning is the process through which animals learn about behaviors and the consequences associated with them. It allows animals to perform behaviors that result in positive outcomes while avoiding behaviors that lead to negative outcomes. The study of instrumental learning can be traced back at least to the late 1800s (Thorndike, 1898, 1911), and has provided many important insights into the causes and consequences of various behaviors. In humans, it is thought to be the basis of voluntary behavior, as well as a driving force behind many public health issues, including obesity (Epstein & Saelens, 2000) and addiction (Miles, Everitt, Dalley, & Dickinson, 2004; Pelloux, Everitt, & Dickinson, 2007). Furthermore, abnormalities in instrumental behaviors have been identified as a feature of many psychiatric disorders (Atkins & Stoff, 1993; Gillan, Pappmeyer, Morein-Zamir, Sahakian, Fineberg, Robbins, & de Wit, 2011). Developing a deeper understanding of instrumental behavior will help us treat these disorders, as well as prevent relapse and promote healthier behavioral alternatives.

### **Actions and Habits**

An important distinction in our understanding of instrumental behavior is between goal-directed actions and habits (e.g., Dickinson, 1985). Actions are goal-directed in that they are sensitive to the value of their outcome or reinforcer. In contrast, habits are not sensitive to the value of their outcome; rather, they are performed automatically and under stimulus control, without the outcome in mind. Actions and habits are often thought of in terms of relationships between the organism's representations of the stimulus (S), the response (R), and the response's outcome (O). In an action, the response

is thought to generate a direct representation of the outcome. For this reason, it is sometimes referred to as being under the control of an R-O association. In a habit, a stimulus that has been associated with the response comes to evoke the response directly, without mediation by a representation of the reinforcer. For this reason, it is referred to as being under the control of an S-R association (Adams & Dickinson, 1981; Balleine & O'Doherty, 2010; Lingawi, Dezfoulli, & Balleine, 2016).

Researchers have developed the reinforcer-devaluation procedure in order to determine whether an instrumental response is a goal-directed action or a habit (Adams, 1981, 1982). During an acquisition phase, rats acquire an instrumental (operant) response to gain access to a reinforcer (e.g., a food pellet). Following the end of the acquisition phase, the rats enter the devaluation phase. In this phase, rats receive food pellets without the response manipulandum available, and then receive an injection of lithium chloride (LiCl) to make them nauseous. This causes them to associate the food pellets with the illness (taste aversion) and reject them (Garcia, Lasiter, Bermudez-Rattoni, Deems, 1985). Once the rats reject the pellets, they enter a test session where the response manipulandum is once again available. Unlike acquisition, the response is put on extinction during the test, meaning that performing the response no longer produces food pellets. Because the test is conducted in extinction, the absence of the reinforcer ensures that the rats cannot simply associate the response with the averted reinforcer through direct experience between the two. Instead, they must use what they learned about the outcome in acquisition and devaluation to influence their rate of responding.

Using this method, Adams (1982) observed that a moderately trained response (100 response-reinforcer pairings) was suppressed after reinforcer devaluation (reinforcer devaluation effect). This result indicated that the response was motivated by the remembered current value of the outcome (action). However, in other rats that received more extensive training with the response (500 response-reinforcer pairings), devaluation had no effect on responding, suggesting that the behavior no longer depended on the remembered value of the outcome (habit). Adams (1982) and Adams and Dickinson (1985) suggested that extended training promotes behavioral autonomy that occurs independently of the value of the outcome. Such behavioral autonomy is the hallmark of a habit. Thus, the reinforcer devaluation method can in principle test whether a given instrumental response is a goal-directed action (sensitive to reinforcer devaluation) or a habit (insensitive to reinforcer devaluation).

A common view is that instrumental behaviors start out as actions, but become autonomous habits with extended practice or training (Adams, 1982; Killcross & Coutureau, 2003). Dickinson (1985) hypothesized that the change in the correlation between behavior rate and reinforcement rate that occurs with extended training may play an important role in the conversion of action into habit. Early in training, behavior rate and reinforcement rate both increase rapidly and are highly correlated with each other. However, with extended training, the reinforcement rate begins to plateau. This means that animals experience the correlation between behavior rate and reinforcement rate over a large range of values during the first few sessions, but over a more restricted range later in training. Dickinson argued that when an animal's rate of behavior varies during the

early stages of training, the animal experiences and stores knowledge about the relationship between the behavior rate and reinforcement rate. Consequently, reinforcer devaluation depresses performance through this knowledge. However, when a response is more extensively trained and the response rate plateaus, the animal no longer experiences the behavior rate-reinforcer rate correlation because there is so little variation in their responding. Therefore, performance of the response is no longer controlled by the value of the reinforcer, resulting in reinforcer devaluation having little effect.

A prediction that comes from this line of reasoning is that reinforcement schedules that create weak behavior rate-reinforcement rate correlations will produce habits more readily than reinforcement schedules that create stronger behavior-reinforcement correlations. Schedules of reinforcement are generally divided into two categories, interval schedules and ratio schedules. Ratio schedules reinforce responses after a certain number of responses have been emitted, while interval schedules only reinforce responses that are made after a certain amount of time. Because animals are reinforced after a certain number of responses on a ratio schedule, ratio schedules create a high correlation between the behavior rate and the reinforcement rate. On the other hand, interval schedules, which only reinforce responses after a certain period of time, have a low correlation between behavior rate and reinforcement rate. According to Dickinson's hypothesis, interval schedules should produce habits more readily than ratio schedules. This prediction was supported by the findings of Dickinson, Nicholas, and Adams (1983), who found that rats demonstrated habitual responding after moderate training on

an interval schedule, but remained goal-directed after equivalent training on a ratio schedule.

Analyses of actions and habits have suggested that behaviors can have both action and habit properties at the same time (Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995). Dickinson has argued that instrumental training engages the two learning processes (S-R and R-O) concurrently. According to Dickinson's dual-process theory, S-R and R-O processes both start developing at the beginning of training. The strength of an instrumental response is defined as the sum of strengths of the R-O and S-R processes. The amount that each process contributes to the overall strength of the response is usually not equal. Early in training, the R-O system contributes more to response strength, but with extended training the R-O process weakens and the S-R process strengthens and contributes the bulk of the strength. This line of theorizing has influenced dual-system theories of the neural basis of instrumental behavior (Balleine and Dickinson, 1998a; Corbit and Balleine, 2003; Killcross and Coutureau, 2003; Yin, Knowlton, & Balleine, 2004; Yin, Ostlund, & Balleine, 2005).

More recent experiments reported by Thrailkill and Bouton (2015) also provide evidence for the coexistence of goal-directed and stimulus controlled properties in a single instrumental response. In their experiments, minimally trained rats reduced their responding following reinforcer devaluation; however, responding was never completely abolished (also observed by Colwill & Rescorla, 1985b; Trask & Bouton, 2014). Thrailkill and Bouton suggested that once the goal-directed component maintaining a behavior was removed by the devaluation procedure, a habit-based component was

revealed, suggesting that habitual properties are present in minimally trained rats and that these properties influence behavior.

Experiment 4 of Thraillkill and Bouton (2015) produced particularly relevant results. This experiment was designed to test how well actions and habits transfer to new contexts. While Pavlovian responses are known to transfer well across contexts (e.g., Bouton & King, 1983; see also Rosas, Todd, & Bouton, 2013 for a review), animals often show a reduction in instrumental responding following a context switch (context-switch effect) (e.g., Bouton, Todd, Vurbic, & Winterbauer, 2011). (Context was defined as different sets of operant chambers that have different sets of cues such as scents, floor grids, and wall patterns). It was unclear if the R-O and S-R components of an instrumental response are affected the same way by a context switch, so Thraillkill and Bouton trained rats using a factorial design to test this. Rats were either moderately or extensively trained to perform a free-operant lever-pressing response. In a devaluation phase, half of the rats received the food outcome paired with a LiCl injection while the other half received food and LiCl unpaired. Rats were then tested for lever pressing (in extinction) in both their training context and a neutral context (rats received context exposures, but no instrumental training in this context). Consistent with existing results, there was a devaluation effect only in the moderately trained rats, as well as a context-switch effect (reduced responding) in the neutral context in all rats. Interestingly, the devaluation effect transferred perfectly across contexts: That is, the difference between the Paired and the Unpaired rats remained consistent across the training and neutral contexts. As a whole, the authors took their results to mean that there is a habitual portion

of the response that develops early in training, which is why reinforcer devaluation does not fully suppress responding. Secondly, the context switch primarily affects the habit component. The logic behind this idea was that the size devaluation effect (which can be interpreted roughly as the degree to which a behavior is goal-directed) remained consistent across contexts. Thus, the goal-directed component of the behavior transferred perfectly across contexts. However, the behavior left over after the devaluation effect, which must be habitual, was weakened by the context switch. Hence, the habit component of the behavior must have been more affected by the context switch than the action component. Likewise, the pure habit that resulted from more extensive practice was also affected by the context switch.

### **Extinction and the Contextual Control of Instrumental Behavior**

Extinction occurs when a response that was previously reinforced is no longer reinforced. During extinction, an animal stops performing a response not because the behavior has been erased or unlearned, but because the current knowledge that the response no longer leads to the reinforcer inhibits or interferes with the expression of the response (Bouton, 1988; Bouton, 2002; Todd, Vurbic, & Bouton, 2014). One of the best examples of this comes from the literature on the “renewal effect” (e.g., Bouton & Bolles, 1979). Following extinction (nonreinforcement) of a trained response, the response often returns (renews) when it is removed from the context in which extinction took place. This phenomenon has been studied for many years in Pavlovian conditioning under several different conditions, the most well-studied and robust of which is ABA renewal.

In ABA renewal, a conditioned stimulus (CS), such as a tone, is paired with an unconditioned stimulus (US), such as shock, in Context A. This causes the animal to begin performing a conditioned response (CR), such as freezing, to the tone. The animal is then placed in Context B, where the tone CS now occurs without the shock US, causing the animal to stop performing a CR to the tone (extinction). Finally, the animal is returned to Context A, where conditioned responding to the CS renews (Bouton & Bolles, 1979; Bouton & King, 1983; Bouton & Peck, 1989). In other words, the animal begins freezing in the presence of the tone again, despite previous extinction training that occurred in the separate context. Two other forms of renewal demonstrate that this effect does not depend merely on return to the acquisition context. In AAB renewal, a response is both conditioned and extinguished in Context A. The response then renews when it is tested in a novel context (Context B) (Bouton & Ricker, 1994; Laborda, Witnauer, & Miller 2011; Tamai & Nakajima, 2000). In ABC renewal, a response is trained in Context A and extinguished in Context B, but then renews when tested in a novel context (Context C) (Bouton & Bolles, 1979; Harris, Jones, Bailey, & Westbrook, 2000; Thomas et al., 2003). More recently, all three forms of renewal have been observed in instrumental (operant) learning paradigms (Bouton et al., 2011, Todd, Winterbauer, & Bouton 2012a; Todd, Winterbauer, & Bouton 2012b).

Theoretically, research on renewal implies that extinction is context specific, and removal from the extinction context is both necessary and sufficient for renewal to occur. ABA renewal can be explained as the test context being excitatory due to its associative history with reinforcement, but AAB and ABC renewal cannot. In particular, AAB



renewal suggests that simply testing the response outside of the extinction context is sufficient for renewal to occur because the excitatory learning acquired in Context A is inhibited by the extinction learning in Context A, which is not the case in ABC renewal. The prevailing explanation for this observation is that extinction results in new learning that is both inhibitory and specific to the context in which it takes place. According to one version of this hypothesis that is specific to instrumental learning, the organism specifically learns to inhibit the extinguished response in the extinction context. When the context is switched, there is a release from the context-specific response inhibition, allowing the behavior to renew (Bouton & Todd, 2014; Todd, 2013).

Evidence for this claim comes from a series of experiments reported by Todd (2013) and Todd et al., (2014), which also experimentally ruled out alternative explanations of renewal. Experiment 1 of Todd (2013) tested the possibility that the context enters into a direct association with the representation of the reinforcer. Rats were trained to perform two responses in two different contexts respectively for the same food pellets, and were then given extinction training of each response in the opposite context. This procedure ensured that each context was equally associated with reinforcement and nonreinforcement. When the responses were tested in each context, each response renewed specifically in its acquisition context (ABA renewal). This result was replicated in ABC (Experiment 2) and AAB (Experiment 3) renewal designs. Because the procedures controlled for exposure to reinforcement and nonreinforcement in the two contexts, the results could not be explained by differences in excitatory or inhibitory associations with the reinforcer in them. In Experiment 4 of Todd (2013), Todd tested the

possibility that the extinction context can function as a negative occasion setter and demonstrated that extinction training in one context did not weaken the renewal of a different response that had been trained there and extinguished in a separate context. If the context had functioned as a negative occasion-setter, extinction of one response in a context should have transferred and inhibited the renewal of the other responses in that context (Holland & Coldwell, 1993; Morell & Holland, 1993). However, no decrement in renewal was observed.

Todd et al. (2014) provided direct evidence of an inhibitory context-response association in extinction using a discriminated operant procedure where rats were only reinforced for responses in the presence of a stimulus. In Experiment 3, a single group of rats learned four stimulus-response combinations (S1R1, S2R2, S3R1, and S4R2) in three different contexts. S1R1 and S2R2 were trained in Context A, S3R1 was trained in Context B, and S4R2 was trained in Context C. Each stimulus-response combination was then extinguished in its acquisition context. Finally, S1R1 and S2R2 were tested in each context. The authors hypothesized that the extinction training of S3R1 in Context B and S4R2 in Context C would suppress AAB (or AAC) renewal of S1R1 specifically in Context B and S2R2 specifically in Context C, which is what happened. In Context B, S1R1 renewal was suppressed compared to S2R2 renewal, while in Context C, S2R2 renewal was suppressed relative to S1R1 renewal. In summary, extinction training of a specific response in a specific context weakened the ability of that response to renew in that context, even when it was occasioned by a different stimulus. This finding is consistent with the hypothesis that an inhibitory association is formed between the

context and the response during extinction (see also Bouton, Trask, & Carranza-Jasso, 2016; Todd, 2013).

### **Contextual Control of Actions and Habits**

Recent work has sought to further characterize the role of the context in instrumental behavior. Bouton, Todd, and León (2014) suggested two possibilities: 1) the context directly evokes (or inhibits) the response (Context-R), and 2) the context might work more hierarchically by signaling (or setting the occasion for) the R-O relationship (Context-(R-O)). Recent findings provide support for both accounts. Thraillkill and Bouton (2015) demonstrated that switching rats from their training context to a neutral context produced a context-switch effect under various circumstances, including after both moderate training (which demonstrably produced an action) and after extensive training (which demonstrably produced a habit). Because habits, which are not influenced by changes in the value of the outcome, were still influenced by a context-switch, it is reasonable to suggest that the context plays the role of the antecedent stimulus for habit and controls the response directly, independent of the value of the outcome, supporting a role for Context-R associations (Bouton & Todd, 2014).

