Spacing extinction trials alleviates renewal and spontaneous recovery

Gonzalo P. Urcelay
University of Cambridge, Cambridge, England

Daniel S. Wheeler
Johns Hopkins University, Baltimore, Maryland

AND

Ralph R. Miller
State University of New York, Binghamton, New York

Studies of extinction in classical conditioning situations can reveal techniques that maximize the effectiveness of exposure-based behavior therapies. In three experiments, we investigated the effect of varying the intertrial interval during an extinction treatment in a fear-conditioning preparation with rats as subjects. In Experiment 1, we found less fear at test (i.e., more effective extinction) when extinction trials were widely spaced, relative to intermediate or massed extinction trials. In Experiment 2, we used an ABA renewal procedure and observed that spaced trials attenuated renewal of conditioned fear relative to massed trials. In Experiment 3, we used a similar design, but instead of changing the physical context at the time of testing, we interposed a retention interval after the extinction treatment to produce a change in the temporal context. The results showed less spontaneous recovery of fear after spaced than after massed extinction trials. These results suggest that extinction is more enduring when the extinction trials are spaced rather than massed. Although the benefits of spacing trials are small when there is no contextual change from extinction to testing, a change in either physical or temporal context following massed extinction trials leads to a recovery from extinction, which is reduced when the trials are spaced.

There is now considerable agreement that exposure-based therapy for anxiety disorders is highly analogous to behavioral extinction in a classical conditioning situation (Bouton, 2002; Craske et al., 2008; Delgado, Olsson, & Phelps, 2006; Hofmann, 2008; Lang, Craske, & Bjork, 1999). In exposure-based treatments, participants with anxiety disorders are repeatedly confronted with fear-inducing stimuli (e.g., a snake) that presumably have been associated with aversive outcomes (e.g., a snake bite). Treatment rests on the assumption that clients have associated various stimuli with the aversive outcome and is consequently designed to decrease the probability that when participants experience any of these fear-inducing stimuli, they will become anxious and thus avoid situations in which these stimuli could be encountered. In an experimental extinction study, an initially neutral cue (e.g., a tone) is initially associated with a biologically significant outcome (e.g., a mild shock) and after repeated pairings, subjects (usually nonhuman animals, but see Hermans et al., 2005, for one example with human participants) show a conditioned response indicative of fear when they are again presented with the cue. During the subsequent extinction treatment, when subjects repeatedly experience the cue alone (without a biologically signifi-

cant outcome), the fear response decreases until it is no longer observable.

There is one characteristic of both extinction and exposure therapy that strengthens the above-mentioned analogy. Both experimental extinction and exposure therapy fail to readily generalize to situations (environments or contexts) different from those in which the treatment was conducted. After experimental extinction and exposure therapy, changes in the context result in recovery from extinction (Bouton, 1993). In the case of exposure therapy, this is important because poor generalization results in relapse (Rachman, 1989). Thus, research (see the General Discussion section) has been conducted to identify variables that attenuate recovery from extinction as a means to gain information about these variables for potential applications to clinical settings. The central focus of the experiments reported here was to analyze the role of the intertrial interval (ITI) between extinction trials (i.e., the spacing of trials) upon recovery from extinction.

In the present series of experiments, we were specifically concerned with a variable that might enhance the effects of extinction treatment and perhaps alleviate renewal and spontaneous recovery: the spacing of extinction trials. From a theoretical perspective, this variable is of interest

---

R. R. Miller, rmiller@binghamton.edu

---

because according to different theories, it is postulated that the behavioral outcome should be different depending on whether extinction trials are spaced or massed. Some views of experimental extinction after Pavlovian conditioning suggest that the level of responding during the extinction treatment should correlate positively with the amount of extinction, thereby predicting that massed extinction trials should produce robust extinction (Rescorla, 2001). With massed extinction trials, the high level of fear from the immediately prior extinction trial (trial n) should summate with fear evoked during the subsequent trial (trial $n+1$). Hence, massed extinction trials should result in stronger responding during extinction and consequently better extinction. In trial-wise models of conditioning, more extinction is also anticipated with massed trials than with spaced trials, provided that total session length is adjusted to the spacing of trials (massed extinction trials in short sessions and spaced extinction trials in long sessions). For example, the Rescorla–Wagner (1972) model predicts that if the same number of extinction trials are implemented in sessions of different lengths, the shorter session (i.e., massed) should result in more extinction than would the longer session (i.e., spaced). The comparator hypothesis (Miller & Matzel, 1988) also anticipates this outcome, because less exposure to context alone (i.e., the massed extinction treatment) should make the context a better comparator stimulus for the extinguished target at the time of testing. In support of these predictions, Rescorla and Durlach (1987) extinguished two cues in two separate contexts with different session lengths and found that massed extinction was better than spaced extinction. However, it is important to note that their use of different session lengths created a confound of trial spacing and context exposure.

On the contrary, Bouton (1993) proposed that during extinction, subjects form a new, inhibitory association between the conditioned stimulus (CS) and the unconditioned stimulus (US). Thus, he anticipated in his proposal that the spacing of trials, which is known to have a strong impact on excitatory conditioning, should have a similar effect on inhibitory learning. In excitatory conditioning, massing training trials has a detrimental effect on the acquisition of behavioral control, which is known as the *trial massing effect* (Barela, 1999). Following Bouton (1993), one might expect that spacing extinction trials would result in greater decrements in conditioned responding than would massing extinction trials. This prediction has also received support in the literature. Westbrook, Smith, and Charnock (1985) observed, in a taste-aversion preparation, better extinction after two nonreinforced exposures when these were separated by 24 h, relative to separation by 30 min. In other words, they saw a benefit of conducting an extinction treatment with spaced trials, and, as opposed to Rescorla and Durlach (1987), they did not confound the total amount of exposure to the context, since the interval between extinction trials was spent outside the conditioning boxes. Therefore, it is possible that the advantage seen by Rescorla and Durlach with massed trials stemmed from the fact that their subjects spent less time in the extinction context with a massed extinction trial regimen.

We followed the lead of Westbrook et al. (1985) and equated context exposure while comparing massed and spaced extinction trials. Under these circumstances, it is possible to isolate the mechanisms responsible for these discrepant findings without appealing to differential exposure to the context as a contributor to the effect of the extinction treatment. Of interest for the present discussion is a recent report by Cain, Blouin, and Barad (2003) in which, using a fear-conditioning preparation with mice as subjects, they observed increased extinction and less recovery from extinction when they conducted massed extinction trials. One important feature of their report is that they kept the session length constant between massed and spaced extinction conditions, which, according to some of the above-mentioned models (Miller & Matzel, 1988; Rescorla & Wagner, 1972), should result in no difference between the groups, and according to others (Bouton, 1993), should result in less extinction. However, Cain et al. did not subject their extinction treatments to renewal tests, and in their Experiment 2, which tested for spontaneous recovery, they used only an 8-day retention interval. In that experiment, Cain et al. observed, 8 days after the extinction treatment, less fear (i.e., more extinction) after the massed trials than after the spaced trials. It is possible, though, that their retention interval was not sensitive to spontaneous recovery. In fact, both the massed and spaced groups froze 8 days after the extinction treatment at levels similar to those of the respective groups at the end of extinction. However, at the end of extinction, extinction treatment in the spaced condition appeared to be ineffectual, and this lack of extinction *during* the extinction treatment seemingly carried over the 8-day retention interval. Consequently, so-called *spontaneous recovery* here may have actually reflected a lack of extinction.

