Reinstatement interacts with interpolation interval to determine the effectiveness of counterconditioning

by

Elizabeth Jewel Powell

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 5, 2013

Keywords: Counterconditioning, interference, reinstatement, spontaneous recovery, differential retrieval

Approved by

Martha Escobar, Chair, Associate Professor of Psychology
M. Christopher Newland, Alumni Professor of Psychology
Francisco Arcediano, Assistant Professor of Psychology

Abstract

Extinction and counterconditioning have been used extensively as applied techniques that attenuate conditioned fear. Relapse through the form of renewal, reinstatement, and spontaneous recovery present challenges for these applications because they decrease the extent to which response attenuation is maintained over a wide range of circumstances and across long time intervals. Previous research has demonstrated that delaying extinction after acquisition of conditioned fear results in less relapse of fear at test than does conducting extinction immediately after acquisition. Although this effect has been repeatedly found in extinction, it has been relatively undocumented in counterconditioning. The present studies assessed this effect in counterconditioning and sought to determine whether differences in immediate and delayed counterconditioning are the result of different memory representations or differential acquisition of the interpolated association. This was assessed through the use of reinstatement procedures, intended to encourage retrieval of either the original or the interpolated phases of training.

Acknowledgments

I would like to express my gratitude to my advisor, Martha Escobar, for her support, patience, and encouragement during the duration of this project and throughout my graduate studies. Her insights and advice have been essential to the completion of this project and for the development of my graduate career.

I would also like to thank my committee members, Chris Newland and Pako Arcediano for their encouragement and advice on this project. I would further like to acknowledge my lab members, Justin Johnson, Caroline Bennett, Blake Arnsdorff, and Jessica Angel for their help with data collection.

I would also like to give special recognition to my colleagues, including Whitney Kimble, Jessica Busler, DJ Amegbletor, and Maggie Davis for their encouragement, support and commiseration during the data collection, writing, and waiting processes.

I'd like to thank my mother for encouraging me constantly through the years and for being so proud and supportive of my graduate career. Lastly, I am grateful for my amazing fiancé, Robby Dunaway, who listened to my thesis proposal and defense presentations so many times that he could have given them himself, and whose encouragement through this process made everything seem more possible to accomplish.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANOVA	Analysis of Variance
CC	Counterconditioning
CR	Conditioned Response
CS	Conditioned Stimulus
ITI	Intertrial Interval
MSE	Mean Squared Error
noUS	Nonpresentation of an Unconditioned Stimulus
US	Unconditioned Stimulus

Introduction

All organisms must learn about their environments in order to adapt and survive. More specifically, organisms must use signals in their environment to predict whether certain events will occur, such as whether food will be available or whether they are about to encounter danger. At the basis of such cue-outcome learning is a process known as Pavlovian conditioning (*cf.* Pavlov, 1927), in which a neutral stimulus is paired with a stimulus that naturally (i.e., unconditionally) elicits some type of measurable response. After repeated pairings with the unconditioned stimulus (US), presentation of the previously neutral stimulus (now called the conditioned stimulus, CS) alone results in a conditioned response (CR), which is generally similar to the original unconditioned response. A number of factors have been found to affect the course and onset of Pavlovian conditioning, including biological predispositions, stimulus salience, temporal factors, and contingency (for reviews, see Domjan, 2005; Rescorla, 1972).

Often in nature, relationships between stimuli change due to various circumstances, such as the passage of time or a change in environmental cues (context), and organisms must learn to adapt to respond to these new contingencies. In cases when incompatible associations are formed to the same CS or US, recall of one of these associations may be impaired. This process is commonly known as associative interference. For example, an organism may learn a relationship between two particular stimuli (e.g., A-B), and later find that the cue (A) leads to a different outcome (e.g., A-C; the so-called A-B, A-C paradigm). If, when presented with Stimulus A, the organism's responding is consistent with original learning (A-B), then proactive interference is

said to have taken place. By contrast, if responding is consistent with the interpolated phase of learning (A-C), then retroactive interference is said to have occurred.

Several Pavlovian conditioning paradigms appear to reflect A-B, A-C stimulus interference (for a review, see Bouton, 1993; Miller & Escobar, 2003). In such Pavlovian stimulus interference paradigms, A is usually a CS, B is usually a US (or US1), and C is a US distinct from B (US2). In some instances, B or C may be the absence of the US (for the purposes of the present paper, the nonpresentation of a US [noUS] will be considered equivalent to a second US). For example, in extinction, a CS is initially paired with a US, and later presented without the US (CS-US then CS-noUS; Pavlov, 1927). In latent inhibition, a CS is initially presented by itself and later paired with a US (CS-noUS then CS-US; Lubow & Moore, 1959). Finally, in counterconditioning, a CS is initially paired with a US (US1), and is later paired with a US that elicits a response of a different emotional valence (US2; Dearing & Dickinson, 1979; Sherrington, 1947). In all of these paradigms, the responses elicited by the original (A-B) association are incompatible with the responses elicited by the interpolated (A-C) association. Each of these paradigms can also lead to the development of proactive and retroactive interference. Specifically, proactive interference is observed early during the interpolated phase of training because acquisition of the second association (A-C) is retarded relative to a control condition in which the initial (A-B) training did not take place. Retroactive interference, however, is typically observed at the completion of interpolated (A-C) training because, with sufficient trials, responding reflects the interpolated (A-C) rather than the original (A-B) association.

In general, interference occurs whenever a given stimulus element has more than one associate, and each of these associates is encoded in a given physical, temporal, or spatial

context. Such contextual information should help disambiguate the situation and aid in retrieval of the cue-outcome relationship that is most consistent with the retrieval context. That is, the similarity between the conditions present at learning and the conditions present at retrieval determine which phase of learning is recalled (for a review, see Bouton, 1993).

Both physical and temporal contexts appear to play an important role in determining which phase of learning is recalled after completion of training (such assessment will hereafter referred to as 'testing'; Bouton, 1993; Bouton, 2004; Bouton & King, 1983; Bouton & Peck, 1989). Indeed, with identical training, renewal of the original association (i.e., proactive interference) may occur if the test context is more congruent with the original association than the interpolated association. Alternatively, if the test context is congruent with the interpolated association more than the original association, then the interpolated association will be renewed at test (i.e., retroactive interference; Archer, Sjoden, Nilsson, & Carter, 1979; Bouton, 2004). Take for example the so-called ABA renewal design, in which subjects learn the original association in Context A, undergo interpolated training in Context B, and then are tested in the context of original learning (Context A). In this design, subjects' responding tends to be consistent with original learning (proactive interference) because the test context evokes memories of original learning. However, subjects who receive similar training, but are then tested in Context B (ABB design), respond consistently with interpolated learning (retroactive interference).