Evidence for the hierarchical role of context in controlling the response has been reported by Trask and Bouton (2014) using a design analogous to one developed by Colwill and Rescorla (1990). In Context A, R1 led to O1 and R2 led to O2 (responses were lever press and chain pull, outcomes were grain pellets and sucrose pellets). In Context B, R1 led to O2 and R2 led to O1. In separate aversion conditioning trials, O2 was paired with LiCl in both contexts. R1 and R2 were then tested in both contexts.

During the test, the rats performed R1 more than R2 in Context A, but R2 more than R1 in Context B. Put more simply, devaluing O2 suppressed R1 in the Context where R1 was associated with O2, and suppressed R2 in the Context where R2 was associated with O2. This finding is consistent with what was observed by Colwill and Rescorla (1990) with simple discriminative stimuli (noise and light cues) and suggests that the rats had learned about specific combinations of context, response, and outcome. In other words, the rats had learned distinct context-(R-O) associations. The results of this experiment, as well as Thraill and Bouton (2015), suggest that the context forms both direct associations with the response, and can help to distinguish between different R-O associations in a hierarchical manner.

### **Present Thesis**

Despite years of interest in the extinction and renewal of instrumental behaviors, as well as actions and habits, it is not known what happens specifically to actions and habits following retroactive interference treatments such as extinction. Retroactive interference is observed when new learning interferes with learning that took place earlier (Bouton, 1993). Extinction is an excellent example of a retroactive interference treatment. During extinction training, the knowledge that the response leads to the outcome is inhibited by the learning that takes place during extinction. However, removal from the extinction context removes the interference, allowing the original learning to control behavior, as is commonly observed in renewal. While it has been shown in a number of settings that both moderately trained and extensively trained instrumental behaviors renew after extinction (Todd et al., 2012b), it is unclear whether the action (R-O) and

habit (S-R) properties that controlled the expression of behavior at the end of acquisition still control the behavior when it recovers or recurs after extinction. Put more simply, after extinction, do actions renew as actions and do habits renew as habits? I also examined what happens to the R-O (action) component of a behavior after it is converted to habit following extensive training. Like extinction, habit formation might be seen as a retroactive interference treatment, with the extended habit training interfering with action training and preventing goal-directed behavior. It is not clear if a behavior can return to R-O control following extended training, or whether the S-R learning in habit may catastrophically interfere with R-O control. I also asked if such renewal is dependent upon a return to the acquisition context (ABA renewal) or if it can also occur in a neutral context (ABC renewal).

The present experiments were designed to assess the renewal of actions and habits following retroactive interference treatments. Experiment 1 evaluated renewal of an action using an ABA renewal design. Experiment 2 asked the same question of habits. Experiment 3 then asked if actions can renew following training sufficient to produce a habit in an ABA renewal paradigm. Experiment 4 expanded on Experiment 3 by asking if the renewal of an action after habit training depended on return to the action acquisition context (ABA renewal) or if it could also be accomplished simply by removal from the habit context (ABC renewal). The results of these experiments increase our current knowledge of actions and habits as well as how contextual cues play a role in controlling goal-directed and habitual behavior.

### **Experiment 1: Renewal of Action after Extinction**

The purpose of Experiment 1 was to determine if actions renew as actions in an ABA renewal paradigm. To do this, as illustrated in Table 1, rats were first trained rats to lever press in Context A to the point where the response should still be an action, and then the reinforcer devaluation procedure was performed in both Context A and Context B. The response was then extinguished in Context B. Finally, the response was tested in extinction in both Context A and Context B in a counterbalanced order. Renewal of the extinguished response was expected to occur in Context A, but the critical question was what would renewal look like in the two groups? Both the Paired and Unpaired rats should increase (renew) responding in Context A by virtue of removal from the extinction context. However, it was unclear if any devaluation effect that was present at the end of devaluation would be present after extinction. In other words, would the rats demonstrate a reinforcer devaluation effect in Context A, which would specifically suggest the ABA renewal of a goal-directed action?

Table 1

*Experiment 1 Design*

Group	Acquisition (A)	Devaluation (A and B)	Extinction (B)	Test (A and B)
Paired	3 Sessions: R+	Pellets → LiCl	3 Sessions: R-	R-
Unpaired	3 Sessions: R+	Pellets / LiCl	3 Sessions: R-	R-

*Note.* A and B denote separate contexts; R = response; + = Reinforced - = Not reinforced. Exposures to the contexts were equated throughout the experiment (not shown).

**Method**

*Subjects*

The subjects were 32 naïve female rats purchased from Charles River Laboratories (St. Constance, Quebec). They were between 75 and 90 days old at the start

of the experiment and were individually housed in a room maintained on a 16:8-h light:dark cycle. Experimentation took place during the light period of the cycle. The rats were food-deprived to 80% of their initial body weights throughout the experiment.

### *Apparatus*

The apparatus was two unique sets of four conditioning chambers (Model ENV-007-VP; Med Associates, St. Albans, VT). Each set of chambers was housed in a separate room in the laboratory, allowing us to run eight animals at a time. Each chamber was housed in its own sound-attenuating chamber. All chambers measured  $31.75 \times 24.13 \times 29.21$  cm (Length  $\times$  Width  $\times$  Height). The side walls consisted of clear acrylic panels, and the front and rear walls were made of brushed aluminum. A recessed food cup was centered on the front wall approximately 2.5 cm above the floor. A retractable lever (Model ENV-112CM, Med Associates) was positioned to the left of the food cup. The lever was 4.8 cm wide and 6.3 cm above the grid floor. It protruded 2.0 cm from the front wall when extended. One light was immediately above the lever, while the other was 3.0 cm to the right of the first light. The chambers were illuminated by 7.5-W incandescent bulbs mounted to the ceiling of the sound attenuation chamber. Ventilation fans provided background noise of 65 dBA. The reinforcer used 45-mg grain food pellet (MLab Rodent Tablets; TestDiet, Richmond, IN).

The two contexts differed in several ways. Context 1 had a staggered grid floor that alternated between thick and thin bars and was scented using Hannaford's distilled white vinegar (Scarborough, Me). Context 2 also had a grid floor, but all of the bars were the same size, and it was scented using Vick's Vaporub (Cincinnati, Oh). It also had a

1.91 cm wide vertical green stripe on the glass pane. The grids of the floor were mounted on the same plane within their respective box and were spaced 1.6 cm apart (center-to-center).

### *Procedure*

*Magazine Training.* On the first day of training, the rats learned to eat pellets from the food cup in each of their assigned contexts (order counterbalanced). This required two magazine training sessions (one in each context) that consisted of 30 noncontingent pellet presentations delivered on a random time (RT) 30-s schedule of reinforcement. These sessions lasted approximately 15 to 20 minutes. Rats were returned to their home cages following completion of each session.

*Instrumental Acquisition.* Rats were then given one daily session in each context (two sessions a day) for three days. In Context A, the lever was inserted into the chamber after a two-min delay. After this, the rats were reinforced for lever-pressing according to a random interval (RI) 30-s schedule. This lasted until the rat had earned 30 reinforcers, at which the point the lever was retracted. Once all of the rats reached 30 reinforcers, they were returned to their home cages. Sessions in Context B were the same duration as those in A, but the lever was not present. This method controlled for exposure to both contexts while ensuring that the rats were not exposed to any response-reinforcer contingency in Context B. Order of the context exposures was counterbalanced across group and day. Furthermore, the sessions were double alternated. For example, on Day 1 half of the rats were exposed to Context A and then Context B, while the other half were exposed to Context B and then Context A. The orders were then reversed on Day 2. The rats were



then given context exposures in the opposite pattern for Days 3 and 4, after which, the pattern would repeat. This method of counterbalancing was used for all stages (except reinforcer devaluation) in all experiments.

*Reinforcer Devaluation.* Reinforcer devaluation began on the day following the third instrumental training session. During this phase of the experiment, rats were given only one session per day in either Context A or Context B, with the lever unavailable throughout the phase. During this stage, the rats were divided into a Paired group, which were given LiCl injections and pellets during the same sessions, and an Unpaired group, which received LiCl injections and pellets during different sessions. The sessions alternated between injection sessions and noninjection sessions. On injection sessions, the Paired rats were given noncontingent pellets on a RT 30-s schedule, while the Unpaired group was given context exposures that were matched for time. At the end of the session, each rat was given a 20ml/kg LiCl (0.15 M) injection. On noninjection sessions, the Unpaired rats were given pellets noncontingently on a RT 30-s schedule, while the Paired rats were given context exposures matched for time. No injections were given after these sessions. At the beginning of the devaluation phase, 50 pellets were delivered to each rat during the appropriate session. Following each injection session, the average number of pellets eaten by the Paired rats was calculated, and this became the new number of pellets delivered to the rats on the next session. This method ensured that all rats received the same number of injections, and roughly the same number of pellets, but only the Paired rats received the injections and pellets on the same day, resulting in only the Paired rats developing a taste aversion to the pellets.

Reinforcer devaluation lasted for three four-day cycles, for a total of twelve days, at which point the Paired rats ate less than an average of 1 pellet. The first four-session cycle followed an ABBA pattern in which there was one injection session and one noninjection session in each context. The full pattern employed during devaluation was A (injection) - B (no injection) - B (injection) - A (no injection). The second cycle was run using an BAAB pattern, while the third cycle returned to the ABBA pattern. There were always 48 hours between successive injections.

*Extinction.* Following the final day of the devaluation phase, the rats were given three daily instrumental extinction sessions in Context B. As was the case in acquisition, the lever was inserted after a two-min delay. However, lever pressing no longer produced any reinforcers. The sessions lasted for 15 minutes. The rats also received 15 min exposures to Context A without the lever to control for context exposure. The sequence of exposures to each context were counterbalanced across day and group as per acquisition.

*Test.* The rats were given a 10-min test session with the lever present in each context. The test was conducted in extinction to ensure that the rats could not directly associate their behavior with the newly-devalued reinforcer. Half of the rats from each group were tested in Context A first and Context B second, while the other half was tested in Context B first and Context A second.

*Consumption Test.* The rats were given a pellet consumption test on the day following the lever test to assess the aversion to the pellets in the Paired group. The rationale for this test was to confirm that the Paired rats still had an aversion to the

pellets. Ten pellets were delivered to each rat on an RT 30-s schedule without the lever available. The number of pellets consumed by each rat was counted and testing was conducted in each context. The order of sessions was counterbalanced as per the main test session, but in the opposite order for each rat.

*Reacquisition Test.* The rats were given a reacquisition test on the day following the pellet consumption test to assess the pellet's ability to support the response. The reacquisition test was conducted to confirm that the pellet no longer served as a reinforcer for the Paired rats. These sessions were similar to the acquisition sessions. The lever was inserted after a two-min delay, and the rats could press to earn pellets on a RI 30-s schedule. However, instead of ending after 30 reinforcers, this test session ended after 30 minutes. Test sessions were conducted in each context, with the order counterbalanced as per the main test session.

## Results

*Acquisition.* The acquisition data are depicted in Figure 1a. The average response rates of both groups increased across sessions, with no differences between the groups. This conclusion was supported by a Devaluation (Paired, Unpaired) by Session repeated measures ANOVA, which found a main effect of Session  $F(2, 56) = 98.18$ ,  $MSE = 1,228.26$ ,  $p < .001$ ,  $\eta^2 = .78$  on response rate, and no Devaluation by Session interaction,  $F < 1$ .

*Devaluation.* The mean proportion of delivered pellets that were consumed on each trial of the devaluation phase is shown in Figure 2. The Unpaired rats ate all of their pellets throughout the phase. The Paired rats significantly reduced the number of pellets

they ate across sessions. This was confirmed by a repeated measures ANOVA, which found a main effect of Session,  $F(5, 75) = 45.12$ ,  $MSE = 2.34$ ,  $p < .001$ ,  $\eta^2 = .75$  on proportion of pellets consumed.

*Extinction.* The results of the lever press extinction phase are displayed in Figure 1b. During extinction, responding decreased over both Sessions and 5-min within-session blocks. These conclusions were confirmed by a Devaluation (Paired, Unpaired) by 5-min Block by Session repeated measures ANOVA, which found main effects of Session,  $F(2, 120) = 53.02$ ,  $MSE = 222.32$ ,  $p < .001$ ,  $\eta^2 = .64$ , and Block,  $F(2, 120) = 76.38$ ,  $MSE = 438.25$ ,  $p < .001$ ,  $\eta^2 = .72$ . There was also a significant Session by Devaluation interaction,  $F(2, 120) = 10.14$ ,  $MSE = 44.53$ ,  $p < .001$ ,  $\eta^2 = .25$ . The effects of Devaluation and the Devaluation by Block interaction, were not statistically significant, largest  $F = 2.17$ .

The Session 1 extinction data are particularly important, as the parameters match those of previous action and habit tests from our lab (Thraillkill & Bouton, 2015). The presence of a devaluation effect on Session 1 of extinction confirm that the response was an action. A Devaluation by 5-min Block repeated measures ANOVA found a main effect of Devaluation during Session 1,  $F(1, 30) = 16.27$ ,  $MSE = 268.67$ ,  $p < .001$ ,  $\eta^2 = .35$ , indicating that there was a significant devaluation effect and that the behavior was goal-directed. There was also a main effect of Block,  $F(2, 60) = 58.84$ ,  $MSE = 360.98$ ,  $p < .001$ ,  $\eta^2 = .66$ , and a significant Block by Devaluation interaction,  $F(2, 60) = 3.44$ ,  $MSE = 21.13$ ,  $p = .038$ ,  $\eta^2 = .10$ .

*Test.* The test data are summarized in Figure 3. The rats responded more in the acquisition context (Context A) than in the extinction context (Context B), which indicated an ABA renewal effect. Their performance also indicated renewal of action. These conclusions were supported by a Context by Devaluation repeated measures ANOVA. The ANOVA found main effects of Context,  $F = 84.12$ ,  $MSE = 75.69$ ,  $p < .001$ ,  $\eta^2 = .74$ , and Devaluation,  $F = 7.15$ ,  $MSE = 75.69$ ,  $p = .012$ ,  $\eta^2 = .19$ , on response rate, as well as a Devaluation by Context interaction,  $F(1, 30) = 9.51$ ,  $MSE = 8.56$ ,  $p = .004$ ,  $\eta^2 = .24$ . Overall, there was an ABA renewal effect in both groups, but the interaction suggests that it was stronger in the Unpaired group than the Paired group. Within context independent samples t-tests were conducted to verify this. The test found reliable evidence of a devaluation effect in Context A,  $t(30) = -3.08$ ,  $p = .004$ ,  $d = 1.09$ , indicating that the behavior was an action in this context. However, there was no difference in response rates between the Paired and Unpaired rats in Context B,  $t(30) = 1.46$ ,  $p = .155$ ,  $d = .52$ . Together, these results suggest that the response renewed in both groups, and that it renewed as an action following return to the acquisition context.