The present series of experiments was designed to investigate the consequences of conducting the extinction treatment with massed or spaced trials, but keeping the total session length constant, as Cain et al. (2003) did. We conducted renewal (Experiment 2) and spontaneous recovery (Experiment 3) tests, the latter with a 22-day retention interval instead of the 8 days employed by Cain et al. Whereas trialwise models of learning anticipate no difference between these two treatments, models like that proposed by Bouton (1993) anticipate better extinction with spaced trials. Of note, Bouton (1993) did not explicitly describe a mechanism to explain why the spacing of extinction trials should be beneficial, but the model clearly states that extinction is better captured as new learning. Consequently, the model implicitly predicts that spacing extinction trials should be better than massing extinction trials, even when the total amount of context exposure is kept constant. Because our intention was to replicate Cain et al.'s findings, we chose to conduct reinforced training in one context and the extinction treatment and testing in a different context. In other words, we used an ABB design (where A stands for the context of training and B for the context of extinction and testing).

EXPERIMENT 1

In Experiment 1, we used four groups of rats. The first group served as a spontaneous forgetting control and, therefore, did not experience any extinction trials during the extinction phase. The three remaining groups all experienced the extinction treatment. The difference between these three groups was the spacing of extinction trials, as measured by the ITI from CS termination to CS onset. One group experienced massed extinction trials (6-sec ITI), the second experienced extinction with an intermediate spacing (120-sec ITI), and the last group experienced spaced trials (600-sec ITI). Of note, all three groups experienced extinction (or exposure to the context, in the case of the control) in a single session of 220 min.

Method

Subjects

The subjects were 24 male (mean weight = 259 g) and 24 female (mean weight = 200 g) naive, young adult Sprague Dawley rats counterbalanced for sex within groups (ns = 12), bred in our colony at SUNY Binghamton. Subjects were individually housed and maintained on a 16:8-h light:dark cycle; experimental sessions occurred roughly midway through the light portion of the cycle. Subjects had free access to food in the home cage. One week prior to the initiation of the experiment, water availability was progressively reduced to 10 min/day, provided approximately 2 h after any scheduled treatment.

Apparatus

The apparatus consisted of 24 operant chambers, each measuring 30.5 × 27.5 × 27.3 cm (l × w × h). The chambers were counterbalanced within groups to the extent possible. All of the chambers had clear Plexiglas ceilings and side walls and metal front and back walls. On one metal wall of each chamber, there was an operant lever and, adjacent to it, a niche (4.5 × 4.0 × 4.5 cm) centered 3.3 cm above the apparatus floor, into which a drop (0.04 cc) of distilled water could be presented by a solenoid valve into a cup. The chamber floors were 4-mm bars in a grid, spaced 1.7 cm apart center to center, connected with NE-2 neon bulbs, which allowed constant-current footshocks to be delivered by means of a high-voltage AC circuit in series with a 1.0-MΩ resistor. Each chamber was housed in an environmental isolation chest, which was dimly illuminated by a houselight (No. 1820 incandescent bulb) mounted on the left wall of the experimental chamber. Ventilation fans in each enclosure provided a constant 76-dB (C-scale) background noise. Three 45-Ω speakers mounted on three different interior walls of each environmental chamber could deliver a tone (3000 Hz) at 8 dB above background (C scale, SPL), a 6-Hz click at 6 dB (C) above background, and a white noise at 8 dB above the ambient background sound. The click train was our target stimulus throughout the three experiments, and the tone was used in Experiment 2 (see the rationale below). The white noise was used in Experiments 2 and 3 as a discriminative signal to facilitate and maintain ongoing operant behavior for water.

Procedure

In this fear-conditioning preparation, we trained water-deprived rats to press a lever for water and measured decrements in leverpressing when the subjects were presented with the target stimulus as an index of how fearful the stimulus was. Therefore, we initially trained subjects to leverpress (shaping) in one context, which we called *Context B*. During subsequent Pavlovian fear conditioning (training), the subjects were placed in a different context, in which they did not have access to water when they pressed the lever. We will call this training location *Context A*. Then subjects were returned to the context in which they learned to leverpress and the extinction treatment was conducted. In this experiment, we also reshaped subjects to leverpress and tested subjects in the latter context; therefore, our design was an ABB design, in which the letters denote the contexts in which Pavlovian training (Context A), extinction (Context B), and testing (Context B) were conducted. Context identities were defined by the physical box in which a treatment was conducted, plus the presence of an odor cue or a distinctive floor, and the houselights. In other words, leverpress shaping (Days 1–5), reshaping (Days 7 and 9), extinction (Day 8), and testing (Day 10) were conducted in one context (B), whereas training (Day 6) was conducted in another context (A). The identity of Context A was accentuated by turning the houselight off and by the placement of a wooden block within the sound-attenuating environmental chest, outside the operant chamber, which contained two drops of mint (methyl salicylate). Conversely, Context B had the houselight on, no odor cue, and a Plexiglas surface over the grid floor. The physical identities of the contexts were further manipulated by using different instances of the apparatus for Contexts A and B.

Shaping. A 5-day period of acclimation to Context B with shaping of leverpress behavior was conducted in daily 60-min sessions. Subjects were shaped to leverpress for water on a variable-interval 20-sec schedule in the following manner. On Days 1 and 2, a fixed-time 120-sec schedule of noncontingent water delivery was in force concurrently with a continuous reinforcement schedule. On Day 3, noncontingent water delivery was discontinued, and subjects were trained on the continuous reinforcement schedule alone. Subjects that made less than 50 responses on this day were shaped through successive approximation later in the same day. On Days 4 and 5, a variable-interval 20-sec schedule was imposed. This schedule of reinforcement prevailed throughout the remainder of the experiment, except for Phase 1, during which no nominal stimulus was presented. Water presentation was always accompanied by the houselight's being turned off for 0.5 sec.

Phase 1. On Day 6, the lever was present, but no water was available. On this day of training in Context A, all subjects experienced four click→shock (CS→US) pairings. The first trial started 2 min after the subject was placed in the training box, and the last trial ended 2 min before the session ended. Trials were separated from each other by a 2-min interval (offset–onset), within the 14-min training session. The footshock started during the last 0.5 sec of the 60-sec CS. Therefore, the clicks and the footshock ended simultaneously.

Reshaping. On Day 7, all subjects experienced one 60-min session to restabilize leverpressing on the variable-interval 20-sec schedule in Context B.

Phase 2. On Day 8, subjects in the control condition were exposed to Context B for 220 min. Subjects in the spaced group experienced 20 exposures to the 60-sec CS without any presentation of the footshock. These presentations were separated from each other by an ITI of 600 sec from CS offset to CS onset. Subjects in the massed group also experienced the 20 nonreinforced presentations of the CS, but in a massed manner (6-sec ITI). Finally, subjects in the intermediate group experienced the 20 CS presentations separated from each other by a 120-sec ITI. Since the total time spent in the context was kept constant for all groups, whether subjects in the massed and intermediate groups experienced the trials either at the beginning or the end of the session could be a potential confound. Therefore, subjects in these two groups were subdivided into two subgroups according to the time of presentation of the nonreinforced trials (early vs. late; see Table 1). Subjects in the early subgroup experienced the first extinction trial 5 min after the session started and the last trial 194 min before the session ended. Subjects in the late subgroup received the first extinction trial 193 min after the session started and the last extinction trial 5 min before the 220-min session ended. Every 30 min, the sound-attenuating enclosures were opened to ensure that the animals were awake. Care was taken that no CS was on during these checks.

Table 1
Design of Experiment 1

Group	Phase 1 Training (Context A)	ITI	Phase 2 Extinction (Context B)	ITI	Test (Context B)
Control	4 X→US	120	context only		X
Massed	4 X→US	120	20 X	6	X
Intermediate	4 X→US	120	20 X	120	X
Spaced	4 X→US	120	20 X	600	X

Note—X was a clicker CS, and US was a footshock. Numbers next to pairings denote the numbers of trials. ITI, intertrial interval, in seconds, for acquisition and extinction trials. The massed and intermediate groups were subdivided into subgroups that experienced trials early or late in the extinction session, which was 220 min long.