Contextual cues are not only provided by environmental stimuli. Time can also serve as a contextual cue controlling behavior (Estes, 1955), and can play an important role in interference (Escobar, Arcediano, Platt, & Miller, 2004; Escobar & Miller, 2003). More specifically, the effects of retroactive interference appear to diminish with time, a phenomenon referred to as

spontaneous recovery (*cf*. Pavlov, 1927). Spontaneous recovery occurs after interpolated learning when, after the passage of time, behavior is again under the control of the original (A-B) association. Spontaneous recovery has been observed in extinction (Rescorla, 2004a), latent inhibition (Kraemer and Spear 1992) and counterconditioning (Bouton and Peck, 1992).

Selective reactivation of a particular training memory can also occur by having subjects re-experience a portion of that training experience. For example, unsignaled presentations of a US given after interpolated training can be used to remind the subject of the CS-US parings, thus reinstating the conditioned response (Bouton & Bolles, 1979; Rescorla & Heth, 1975). In the case of extinction, reinstatement of the US memory before testing would result in behavior consistent with original learning (i.e., proactive interference; Bouton & Bolles, 1979; Rescorla & Heth, 1975). Conversely, in the case of latent inhibition, reinstatement of the US memory before testing would result in behavior consistent with interpolated learning (i.e., retroactive interference; Kasprow, Catterson, Schachtman, & Miller, 1984). In counterconditioning, reinstatement of the US1 memory results in behavior consistent with original learning (e.g., Brooks, Hale, Nelson, & Bouton, 1995).

Extinction and counterconditioning have been used extensively in applied practice as techniques that attenuate maladaptive or undesirable behaviors (e.g., Craske & Mystkowski, 2006; Wolpe, 1968). Renewal, reinstatement, and spontaneous recovery present challenges for these applications because the intention is that response attenuation be maintained over a wide range of circumstances and across long time intervals. Various strategies have been suggested to increase retention of extinction and attenuate the return of the original response (Miller & Laborda, 2011). The timing of extinction in relation to acquisition appears to be an important determinant of the effectiveness of extinction. Various reports suggest that delaying extinction

results in less renewal (Huff, Hernandez, Blanding, & LaBar, 2009; Woods & Bouton, 2008, Experiment 1; but see Myers, Ressler, & Davis, 2006), and less spontaneous recovery (Chang & Maren, 2009; Huff et al., 2009; Johnson, Escobar, & Kimble, 2010; Maren & Chang, 2006; Rescorla, 2004b; Stafford, Maughan, Ilioi, & Lattal, 2013; Woods & Bouton, 2008, Experiments 2 & 4; but see Archbold, Bouton, & Nader, 2010; Johnson et al., 2010; Myers, et al., 2006; Norrholm et al. 2008; Schiller et al. 2008) than conducting extinction immediately after acquisition. For example, Maren and Chang (2006) observed less spontaneous recovery of conditioned fear (as measured by freezing) in subjects undergoing extinction 24 h after acquisition (delay condition) than in subjects who received extinction training 15 minutes after acquisition (immediate condition). Maren and Chang (2006) hypothesized that, because inducing fear prior to extinction has been shown to decrease the long-term effectiveness of extinction (Lovibond, Davis and O'Flaherty 2000; Morris, Furlong and Westbrook 2005), the increase in spontaneous recovery observed in the Immediate condition was related to the recent experience with shock. Consistent with this view, Maren and Chang observed reduced spontaneous recovery in the Immediate condition by decreasing the number of CS-US pairings and conducting extinction in a context different from the acquisition context (Experiment 4). Conversely, arousing fear by presenting the US in a novel context 15 min prior to extinction increased spontaneous recovery in the delayed extinction condition (Experiment 5).

For the present studies, an aversive-to-appetitive (i.e., CS-shock then CS-food) counterconditioning procedure was selected because it is structurally similar to extinction in that original learning (CS-shock trials) involves the acquisition of fear-motivated behavior, while interpolated learning (CS-food trials) involves a reduction or elimination of those fear responses. However, the process by which this fear attenuation is achieved is different because extinction

involves passive dissipation of fear, whereas counterconditioning involves the active acquisition of a new response. Thus, the two phases of counterconditioning training involve arousal, but this arousal has a different emotional valence and elicits different behavioral and physiological correlates. Consequently, it is unclear whether arousal evoked during the interpolated phase of counterconditioning will be equivalent to dissipation of fear in extinction. Counterconditioning provides the opportunity of reinstating original or interpolated learning by presenting a US associated with a distinct internal context (e.g. reinstating fear through the presentation of shock). Thus, reinstatement was selected as the memory retrieval procedure for these studies to determine whether the discrepancy in performance at test between Immediate and Delayed conditions can be reduced or eliminated. Such a manipulation is not possible in extinction or latent inhibition, in which the nonpresentation of the outcome as one of the phases of the interference procedure makes it difficult to use 'noUS' as a reinstating event.

The present studies had the following specific goals: (a) to determine whether the effects observed with immediate vs. delayed interpolation interval in extinction are generalizable to counterconditioning and, if so, (b) to determine whether these effects are due to differential acquisition of interpolated information in the immediate vs. delayed condition, or to differential retrieval of one vs. the other memory. This last goal can was assessed with a reinstatement procedure because in counterconditioning, the presentation of either US can differently determine which phase of learning should be retrieved. Based on the extinction data previously collected in our laboratory (Johnson et al., 2010), I hypothesized that subjects who experienced delayed counterconditioning would exhibit less spontaneous recovery of fear than subjects who underwent immediate counterconditioning. I also anticipated that subjects exposed to unsignaled presentations of US1 or US2 before test would respond (at test) consistently with the congruent

phase of training, regardless of whether they received immediate or delayed counterconditioning. In other words, I anticipated that reinstating memories of one of the phases of training should reduce any differential effects of the interpolation interval.