*Consumption Test.* The consumption test results are shown in Figure 4. The results established that the Paired rats maintained their aversion to the pellets. This was confirmed with a Context by Devaluation repeated measures ANOVA, which found a main effect of Devaluation,  $F(1, 30) = 2,150.54$ ,  $MSE = 18.60$ ,  $p < .001$ ,  $\eta^2 = .986$ , and a marginal effect of Context,  $F(1, 30) = 4.12$ ,  $MSE = .03$ ,  $p = .051$ ,  $\eta^2 = .121$  on proportion of pellets consumed. A paired t-test between the Paired rats in Contexts A and Context B was conducted to further analyze the marginal effect of context. This revealed a

marginally significant effect of Context,  $t(15) = -2.03$ ,  $p = .06$ ,  $d = .70$ . This marginal effect was likely driven by two rats that ate four and five pellets in Context B, respectively. Given that these rats ate 0 pellets in Context A, it is possible that there was an equipment malfunction during their Context B session, and not all of the pellets were delivered. This would result in the rats appearing to have eaten more pellets than they actually did. For these reasons, the marginal effect of Context likely should not be interpreted as different levels of aversion toward the outcome in the two contexts.

*Reacquisition Test.* The reacquisition data for Contexts A and B are shown in panels A and B of Figure 5 respectively. Separate Devaluation by 1-min Block repeated measures ANOVAs confirmed that the reinforcer was valuable enough to support the response in both contexts only for the Unpaired rats. The Context A analysis revealed no effect of Block,  $F = 1.12$ , but a main effect of Devaluation,  $F(1, 30) = 137.92$ ,  $MSE = 137.92$ ,  $p < .001$ ,  $\eta^2 = .821$  on responding, as well as a Block by Devaluation interaction,  $F(29, 870) = 4.12$ ,  $MSE = 76.25$ ,  $p < .001$ ,  $\eta^2 = .12$ . In Context B, the analysis found significant effects of Devaluation,  $F(1, 30) = 105.14$ ,  $MSE = 105.14$ ,  $p < .001$ ,  $\eta^2 = .778$ , Block,  $F(29, 870) = 4.93$ ,  $MSE = 72.50$ ,  $p < .001$ ,  $\eta^2 = .14$  and a Devaluation by Block interaction,  $F(29, 870) = 10.47$ ,  $MSE = 10.47$ ,  $p < .001$ ,  $\eta^2 = .26$ .

Between subject analyses of the first minute of the reacquisition test were performed to determine if the behavior was goal directed at the beginning of the session. Between subjects t-tests determined that there were significant devaluation effects during the first minute of the reacquisition tests in Context A,  $t(30) = -4.69$ ,  $p < .001$ ,  $d = 1.66$ ,

and in Context B,  $t(30) = -2.85$ ,  $p = .008$ ,  $d = 1.01$ . The results of these tests suggest that the behavior remained goal-directed in both contexts.

### **Discussion**

Using parameters of lever-press training commonly employed by the Bouton laboratory (e.g., Thrailkill & Bouton 2015), training resulted in a lever press response that was an action, as confirmed by the devaluation effect observed at the beginning of extinction. Although the action was extinguished in Context B, it renewed as an action when returned to Context A, which was revealed by the reemergence of a devaluation effect during the Context A test. Using an ABA renewal design, the experiment found evidence that behaviors that are goal-directed renew as such following extinction in a separate context. This result extends the renewal effect to include specific goal-directed actions and response-outcome associations.

### **Experiment 2: Renewal of Habits**

Experiment 2 tested whether habit renews as a habit following extinction in a separate context. To do this, rats received 12 sessions of training (as opposed to three) in Context A, which has been shown to produce habitual responding in the laboratory (e.g., Thrailkill & Bouton, 2015). The rats then received reinforcer devaluation (Paired and Unpaired treatments) in both Contexts A and B, and then four sessions of extinction in Context B. As in Experiment 1, the rats were tested in extinction in both Context A and Context B. The question here was what would happen in the Paired rats during the Context A test. Both groups should undergo renewal of responding due to the removal from the extinction context, but if the renewed behavior is still a habit, there should be no

devaluation effect present during the test (similarity in responding of devalued and not-devalued groups is the definition of habit).

Table 2

*Experiment 2 Design*

Group	Acquisition (A)	Devaluation (A and B)	Extinction (B)	Test (A and B)
Paired	12 Sessions: R+	Pellets → LiCl	4 Sessions: R-	R-
Unpaired	12 Sessions: R+	Pellets / LiCl	4 Sessions: R-	R-

*Note.* A and B denote separate contexts; R = response; + = Reinforced; - = Not reinforced. Exposures to the contexts were equated throughout the experiment (not shown).

**Method**

*Subjects and Apparatus*

The subjects were 16 naïve female rats from the same supplier as those in Experiment 1. Housing, maintenance, and the apparatus were also the same.

*Procedure*

The procedure was exactly the same as that of Experiment 1 except that were 12 training sessions instead of three (to produce a habit), and there were four extinction sessions instead of three. Contexts were again matched for exposure and context exposures were counterbalanced across group and session.

**Results**

*Acquisition.* The acquisition data are depicted in Figure 6a. All rats acquired the lever pressing response. This was confirmed by a Session by Devaluation (Paired, Unpaired) repeated measures ANOVA, which yielded a significant effect of Session,  $F(11, 154) = 56.91, MSE = 1,239.09, p < .001, \eta^2 = .80$  on response rate. There were no effects of Devaluation or Devaluation by Session interaction, largest  $F = 1.40$ .



*Devaluation.* The devaluation data are displayed in Figure 7. The Paired rats acquired an aversion to the food pellets. This conclusion was confirmed by a repeated measures ANOVA, which found a main effect of Session among the Paired rats,  $F(5, 30) = 25.51$ ,  $MSE = 1.11$ ,  $p < .001$ ,  $\eta^2 = .81$ . The Unpaired rats ate all of the pellets given to them on every trial.

*Extinction.* The extinction data are depicted in Figure 6b. Response rates decreased in both groups both across and within sessions. A repeated measures ANOVA confirmed that there was a significant decrease in response rate both across Session,  $F(3, 42) = 15.68$ ,  $MSE = 182.31$ ,  $p < .001$ ,  $\eta^2 = .53$ , and 5-min Block,  $F(2, 28) = 57.18$ ,  $p < .001$ ,  $\eta^2 = .80$ . There was also a significant Devaluation by Session by Block interaction,  $F(3, 42) = 20.47$ ,  $p < .001$ ,  $MSE = 238.02$ ,  $\eta^2 = .59$ ;  $F(6, 84) = 2.23$ ,  $p = .048$ ,  $MSE = 19.89$ ,  $\eta^2 = .14$ , as well as a main effect of Devaluation,  $F(1, 14) = 5.80$ ,  $p = .031$ ,  $MSE = 507.00$ ,  $\eta^2 = .29$ .

As was the case in Experiment 1, the Session 1 extinction data are particularly important for assessing whether the response was habitual. A Devaluation by Block repeated measures ANOVA found a main effect of Devaluation,  $F(1, 14) = 25.45$ ,  $p < .001$ ,  $MSE = 25.45$ ,  $\eta^2 = .65$ , suggesting that the behavior was goal-directed despite the amount of training the rats had received. There was also a main effect of Block,  $F(2, 28) = 16.29$ ,  $MSE = 359.75$ ,  $p < .001$ ,  $\eta^2 = .53$ , but no Block by Devaluation interaction,  $F = 1.41$ . Interestingly, the response rates from the Paired rats increased from Session 1 of extinction to Session 2. A paired t-test explored differences in response rates within the Paired rats during the first block of Sessions 1 and 2 of extinction. The analysis revealed

a marginal increase from Session 1 Block 1 to Session 2 Block 1,  $t(7) = -2.22$ ,  $p = .062$ ,  $d = 0.69$ . Furthermore, an analysis of the Session 2 extinction data suggests that there was no difference in response rates between the Paired and the Unpaired groups at the time,  $F(1, 14) = 2.12$ ,  $MSE = 61.20$ ,  $p = .167$ ,  $\eta^2 = .132$ , suggesting that the behavior was now habitual. Furthermore, the behavior also appeared habitual during Session 3 of extinction,  $F(1, 14) = 0.05$ ,  $MSE = 1.76$ ,  $p = .825$ ,  $\eta^2 = .004$ .

*Test.* The test data are depicted in Figure 8. There was a strong renewal effect, and strong evidence of habit. This was confirmed by Devaluation by Context ANOVA, which revealed a significant effect of Context,  $F(1, 14) = 52.60$ ,  $MSE = 808.02$ ,  $p < .001$ ,  $\eta^2 = .79$ , but not of Devaluation,  $F < 1$ , on response rate. There was also no Context by Devaluation interaction,  $F < 1$ . The pattern suggests that the rats responded more in the renewal context than the extinction context and that the difference between the Paired and Unpaired rats remained constant, and null, across contexts. Within-context independent t-tests were conducted to confirm this. In Context A, there was no difference in response rates between the Paired rats and the Unpaired groups,  $t(14) = .33$ ,  $p = .748$ ,  $d = .16$ , nor was there a difference in Context B,  $t(14) = -1.25$ ,  $p = .232$ ,  $d = .60$ . These results suggest that the response renewed in both groups upon return to the acquisition context, and that the response renewed as a habit.

*Consumption Test.* The data from the consumption test can be found in Figure 9. In the consumption test, the Paired rats rejected the vast majority of the pellets they received in both contexts, while the Unpaired rats ate all of the pellets they received. A Devaluation by Context repeated measures ANOVA found a significant effect of

Devaluation,  $F(1, 14) = 222.81$ ,  $MSE = 6.04$ ,  $p < .001$ ,  $\eta^2 = .94$ , on proportion of pellets consumed, but no effect of Context or a Devaluation by Context interaction,  $F < 1$ .

*Reacquisition Test.* The reacquisition data for Contexts A and B are shown in Figures 10a and 10b respectively. In Context A, there was no main effect of Block,  $F = 1.29$ , but there was a main effect of Devaluation,  $F(1, 14) = 56.22$ ,  $MSE = 131,970.22$ ,  $p < .001$ ,  $\eta^2 = .80$ , and a Block by Devaluation interaction,  $F(29, 406) = 9.15$ ,  $MSE = 298.41$ ,  $p < .001$ ,  $\eta^2 = .395$ . In Context B, there were main effects of Block  $F(29, 406) = 3.33$ ,  $MSE = 55.57$ ,  $p < .001$ ,  $\eta^2 = .19$ , and Devaluation,  $F(1, 14) = 75.22$ ,  $MSE = 73,705.63$ ,  $p < .001$ ,  $\eta^2 = .84$ , as well as a significant Devaluation by Context interaction,  $F(29, 406) = 8.11$ ,  $MSE = 135.47$ ,  $p < .001$ ,  $\eta^2 = .367$ . These results suggest that the pellet was reinforcing enough to support the response in both contexts only in the Unpaired rats. Between subject analyses of the first minute of the Reacquisition test in each context were conducted to determine if the behavior was an action or a habit. Independent samples t-test confirmed that there was no difference in responding between the Paired and the Unpaired rats in Context B,  $t(14) = -1.41$ ,  $p = .181$ ,  $d = 1.15$ , or in Context A,  $t(14) = .19$ ,  $p = .852$ ,  $d = .10$ . These results further suggest that the response was a habit.

## Discussion

Following acquisition of a habit in Context A, and extinction of the response in Context B, the response renewed as a habit upon return to Context A. Both the Paired and the Unpaired groups increased their responding when returned to the acquisition context and importantly, there was no evidence of a devaluation effect, suggesting that the

behavior renewed as a habit. This result extends the findings of Experiment 1 to include habit as a form of instrumental behavior that is subject to renewal. Together, the results of Experiments 1 and 2 suggest that both actions (Experiment 1) and habits (Experiment 2) can be inhibited by extinction but retain their status when they renew in an ABA renewal paradigm.

An issue in this experiment is the difference in response rates between the Paired rats and the Unpaired rats on Session 1 of extinction. The amount of training given to the rats should have been sufficient to create a habit, which would have been demonstrated by a lack of devaluation effect. However, there was a clear devaluation effect throughout Session 1 of extinction. One possible explanation for this result is that when the context changed, the habit component of the behavior was weakened (Thrailkill & Bouton, 2015). Recall that extinction occurred in Context B after habit training had occurred in Context A. However, there was an increase in responding in the Paired rats from Session 1 to Session 2 of extinction, and there was no devaluation effect present on those days, which is difficult to explain.

An important point in Experiments 1 and 2 is that the rats were tested in their acquisition context, as opposed to a neutral context. This makes it unclear if removal from the extinction context is the key mechanism underlying the renewal of R-O and S-R, or whether return to the acquisition context is necessary. Existing evidence suggests that removal from the extinction context should be sufficient to renew goal-directed behavior, but not necessarily habitual behavior. This is because habitual behavior seems

to be more context specific than goal-directed behavior (Thrailkill & Bouton, 2015), and may therefore renew primarily when the rat is returned to the habit acquisition context.

### **Experiment 3: Renewal of Action after Conversion to Habit**

Experiment 3 tested the possibility of ABA renewal of an action after the response was converted to a habit. The extended training necessary to produce a habit can be thought of as a retroactive interference paradigm that, like extinction, may interfere with action learning rather than erase it. It may therefore be possible to create circumstances that promote the expression of R-O properties of behavioral control after habit training by creating distinct action and habit acquisition contexts. To do this, rats first lever pressed in Context A for 3 sessions (action training), followed by 12 sessions of training of the same response in Context B (habit training). We then performed reinforcer devaluation in both contexts, and tested the response in extinction in both contexts. No effect of devaluation was expected in Context B, as the behavior should be a habit in the habit training context. However, when returned to Context A, the action training context, the behavior might return to R-O control and renew as an action, which would manifest as a devaluation effect in the Paired rats.

Table 3

#### *Experiment 3 Design*

Group	Action Acquisition (A)	Habit Acquisition (B)	Devaluation (A and B)	Test (A and B)
Paired	3 Sessions: R+	12 Sessions: R+	Pellets → LiCl	R-
Unpaired	3 Sessions: R+	12 Sessions: R+	Pellets / LiCl	R-

*Note.* A and B denote separate contexts; R = response; + = Reinforcement. Exposures to the contexts were equated throughout the experiment (not shown).

### **Method**

### *Subjects and Apparatus*

The subjects were 16 naïve female Wistar rats of the same age as the previous experiments. Housing and maintenance was the same. The apparatus was identical to that of Experiment 2.

### *Procedure*

The procedure was the same as in the previous experiments except as noted. Rats were first given 3 sessions of RI 30-s training in Context A, followed by 12 sessions of training in Context B. Devaluation, test, consumption, and reacquisition then occurred, and followed procedures identical to those in the previous experiments. There was no extinction phase in this experiment. Contexts were again matched for exposure and context exposures were counterbalanced across group and sessions, as in the previous experiments.

## **Results**

*Acquisition and Devaluation.* One rat was removed from the analyses because it was an outlier showing unusually high responding during the test (Context A,  $z = 1.77$ , Context B,  $z = 2.13$ ). The acquisition data are depicted in Figure 11. All of the rats acquired the lever press response in both contexts. The overall rate of lever pressing increased across sessions, which was confirmed by a Devaluation by Session repeated measures ANOVA. The analysis found a main effect of Session,  $F(14, 182) = 44.80$ ,  $MSE = 1,642.52$ ,  $p < .001$ ,  $\eta^2 = .78$ , with no effect of Devaluation or a Devaluation by Session interaction,  $F < 1$ . Separate ANOVAs indicated that there was a main effect of

Session in both Context A,  $F(2, 26) = 41.10$ ,  $MSE = 481.10$ ,  $p < .001$ ,  $\eta^2 = .76$ , and Context B,  $F(11, 143) = 27.16$ ,  $MSE = 985.89$ ,  $p < .001$ ,  $\eta^2 = .68$ .