Reshaping. On Day 9, all subjects experienced one 60-min session in Context B to restabilize leverpressing on the variable-interval 20-sec schedule.

Testing. On Day 10, in Context B, suppression of baseline leverpressing during presentation of the CS was assessed in all groups. Each subject received three nonreinforced 60-sec presentations of the CS during a 16-min session with the onsets occurring at 4, 8, and 12 min into the session. The response rate (number of leverpresses/min) during each 60-sec period preceding each CS exposure (pre-CS score) and the rate during each 60-sec CS exposure (CS score) was recorded.

Data Analysis

A suppression ratio (Annau & Kamin, 1961) of each subject was calculated by the formula $A/(A + B)$, where A is the mean rate of leverpressing during the three 60-sec CSs (pooled) and B is the mean rate of leverpressing during the three 60-sec pre-CS periods (also pooled). This ratio was used as an index of the subject's fear to the presentations of the target CS. Thus, a single ratio was calculated for each subject. Ratios typically range from 0–.5, with lower values representing more fear and higher values less fear. All experiments were analyzed with ANOVAs, unless otherwise stated. We report effect sizes calculated using the algorithm provided by Myers and Well (2003, p. 210).

Results and Discussion

As can be observed in Figure 1, control subjects that did not experience any of the extinction trials strongly suppressed responding in the presence of the target CS. However, among the subjects that experienced extinction, there were differences depending on the trial spacing during extinction. Specifically, as the trial spacing increased, the extinction treatment seemed to be more effective. These impressions were confirmed by the following analyses.

A one-way ANOVA with treatment (control, massed, intermediate, or spaced) as the independent variable comparing leverpresses during the minute before the first CS presentation indicated no reliable differences between groups [$F(3,44) = 0.63$, $MS_e = 68.86$, $p = .59$]. A similar ANOVA was conducted on the suppression ratios in order to determine whether there were differences among groups. This ANOVA proved significant [$F(3,44) = 15.69$, $MS_e = 0.012$, $p < .001$, Cohen's $f = 0.95$]. A planned comparison using the overall error term from the ANOVA demonstrated that the massed group suppressed less than did the control group [$F(1,44) = 17.5$, $p < .001$, Cohen's $f = 0.82$], thereby confirming that the extinction treatment with massed extinction trials was effective. A

second comparison showed that the spaced group suppressed even less than did the massed group [$F(1,44) = 4.49$, $p < .05$, Cohen's $f = 0.38$]. The intermediate group did not differ from either the massed or the spaced group, but, descriptively, suppression in the intermediate group was intermediate relative to the massed and spaced groups. Consistent with this result, a regression analysis indicated a direct linear trend in the groups that experienced the extinction treatment [$F(1,34) = 7.06$, $MS_e = 0.017$, $p < .05$, Cohen's $f = 0.41$]. To determine whether early or late extinction subgroups in the massed and intermediate groups differed, we conducted simple t tests, which indicated that within groups, there were no reliable differences between these subgroups [largest $t(10) = 1.39$, $p = .19$].

This experiment demonstrated that the trial spacing during extinction influenced the effectiveness of the extinction treatment, with greater spacing of trials resulting in more robust extinction. Moreover, we detected a linear relationship between the spacing of extinction trials and the amount of extinction observed during testing. There is, however, an alternative explanation that deserves discussion. For example, one could argue that a change in the ITI from training to extinction might have reduced extinction in the massed group, because this group received train-
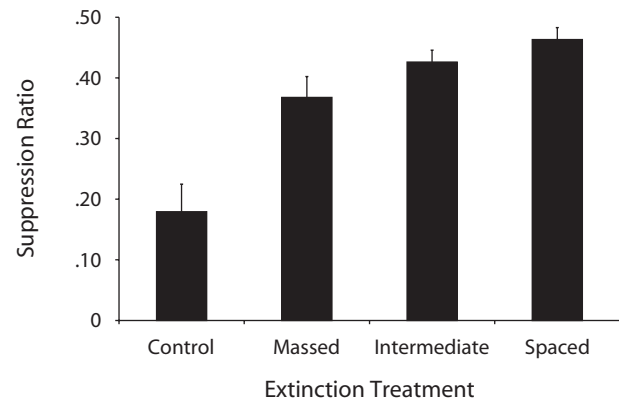

Figure 1. Mean suppression ratios to the clicker at test in Experiment 1. The control group did not receive any extinction trials, and the massed, intermediate, and spaced groups received 20 extinction trials, separated by 6, 120, and 600 sec, respectively. Lower values denote more suppression, and larger values denote less suppression (i.e., more extinction).

ing with a 120-sec ITI and extinction with a 6-sec ITI. Although a generalization decrement due to this change in ITI is a possibility, it cannot explain the inverse relationship between trial spacing during extinction and suppression at test. A generalization decrement account would anticipate strong suppression in the intermediate group and weaker suppression in the massed and spaced groups, which is inconsistent with the linear trend observed here. A similar argument could be raised when looking at the transfer from extinction to test. In this experiment and the remaining ones, we tested subjects by presenting the target CS three times and assessing conditioned suppression on the basis of the average of the three test trials, which gives a better and more stable estimate of each subject's suppression to the CS than would a single test trial. However, the ITI during testing (180 sec) was different from those during training and extinction. To assess this possibility, we conducted a similar analysis, but only on the first trial suppression ratios, and the pattern was similar to the one presented in Figure 1. In summary, our pattern of data does not seem to be the result of these generalization effects but, rather, a true effect of the ITI during extinction.

In our introduction, we noted the similarities between experimental extinction and exposure therapy. We also noted that recovery has been observed from both experimental extinction and exposure therapy, following manipulations such as a change in the physical context from extinction to test (ABA or ABC renewal) and a change in the temporal context achieved by interposing a retention interval between extinction and test (spontaneous recovery). In emphasizing the similarities between experimental extinction and therapy, we also noted that these two forms of recovery are the ones seemingly most closely related to relapse in humans. Relapse in humans after successful treatment can occur when patients encounter a fear-eliciting stimulus outside the therapist's office or when an interval occurs between the cessation of treatment and an encounter with a fear-eliciting stimulus. In Experiment 1, we used an ABB design, in which extinction was assessed in the same physical context in which it happened. If one were to oversimplify the therapist's situation, our Experiment 1 was perhaps more akin to an assessment in the therapist's office. The next two experiments were conducted to replicate these findings and to assess whether the spacing of trials during extinction can alleviate (or increase) the recovery from extinction seen after changes in physical (ABA renewal, Experiment 2) or temporal (spontaneous recovery, Experiment 3) contexts.

EXPERIMENT 2

In Experiment 2, we wanted to determine whether extinction with massed or spaced trials would alleviate recovery from extinction when the test was conducted in the training context (ABA) rather than in the extinction context (ABB). Because the intermediate group in Experiment 1 did not differ from either the massed group or the spaced group, in Experiment 2 we used the same parameters as in Experiment 1, but we did not include groups that experi-

enced extinction with intermediate spacing. Six groups of rats experienced Pavlovian fear conditioning. During extinction, one group served as forgetting controls, another group experienced massed extinction trials, and a third group experienced spaced extinction trials. These three groups were subsequently tested in the extinction context, and thus served both as renewal controls and as a potential replication of the findings from Experiment 1. Three additional groups experienced similar treatments but were tested in the context in which training occurred (ABA design). These were the groups that could potentially show recovery from extinction due to renewal. In other words, we used a 3×2 factorial design in which the treatment during extinction (control, massed, or spaced) and the test context (ABB or ABA) were the main factors. One aspect of renewal that has been debated is whether ABA renewal is to some degree the result of fear of the training context summating with fear of the CS at test or whether this effect is completely due to testing being conducted outside of the extinction context. That is, when we compare an ABB group with an ABA group, we are assessing extinction in the ABA group outside the extinction context, but we are doing so by comparing performance in a context in which reinforced training was experienced (Context A) with performance in a context in which subjects never experienced reinforcement (Context B). To control for the latter variable, we used two distinctive contexts as in Experiment 1, but in addition to training the target CS in Context A, we trained a second CS in Context B in order to equate exposure to the signaled reinforcer in both contexts, thereby obviating the context summation interpretation of the ABA renewal effect. We also equated the ABA and ABB groups in their total exposure to both Contexts A and B throughout the experiment, because in prior research, differential context exposure has been found to confound renewal effects (Lovibond, Preston, & Mackintosh, 1984).