Experiment 1

Extinction appears to produce less spontaneous recovery when the interval between acquisition and extinction is long than when extinction immediately follows conditioning (Chang & Maren, 2009; Huff et al., 2009; Johnson et al., 2010; Maren & Chang, 2006; Rescorla, 2004b; Stafford et al., 2013; Woods & Bouton, 2008). That is, delayed extinction results in more retroactive interference than does immediate extinction. Although these effects have been repeatedly reported in extinction, there is little to no evidence that delaying interpolation will have a similar impact on other Pavlovian stimulus interference paradigms, such as counterconditioning. Therefore, Experiment 1 explored the effects of manipulating the interpolation interval on counterconditioning. An auditory CS was paired with a mild footshock in Phase 1; then, either 12 minutes (Condition Immediate) or 24 hours (Condition Delayed) later, this auditory CS was paired with an appetitive US (i.e., an aversive-appetitive counterconditioning procedure). If interpolation interval affects counterconditioning in a manner similar to extinction, more retroactive interference should be observed when the interpolated phase is delayed after original learning than when interpolated learning immediately follows original learning. Stated differently, delaying counterconditioning should result in less spontaneous recovery of conditioned fear than does immediate counterconditioning.

Methods

Subjects. The subjects were 32 male Sprague-Dawley rats (weight range: 319.2 – 406.5 g; Charles River Laboratories, Wilmington, MA), randomly assigned to one of four groups: Counterconditioning-Immediate (CC-Immediate), Control-Immediate, Counterconditioning-

Delayed (CC-Delayed), and Control-Delayed (*n*s = 8). All animals had served as subjects in a previous study, but were naïve to the modified experimental apparatus and stimuli used in the present study. Subjects were pair-housed in standard plastic cages with wire lids in a vivarium maintained on a 12 hr light/12 hr dark cycle (lights on at 6:00 am). All experimental manipulations occurred during the light portion of the cycle. All cagemates were assigned to different treatment groups. Access to food was gradually decreased to maintain subjects at approximately 85% of their free feeding weight over the week before initiation of the first experimental session. Subjects were fed approximately one hour after daily experimental sessions, or at the equivalent time during retention intervals (water was available *ad libitum*).

Apparatus. The apparatus consisted of eight Med Associates standard rat operant chambers, each measuring 30.5 x 24.1 x 21.0 cm (l x w x h). The side walls of the chamber were made of aluminum sheet metal, and the front and back walls and the ceiling of the chamber were made of clear polycarbonate. The floor was constructed of 4.8 mm stainless steel rods, spaced 1.6 cm center-to-center. These rods could be electrified with a scrambled footshock. All chambers were equipped with a speaker that could produce an 80-dB (A-scale) white noise). The USs consisted of delivery of a 1.0-s, 0.60-mA, footshock, or two 45-mg sucrose pellets into a cup located inside a niche (5.1 x 5.1 x 5.1 cm, l x w x h), which opened 1.5 cm above the grid floor. Infrared photobeams were used to detect head entries into this niche. Each chamber was housed in a melamine sound-attenuation cubicle. All sound attenuation cubicles were equipped with an exhaust fan, which provided a constant, 70 dB (A-scale), background noise.

Procedure. Table 1 presents a schematic of the experimental design for this and the subsequent studies. On Day 1, all subjects were trained to introduce their heads into the food cull to obtain the sucrose reinforcer, using an autoshaping procedure. Reinforcement was delivered

on a Fixed Time 4 min (FT-4') schedule in a 60-min session. Only subjects that met the criteria of producing at least 50 head entries throughout the session and consumed at least 20 of the 30 pellets delivered throughout the session proceeded to the acquisition phase of training.

On Day 2 (Delayed condition) or 3 (Immediate condition), subjects in the CC condition received three pairings of CS X and shock (US1), with an intertrial interval (ITI) of 10 min. Shock delivery coincided with CS X termination. Subjects in the Control condition received three explicitly unpaired presentations of each CS X and shock with a mean ITI of 5 min (CS presentations were scheduled to coincide with those of the CC condition). Subjects were allowed to remain in the chambers for 2 min after the last conditioning trial or stimulus presentation. Session duration was 25 minutes. Upon completion of the aversive conditioning session, all subjects were removed from the test chambers and returned to their home cages, where they spent their scheduled interval. All subjects were then returned to the test chambers to receive the counterconditioning trials. Twelve minutes (Immediate conditions) or 24 h (Delay Condition) after the last X-US trial, all subjects received 20 X-sucrose pairings (US2 = sucrose pellets), occurring with a mean ITI of 3(±1) min in a 70-min session.

All subjects spent a 48 hour retention interval in their home cages, and upon completion of that interval (i.e., on Day 5), they were returned to the test chambers and tested for head entry responses to CS X. During testing, all subjects received 4 presentations of CS X, which occurred at 5, 15, 25, and 35 minutes into the session.

Statistical Analyses. In this and all subsequent experiments, response ratios were calculated by comparing the number of head entries during the 30-s CS presentation against the number of head entries during a 90-s baseline period using the formula A/[A+(B÷3)], where A represents responding during the CS period and B represents responding during the baseline

period. Response ratios were analyzed in blocks of five trials during counterconditioning and blocks of two trials during testing. When using response ratios as a dependent measure, subjects must respond during the baseline period; otherwise, the fraction cannot be computed. Because fear levels were very high during the beginning of the counterconditioning sessions due to recent aversive training, several subjects' scores were incalculable during the first block of trials. This posed a problem for our repeated measures analyses, which cannot be calculated if one repeated factor is missing. Thus, for any subject missing one value, the missing value was replaced with the mean of the adjacent two scores; or, if the missing value was the first block of training, it was replaced with the value from the second block of training. For all analyses, statistical significance was established at $p < .05$.

Results and Discussion

Consistent with comparable studies in the extinction literature, immediate counterconditioning resulted in more spontaneous recovery of conditioned fear at test than delayed counterconditioning.

The development of counterconditioning was assessed using a 2 (condition: CC *vs.* control) x 2 (delay: immediate vs. delayed) x 5 (block of 4 trials) analysis of variance (ANOVA), which revealed a main effect of condition, $F(1, 28) = 24.12$, $MSE = 0.09$, as well as a main effect of block and a Condition x Block interaction, Fs(4, 112)= 23.11 and 5.65, $MSE = 0.02$. That is, the CS-shock pairings resulted in retarded acquisition of the CS-sucrose association in the CC condition; the same number of footshock presentations did not have this effect in the Control condition, in which the shock was not associated to the CS (see Figure 1). Thus, CS-shock presentations in Phase 1 proactively interfered with the acquisition or expression of the CS-sucrose association trained in Phase 2. Neither the main effect of delay, nor any interactions with

this factor were statistically significant (all Fs < 0.49), suggesting that counterconditioning progressed equivalently in both delay conditions. That is, whether or not interpolation was delayed did not have an effect on acquisition of the CS-sucrose contingency.