The devaluation results are shown in Figure 12. During devaluation, the proportion of pellets consumed by the Paired rats significantly decreased across sessions. A repeated measures ANOVA found a main effect of Session,  $F(5, 35) = 77.50$ ,  $MSE = 1.51$ ,  $p < .001$ ,  $\eta^2 = .92$  on proportion of pellets consumed among the Paired rats. The Unpaired rats ate all of their pellets during each devaluation trial.

*Test.* The mean response rates of the groups during the tests can be found in Figure 13a. A Context by Devaluation repeated measures ANOVA was used to assess the response rates of both groups in both contexts. The analysis found a main effect of Context,  $F(1, 13) = 25.64$ ,  $MSE = 294.51$ ,  $p < .001$ ,  $\eta^2 = .66$ , and a nonsignificant effect of Devaluation,  $F(1, 13) = 3.59$ ,  $MSE = 274.59$ ,  $p = .08$ ,  $\eta^2 = .22$ . There was no Context by Devaluation interaction,  $F(1, 13) = 1.45$ ,  $MSE = 16.68$ ,  $p = .25$ ,  $\eta^2 = .10$ . Within-context independent samples t-tests were used to isolate the devaluation effects. In Context A, the Paired group responded significantly less than the Unpaired group,  $t(13) = -2.38$ ,  $p = .034$ ,  $d = 1.23$ , indicating the presence of a devaluation effect and thus action. Importantly, there was no difference between Paired and Unpaired groups in Context B,  $t(13) = -1.25$ ,  $p = .234$ ,  $d = .63$ , suggesting habitual responding there.

The median response rates during the test are also displayed in Figure 13b. While the mean response rates of the Paired and Unpaired rats did not statistically differ in Context B, a visual inspection of the means suggested a slight devaluation effect, created by an individual rat with a particularly high response rate. Medians are less affected by

such extreme scores, and the pattern of the medians clearly suggests habit in Context B and action in Context A. Mann-Whitney U tests were conducted to compare the groups in Context A and Context B. There was no difference in responding between the Paired and Unpaired groups in Context B,  $U = 17.5$ ,  $p = .371$ , but there was a significant difference in Context A,  $U = 2$ ,  $p = .004$ . The results of this analysis continue to support the idea that the behavior was an action in Context A but a habit in Context B.

*Consumption Test.* The consumption data are shown in Figure 14. There was a main effect of Devaluation,  $F(1, 13) = 5,297.06$ ,  $MSE = 8.09$ ,  $p < .001$ ,  $\eta^2 = .99$ , on proportion of pellets consumed. There was no effect of Context or Context by Devaluation interaction,  $F < 1$ . These data confirm that the taste aversion was still present in the Paired rats in both contexts, but there was no such aversion in the Unpaired rats.

*Reacquisition Test.* The reacquisition data are summarized in Figure 15. In Context B, there was a marginally significant effect of Block,  $F(29, 377) = 1.42$ ,  $MSE = 5.06$ ,  $p = .075$ ,  $\eta^2 = .01$ , as well as a significant Devaluation by Block interaction,  $F(29, 377) = 7.95$ ,  $MSE = 301.92$ ,  $p < .001$ ,  $\eta^2 = .38$ . There was also a significant effect of Devaluation,  $F(1, 13) = 42.69$ ,  $MSE = 115,684.69$ ,  $p < .001$ ,  $\eta^2 = .77$ . An independent samples t-test did not find a significant devaluation effect during the first minute of the Reacquisition test,  $t(13) = -1.14$ ,  $p = .273$ . In Context A there was a main effect of Devaluation,  $F(1, 13) = 179.41$ ,  $MSE = 101,637.73$ ,  $p < .001$ ,  $\eta^2 = .93$ . There was no main effect of Block,  $F = 1.32$ , but there was a Devaluation by Block interaction,  $F(29, 377) = 6.32$ ,  $MSE = 247.01$ ,  $p < .001$ ,  $\eta^2 = .33$ . An independent samples t-test of the first minute did not find a devaluation effect,  $t(13) = -1.01$ ,  $p = .329$ ,  $d = .52$ . These data



indicate that the pellets were still reinforcing enough to support the response in both contexts only in the Unpaired rats.

### **Discussion**

Experiment 3 demonstrated renewal of goal-directed responding in Context A following an amount of training sufficient to produce a habit in Context B (Thrailkill & Bouton, 2015). That is, based on the results of a reinforcer devaluation test, following limited training in Context A, and extended training in Context B, the response renewed as an action in Context A and remained a habit in Context B. These results continue to support the notion that habits are context specific (Thrailkill & Bouton, 2015), as a return to the action acquisition context renewed the R-O responding. Perhaps more important, they suggest that the conversion of an action into a habit does not destroy the original action learning, which can still be shown in the action context. However, it is unclear from this experiment if return to the action acquisition context was necessary to renew action responding, or whether mere removal from the habit context would be sufficient to renew R-O responding. Experiment 4 was therefore designed to assess this question.

#### **Experiment 4: Returning a Habit to its Action Context versus a Neutral Context**

The purpose of Experiment 4 was to determine if the renewal of action after habit training requires return to the action acquisition context (ABA renewal), or if it can also be accomplished by simple removal from the habit training context and placement in a neutral third context (ABC renewal). If the context plays part of the role of the stimulus in eliciting habitual behavior (Thrailkill & Bouton, 2015), then removing an animal from the habit-training context should weaken habitual control, possibly resulting in the

renewal of action. However, it is unclear if returning the rat back to the action-acquisition context would add to the renewal of R-O responding compared to a neutral context. That is, is ABA renewal greater than ABC renewal?

In this experiment, rats received three sessions of training in Context A followed by 12 sessions of habit training (of the same response) in Context B as in Experiment 3. Reinforcer devaluation was then conducted in both Context A and Context B, as well as an associatively neutral Context C. Half of the rats were then tested in Context A and Context B, and half of the rats in Context C and Context B. As in Experiment 3, we would expect the rats' behavior to be a habit in Context B, and an action in Context A. Of more interest was what will happen when the test occurs in neutral context (Context C). Switching from the habit learning context should weaken habit (Thrailkill & Bouton, 2015), perhaps allowing transfer from S-R to R-O control, causing the behavior to renew as an action there as well. Also of interest was whether rats in this group would show as strong an action as the rats that were returned to the action acquisition context. Previous research has shown ABA renewal to be more robust than ABC renewal (Todd et al., 2012b, Experiment 3).

Table 4

*Experiment 4 Design*

Group	Action Acquisition (A)	Habit Acquisition (B)	Devaluation (A, B and C)	Test (A/C and B)
Paired ABA	3 Sessions (A): R+	12 Sessions (B): R+	Pellets → LiCl	(A and B): R-
Unpaired ABA	3 Sessions (A): R+	12 Sessions (B): R+	Pellets / LiCl	(A and B): R-
Paired ABC	3 Sessions (A): R+	12 Sessions (B): R+	Pellets → LiCl	(C and B): R-
Unpaired ABC	3 Sessions (A): R+	12 Sessions (B): R+	Pellets / LiCl	(C and B): R-

*Note.* A, B, and C denote separate contexts; R = response; + = Reinforced; - = Not reinforced. Exposures to the contexts were equated throughout the experiment (not shown).

**Method***Subjects*

Experiment 4 consisted of two replications of 24 female Wistar rats each. These are referred to Experiments 4a and 4b. The age and housing conditions of the rats were identical to previous experiments.

*Apparatus*

*Experiment 4a.* The apparatus used in this experiment was identical to the previous experiments with the exception that a third set of four operant chambers was used. These chambers were of the same dimensions as the other chambers and housed in another room in the laboratory, with a similarly spaced grid floor. There were no markings on the walls. Mr. Clean's citrus scented cleaning solution (Cincinnati, OH) was used to give the third context a unique scent.

*Experiment 4b.* Due to the fact that the boxes from the previous experiments were being used by another experimenter at the time of running this experiment, two new sets of boxes were used for this experiment. The two new sets of chambers measured  $31.75 \times 24.13 \times 29.21$  cm (length  $\times$  width  $\times$  height). The sidewalls consisted of clear acrylic plastic, and the front and rear walls were made of brushed aluminum. A recessed food cup was centered on the front wall approximately 2.5 cm above the floor. A retractable lever (model ENV-112CM, Med Associates) was positioned to the left of the food cup. The lever was 4.8 cm wide and 6.3 cm above the grid floor. It protruded 2.0 cm from the front wall when extended. The chambers could be illuminated by 7.5-W incandescent bulbs mounted to the ceiling of the sound-attenuation chamber. Ventilation fans provided background noise of 65 dBA. In Context 1, the floor consisted of 0.5 cm diameter stainless steel floor grids spaced 1.6 cm apart (center-to-center) and mounted parallel to the front wall. The ceiling and side wall had black horizontal stripes, 3.8 cm wide and 3.8 cm apart. Context 1 was scented using Hannaford's distilled white wine vinegar. In Context 2, the floor consisted of alternating stainless-steel grids with different diameters (0.5 and 1.3 cm, spaced 1.6 cm apart). The ceiling and side wall were covered with dark dots (2.0 cm in diameter). Context 2 was scented using Vicks Vaporub.

Context 3 was identical to the sets of chambers used in the previous experiments, meaning that it was smaller than the other two chambers used in this experiment, and it did not have any distinct markings on the back panel. It was scented using Mr. Clean's citrus scented cleaning solution.

### *Procedure*

*Acquisition.* Rats were given 3 days of training in Context A and 12 training sessions in Context B, as in Experiment 3. Context assignments were counterbalanced across rats in order to ensure that one third of the rats received action training in each set of boxes, one third of the rats received habit training in each set of boxes, and one third of the rats were given neutral exposures to each set of boxes. This was also divided equally between the Paired and the Unpaired rats. Rats were given exposure to all three contexts every day of training following a method similar to the ones used in the previous experiments. Order of context exposure cycled through a pattern that ensured that exposure to the contexts never occurred in the same order on two consecutive days (e.g., ACB, BAC, CBA, BCA, ABC, CAB). This cycle was completed twice.

*Devaluation.* Devaluation was conducted in all three contexts. The order of reinforcer devaluation in each Context was counterbalanced so that one sixth of the rats were given reinforcer devaluation in Context A, then B, then C, while another sixth were given reinforcer devaluation in Context B, then A, then C, etc. An example of the pattern used for this experiment was A (injection) – A (no injection) – B (injection) – B (no injection) – C (injection) – C (no injection). This cycle was performed until the Paired rats ate an average of less than one pellet. This meant that three cycles were used in Experiment 4a, whereas only two were used in Experiment 4b. Rats were only given one session per day during devaluation.

*Test.* All rats were tested in Context B, while half were tested in Context A (Groups Paired ABA and Unpaired ABA), and the other half were tested in Context C (Groups Paired ABC and Unpaired ABC). This meant that all of the rats were tested in

their habit training context (Context B), but only half of the rats were also tested in their action training context (Context A). The other half were also tested in a familiar, but neutral context (Context C).

*Consumption and Reacquisition.* Consumption tests were conducted in all three contexts, while the reacquisition test for each rat was conducted in the two contexts in which it had been tested (A and B for ABA groups, B and C for ABC groups). This occurred on the day following the consumption test. Consumption testing occurred in the opposite order as the extinction test. Reacquisition testing lasted for 15 minutes. Running order of the reacquisition test was the same as the extinction test day.

## Results

*Acquisition.* The acquisition data are depicted in Figure 16. Both Paired and Unpaired rats increased lever pressing across sessions, which was confirmed by an Experiment by Devaluation by Group (ABA, ABC) by Session repeated measures ANOVA, which found a main effect of Session,  $F(14, 560) = 64.04$ ,  $MSE = 4,762.04$ ,  $p < .001$ ,  $\eta^2 = .62$  on response rate, with no main effects of Devaluation, Group, or Experiment, or any significant interactions, largest  $F = 1.56$ .

*Devaluation.* Because the rats in Experiment 4a took three cycles to fully devalue the pellet and the rats in Experiment 4b took two cycles, the devaluation data were analyzed separately for each experiment. The Experiment 4a devaluation data are depicted in Figure 17a. A repeated measures ANOVA confirmed that the Paired rats consumed significantly fewer pellets across sessions,  $F(8, 80) = 39.93$ ,  $MSE = 1.68$ ,  $p < .001$ ,  $\eta^2 = .80$  throughout the devaluation phase. There was no effect of Group (ABA vs

ABC), or Session by Group interaction, largest  $F = 3.15$ . The Unpaired rats consumed all of the pellets given to them. The Experiment 4b devaluation data are presented in Figure 17b. A repeated measures ANOVA confirmed that the Paired rats consumed significantly fewer pellets across Sessions,  $F(5, 50) = 77.19$ ,  $MSE = 1.87$ ,  $p < .001$ ,  $\eta^2 = .89$  throughout devaluation. There was no effect of Group,  $F(1, 10) = 1.08$ ,  $MSE = .04$ ,  $p = .323$ ,  $\eta^2 = .10$ , but there was a Session by Group interaction,  $F(5, 50) = 7.22$ ,  $MSE = .18$ ,  $p < .001$ ,  $\eta^2 = .42$ . This was caused by a large difference between the Paired ABA and the Paired ABC rats during Session 4, which did not exist in the final two sessions. An Experiment by Group ANOVA was conducted to analyze any differences in the overall level of pellet rejection on the last day of devaluation. The results indicated that there were no effects of Experiment or Test Context, nor was there a significant interaction between the two,  $F < 1$  on pellet consumption at the end of reinforcer devaluation. The Unpaired rats consumed all of the pellets given to them during the devaluation phase.

*Test.* One rat was removed from the analyses due to extremely high response rates during the test (Context A,  $z = 3.52$ , Context B,  $z = 3.68$ ). The test data for groups Paired ABA and Unpaired ABA can be found in Figure 18a, while the test data for groups Paired ABC and Unpaired ABC are displayed in Figure 18b. A Test Context (A/C vs B) by Group (ABA vs ABC) by Devaluation by Experiment repeated measures ANOVA was conducted to analyze differences in response rates. The analysis found a significant effect of Context,  $F(1, 39) = 19.31$ ,  $MSE = 526.26$ ,  $p < .001$ ,  $\eta^2 = .33$ . There were also significant effects of Devaluation,  $F(1, 39) = 10.27$ ,  $MSE = 590.06$ ,  $p = .003$ ,  $\eta^2 = .21$ ,

and Experiment,  $F(1, 39) = 14.77$ ,  $MSE = 849.03$ ,  $p < .001$ ,  $\eta^2 = .28$ . No other factors approached significance, largest  $F = 2.35$ .