Method

Subjects

The subjects were 36 male (mean weight = 300 g) and 36 female (mean weight = 214 g) naive, young adult Sprague Dawley rats, bred in our colony at SUNY Binghamton. They were randomly assigned to one of the six groups (ns = 12), counterbalanced for sex within groups. The subjects were housed, water deprived, and maintained in the same way as the subjects in Experiment 1.

Apparatus and Stimuli

Since the extinction phase was 220 min long, we could only run two squads of 12 subjects in each room on a given day. Therefore, the study was conducted in three different experimental rooms with twelve chambers in each room. Twenty-four subjects were run in each room, counterbalanced across groups. All three rooms were identical to that used in Experiment 1.

For any given subject, two contexts (A and B) were used in this experiment. The identities of the two contexts were defined in the same way as those in Experiment 1, with the exception that the acrylic floor plates in one of the contexts were replaced with a second odor cue (a banana scent), due to the fact that subjects received footshocks in both contexts (the floor plates would impede the rats from receiving shocks). Because one of our critical manipulations involved comparing subjects tested in one context with subjects tested in a different context, the two contexts were counterbalanced within

Table 2
Design of Experiment 2

Group	Phase 1 (Contexts A & B)	Phase 2 (Contexts A & B)	Test
Control-ABB	4 X→US(A) 4 Y→US(B)	Context (A), Context (B)	X (B)
Massed-ABB	4 X→US(A) 4 Y→US(B)	20 X massed (B), Context (A)	X (B)
Spaced-ABB	4 X→US(A) 4 Y→US(B)	20 X spaced (B), Context (A)	X (B)
Control-ABA	4 X→US(A) 4 Y→US(B)	Context (A), Context (B)	X (A)
Massed-ABA	4 X→US(A) 4 Y→US(B)	20 X massed (B), Context (A)	X (A)
Spaced-ABA	4 X→US(A) 4 Y→US(B)	20 X spaced (B), Context (A)	X (A)

Note—X and Y were clicker and tone CSs, counterbalanced, and US was a footshock. Numbers next to pairings denote the number of trials. Letters inside parentheses denote the context (A or B) in which acquisition, extinction, or test took place. The physical identities (see the text) of Contexts A and B were counterbalanced. Massed groups were subdivided into subgroups that experienced trials early or late in the extinction session, which was 220 min long.

groups, so that for half of the subjects, Context A was identified by the mint scent, whereas for the other half, Context A was identified by the banana scent. CS X was always trained in Context A, whereas CS Y was always trained in Context B.

Procedure

The design of Experiment 2 is summarized in Table 2. In order to equate subjects for exposure to both contexts and to obtain good leverpressing in both contexts during testing, all subjects experienced shaping (Days 1–5) and reshaping (Day 9) session in both contexts (A and B). Therefore, during each daily shaping and reshaping session, subjects experienced two 30-min sessions, one in each context, separated by 60 min. The order was reversed each day.

Shaping. Prior to Phase 1 of training, a 5-day period of acclimation to the apparatus and shaping of leverpress behavior was conducted in two daily 30-min sessions, one in each of the two different contexts. Subjects were shaped to leverpress for water on a variable-interval 20-sec schedule in the same way as were those in Experiment 1, with the differences being that the session length was shortened to 30 min and that subjects experienced two sessions per day, one in each context. Water presentation was always accompanied by 0.5 sec of the white noise.

Phase 1. On Day 6, the lever was present, but water was not available. On this day, all subjects experienced eight CS→US pairings. Four X→US trials were experienced in Context A, and four Y→US trials were experienced in Context B. In each of the two training sessions, the first trial started 2 min after the subject was placed in the training box, and the last trial ended 2 min before the session ended. Trials were separated from each other by a 2-min interval (offset–onset), within a 14-min training session. The footshock was presented during the last 0.5 sec of the 60-sec CS. Therefore, the stimulus (the clicks or the tone) and the footshock terminated simultaneously. The two sessions were separated by 60 min, and the order of sessions was counterbalanced, so that for half of the subjects in each group, X was trained first and Y afterward, whereas for the remaining half, this was reversed.

Phase 2. On Days 7 and 8, subjects in the control condition were exposed to each context for 220 min. That is, on Day 7, half of the subjects in each group were exposed to Context A, and the other half were exposed to Context B. This was reversed on Day 8. All subjects in the extinction groups experienced the extinction treatment in Context B, either on Day 7 or on Day 8. On the remaining day, they

were exposed to Context A. All other procedural details were similar to those in the extinction session (Phase 2) of Experiment 1.

Reshaping. On Day 9, all subjects experienced two half-hour sessions to restabilize leverpressing on the variable-interval 20-sec schedule. One session was conducted in Context A and the other in Context B, with the order counterbalanced within groups.

Testing. On Day 10, suppression of baseline responding during presentation of the CS was assessed in all groups. Each subject received three nonreinforced 60-sec presentations of the CS during a 16-min session with the onset occurring 4, 8, and 12 min into the session. Subjects assigned to the ABA condition were tested in Context A, the training context, whereas subjects assigned to the ABB condition received testing in Context B, the extinction context.

Data Analysis

A suppression ratio was calculated in the same way as in Experiment 1. The ratios were compared using a 2 (test condition: ABB
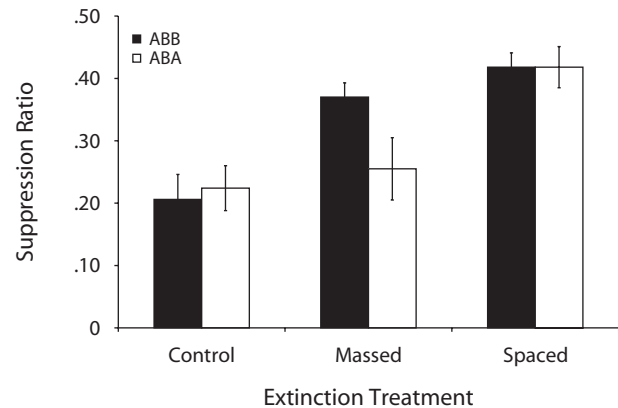

Figure 2. Mean suppression ratios to the target CS in the testing phase of Experiment 2. The control groups did not receive any extinction trials, and the massed and spaced groups received 20 extinction trials. Subjects were tested in the extinction context (ABB design, black bars) or in the acquisition context (ABA design, white bars). Lower values denote more suppression, and larger values denote less suppression (i.e., more extinction).

or ABA) × 3 (extinction condition: control, massed, or spaced) ANOVA. One subject in the massed-ABA group was excluded from all analyses because the rat did not leverpress during the test session.