Spontaneous recovery of conditioned fear was assessed using a 2 (condition) x 2 (delay) x 2 (block of two trials) ANOVA, which revealed a main effect of condition and a Condition x Delay interaction, $F(1, 27) = 51.07$ and 4.80, respectively, $MSE = 0.05$, as well as a main effect of delay, $F(1, 27) = 4.83$, $MSE = 0.05$. These data suggest that delaying counterconditioning resulted in less spontaneous recovery than did providing counterconditioning immediately after fear conditioning. These results are consistent with previous extinction studies, in which delayed extinction resulted in less spontaneous recovery of conditioned fear at test than did immediate extinction (Chang & Maren, 2009; Huff et al., 2009; Johnson, Escobar, & Kimble, 2010; Maren & Chang, 2006; Rescorla, 2004b; Stafford et al., 2013; Woods & Bouton, 2008, Experiments 2 & 4; but see Archbold, Bouton, & Nader, 2010; Myers, et al., 2006; Norrholm et al. 2008; Schiller et al. 2008).

Because counterconditioning progressed equivalently in both the immediate and delayed conditions, it would seem that the differences in fear recovery at test between the immediate and delayed conditions was not due to differential acquisition of the interpolated association. Alternatively, it is likely that retrieval cues provided by a distinct temporal context in the Delayed conditions resulted in behavior consistent with the counterconditioning phase, while such cues were unavailable in the Immediate conditions, causing subjects in the Immediate conditions to revert to the first-learned association (See Bouton, 1993, Wheeler, Stout, & Miller, 2004; and General Discussion). It should be noted, however, that others have found differences in the rate of extinction in Immediate and Delayed extinction, with the Immediate condition

exhibiting faster extinction (faster acquisition of the interpolated association) than the Delayed condition (e.g., Woods & Bouton, 2008; but see Chang & Maren, 2009; Maren & Chang, 2006; Myers, et al., 2006).

Experiment 2

In Experiment 1, as in the extinction literature, we observed less spontaneous recovery of fear in the delayed interpolation (in this case, counterconditioning) groups than in the immediate interpolation groups. One possible explanation for this observation is that the temporal contexts of the original and interpolated phases are ambiguous in immediate training because of their close temporal proximity, while the temporal contexts of delayed training are more distinct (due to their relative temporal distance as compared with that of immediate training). A possible way to increase the distinction between the two phases of training would be to differentially prime one or another phase of the procedure. For example, a physical cue can serve as a 'reminder' of a specific phase of learning. The reinstatement procedure uses presentations of the US to provide such reminders for a given phase of conditioning. A problem with attempting reinstatement in the extinction paradigm is that the only phase of training that can be reinstated is the acquisition (CS-US) phase, not the extinction (CS-noUS) phase. The counterconditioning paradigm precludes this problem because US1 and US2 can be used to reinstate the original or interpolated phases of learning, respectively. In Experiment 2, animals were presented with US1 shortly before testing. We expected that unsignaled presentations of US1 would induce subjects to produce behavior consistent with original learning (i.e., conditioned fear), regardless of delay condition. Such findings would be consistent with a study by Brooks et al. (1995, Experiment 2) in which unsignaled presentations of shock after aversive-appetitive counterconditioning resulted in responding consistent with original learning (X-shock pairings). Based on the results of Experiment 1, subjects in the Immediate conditions were expected to exhibit fear due to the

effects of spontaneous recovery alone, while subjects in the Delayed conditions were not expected to exhibit as much fear due to spontaneous recovery. However, exposure to US1 shortly before testing should reinstate the memory of the Phase 1 aversive conditioning, as observed in the Brooks et al. study, resulting in increased conditioned fear across both conditions. In other words, reinstating aversive learning should result in increased fear in delayed counterconditioning, while little to no changes should be observed in the Immediate Counterconditioning group.

Methods

Subjects, apparatus, and procedure. The subjects were 32 male Sprague-Dawley rats (weight range: 226.9 - 381.4g; Charles River Laboratories, Wilmington, MA), randomly assigned to one of four groups: CC-Immediate, Control-Immediate, CC-Delay, and Control-Delay (*n*s = 8). All animals had served as subjects in a previous study, but were naïve to the modified experimental apparatus and stimuli used in the present study. Subjects were housed and maintained as described in Experiment 1. The experimental apparatus was the same used in Experiment 1.

The training procedures for Experiment 2 were identical to those described for Experiment 1, with the following exceptions. On the day following counterconditioning treatment (Day 4), all subjects received two presentations of US1 (0.6-mA, 0.5-s footshock) in a 14-min session. Delivery of the first footshock occurred 2 minutes into the session, followed 10 minutes later by a second footshock. Subjects were allowed to remain in the chambers for 2 minutes after delivery of the second footshock, and then they were removed and returned to their home cages.

On Day 5, all subjects were tested for head entry responses to CS X as described in Experiment 1, except that the first CS X presentation occurred 12 min rather than 5 min into the test session. Recent shock treatment may severely disrupt baseline behavior, but prior pilot research suggested that subjects would respond at the expected baseline levels approximately 10 min into the session. Thus, delaying the first CS presentation was expected result in a stable baseline (CS X presentations occurred at 12 and 22 minutes into the session).

Results and Discussion

Experiencing shock prior to testing resulted in reinstatement of fear in the Delayed condition. That is, experiencing the shock again attenuated the benefits of delayed conditioning. The development of counterconditioning was assessed using a A 2 (condition: CC *vs.* Control) x 2 (delay: immediate vs. delayed) x 5 (block of 5 trials) ANOVA, which revealed a main effect of condition, $F(1, 24) = 9.25$, $MSE = 0.17$, as well as a main effect of block and an interaction between condition and block, Fs(4, 96)=4.62 and 2.82, $MSE = 0.03$. That is, the CS-US1 pairings resulted in retarded acquisition of the CS-US2 association in the CC condition relative to the control condition (see Figure 2). Also, as conditioning progressed, responding increased across all groups, but this response increase was larger in the counterconditioning groups than the control groups. As in Experiment 1, counterconditioning progressed equivalently in both delay conditions (neither the main effect of delay nor the interactions between this and the other factors reached statistical significance, all Fs < 1.57). That is, whether or not interpolation was delayed did not have an effect on acquisition of appetitive conditioning.