Within context analyses were conducted to gain a better understanding of the data. An independent samples t-test confirmed that there was a significant devaluation effect in Context A in the groups tested there,  $t(21) = -2.56$ ,  $p = .019$ ,  $d = 1.06$ , as well as in Context C,  $t(22) = -2.68$ ,  $p = .014$ ,  $d = 1.06$ . However, there was no significant devaluation effect in Context B (even when collapsing over group),  $t(45) = -1.71$ ,  $p = .094$ ,  $d = .05$ . These results are consistent with the view that the behavior was a habit in the habit context, but renewed as an action in both the action training context (Context A) and in the neutral context (Context C).

We also conducted a Group (ABA vs ABC) by Devaluation repeated measures ANOVA to isolate response rates during the crucial Context A and C tests to gain a better understanding of the potential role of returning to the action acquisition context compared to a neutral context. This set of analyses found an overall effect of Devaluation,  $F(1, 47) = 13.49$ ,  $MSE = 443.40$ ,  $p = .001$ ,  $\eta^2 = .24$ , but no effect of Group, or Group by Devaluation interaction, largest  $F = 2.73$ . Overall, these analyses indicate that the response was an action in Contexts A and C, but was a habit in Context B. Furthermore, the action seemed to be equally strong in Context A and Context C.

*Consumption Test.* The data from the consumption test for the ABA and ABC rats are depicted in Figures 19a and 19b respectively. A Context by Group (ABA vs ABC) by Devaluation repeated measures ANOVA yielded a main effect of Devaluation,  $F(1, 39) = 3,038.02$ ,  $MSE = 28.05$ ,  $p < .001$ ,  $\eta^2 = .99$  on pellet consumption. There were no effects



of Context or of Group, nor was there any interaction, largest  $F = 1.08$ . These results confirm that the aversion to the pellets was maintained in the Paired rats while the Unpaired rats had no such aversion.

*Reacquisition Test.* The reacquisition data are shown in Figure 20. In Context B, a Block by Devaluation repeated measures ANOVA found that there was a marginally significant effect Block,  $F(14, 602) = 1.56$ ,  $MSE = 137.31$ ,  $p = .087$ ,  $\eta^2 = .04$ , but there was a Devaluation by Block interaction,  $F(14, 602) = 13.18$ ,  $MSE = 1,163.11$ ,  $p < .001$ ,  $\eta^2 = .24$ . There was also a significant effect of Devaluation,  $F(1, 43) = 30.07$ ,  $MSE = 97,185.56$ ,  $p < .001$ ,  $\eta^2 = .41$ . There was no difference in responding between the Paired and Unpaired rats during the first minute,  $t < 1$ , suggesting that the response was still a habit in Context B.

In Context A, a Block by Devaluation repeated measures ANOVA found no effect of Block,  $F < 1$ , but did find a Block by Devaluation interaction,  $F(14, 294) = 6.20$ ,  $MSE = 444.92$ ,  $p < .001$ ,  $\eta^2 = .23$ , as well as a significant effect of Devaluation,  $F(1, 21) = 18.62$ ,  $MSE = 40,754.78$ ,  $p < .001$ ,  $\eta^2 = .47$ . There was no difference in response rates during the first minute between the Paired and Unpaired rats,  $t < 1$ , suggesting possible spontaneous recovery of habit.

In Context C, a Block by Devaluation repeated measures ANOVA found no effect of Block,  $F < 1$ , but there was a Block by Devaluation interaction,  $F(14, 280) = 6.78$ ,  $MSE = 481.67$ ,  $p < .001$ ,  $\eta^2 = .25$  as well as a main effect of Devaluation,  $F(1, 20) = 28.32$ ,  $MSE = 62,357.34$ ,  $p < .001$ ,  $\eta^2 = .59$ . There was a significant difference in response rates between the Paired and Unpaired rats during the first minute of the

reacquisition test,  $t(22) = -2.29$ ,  $p = .032$ ,  $d = .94$ , suggesting that the response was still an action. These results suggest that the pellet was reinforcing enough to support the response in all contexts only in the Unpaired rats.

### **Discussion**

Rats that were returned to their action training context (A), or a neutral context (C) showed renewed action after habit training in Context B. The results suggest that neutral contexts are as effective as the action acquisition context at replacing habitual control with action control. Given this pattern of results, it seems that any change in context is sufficient to reduce the S-R control and renew R-O responding. Furthermore, the fact that the size of the devaluation effect was approximately equal in Groups ABA and ABC demonstrates the ability of actions to transfer across context, i.e., from Context A (where it was trained) to Context C. Overall, this experiment supports the notion that habits are context specific and that actions transfer, apparently intact, across contexts.

### **General Discussion**

The present experiments examined the consequences of inhibiting R-O (action) and S-R (habit) associations following retroactive interference treatments. In Experiment 1, rats were trained to perform an instrumental response under conditions known to promote an action in Context A, received reinforcer devaluation in Contexts A and B, and extinction training of the response in Context B. Finally, the response was tested in extinction in both contexts. The results of the test indicated that actions renew as actions in an ABA renewal paradigm, suggesting that extinction in Context B did not erase or qualitatively change the R-O association that controlled the behavior in Context A. While

renewal of actions has been previously shown (Cohen-Hatton & Honey, 2013), reinforcer devaluation was done in the prior experiment using an unusual procedure that could have easily modified the R-O association (which was not affected using LiCl reinforcer devaluation here). Experiment 2 tested renewal of habits using the same ABA renewal design and demonstrated that habits likewise renew as habits under these circumstances. This result indicates that any inhibition of the response that formed during extinction in Context B did not erase or change the S-R association in Context A. Together, the results of Experiments 1 and 2 suggest that actions and habits both retain their status when they renew after extinction.

Experiment 3 asked what happens to an action following retroactive interference cause by its conversion into habit through additional training. By moderately training a response in Context A, overtraining it in Context B, and performing reinforcing devaluation in both contexts, I found that a response can renew as an action in the action acquisition context following habit training, while remaining a habit in the habit acquisition context. This result suggests that S-R associations that form during overtraining do not erase previous R-O associations. Upon return to the action acquisition context, the S-R component of the behavior was attenuated, allowing the response to renew to its goal-directed status. The results of Experiment 4 in turn suggest that the action status of a behavior renews both when returned to the action acquisition context and to a neutral context. Together, these results suggest that actions and habits maintain their associative properties following retroactive interference paradigms such as

extinction and additional training, and that these properties once again control performance, or renew, when the context is changed.

The results of these experiments provide new insight to motivational processes in the renewal of instrumental behavior. While previous accounts of renewal have addressed extinction of instrumental learning in a general way, I am unaware of any account that directly addresses the renewal of R-O and S-R properties following either extinction or additional training. The results of these experiments suggest that R-O associations are subject to very similar laws that govern general instrumental renewal. Extinction seems to inhibit the R-O association only in a specific context, and release from the context-specific inhibition allows the R-O association to renew, as is suggested by Experiment 1. Renewal of S-R associations (Experiment 2) seem to work in a similar way; however, it may turn out that release from context-specific inhibition may not be sufficient to renew the S-R association. Instead, it may be necessary to return to the habit acquisition context. This idea is supported particularly by the results of Experiment 4, where removal from the habit training context appears to have renewed the R-O associations regardless of whether the rat was returned to its action acquisition context (A) or a neutral context (C). This finding supports the hypothesis that habits are specifically linked to their training context (B), while actions are able to transfer across contexts A and C unchanged.

Generally, the results are in accord with the notion that the context is a part of the stimulus supporting habit responding (Thrailkill & Bouton, 2015). When an animal is removed from the habit training context, the contextual S is removed, reducing S-R control and evidently promoting more of the original, first-learned goal-directed behavior

(Experiment 4). This allows actions to transfer across contexts more effectively than habits. A way to assess this is by comparing the size of devaluation effects across different contexts, which can be thought of as the goal-directed component of the behavior. In Experiment 4, the size of the devaluation effect was very similar in both the action training context and the neutral context, suggesting that the action transferred between them very well. This is similar to observations by Thraillkill and Bouton (2015), and supports their idea that actions transfer across contexts better than habits.

An important issue to address in Experiment 2 is the response rates of the Paired rats in extinction. Following habit training in Context A, the rats received extinction in Context B. On Session 1 of extinction, there was a clear devaluation effect, suggesting that the response was an action, even though it had been given extended training. However, the Paired rats increased their responding on Session 2, and converged with the Unpaired rats, suggesting that the behavior was then a habit. The extinction data from Sessions 3 and 4 also suggest the behavior was a habit, as did the Context A test data. The unique action performance on Session 1 raises several questions. First, it is inconsistent with the results of Thraillkill and Bouton (2015, Experiment 4), who reported observing a weakened habit when it was switched to a neutral context following 12 sessions of 30 reinforcers on a VI 30-s schedule and 6 cycles of devaluation. The parameters of the first extinction session of Experiment 2 were the same as in the test in Thraillkill and Bouton (2015), suggesting that I should have also observed a habit in Context B on Day 1. While the results in both the present Experiment 2 and Thraillkill and Bouton (2015, Experiment 4) showed reduced responding following a context

switch, the rats in the present Experiment 2 seemed to lose both S-R associations, and gain R-O associations, whereas Thraillkill and Bouton's rats seemed to lose only S-R. One possible explanation for the difference is that Thraillkill and Bouton tested the rats in extinction in both the habit training context and the neutral context on their test day. In contrast, in Experiment 2, the rats were given an extinction session in the extinction (neutral at the time) context, as well as a mere exposure session to the habit training context. It is possible that extinction and exposure in the habit training contexts had different effects on the response, which influenced what was observed in Context B.

Another interesting pattern in Experiment 2 extinction was that the response seemed to return to habit status after the first day of extinction (i.e., on Sessions 2, 3, and 4). One possible explanation for this is that the habit spontaneously recovered to some extent after Session 1. The idea that habits, and potentially actions, can spontaneously recover is in line with the idea that actions and habits renew as actions and habits (Experiments 1 and 2). Both ideas are consistent with the idea that retroactive interference paradigms do not erase existing R-O and S-R associations.

There is still a question as to why habits are more sensitive to contextual cues than actions. One potential perspective on this is that habits are arguably learned after actions. Nelson (2002) has provided important evidence as to differences in context-sensitivity between first and second learned associations, specifically that the second learned association to a CS is more context-specific than the first. In his Experiment 1, Nelson first trained stimulus T as a conditioned inhibitor in Context A by compounding it with stimulus D, which had been trained as a conditioned exciter, during nonreinforced

trials (D+, TD-). Then Nelson gave stimulus T excitatory conditioning trials in Context A (T+). Finally, Nelson tested stimulus T in both Context A and Context B. He found that T had become excitatory in Context A; however, this excitation was attenuated in Context B. Importantly, the context switch had more of an effect on the response on the group for which T had first been trained as an inhibitor compared to an excitatory training only group or a no pre-training control group. This suggests that the second learned association was more context specific than the first. In Experiment 2, Nelson demonstrated that this was also true when the second-learned association was inhibitory rather than excitatory. Nelson argued that the additional training made the CS ambiguous, and animals used the context to disambiguate the meaning of the CS. This idea is based on one proposed by Bouton (1993), which posits that during extinction, the meaning of the CS becomes ambiguous, and the context helps to disambiguate its meaning, resulting in context-specific inhibition.

At the root of the present experiments is the notion that behaviors can have both R-O and S-R components that interact to regulate levels of outcome sensitivity. Various dual-process theories have been proposed to explain how R-O and S-R learning interact (Dickinson et al., 1995; de Wit & Dickinson, 2009). A commonly-accepted principle is that behaviors develop R-O associations early in training, and S-R associations develop later in training. However, based on recent evidence, it may be more likely that R-O and S-R associations both develop early in training (Thrailkill & Bouton, 2015). This raises multiple questions as to how behaviors become habitual. Do actions and habits develop at the same time, but actions stop developing earlier than habits, allowing habits to control

behavior with extended training? Do habits form slower, but ultimately stronger associations than actions? Gaining a better understanding of the development of specific R-O and S-R strategies will be important in helping to inhibit bad habits such as overeating and drug use, as well as promoting good ones such as exercise.

A neurobiological explanation for the transition between R-O and S-R strategies may be provided by Gremel and Costa (2013), who have also reported observing behaviors that are actions in one context but habits in another. The authors propose that competing orbitofrontal and striatal circuits control context-induced shifts between habits and actions. To test this, Gremel and Costa trained a lever press response in mice on a random ratio and random interval schedule in different contexts to promote the use of R-O and S-R strategies, respectively (Dickinson et al., 1983). Following this training, they employed the specific satiety method of reinforcer devaluation to test for a devaluation effect in each context, and reported observing an action in the random-ratio context, but a habit in the random-interval context. The authors then lesioned either the dorsomedial striatum (DMS), dorsolateral striatum (DLS), or the orbitofrontal cortex (OFC) of mice subjected to the same paradigm to evaluate the role of each structure. Gremel and Costa (2013) reported that DMS-lesioned mice were habitual in both contexts, whereas DLS-lesioned mice were goal-directed in both contexts, suggesting a role for the DMS in action and the DLS in habit (see also Yin et al., 2005; Yin et al., 2004; Yin, Knowlton, & Balleine, 2006). When the authors lesioned the OFC, the mice failed to reduce lever pressing after devaluation in either context, suggesting habitual behavior. Furthermore, the authors selected individual neurons in these corticostriatal regions to record during



each test phase. Shifts in activity of OFC and DMS neuron ensembles during the test correlated with the execution of goal-directed behaviors, but not the execution of habits. This result suggests that shifting back to goal-directed behavior after habits are established corresponds to a shift in the activity of specific corticostriatal ensembles in these regions. Finally, chemogenetic lesioning of OFC neurons inhibited goal-directed behavior. In another experiment, selective depletion of CB1 receptors in the OFC-dorsal striatum (DS) pathway also prevented habit formation (Gremel, Chancey, Atwood, Luo, Neve, Ramakrishnan, Deisseroth, Lovinger, & Costa, 2016). Together, the evidence suggests that the OFC may play a role in the balance between actions and habits, and that CB1 receptors are involved as well.