Results and Discussion

Figure 2 illustrates the results of Experiment 2. Groups that were tested in the context in which extinction took place (ABB) displayed a pattern similar to that observed in Experiment 1. Specifically, the control group showed substantial suppression. Suppression in the massed group was weaker (as depicted by a higher bar in the figure) than in the control group, which is indicative of extinction. The spaced group showed less suppression than did the massed group, which is indicative of more effective extinction, although that difference was not reliable. Regardless, this pattern is similar to that observed in Experiment 1. Groups tested in the training context (ABA) showed a different pattern. The massed group did not differ appreciably from the control group, suggesting a renewal effect when extinction was conducted with massed trials. That is, when subjects in the massed condition were tested in the training context, they suppressed more than when tested in the extinction context (ABB), which clearly indicates the presence of a renewal effect. Interestingly, renewal was not apparent for subjects that received extinction with spaced trials. The spaced groups did not seem to differ appreciably, regardless of the context in which testing occurred. These impressions were quantified by the following statistical analyses.

A 2 × 3 ANOVA with test condition (ABB or ABA) and extinction (control, massed, or spaced) as factors conducted on the basal leverpresses (60 sec before presentation of the first CS) indicated no differences between groups [largest $F(2,65) = 1.09, p > .34$]. A similar ANOVA conducted on the suppression ratios revealed a main effect of extinction [$F(2,65) = 16.27, MS_e = 0.015, p < .001$], no main effect of test [$F(1,65) = 1.21$], and a nonsignificant interaction [$F(2,65) = 2.01, MS_e = 0.015, p = .14$, Cohen's $f = 0.16$]. Despite the nonsignificant interaction, we conducted planned comparisons using the overall error term of the 2 × 3 ANOVA to determine whether differences among groups were evident in the data. Comparisons of subjects tested in the extinction context (ABB) revealed less suppression with massed extinction relative to the forgetting controls [$F(1,65) = 10.66, p < .01$, Cohen's $f = 0.63$], suggesting that in the massed group, extinction was evident when subjects were tested in the extinction context. This result replicates that of Experiment 1. However, subjects in the spaced condition did not differ reliably from their massed counterparts [$F(1,65) = 0.89, p = .34$]. Although this difference was not significant in this experiment, it was significant in Experiment 1, and the ordinal direction of the difference in means in Experiment 2 was in the same direction as in Experiment 1: More extinction was apparent after spaced trials than after massed trials. Analysis of the groups tested in the training context (ABA) revealed a different pattern. The comparison between the massed-ABA and massed-ABB groups was reliable [$F(1,65) = 5.02$,

$p < .05$, Cohen's $f = 0.40$]. Subjects tested in the training context (ABA) suppressed more than did those tested in the extinction context (ABB), which shows renewal in the massed condition. No such renewal effect was apparent when the spaced-ABA and spaced-ABB groups were compared [$F(1,65) = 0.00, p = .99$]. To make sure that early rather than late extinction in the massed subgroups did not affect the test data, we conducted two t tests. The results of the t test between early and late massed subgroups tested in the extinction context (ABB) were not significant [$t(10) = 0.21, p = .83$]; neither were those with subjects tested in the training context (ABA) [$t(9) = 1.47, p = .17$]. Our results do not seem to be attributable to any difference in the amount of time between the last extinction trial and test.

Overall, these comparisons suggest that when groups were tested in the extinction context (ABB), a pattern emerged similar to that observed in Experiment 1. When testing was conducted in the training context (ABA), subjects that experienced massed extinction trials showed a recovery from extinction—a renewal effect. No such recovery was apparent in subjects that experienced highly spaced extinction trials. Thus, on the basis of these results, one might conclude that, at least with the present parameters, spacing trials during the extinction treatment alleviated the renewal effect. But another way of viewing these results is that when testing is conducted in the extinction context, a small (nonsignificant in this experiment, but significant in Experiment 1) benefit of spacing extinction trials is observed. When one returns to the training context for testing, this small difference becomes robust. This result is consistent with the view that the effect of trial spacing during extinction becomes more evident when one tests for extinction under conditions that produce recovery from extinction, such as renewal. We will revisit this point in the General Discussion section.

EXPERIMENT 3

Experiment 2 demonstrated that, at least with the parameters used here, renewal can be alleviated if trials are widely spaced. Experiment 3 was intended to test whether spaced extinction trials also alleviate another form of recovery from extinction—namely, spontaneous recovery. We anticipated that, if the results observed in Experiment 2 are reliable, then we should see a small benefit of spacing over massing trials when testing is conducted immediately after the extinction treatment but a larger benefit of spaced trials when testing is delayed, a procedure that ordinarily results in spontaneous recovery from extinction.

In Experiment 3, we used a design similar to that of Experiment 1; that is, training was conducted in Context A, whereas extinction and testing were conducted in Context B, constituting an ABB design. The extinction phase included forgetting controls (no extinction), massed extinction, and spaced extinction. In contrast to Experiment 2, in which we manipulated spatial contexts, the critical manipulation in Experiment 3 was the interval between extinction and test. Bouton (1993) proposed that the passage of time can be viewed as a change in the temporal

Table 3
Design of Experiment 3

Group	Phase 1 (Context A)	Phase 2 (Context B)	Retention Interval (days)	Test (Context B)
Control-short	4 X→US	Context	2	X
Massed-short	4 X→US	20 X massed	2	X
Spaced-short	4 X→US	20 X spaced	2	X
Control-long	4 X→US	Context	22	X
Massed-long	4 X→US	20 X massed	22	X
Spaced-long	4 X→US	20 X spaced	22	X

Note—X was a clicker CS, and US was a footshock. Numbers next to pairings denote the number of trials. Massed groups were subdivided into subgroups that experienced trials early or late in the extinction session, which was 220 min long. The retention interval was either 2 or 22 days and was spent in the home cages.

context, which should impact extinction in the same way as a change in the spatial context. Subjects in each of the three aforementioned extinction conditions were tested either 2 or 22 days after the extinction treatment. The study was implemented with a 3 × 2 design, in which the first factor was the nature of extinction (control, massed, or spaced), and the second factor was the retention interval interposed between extinction and testing (2 days [short] or 22 days [long]). The design is summarized in Table 3. On the basis of analogy with Experiment 2, we expected to observe more recovery from extinction after 22 days in subjects extinguished with massed trials than in subjects extinguished with spaced trials. In other words, we expected the extinction treatment to be more enduring in the spaced condition than in the massed condition.

Method

Subjects

The subjects were 36 male (mean weight = 257 g) and 36 female (mean weight = 198 g) naive, young adult Sprague Dawley rats, counterbalanced for sex within groups (ns = 12), bred in our colony at SUNY Binghamton. Subjects were housed, water deprived, and maintained in the same way as the subjects in Experiments 1 and 2.

Apparatus and Stimuli

The apparatus and stimuli were the same as those used in Experiment 2. The only difference was that, as in Experiment 1, we used a Plexiglas surface over the grid floor, instead of an additional odor cue, to help differentiate the two contexts. In other words, we used the three experimental rooms of Experiment 2, but the context identities were determined exactly as in Experiment 1, and we used the ABB design.

Procedure

Shaping (Days 1–5), reshaping (Days 8 and 29 for the short and long groups, respectively), extinction, and testing were conducted in one context (Context B), whereas training was conducted in another context (Context A).

Shaping. Acclimation was conducted on Days 1–5 in the same manner as in Experiment 1. However, water presentation was always accompanied by 0.5 sec of the white noise, instead of the houselight being turned off.

Phase 1. This phase was conducted on Day 6 exactly as in Experiment 1. In Experiment 2, we failed to obtain a significant difference between the massed and spaced groups when the test was conducted in the extinction context (ABB). Because fear extinction has been observed to correlate with the amount of responding to the fear eliciting stimuli (Rescorla, 2001), we reasoned that increasing overall

fear might increase sensitivity to detect differences between these groups. Therefore, we increased the shock level from the 0.5 mA of Experiment 1 to 0.7 mA.