The occurrence of reinstatement was assessed by comparing the last block of counterconditioning to the first test block. A 2 (condition) x 2 (delay) x 2 (test: pre- *vs*. post-reinstatement) ANOVA revealed a main effect of condition, $F(1, 18) = 8.75$, $MSE = 0.08$, as well

as a Test x Condition interaction, $F(1, 21) = 5.52$, $MSE = 0.06$. This reflects the observation that, following the shock US reinstatement presentations, both the immediate and delayed counterconditioning groups exhibited increased suppression, suggesting that two unsignaled presentations of shock were sufficient to reinstate the conditioned fear acquired in Phase 1.

To further assess reinstatement of conditioned fear, a 2 (condition) x 2 (delay) ANOVA was conducted on the data recorded during testing, which revealed main effects of condition and delay, Fs(1, 18) = 13.90 and 4.67, respectfully, $MSE = 0.07$, but no condition by delay interaction (F=.36), suggesting that, although reinstatement resulted in more fear in the immediate groups, this reduction affected the CC-Immediate and CC-Delay groups similarly. That is, reinstating shock cued memories of original learning resulted in behavior consistent with original learning in both counterconditioning groups (i.e., retroactive interference was reduced).

Experiment 3

Experiment 2 was designed to explore the effects of reinstating the original learning phase in a counterconditioning paradigm. Experiment 3 was designed as a counterpart of Experiment 2, but rather than reinstating original learning, presentations of US2 were used to reinstate the interpolated learning phase (counterconditioning training). This manipulation should make the test context more similar to that of interpolated learning (i.e., the counterconditioning phase). Consequently, reinstatement should cause all subjects to respond consistent with interpolated learning (all subjects should show evidence of retroactive interference).

Because the delayed conditions exhibit responding consistent with the counterconditioning phase even after a 48 h retention interval (see Experiment 1), reinstatement of the counterconditioning association should be observable only in the immediate condition. Therefore, Experiment 3 investigated reinstatement of interpolated learning in the immediate condition only. Counterconditioning was assessed against a control condition receiving no reinstatement US2 presentations.

Methods

Subjects and apparatus. The subjects used were 32 male Sprague-Dawley rats (weight range: 291.1 – 384.3 g; Charles River Laboratories, Wilmington, MA), randomly assigned to one of four groups, CC-Immediate-R+, Control-Immediate-R+, CC -Immediate-NoR, and Control-Immediate-NoR (*n*s = 8). All animals had served as subjects in a previous study, but were naïve to the modified experimental apparatus and stimuli used in the present study. Subjects were housed and maintained, and the apparatus were the same as described in Experiments 1 and 2.

Procedure. All groups received training identical to that described for the Immediate condition of Experiment 1, with the following exceptions. On the day following counterconditioning training (Day 3), subjects in the Reinstatement (R+) condition received 12 presentations of US2 (two 45-mg sucrose pellets) in a 23-min session, while subjects in the No Reinstatement control condition (NoR) received equivalent context exposure. Pellet delivery occurred 2, 6, 9, 11, 14, 18, and 21 min into the session. Subjects were allowed to remain in the chambers for two minutes after delivery of the last US2 presentation. On Day 5, all subjects were tested for responding to CS X, following the procedure described in Experiment 2.

Results and Discussion

Presentations of sucrose did not appear to reinstate the X-Sucrose association in subjects that had received immediate counterconditioning. That is, we observed no fear reduction in the R+ condition as compared to the NoR condition.

Counterconditioning training was assessed using a 2 (condition: CC vs. Control) x 2 (reinstatement: Reinstatement vs. NoR) x 2 (block of five trials) ANOVA, which revealed a main effect of condition, $F(1, 23) = 8.06$, $MSE = 0.20$, and a main effect of block, $F(4, 92) = 8.52$, $MSE = 0.05$, but no other main effects or interactions (all Fs < 2.29). As in Experiments 1 and 2, acquisition of the CS-US2 association was retarded in subjects who received CS-US1 relative to subjects in the control condition, in which the shock was not associated to the CS (see Figure 3). Since no difference in manipulation had yet occurred between the Reinstatement and No-R conditions, prior to the counterconditioning phase, counterconditioning progressed equivalently in both conditions, as expected.

Reinstatement of conditioned fear was assessed using a 2 (condition: CC vs. Control) x 2 (reinstatement: Reinstatement vs. NoR) ANOVA, which revealed a main effect of condition,

$F(1, 21) = 33.43$, $MSE = 1.94$, but no effect of reinstatement condition, (or other main effects or interactions; all Fs < 0.05). That is, additional presentations of sucrose were not sufficient to produce responding consistent with the CS-sucrose association. This suggests that, regardless of reinstatement condition, animals continued to exhibit responding consistent with the original association (i.e., fear). It is unclear why reinstatement was insufficient to produce responding at test consistent with the reinstated phase, unlike in Experiment 2, in which the CS-shock association was reinstated. Possible explanations for this effect will be addressed in the General Discussion.

General Discussion

Numerous reports have shown that delaying extinction results in less recovery of conditioned fear than conducting extinction immediately after acquisition (Chang & Maren, 2009; Huff et al., 2009; Johnson, Escobar, & Kimble, 2010; Maren & Chang, 2006; Rescorla, 2004b; Stafford et al., 2013; Woods & Bouton, 2008, Experiments 2 & 4; but see Archbold, Bouton, & Nader, 2010; Myers, et al., 2006; Norrholm et al. 2008; Schiller et al. 2008). Although this effect has been repeatedly found in the extinction paradigm, it has been unclear, until present, whether similar effects would be observed in the equally relevant counterconditioning paradigm. In Experiment 1, we observed that, consistent with the extinction literature, counterconditioning results in less spontaneous recovery when interpolated (X-Sucrose) learning is delayed after original (X-Shock) learning than when it immediately follows original learning.

We had initially proposed differential acquisition of the interpolated association as a possible explanation for the difference between Immediate and Delayed extinction and counterconditioning. However, we found no differences in the rate of counterconditioning between Immediate and Delayed conditions (but see Woods & Bouton, 2008), which would suggest that differences between Immediate and Delayed conditions at test are more likely due to differential retrieval of one phase of learning over the other.