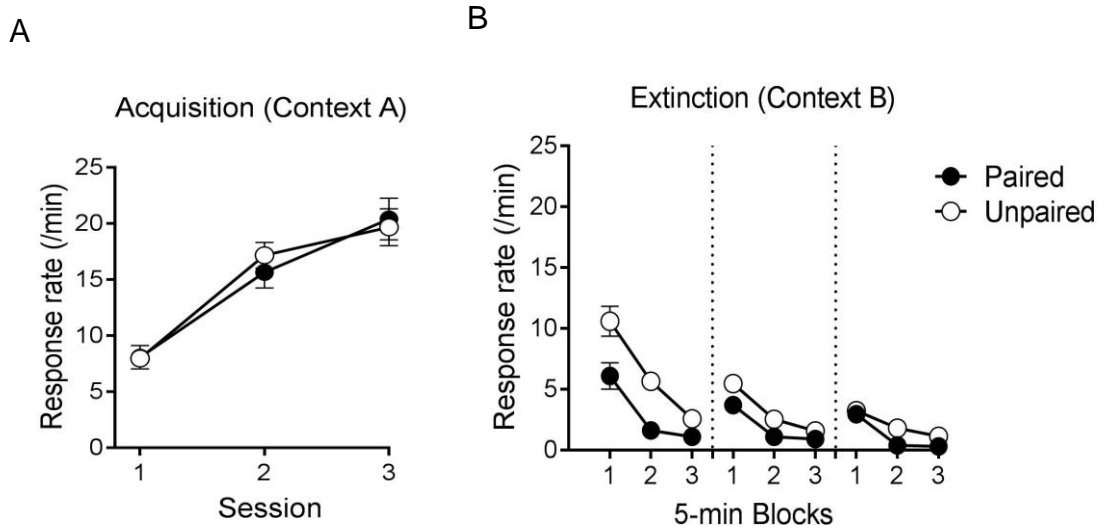
The results of the present experiments lend support to dual-process theories of habit formation that suggest actions and habits develop separately and at the same time (Dickinson et al., 1995; de Wit and Dickinson, 2009; Gremel and Costa, 2013) as opposed to accounts that suggest actions convert into habits following extended training (Dickinson, 1985). An important tenet of this idea is that responses are not purely actions or habit, but contain both R-O and S-R properties that are best expressed under different conditions (e.g., different contexts). An important point of inquiry in the field will be to define the sets of cues that best promotes each type of behavior. Context seems to be an important cue in the transition between actions and habits, but other factors, such as surprise (Rey, Thraillkill, & Bouton, unpublished) and predictability of the reinforcer (Thraillkill, Trask, Vidal, Alcalá & Bouton, 2018) seem to play a role as well. The current experiments expand upon dual-process theories by providing evidence that R-O

associations remain intact after habit learning and are able to control behavior under the correct circumstances. Dual-process theories will need to adapt to account for the context specificity of habits as well as the ability of actions to transfer across contexts.

These experiments also provide insight into why people often relapse after being treated for problematic habitual behavior. The results of Experiment 2 suggest that a habit that is treated in a clinic is subject to renewal as a habit upon return to the acquisition context. This finding could help explain the high levels of relapse following treatment, as people often acquire a response in one context, seek treatment for it in a different context, and then return to the original context. However, the results of Experiments 3 and 4 suggest that it is possible to renew the action components of the behavior through a context switch after habits are learned. While it is not reasonable or feasible for every person with a habitual eating disorder or a drug habit to move to a new context, it may still be possible to renew the goal-directed component of the behavior by making the acquisition context more similar to the extinction/treatment context or to a neutral context. This may involve incorporating stimuli from the treatment context into the acquisition context, or instituting new behaviors into the acquisition context.

In summary, the present experiments have revealed that R-O and S-R associations can remain following retroactive interference treatments such as extinction and additional habit training. Furthermore, they provide additional evidence for the context specificity of habits, and for the ability of actions to transfer across contexts. They also add to the idea that behaviors can simultaneously have S-R and R-O properties that can each be expressed under the proper circumstances.

*Figure 1: (A) Experiment 1 acquisition data. Mean response rates in both groups during acquisition in Context A. (B) Experiment 1 extinction data. Mean response rates of both groups during extinction in Context B. Data are presented both within session (in 5-minute blocks) and across sessions (sessions are separated by dotted lines). Error bars represent standard error of the mean (SEM).*



*Figure 2: Experiment 1 devaluation data. Mean proportion of pellets consumed by both groups during devaluation. Error bars represent standard error of the mean (SEM).*

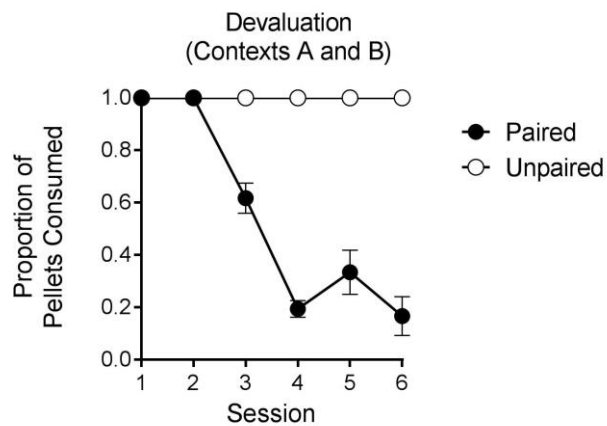


Figure 3: Experiment 1 test data. Mean response rates of both groups in both Contexts A and B. Response rates reflect the overall rate during the 10-minute session. Error bars represent standard error of the mean (SEM).

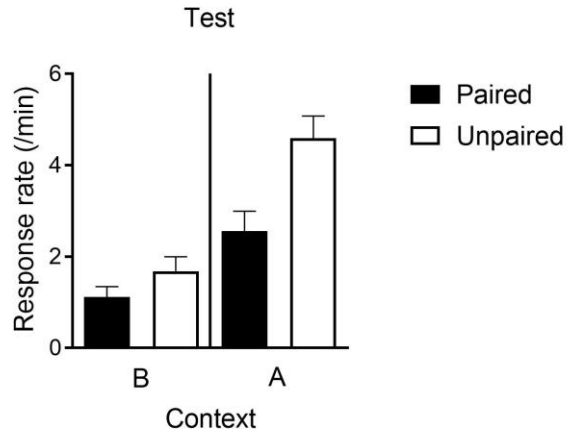


Figure 4: Experiment 1 consumption data. Mean proportion of pellets consumed by each group in both Contexts A and B. Error bars represent standard error of the mean (SEM).

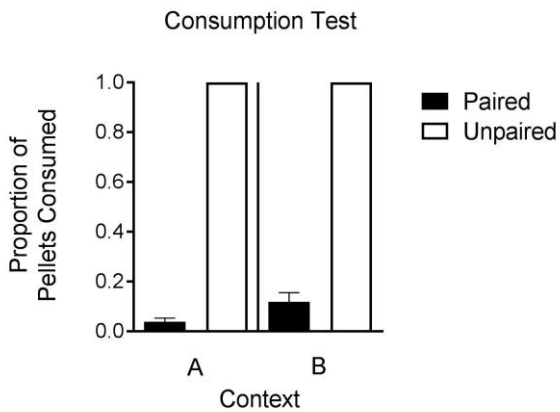
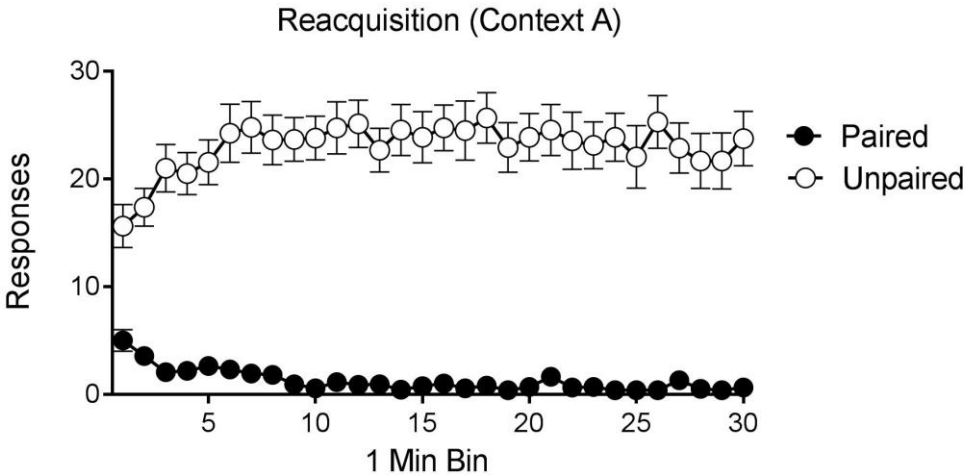


Figure 5: Experiment 1 reacquisition data in (A) Context A and (B) in Context B. Mean number of responses by each group in each context depicted in 1-minute blocks. Error bars represent standard error of the mean (SEM).

A



B

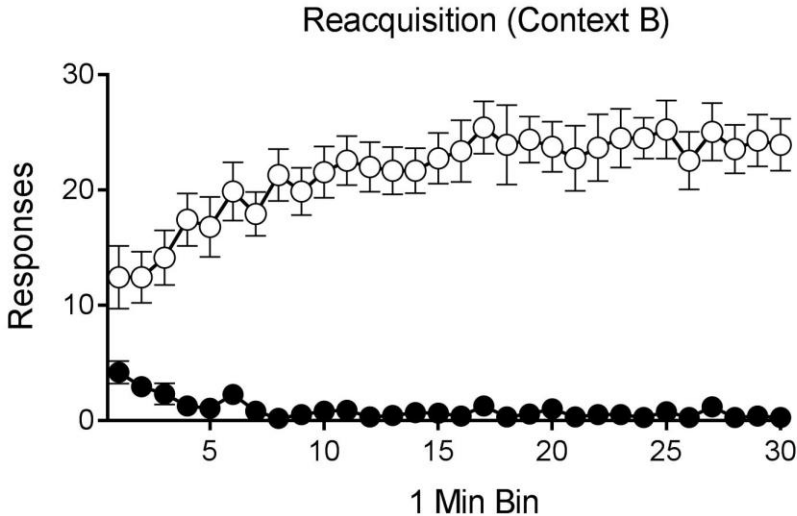


Figure 6: (A) Experiment 2 acquisition data. Mean response rates in both groups during acquisition in Context A. (B) Experiment 2 extinction data. Mean response rates of both groups during extinction in Context B. Data is presented both within session (in 5-minute blocks) and across sessions (sessions are separated by dotted lines). Error bars represent standard error of the mean (SEM).

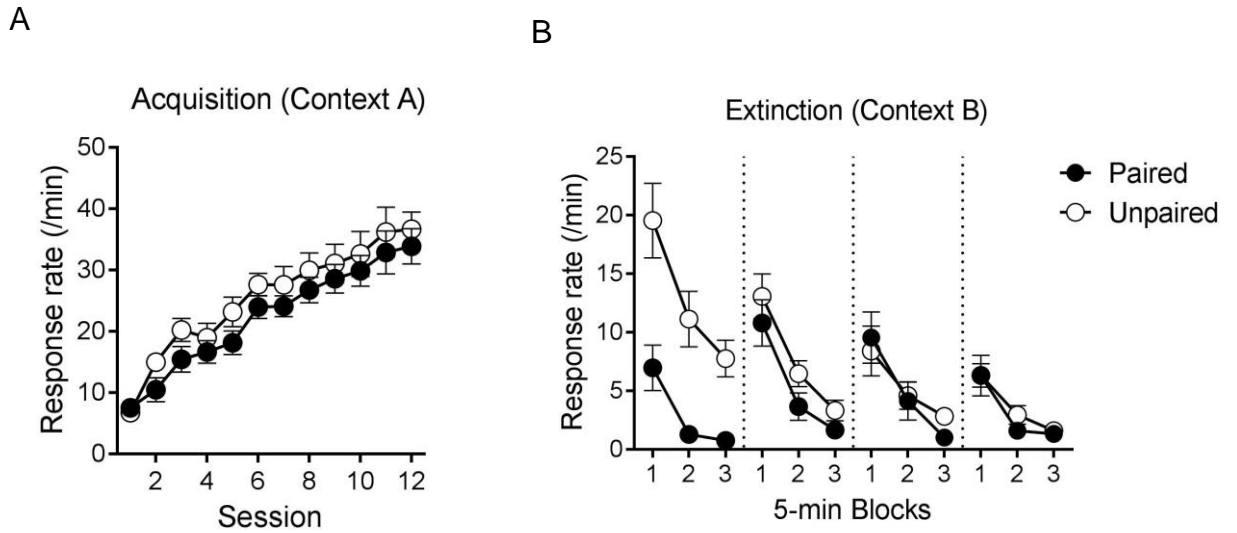


Figure 7: Experiment 2 devaluation data. Mean proportion of pellets consumed by both groups during devaluation. Data is collapsed over Contexts A and B. Error bars represent standard error of the mean (SEM).

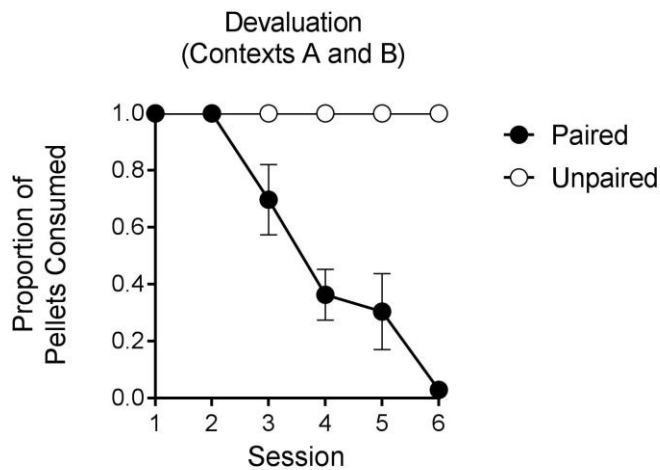


Figure 8: Experiment 2 test data. Mean response rates of both groups in both Contexts A and B. Response rates reflect the overall rate during the 10-minute session. Error bars represent standard error of the mean (SEM).

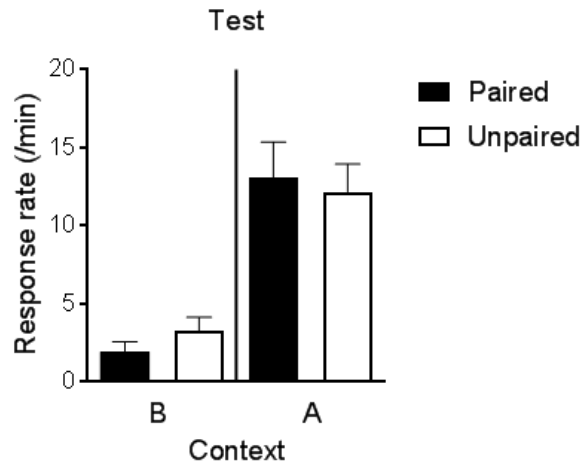


Figure 9: Experiment 2 consumption data. Mean proportion of pellets consumed by each group in both Contexts A and B. Error bars represent standard error of the mean (SEM).