Phase 2. On Day 7, the subjects in the control, massed, and spaced conditions received extinction treatments similar to those in the respective conditions in Experiments 1 and 2.

Short reshaping. This single 60-min session on Day 8 was similar to that in Experiment 1, and it was intended to restabilize leverpressing behavior. Only subjects in the short condition received this treatment.

Short testing. On Day 9, suppression during presentations of the CS relative to baseline responding was assessed in the short groups. This was done in the same way as in the previous two experiments. Testing was conducted in Context B.

Retention interval. During the retention interval (Days 8–28), subjects in the long groups remained in their home cages and were handled for 30 sec three times a week.

Long reshaping. On Day 29, subjects in the long condition experienced one 60-min session to restabilize leverpressing on the variable-interval 20-sec schedule.

Long testing. On Day 30, suppression during presentations of the CS relative to baseline responding was assessed in the long groups.
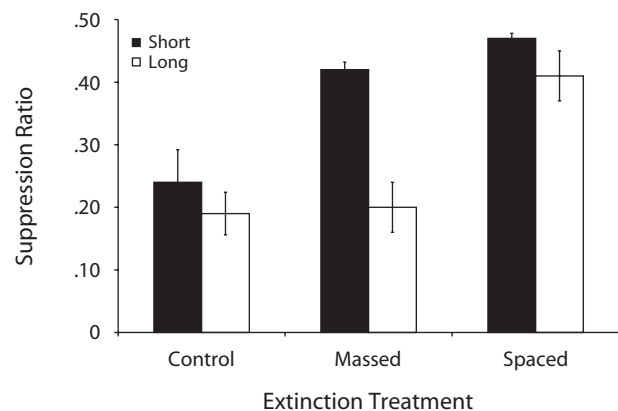

Figure 3. Mean suppression ratios to the target CS (clicker) in the testing phase of Experiment 3. The control groups did not receive any extinction trials, and the massed and spaced groups received 20 extinction trials. Subjects were tested in the extinction context (ABB design) either after a short retention interval (2 days, black bars) or a long retention interval (22 days, white bars). Lower values denote more suppression, and larger values denote less suppression (i.e., more extinction).

Procedurally, this was identical to the test in the short groups. Thus, testing was conducted in Context B.

Results and Discussion

The results of Experiment 3 are depicted in Figure 3. The three groups tested after a short retention interval (short) displayed a pattern similar to that observed in Experiments 1 and 2. That is, the massed-short group suppressed less than the control-short group did, and the spaced-short group displayed even less suppression than did the massed-short group, suggesting that extinction was more robust in the spaced-short group. A different pattern was observed when the subjects were tested after a long retention interval. These groups behaved similarly to the groups in Experiment 2 that were tested in the training context (ABA): Extinction was apparent in the spaced condition relative to the control condition. Moreover, in the massed condition with a long retention interval, more fear was exhibited than in the massed condition tested soon after extinction training, which indicates spontaneous recovery. In the groups that experienced spaced extinction trials, such recovery was not observed. These observations were corroborated with the following statistical analyses.

Because of an absence of leverpressing during the testing session, 1 subject in the massed-long group was excluded from all of the analyses. A 2×3 ANOVA with retention interval (short or long) and the extinction treatment (control, massed, or spaced) as factors conducted on the mean leverpresses during the minute before the first CS presentation indicated no reliable differences between groups [largest $F(2,65) = 1.00, p = .37$]. A similar ANOVA was conducted on the mean suppression to the CS across the three test trials. This analysis yielded a main effect of retention interval [$F(1,65) = 14.21, MS_e = 0.014, p < .001$], a main effect of the extinction treatment [$F(2,65) = 20.81, MS_e = 0.014, p < .001$], and an interaction [$F(2,65) = 3.30, MS_e = 0.014, p < .05$, Cohen's $f = 0.25$]. Planned comparisons using the overall error term from the ANOVA revealed that among subjects tested with a short retention interval, those that experienced a massed extinction treatment (massed-short) suppressed less than did the controls, which experienced no extinction treatment (control-short) [$F(1,65) = 13.17, p < .001$, Cohen's $f = 0.71$], suggesting that extinction was reliable soon after massed extinction trials. Extinction between the massed and spaced groups tested soon after the extinction treatment was not statistically different [$F(1,65) = 0.99, p = .32$], although Figure 3 suggests that the spaced-short group tended to suppress less than did the massed-short group, which is consistent with the results of the previous two experiments. A comparison critical to determine whether spontaneous recovery was observed in the massed groups (massed-short or massed-long) indicated that when a long retention interval was imposed between extinction and testing, subjects suppressed more of the presentation of the target CS [$F(1,65) = 17.87, p < .001$, Cohen's $f = 0.83$]. Spontaneous recovery from extinction was not observed in groups that received spaced extinction trials, since the spaced-long group did not suppress more than did the spaced-short group [$F(1,65) = 1.45, p = .23$].

Finally, a simple t test between the early and late massed subgroups tested after a short retention interval found that these two subgroups did not differ [$t(10) = 0.91, p = .37$]. A similar test between the two subgroups tested after the long retention interval also yielded an absence of reliable differences [$t(9) = 0.51, p = .61$]. Thus, our results in the groups extinguished with massed trials do not seem to differentially depend on the portion of the extinction session on which the subject experienced massed extinction. In summary, in this experiment, we observed a trend toward more extinction when trials were spaced in the session. But more important, we observed a spontaneous recovery effect after massed extinction trials. In contrast, this effect was attenuated when extinction trials were greatly spaced.

GENERAL DISCUSSION

The present experiments were conducted to assess the effect of the interval between extinction trials on two types of recovery from extinction: renewal and spontaneous recovery. Under conditions that do not result in appreciable recovery from extinction (the ABB design with no retention interval between extinction and test), the results of all three experiments suggested a small benefit of extinction with spaced trials relative to massed trials (this was statistically significant only in Experiment 1). In Experiment 2, we observed a renewal effect after massed extinction trials when testing was conducted in the context in which training occurred (ABA renewal), but this effect was attenuated if extinction was conducted with spaced trials. Furthermore, a similar pattern was observed in Experiment 3, but in this experiment, instead of changing the physical context between extinction and testing (renewal), we interposed a long retention interval between extinction and testing (presumably changing the temporal context). Here, we observed spontaneous recovery after massed extinction trials, but this effect was attenuated if extinction was conducted with spaced trials. In all three experiments, there was a direct relationship between the spacing of extinction trials and the effectiveness of extinction observed during testing. Larger intervals between extinction trials consistently produced a tendency toward less conditioned suppression (more effective extinction) at the time of testing.

During the last 25 years, considerable knowledge has been gained regarding the processes involved in experimental extinction. Perhaps the most widely accepted tenet regarding extinction is that it does not erase (or destroy) the original memory of the stimulus. In other words, extinction does not cause unlearning (Bouton, 2002; Rescorla, 2001; but see Barad, 2006, and Stout & Miller, 2007, for views that accept a degree of unlearning). Instead, it is now generally accepted that an extinction treatment leads (at least in part) to new learning and that this new learning interferes with the retrieval of the original fear memory, provided that appropriate retrieval cues from extinction are presented at the time of testing (Bouton, 1993). This view relies on observations of several types of recovery of excitatory responding

after an extinction treatment without any further training of the original memory. For example, Pavlov (1927) observed a recovery from extinction after interposing a retention interval between an extinction treatment and testing, which he named *spontaneous recovery*. A similar recovery from extinction is observed when testing is conducted in a context different from that in which extinction was conducted. This has been called *renewal* (Bouton & Bolles, 1979). A third observation, which suggests that extinction does not destroy the original fear memory, is *reinstatement* (Rescorla & Heth, 1975). Reinstatement is a recovery from extinction that results from exposing subjects prior to testing to the biologically significant outcome used during excitatory training (e.g., an electric shock), provided that testing occurs in the same context in which subjects reexperienced the outcome. Finally, in some circumstances, it has been observed that *reacquisition* following extinction proceeds faster relative to acquisition with a new stimulus (Napier, Macrae, & Kehoe, 1992). Taken together, these observations suggest that during extinction, the original memory was not entirely erased, or otherwise no recovery after extinction would be possible.