We propose that the differences in recovery of conditioned fear in immediate and delayed counterconditioning are instead due to the presence of different retrieval cues at test. During immediate counterconditioning, the original and interpolated phases share very similar temporal

contexts (i.e. they are in close temporal proximity), which results in an ambiguous CS. Alternatively, delaying counterconditioning creates a distinct temporal context for original learning as well as interpolated learning, thus the CS becomes less ambiguous in a specific temporal context (For a review, see Bouton, 1993). Because the temporal context of testing is most similar to that of the interpolated phase in delayed counterconditioning (and extinction), Delayed subjects respond consistently with this phase. In the Immediate conditions, however, the temporal contexts become less informative, and each phase is equally likely to be retrieved, resulting in a greater net amount of conditioned fear.

In Experiments 2 and 3, we attempted to explicitly cause subjects to retrieve a particular phase of training in order perform consistently with that phase of training. In Experiment 2, fear was reinstated by delivering unsignaled presentations of shock on the day prior to testing. This treatment resulted in expression of fear in both the Immediate and Delayed conditions, thus suggesting that delayed counterconditioning results in the acquisition of learning experiences that can be independently cued for recall.

The counterpart of reinstating the original (aversive) association is to reinstate the interpolated (appetitive) association. In Experiment 3, we attempted to attenuate fear in the Immediate conditions by delivering unsignaled presentations of sucrose on the day prior to testing. However, interpolated learning was not successfully reinstated through unsignaled presentations of sucrose (see Figure 3B). We propose that there may be several possible explanations for the lack of reinstatement in Experiment 3.

In Experiment 2, it would seem that reinstatement of shock resulted in an internal state of fear, providing an internal context which was most similar to that of aversive conditioning. Because the internal context was most similar to aversive conditioning, memories of aversive

conditioning became more readily available than memories of counterconditioning, thus leading to responding consistent with aversive conditioning at test. Conversely, appetitive conditioning was not successfully reinstated in Experiment 3. This suggests that either sucrose was not a sufficient cue for retrieving memories of interpolated training, or that immediate counterconditioning produces a memory that is not independent of the original association, and thus cannot be independently cued for recall. Therefore, it remains unclear whether differential retrieval is sufficient to fully explain the differential effects of interpolation interval in Experiment 1.

It is possible that aversive conditioning has more biological relevance and is therefore more easily remembered than appetitive conditioning. Although reinstatement of sucrose should remind subjects of the appetitive association, it is possible that situations involving aversive associations are more salient than those involving appetitive associations. Such a differential weight for aversive and appetitive associations could result in a more distinguishable internal context when fear is evoked than when appetitive associations are involved. Additionally, it may be possible that the number of sucrose presentations during reinstatement sessions in these studies were not sufficient to provide equivalent levels of salience as the reinstatement of shock sessions. As evidenced by our design, approximately 20 appetitive trials were necessary to counter the effects of only three aversive trials. Future studies including additional appetitive reinstatement trials should prove beneficial in addressing this concern.

Lastly, it is possible that original learning is easier to reinstate than is interpolated learning. According to Nelson (2002), second-learned associations are more susceptible to contextual changes than are first-learned associations. If this is the case, then first-learned associations should be more likely to persist despite being exposed to internal conditions which

are dissimilar to those of original learning. Second-learned associations, alternatively, should dissipate after cuing of the original context is provided. The reinstatement in these experiments provided subjects with an internal context intended to evoke memories of either original or interpolated learning (Experiments 2 and 3, respectively). Based on Nelson's findings (2002), it is likely that reinstatement of shock (Experiment 2) was successful in producing behavior consistent with original learning, because, although the original association was less context specific, providing contextual information added to the tendency to respond consistent with original learning. Alternatively, because the original association was not strongly associated to a context, it persisted at test despite the provision of the interpolated internal context in Experiment 3.

To summarize, counterconditioning is similar to extinction in that it is less subject to fear relapse when interpolated learning is delayed rather than when it immediately follows conditioning. This attenuation in conditioned fear observed in delayed counterconditioning can be reduced, however, when original learning is reinstated prior to test (at least when original learning consists of aversive conditioning). In this study, we were unable to reduce the relapse of conditioned fear at test after immediate counterconditioning through reinstatement of interpolated learning. However, this may be due to the type of or saliency level of the US used in counterconditioning. In order to determine whether the effects of immediate counterconditioning could be enhanced, future studies should involve the use of appetitive-to-aversive counterconditioning with reinstatement of both phases of learning. This would shed light on concerns regarding differences in saliency, as well as issues regarding first-learned vs. second learned associations and context dependency. Most importantly, differential retrieval at test plays

an important role in the treatment of conditioned fear, in paradigms using extinction and counterconditioning, especially when treatment is delayed.

References

Archbold,G. E. B., Bouton, M. E. & Nader, K. (2010). Evidence for the persistence of contextual fear memories following immediate extinction. *European Journal of Neuroscience, 31,* 1303-1311.

Archer, T., Sjödén, P. O., & Nilsson, L. G. (1979). Role of exteroceptive background context in taste-aversion conditioning and extinction. *Learning & Behavior,7,* 17-22.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99.

Bouton, M. E. (2004). Context and Behavioral Processes in Extinction. *Learning & Memory, 11,* 485-494.

Bouton , M. E. & Bolles, R. C. (1979) Role of Conditioned Contextual Stimuli in Reinstatement of Extinguished Fear. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 368-378.

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. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior, 20,* 313-321.

Brooks, D. C. & Bouton, M. E. (1994). A Retrieval Cue for Extinction Attenuates Response Recovery (Renewal) Caused by a Return to the Conditioning Context. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 366-379.

Brooks, D.C., Hale, B., Nelson, J. B., & Bouton, M. E. (1995). Reinstatement after counterconditioning. *Animal Learning & Behavior, 23,* 383-390.

Chang C-H, Maren S. (2009). Early extinction after fear conditioning yields a context-independent and short-term suppression of conditioned freezing in rats. *Learning & Memory, 16,* 62-68.

Craske, M. G. & Mystkowski, J. L. (2006). Exposure Therapy and Extinction: Clinical Studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen, Debora (Eds.), Fear and learning: From basic processes to clinical implications. (pp. 217-233). Washington, DC, US: American Psychological Association.