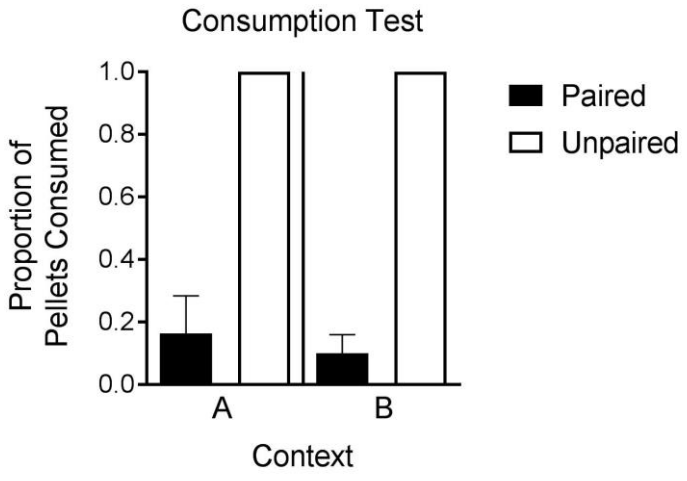
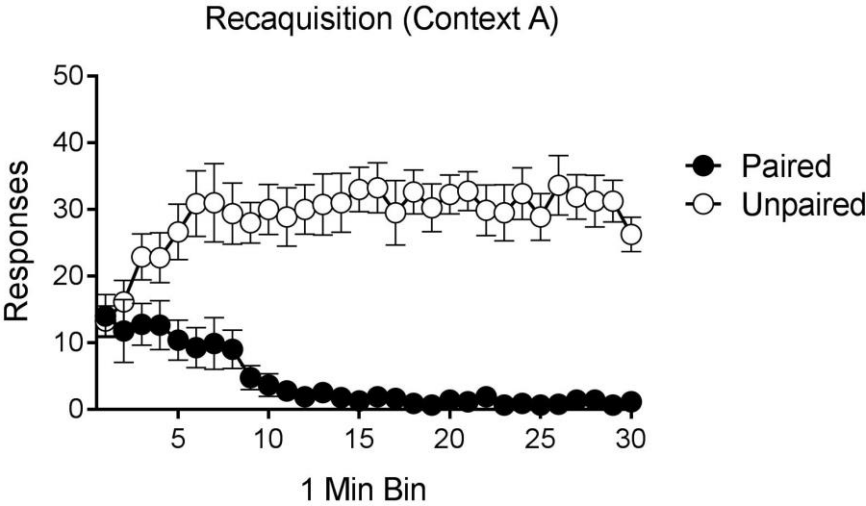


Figure 10: Experiment 2 reacquisition data. Mean number of responses by each group in (A) Context A and (B) Context B, depicted in 1-minute blocks. Error bars represent standard error of the mean (SEM).

A



B

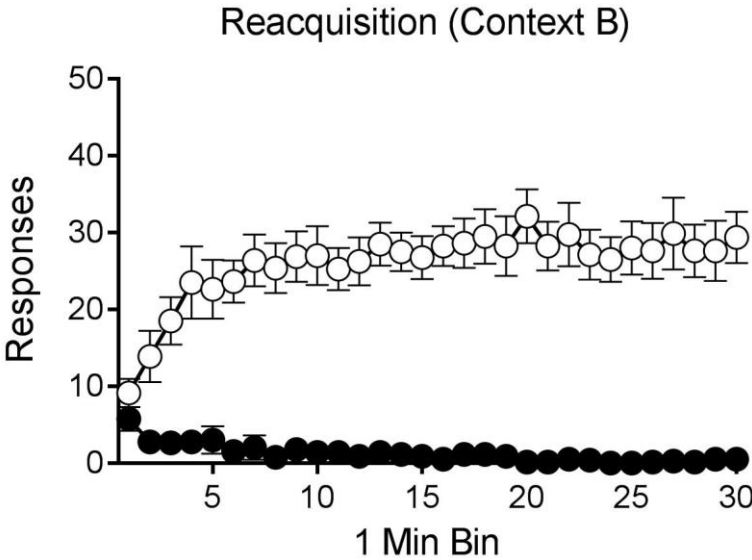




Figure 11: Experiment 3 acquisition data. Mean response rates in both groups during acquisition in Contexts A and B. Error bars represent standard error of the mean (SEM).

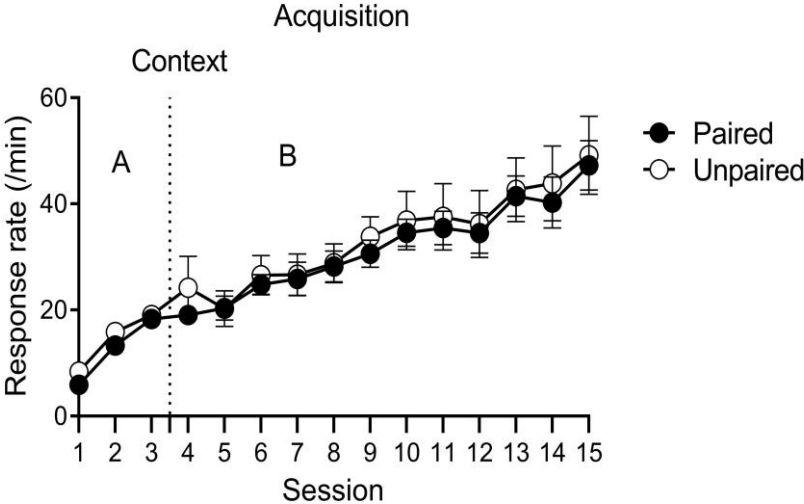


Figure 12: Experiment 3 devaluation data. Mean proportion of pellets consumed by both groups during devaluation. Data is collapsed over Contexts A and B. Error bars represent standard error of the mean (SEM).

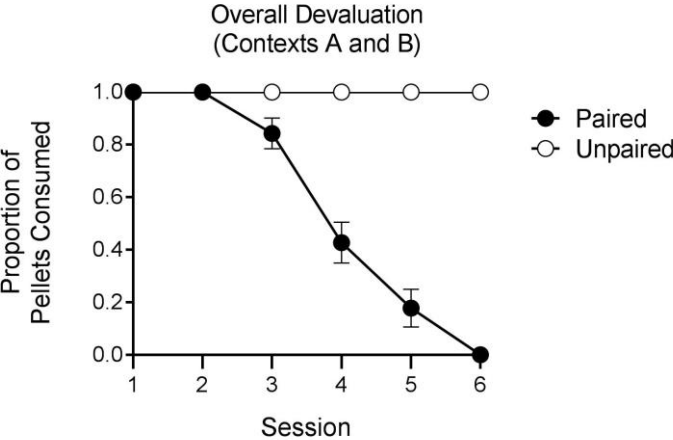


Figure 13: Experiment 3 test data. (A) Mean response rates of both groups in both Contexts A and B. (B) Median response rates of both groups in both contexts. Response rates reflect the overall rate during the 10-minute session. Error bars represent standard error of the mean (SEM).

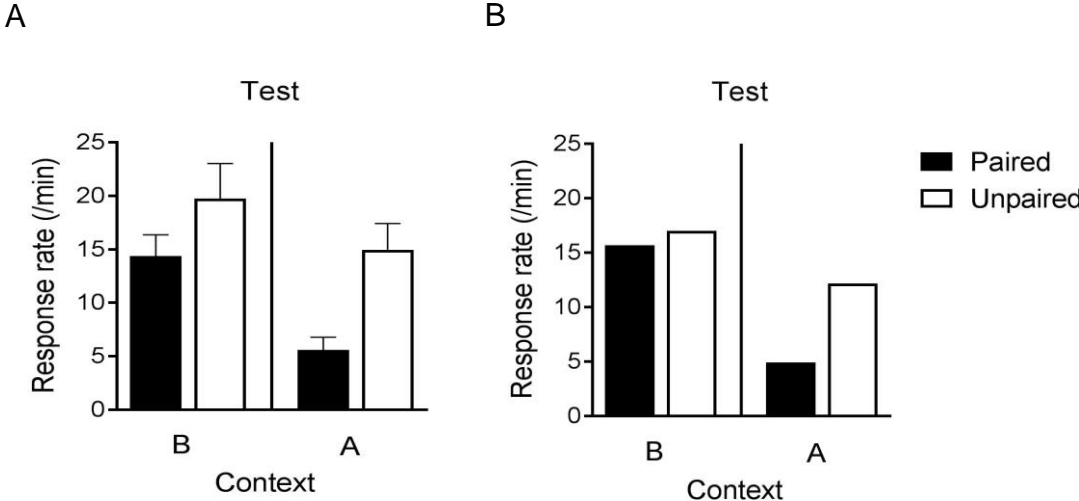


Figure 14: Experiment 3 consumption data. Mean proportion of pellets consumed by each group in both Contexts A and B. Error bars represent standard error of the mean (SEM).

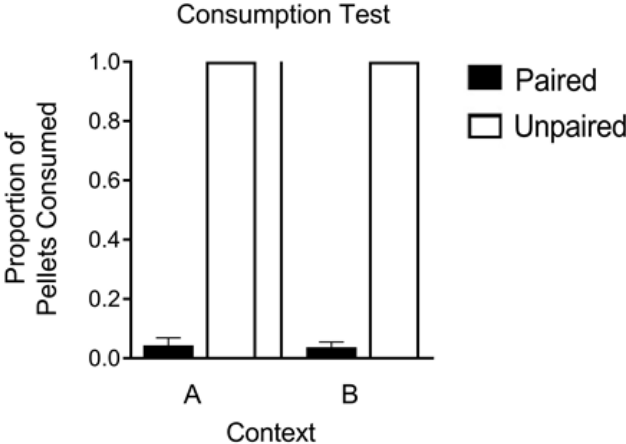


Figure 15: Experiment 3 reacquisition data. Mean number of responses by each group in (A) Context A and (B) Context B, depicted in 1-minute blocks. Error bars represent standard error of the mean (SEM).

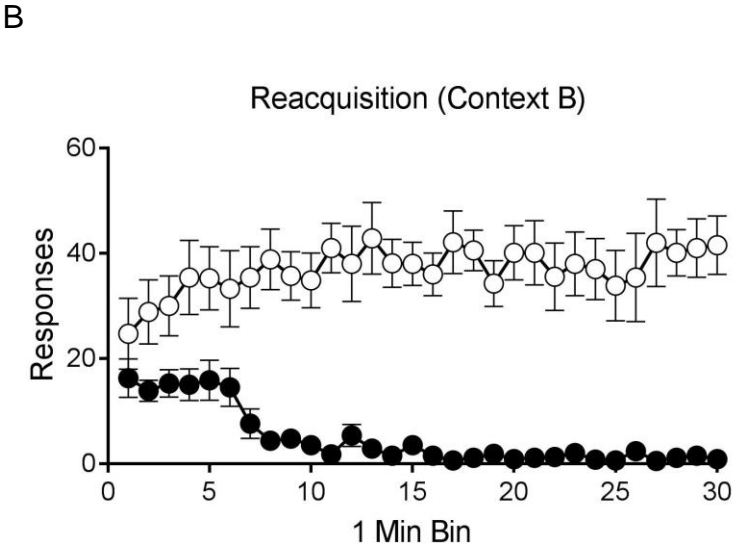
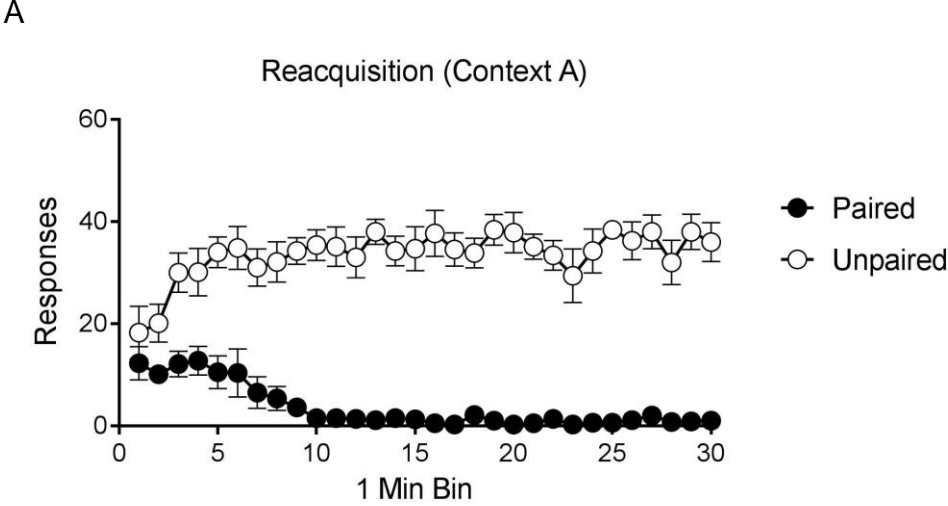


Figure 16: Experiment 4 acquisition data collapsed across Experiments 4a and 4b. Mean response rates of responding in all groups during acquisition in Contexts A/C and B. Error bars represent standard error of the mean (SEM).

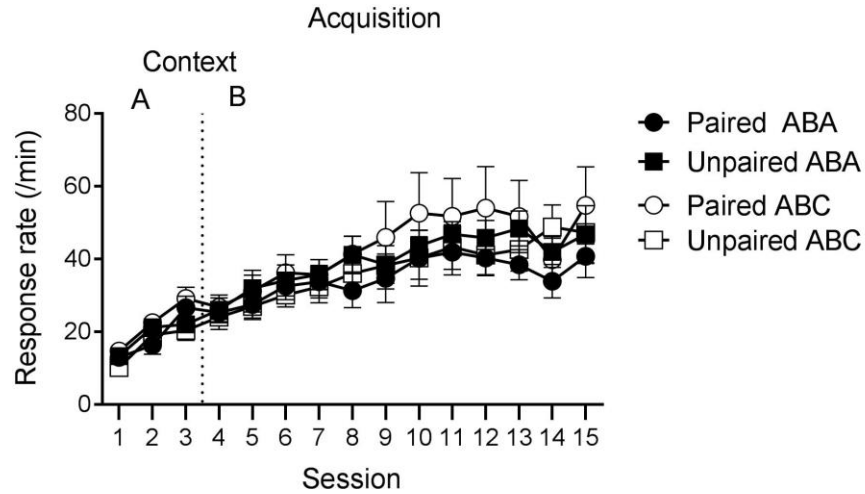
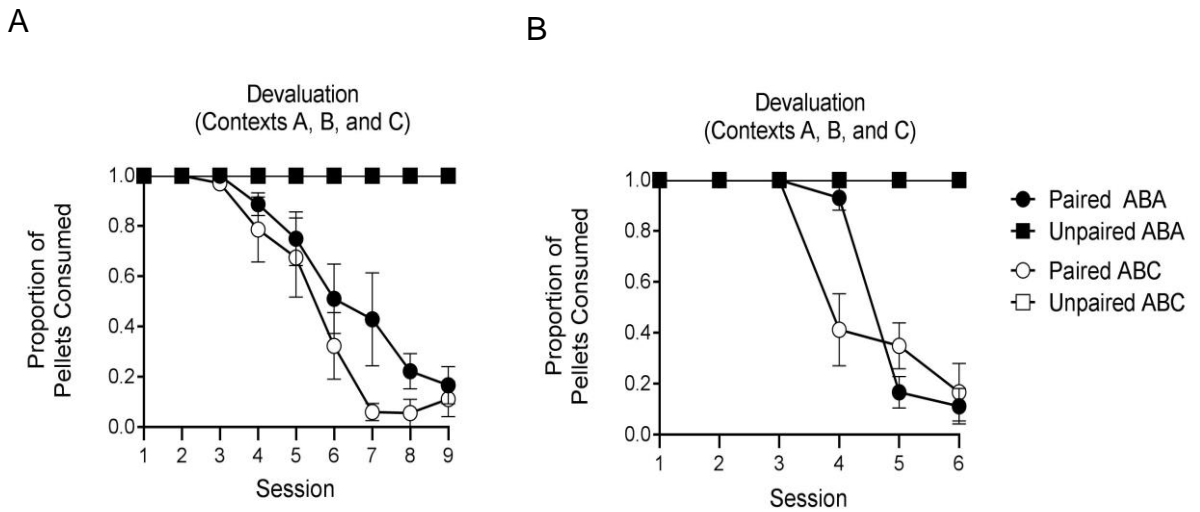
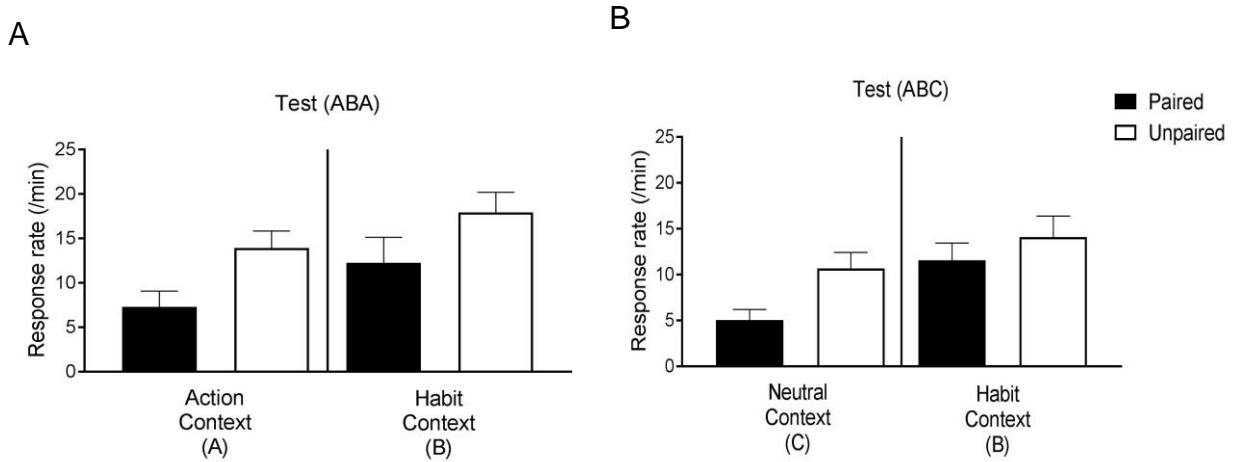