One of the main problems that exposure therapies face is that the treatment is not as enduring as would be desired. In other words, after successful finalization of treatment, fearful reactions to exposed stimuli often reemerge. This has been called *relapse* or *return of fear* (ROF; Rachman, 1989), and it is perhaps the premier challenge that behavior therapy encounters. A closer look at the circumstances under which ROF occurs suggests similarities between ROF and the observations of recovery from extinction delineated above. For example, ROF generally occurs after a period of time from the completion of treatment (Rachman, 1989), which is similar to the observation of spontaneous recovery from extinction. Moreover, one can think of the therapist's office as a specific context in which extinction takes place; therefore, it is not surprising that a renewal-like effect is observed when subjects encounter the previously fear-inducing stimulus outside the therapist's office (e.g., Mystkowski, Craske, & Echiverri, 2002). The research presented in this article was conducted with rats, but, given the clear analogies between different forms of recovery in experimental extinction and the relapse that is often seen after exposure therapy, it appears reasonable to think that identifying ways to make experimental extinction less vulnerable to recovery speaks to the problem of relapse from exposure therapy.

The literature assessing the effect of interval between extinction trials with humans as participants has provided equivocal results. For example, Howat and Grant (1958) used a human eyelid conditioning preparation in which, after training to a criterion, participants experienced extinction with massed or spaced trials and were tested either immediately (20 min) or a day later (24 h). Participants who experienced spaced extinction trials reached a criterion for extinction faster than did participants who experienced massed extinction trials. Regardless of the spacing of extinction trials, all participants showed recovery from extinction when testing occurred either 20 min or 24 h later. More recently, Rowe and Craske (1998a) investigated the effects of massing exposure trials or using an expanding ITI exposure treatment (i.e., intervals between extinction trials doubling every session) on participants with a fear of spiders. They observed faster fear reduction with massed trials than with expanding ITI trials. However, massed exposure resulted in recovery from extinction when participants were reexposed to the spider 1 month later, and this ROF was not observed in participants who experienced expanding ITI trials. But subsequent replication of this study did not provide full support for these observations (Lang & Craske, 2000). In the latter study, total exposure to the CS during extinction was longer, and this may have obscured possible differences between the massed and spaced groups. However, faster extinction in massed groups was observed in both studies. Finally, Tsao and Craske (2000) tested massed, uniformly spaced and expanding ITI exposure and found less ROF in both spaced groups relative to the massed exposure group with a 10-day retention interval. Overall, the human literature suggests that spacing exposure to fearful stimuli alleviates ROF.

The literature with nonhuman subjects contains discrepant findings when the difference between massing and spacing of extinction trials has been assessed. These studies have adopted two forms. In some studies, massed extinction was implemented by giving a single long exposure to the stimulus (i.e., maximal massing). Under these circumstances, either no difference was observed (Martasian, Smith, Neill, & Rieg, 1992), or a benefit of spacing was observed, but this effect disappeared when the extinction treatment was extended (Martasian & Smith, 1993). However, this implementation of the massing of extinction trials introduces two potential problems: (1) In practice, it suggests that a therapist would have to conduct one long session in which complete extinction would have to be achieved; and (2) the neuromolecular processes involved in extinction are thought to start when a given extinction trial actually ends (Pedreira, Pérez-Cuesta, & Maldonado, 2004). Nevertheless, many other extinction studies have varied the ITI while keeping the total amount of CS exposure and the total number of CS presentations constant, just as we did in the present research. We will now discuss those studies.

Benefits of spaced relative to massed extinction treatments have been observed in operant responses in the *Lymnaea stagnalis* snail (Sangha, Scheibenstock, Morrow, & Lukowiak, 2003). Westbrook et al. (1985) found that a long ITI between extinction trials (spaced) produced more long-term loss of a conditioned taste aversion than did a short ITI (massed). However, massed presentations during extinction facilitated the rate of extinction observed during the extinction treatment. Additionally, Morris, Furlong, and Westbrook (2005) reported more robust extinction with spaced than with massed trials in a fear-conditioning preparation, which is consistent with the results of the present experiments. Perhaps most relevant to the present discussion, Cain et al. (2003) reported, in a fear-conditioning preparation, faster extinction when extinction trials were massed (6 sec), relative to interme-

diate (60 sec) and widely spaced (600 sec). Moreover, massing trials during extinction also alleviated spontaneous recovery from extinction when testing was conducted 8 days later. In addition, in a parallel experiment in an appetitive preparation, researchers consistently failed to find differences between massed and spaced extinction on spontaneous recovery (Moody, Sunsay, & Bouton, 2006). Specifically, in several experiments (Moody et al.'s Experiments 1, 2, and 3) with different parametric variations, it was found that extinction proceeded faster with massed trials, but recovery from extinction was equal when trials were massed or spaced. Moreover, in one experiment (Moody et al.'s Experiment 5B) in which they compared massed with spaced extinction after a reinstatement test (i.e., recovery from extinction after presentation of the US alone before testing with the extinguished cue), a massed but not a spaced extinction treatment showed reinstatement (recovery from extinction), just as we observed in our Experiments 2 and 3.

The discrepancies between the data from Cain et al. (2003) and the present data are disturbing, since our procedure and parameters were similar to theirs. The main differences between our procedure and Cain et al.'s are that they used mice, whereas we used rats, and that they used a longer CS duration (120 sec, as opposed to the 60-sec CS used here). One feature of their experiment that is intriguing is the fact that their spaced groups never showed any behavior indicative of extinction during the extinction session. That is, freezing only decreased during the extinction treatment with massed trials, not with spaced trials. It is possible that in mice, these spaced trials acted more like reminders of the fear memory than like effective extinction trials. Consistent with this speculation are data suggesting that the amount of exposure during a reminder treatment can lead either to reconsolidation (short exposure) or extinction (long exposure; Lee, Milton, & Everitt, 2006). Another difference between our Experiment 3 and Cain et al.'s Experiment 2, which assessed the effects of spacing or massing of extinction trials on spontaneous recovery, is that they used an 8-day-long retention interval between extinction and test, whereas we used a 22-day-long retention interval. In fact, the benefit of spaced relative to massed extinction trials that we observed with a short retention interval was relatively small, although the ordinal direction was the same in all three experiments. Anyway, our results are clear when testing was conducted in the acquisition context (renewal; Experiment 2) and after a 21-day retention interval (spontaneous recovery; Experiment 3). In both instances, recovery from extinction was far more robust when extinction was conducted with massed trials than with spaced trials. Spaced extinction trials appear to make extinction more resistant to recovery.

Perhaps the best way to reconcile these discrepant findings is to assume that the effects of massed extinction trials are observed during the extinction treatment as faster response cessation (Moody et al., 2006) but that spacing trials during the extinction treatment produces extinction that is more durable when assessed under circumstances known to undermine extinction, as we observed in Experiments 2 and 3 with ABA renewal and spontaneous recov-

ery manipulations, respectively. Results in taste aversion experiments (Berman, Hazvi, Stehberg, Bahar, & Dudai, 2003) and contextual fear conditioning (Li & Westbrook, 2008) are consistent with this interpretation. In fact, this conclusion can be derived from a memory model that emphasizes the short-term and long-term consequences of massing and spacing trials during extinction (Bjork & Bjork, 1992, 2006). In the framework of this model, massing extinction trials facilitates retrieval from the previous trial (due to recency); thus, faster extinction is observed. Spaced trials lack this benefit, but, if during spaced trials the memories are successfully retrieved, extinction is more durable. The results with humans described above partially confirm these predictions. Our results are at least partially in agreement with this theoretical proposal, because spacing trials alleviated renewal and spontaneous recovery. Although we did not observe an advantage of massed trials in Experiment 1, it is possible that the 2-day retention interval between extinction and testing was sufficient to eliminate the short-term advantage of trial massing.