Dearing, M. F. & Dickinson, A. (1979).Counterconditioning of shock by a water reinforcer in rabbits. *Animal Learning and Behavior, 7,* 360-366.

Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *The Annual Review of Psychology, 56,* 179–206.

Escobar, M. & Miller, R. R. (2003). Timing in retroactive interference. *Learning & Behavior, 31,* 257-272.

Escobar, M., Arcediano, F., Platt, T. L., & Miller, R. R. (2004). Interference and time: A brief review and an integration. *Reviews in the neurosciences, 15,* 415-438.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62,* 145-154.

Huff, N. C., Hernandez, J. A., Blanding, N. Q., & LaBar, K. S. (2009). Delayed Extinction Attenuates Conditioned Fear Renewal and Spontaneous Recovery in Humans. *Behavioral Neuroscience, 123,* 834–843.

Johnson, J. S., Escobar, M., & Kimble, W. L. (2010). Long-term maintenance of immediate or delayed extinction is determined by the extinction-test interval. *Learning & Memory, 17,* 639-644.

Kasprow, W. J., Catterson, D., Schachtman, T. R., & Miller, R. R. (1984). Attenuation of latent inhibition by post-acquisition reminder. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 36B,* 53-63.

Lubow, R.E. & Moore, A.U. (1959). Latent inhibition: The effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of comparative and physiological psychology, 52,* 415-419.

Maren, S., & Chang, C-H. (2006). Recent fear is resistant to extinction. *Proceedings of the National Academy of Sciences, 103,* 18020-18025.

Miller, R. R. & Barnet, R. C. (1993). The Role of Time in Elementary Associations. *Current Directions in Psychological Science, 2,* 106-111.

Miller, R. R. & Escobar, M. (2003). Timing in retroactive interference. *Animal Learning & Behavior, 31,* 257-272.

Miller, R. R. & Laborda, M. A. (2011). Preventing recovery from extinction and relapse: A product of current retrieval cues and memory strengths. Current Directions in Psychological Science, 20, 325-329.

Myers, K. M., Ressler, K. J., & Davis, M. (2006). Different mechanisms of fear extinction dependent on length of time since fear acquisition. *Learning & Memory, 13,* 216-223.

Nelson, J. B. (2002). Context Specificity of Excitation and Inhibition in Ambiguous Stimuli. *Learning and Motivation 33,* 284–310.

Norrholm, S. D., Vervliet, B., Jovanovic, T., Boshoven, W., Myers, K. M., Davis M., Rothbaum, B., & Duncan, E. J. (2008). Timing of extinction relative to acquisition: A parametric analysis of fear extinction in humans. *Behavioral Neuroscience, 122,* 1016-1030.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.

Rescorla, R. A. (1972) Informational Variables in Pavlovian Conditioning. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (pp. 1-46). New York: Academic Press.

Rescorla, R. A. (2004a). Spontaneous Recovery. *Learning & Memory,11,* 501-509.

Rescorla, R. A. (2004b). Spontaneous recovery varies inversely with the training-extinction interval. *Learning & Behavior, 32,* 401-408.

Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of Fear to an Extinguished Conditioned Stimulus. *Journal of Experimental Psychology: Animal Behavior Processes,* 104, 88-96.

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

Schiller, D., Cain, C. K., Curley, N. G., Schwartz, J. S., Stern, S. A., LeDoux, J. E., & Phelps, E. A. (2008). Evidence for recovery of fear following immediate extinction in rats and humans. *Learning & Memory, 15,* 394-402.

Sherrington, C.S. (1947). *The integrative action of the central nervous system.* Cambridge: Cambridge University Press.

Stafford, J. M., Maughan, D. K., Ilioi, E. C., & Lattal, K. M. (2013). Exposure to a fearful context during periods of memory plasticity impairs extinction via hyperactivation of frontal-amygdalar circuits. *Learning & Memory, 20,* 156-163.

Woods, A. M. & Bouton, M. E. (2008). Immediate extinction causes a less durable loss of performance than delayed extinction following either fear of appetitive conditioning. *Learning & Memory, 15,* 909-920.

Wolpe, J. (1958). Psychotherapy by Reciprocal Inhibition. Stanford: Stanford University.

Table 1

Experimental Designs

Group	Original training (A-B)	Interpolation delay	Interpolated training (A-C)	Reinstatement	Test
Experiment 1					
CC-Immediate	X-US1	12 min	X-US2	---	X
Control-Immediate	X / US1	12 min	X-US2	---	X
CC-Delayed	X-US1	24 h	X-US2	---	X
Control-Delayed	X / US1	24 h	X-US2	---	X
Experiment 2					
CC-Immediate	X-US1	12 min	X-US2	US1	X
Control-Immediate	X / US1	12 min	X-US2	US1	X
CC-Delayed	X-US1	24 h	X-US2	US1	X
Control-Delayed	X / US1	24 h	X-US2	US1	X
Experiment 3					
CC-Immediate-R+	X-US1	12 min	X-US2	US2	X
Control-Immediate-R+	X / US1	12 min	X-US2	US2	X
CC-Immediate-NoR	X-US1	12 min	X-US2	Ctxt Exposure	X
Control-Immediate-NoR	X / US1	12 min	X-US2	Ctxt Exposure	X

Figure 1. This figure represents mean response ratios collected during blocks of counterconditioning (Panel A) or spontaneous recovery test trials (Panel B). Lower ratios reflect more response suppression (more fear), which is consistent with original (aversive) learning, whereas higher ratios reflect less suppression (less fear) which is consistent with interpolated (appetitive/ counterconditioning) training.