Figure 17: Experiment 4a devaluation data. Mean proportion of pellets consumed by all groups during devaluation for (A) Experiment 4a and (B) Experiment 4b. Data is collapsed over Contexts A, B, and C. Error bars represent standard error of the mean (SEM).



*Figure 18: Experiment 4 test data. Mean response rates for (A) Groups Paired ABA and Unpaired ABA and (B) Groups Paired ABC and Unpaired ABC. Response rates reflect the overall rate during the 10-minute session. Error bars represent standard error of the mean (SEM)*



*Figure 19: Experiment 4 consumption data. Mean proportion of pellets consumed by (A) Groups Paired ABA and Unpaired ABA collapsed across Experiments 4a and 4b. (B) Experiment 4 consumption data for Groups Paired ABC and Unpaired ABC collapsed across Experiments 4a and 4b. Error bars represent standard error of the mean (SEM).*

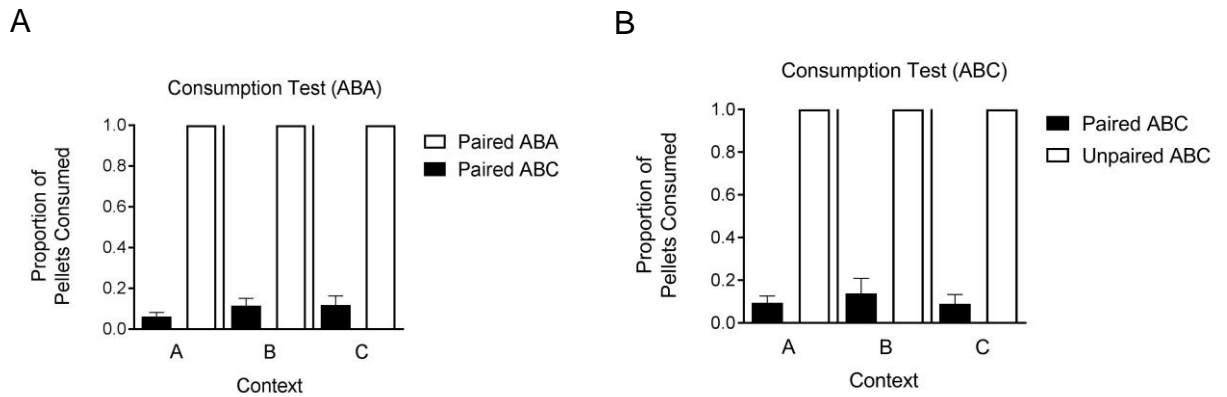
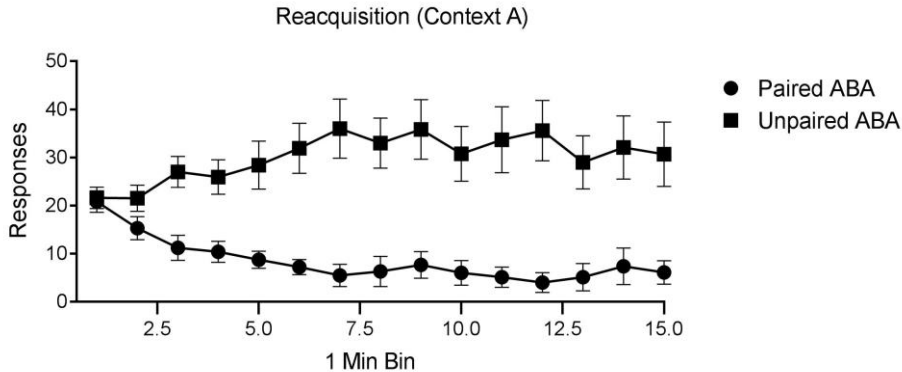
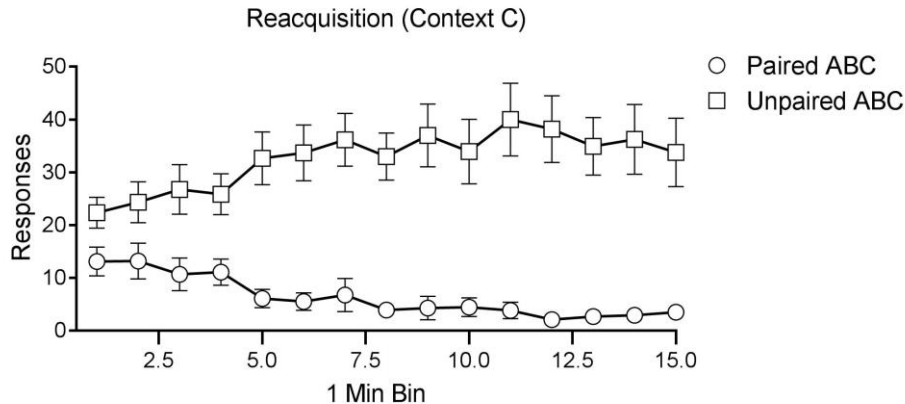
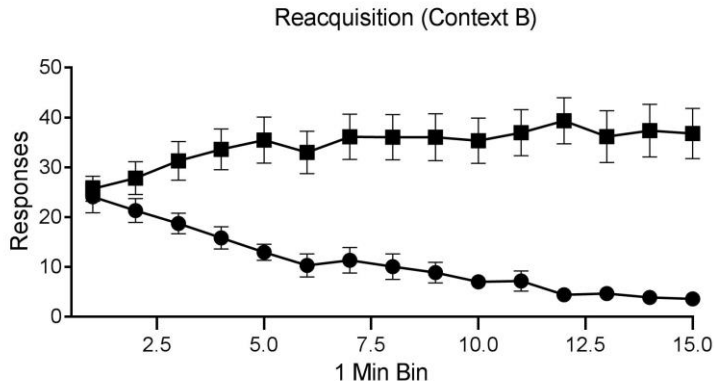


Figure 20: Experiment 4 reacquisition data. Mean number of responses by each group in (A) Context A, (B) Context B, and (C) Context C, depicted in 1-minute blocks. Error bars represent standard error of the mean (SEM).

A



B



## References

- Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, 34, 77–98.
- Atkins, M. S., & Stoff, D. M. (1993). Instrumental and hostile aggression in childhood disruptive behavior disorders. *Journal of Abnormal Child Psychology*, 21(2), 165-178.
- Balleine BW, Dickinson A. (1998a). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Balleine, B. & O’Doherty, J. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal directed and habitual action. *Neuropsychopharmacology*, 35, 48-69.
- Bouton, M. E. (1988). Context and ambiguity in the extinction of emotional learning: Implications for exposure therapy. *Behaviour research and therapy*, 26(2), 137-149.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological bulletin*, 114(1), 80.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological psychiatry*, 52(10), 976-986.
- Bouton, M. & Bolles. R. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, 10, 445-466.
- Bouton, M.E., King, D.A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes* 9, 248–265.
- Bouton, M.E., Peck, C.A. (1989). Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior* 17, 188–198.
- Bouton, M.E., Ricker, S.T., 1994. Renewal of extinguished responding in a second context. *Animal Learning & Behavior* 22, 317–324.
- 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.

- Bouton, M. & Todd, T. (2014). A fundamental role for context in instrumental learning and extinction. *Behavioural Processes*, 104; 13-19.
- Bouton, M.E., Todd, T.P., León, S.P. (2014). Contextual control of a discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition* 40, 92–105.
- Bouton, M., Todd, T., Vurbic, D., Winterbauer, N. (2011). Renewal after the extinction of free operant behavior. *Learning & Behavior*, 29, 57-67.
- Colwill, R. M., & Rescorla, R. A. (1985b). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 520-536.
- Colwill, R.M., Rescorla, R.A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning & Behavior* 18, 71–82.
- Corbit LH, Balleine BW. (2003). The role of prefrontal cortex in instrumental conditioning. *Behavior Brain Res* 146:145–157.
- Dickinson, A. (1994). Instrumental conditioning. In N. J. Mackintosh (Ed.), *Animal learning and cognition: Handbook of perception and cognition series* (2nd ed., pp. 45–79). San Diego, CA: Academic Press.
- Dickinson, A., Nicholas, D. J., & Adams, C. D. (1983). The effect of instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, 35, 35–51.
- Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning and Behavior*, 23, 197-206.
- Epstein, L. H., & Saelens, B. E. (2000). Behavioral economics of obesity: food intake and energy expenditure. *Reframing health behavior change with behavioral economics*, 293-311.
- Garcia, J., Lasiter, P., Bermudez-Rattoni, F., Deems, D. (1985). A general theory of aversion learning. *Annals of the New York Academy of Science*, 443, 8-21.
- Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A.,



- Robbins, T. W., & de Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry*, *168*(7), 718-726.
- Gremel, C. M., Chancey, J. H., Atwood, B. K., Luo, G., Neve, R., Ramakrishnan, C., Deisseroth, K., Lovinger, D.M., & Costa, R. M. (2016). Endocannabinoid modulation of orbitostriatal circuits gates habit formation. *Neuron*, *90*(6), 1312-1324.
- Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature communications*, *4*, 2264.
- Harris, J.A., Jones, M.L., Bailey, G.K., Westbrook, R.F. (2000). Contextual control of overconditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavior Processes* *26*, 174–185.
- Holland, P. C., & Coldwell, S. F. (1993). Transfer of inhibitory stimulus control in operant feature-negative discriminations. *Learning and Motivation*, *24*(4), 345-375.
- Killcross AS, Coutureau E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* *13*:400–408.
- Laborda, M.A., Witnauer, J.E., Miller, R.R. (2011). Contrasting AAC and ABC renewal: the role of context associations. *Learning & Behavior* *39*, 46–56.
- Lingawi, N. W., Dezfouli, A., & Balleine, B. W. (2016). Physiological Mechanisms of Habit Formation. *The Wiley Handbook on the Cognitive Neuroscience of Learning*, 411.
- Miles, F. J., Everitt, B. J., Dalley, J. W., & Dickinson, A. (2004). Conditioned activity and instrumental reinforcement following long-term oral consumption of cocaine by rats. *Behavioral Neuroscience*, *118*(6), 1331-1339.
- Morell, J. R., & Holland, P. C. (1993). Summation and transfer of negative occasion setting. *Animal Learning & Behavior*, *21*(2), 145-153.
- Nelson, J. B. (2002). Context specificity of excitation and inhibition in ambiguous stimuli. *Learning and Motivation*, *33*(2), 284-310.
- Pelloux, Y., Everitt, B. J., & Dickinson, A. (2007). Compulsive drug seeking by rats under punishment: effects of drug taking history. *Psychopharmacology*, *194*(1), 127-137.

- Rosas, J. M., Todd, T. P., & Bouton, M. E. (2013). Context change and associative learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(3), 237-244.
- Rey, C., Thrailkill, E.A., Bouton, M.E. (2018). Breaking a habit using a surprising reinforcer. Unpublished manuscript.
- Tamai, N., Nakajima, S. (2000). Renewal of formerly conditioned fear in rats after extensive extinction training. *International Journal of Comparative Psychology* 13, 137-146.
- Thomas, B.L., Larsen, N., Ayres, J.J.B. (2003). Role of context similarity in ABA, ABC, and AAB renewal paradigms: implications for theories of renewal and for treating human phobias. *Learning and Motivation* 34, 410–436.
- Todd, T.P., Winterbauer, N.E., Bouton, M.E. (2012a). Contextual control of appetite: renewal of inhibited food-seeking behavior in sated rats after extinction. *Appetite* 58, 484–489.
- Todd, T.P., Winterbauer, N.E., Bouton, M.E. (2012b). Effects of the amount of acquisition and contextual generalization on the renewal of instrumental behavior after extinction. *Learning & Behavior* 40, 145–157.
- Todd, T.P. (2013). Mechanisms of renewal after the extinction of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes* 39,193–207.
- Todd, T.P., Vurbic, D., Bouton, M.E. (2014a). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiology of Learning and Memory* 108, 52–64.
- Todd, T.P., Vurbic, D., Bouton, M.E. (2014b). Mechanisms of renewal after the extinction of discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition* 40(3), 355-368.
- Thrailkill, E. & Bouton, M. (2015). Contextual control of instrumental actions and habit. *Journal of Experimental Psychology: Animal Learning and Cognition*, 41, 69-80.
- Thrailkill, E. A., Trask, S., Vidal, P., Alcalá, J. A., & Bouton, M. E. (2018). Stimulus control of actions and habits: A role for reinforcer predictability and attention in the development of habitual behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44(4), 370.
- Trask, S., & Bouton, M. E. (2014). Contextual control of operant behavior: evidence for hierarchical associations in instrumental learning. *Learning & behavior*, 42(3), 281-288.

Thorndike, E. L. (1898). Animal intelligence: an experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2(4).

Thorndike, E. L. (1911). *Animal intelligence: experimental studies*. New York: Macmillan.

Yin HH, Knowlton BJ, Balleine BW. (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, 19:181–189.

Yin, H. H., Knowlton, B. J. & Balleine, B. W. (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavior and Brain Research*. 166, 189–196.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience* 22:513–523.