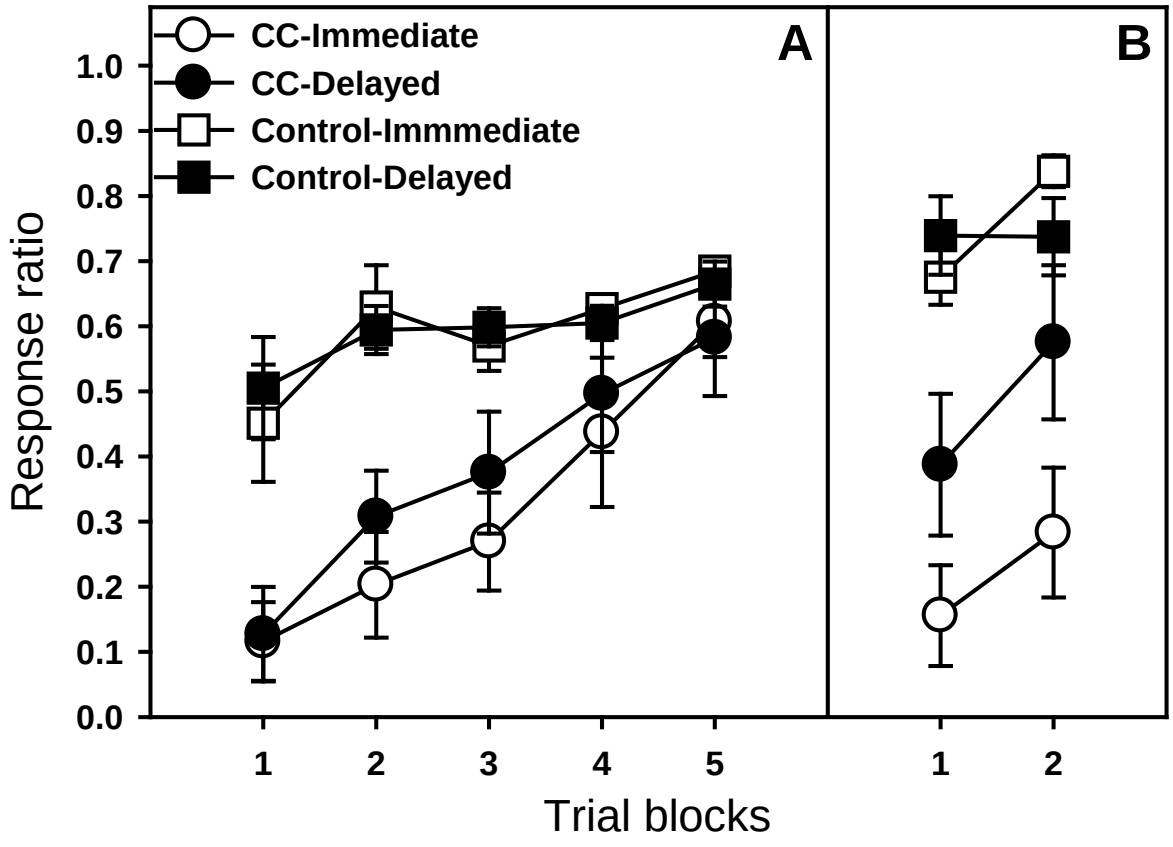

Figure 2. This figure represents mean response ratios collected during blocks of counterconditioning (Panel A) or test trials (Panel B; collected 24 h after reinstatement of shock). Lower ratios reflect more response suppression (more fear), which is consistent with original (aversive) learning, whereas higher ratios reflect less suppression (less fear) which is consistent with interpolated (appetitive/ counterconditioning) training.

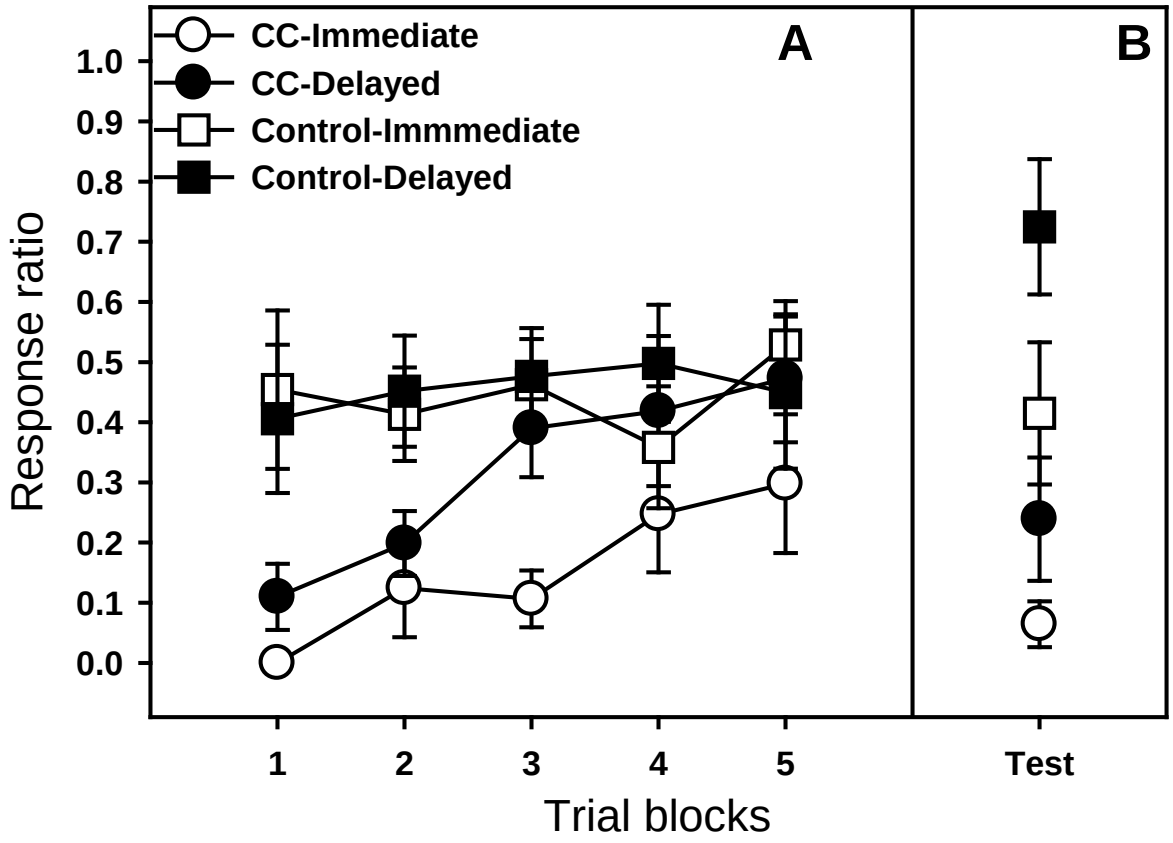

Figure 3. This figure represents mean response ratios collected during blocks of counterconditioning (Panel A) or test trials (Panel B; collected 24 h after reinstatement of sucrose). Lower ratios reflect more response suppression (more fear), which is consistent with original (aversive) learning, whereas higher ratios reflect less suppression (less fear) which is consistent with interpolated (appetitive/ counterconditioning) training.