Our findings are also consistent with the view that extinction does not erase the original learning but, rather, creates new inhibitory learning that is dependent on the physical and temporal attributes of contextual information (Bouton, 1993). According to this approach, extinction creates a new inhibitory (context-dependent) association between the previously reinforced CS and the US. If extinction is new learning, and (excitatory) learning depends on the spacing of trials (Barela, 1999), extinction is also apt to be impaired by massed extinction trials. Thus, one would expect that massed extinction trials would be more vulnerable to recovery from extinction than spaced extinction trials would, as we observed in the present series of experiments. Other learning theories, such as the Rescorla–Wagner model (Rescorla & Wagner, 1972), Mackintosh's (1975) attentional model, the Pearce–Hall model (Pearce & Hall, 1980), Pearce's (1987, 1994, 2002) configural model, and the comparator hypothesis (Denniston, Savastano, & Miller, 2001; Miller & Matzel, 1988; see Stout & Miller, 2007, for a formal version) all assume that the training context drives spacing effects (in both acquisition and extinction), given that their trialwise approach precludes alternative explanations of real-time phenomena. Thus, because the total context exposure in these experiments was the same for massed and spaced extinction groups, these theories anticipate no differences based on the spacing of extinction trials.

Wagner's (1981) SOP model is a real-time formulation that, although it, too, relies in part on context to explain spacing effects, includes a decay parameter that determines the flow of representations from an active state, A1, to another active state, A2, and from the latter to an inactive state, I. According to SOP, when CS and US are both in A1, an excitatory association between them is learned, but when the CS is in state A1 and the US is in state A2, an inhibitory association is formed. The latter mechanism drives extinction, because during an extinction trial, the CS is represented in state A1 and the US's representation moves from I to A2, leading to inhibitory learning. CS massing

during extinction is detrimental for extinction, because the CS representation of trial n should still be in A2 when the next CS is presented during trial $n+1$ and less available to be activated in A1 (so-called *self-generated priming*). However, this correct prediction is dependent on decay rates of the US, which should also be activated in A2 on trial n. If the US is still in A2 during trial $n+1$, then SOP predicts that massed extinction should be better. Clearly, the predictions of the model depend on the relative decay rates of CS and US from A2 to the inactive state. Moreover, a recent revision of the model (Dickinson & Burke, 1996), aimed at explaining retrospective revaluation phenomena, such as recovery from overshadowing after extinction of the overshadowing cue (Kaufman & Bolles, 1981), has difficulty generating even this correct prediction. In the original SOP model, no excitation developed when the representations of two stimuli (e.g., CS and US) were in A2, but Dickinson and Burke proposed that an excitatory link between the two stimuli would develop when the two stimuli are in A2. Consequently, massing extinction with a slow rate of decay (which anticipates the present data in the original version of the model) would cause the CS and US to be simultaneously represented in A2 and, thus, would increase excitation between the two, which would counteract the self-generated priming that allows better spaced extinction relative to massed extinction. Thus, the original model anticipates our data but with parameters that make the revised version anticipate little difference or even make the opposite prediction. Moreover, neither of the two versions anticipates the recovery from extinction observed in our Experiments 2 and 3, which is the phenomenon that prompted us to conduct these experiments.

In the present series of experiments, we were not able to collect data during the extinction treatment, because subjects became sated after approximately 30 min and consequently decreased their leverpressing. Without any baseline responding, it was impossible to calculate a suppression ratio, which was our critical dependent measure. Although online extinction data might be informative in terms of when responding stopped (response cessation) during massed and spaced extinction trials, response cessation likely does not directly reflect resistance to renewal and spontaneous recovery, which are analogous to ROF. Moreover, the effectiveness of extinction treatments might differentially be reflected when looking at response cessation *during* extinction or when looking at transfer of extinction to a test situation (Drew, Yang, Ohyama, & Balsam, 2004). We believe that this is an important aspect of extinction that has not received enough attention and might well be critical for behavioral therapy, because the therapist's decision to terminate treatment is usually based on fear observed at the therapist's office. Perhaps a few evaluations outside of the therapist's office would be advisable if termination of treatment is imminent, since this seems to be a more reliable test of the effects of the therapy upon different forms of relapse. This is an issue that Rowe and Craske (1998a, 1998b) have previously shown to be significant when they found in humans that indices of fear activation during exposure were poor predictors of ROF.

In summary, our results have identified another variable that may be relevant for extinction to have long-lasting effects—trial spacing during extinction trials. Several variables that are likely to alleviate recovery from extinction have already been identified. One of these variables is extinction in multiple contexts (Chelonis, Calton, Hart, & Schachtman, 1999; Gunther, Denniston, & Miller, 1998; Neumann, 2006). Another variable that has proven to alleviate recovery from extinction is extinction in the presence of another excitatory stimulus (Rescorla, 2000, 2006; Thomas & Ayres, 2004; but see Urcelay, Lipatova, & Miller, 2008, and Vervliet, Vansteenwegen, Hermans, & Eelen, 2007, for limits to this strategy). Extinction with a large number of extinction trials (massive extinction) has also proven to alleviate recovery from extinction (Denniston, Chang, & Miller, 2003; Tamai & Nakajima, 2000). In addition to these ways to attenuate recovery from extinction, we have presented data here that suggest that spaced extinction trials also alleviate recovery from extinction relative to massed extinction trials. Of note, this effect is not large when assessed immediately after an extinction treatment in the extinction context (ABB; Experiments 1, 2, and 3). However, under circumstances that thwart extinction (ABA renewal in Experiment 2 and spontaneous recovery in Experiment 3), the benefits of spacing extinction trials are clear and reliable. Ideally, as Drew et al. (2004) have argued, extinction should be assessed by looking at fear cessation both during extinction treatment and during transfer tests outside of the extinction session. If we accept the analogy between experimental extinction and exposure-based behavioral therapies, the latter should be a better test of the effectiveness of treatment. Finally, although there are large differences in the timescales of our experiments and the scheduling of therapy sessions, the data presented here and in most of the literature using human participants seem to agree that spaced extinction trials attenuate recovery from extinction. This conclusion is relevant because, although most therapies involve weekly meetings (Heimberg & Becker, 2002), some intensive treatments have been designed in which clients receive massive exposure therapy over a period of 5 days with follow-up phone calls 1 week, 1 month, and 6 months later (Child Anxiety Network, 2001). Unfortunately, more research is needed to elucidate the underlying processes of extinction and the specific influence of parametric variation of the spacing of extinction trials.

AUTHOR NOTE

REFERENCES

Annau, Z., & Kamin, L. J. (1961). The conditioned emotional response as a function of intensity of the US. *Journal of Comparative & Physiological Psychology*, **54**, 428-432.

BARAD, M. (2006). Is extinction of fear erasure or inhibition? Why both, of course. *Learning & Memory*, **13**, 108-109.

BARELA, P. B. (1999). Theoretical mechanisms underlying the trial-spacing effect in Pavlovian fear conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 177-193.

BERMAN, D. E., HAZVI, S., STEHBERG, J., BAHAR, A., & DUDAI, Y. (2003). Conflicting processes in the extinction of conditioned taste aversion: Behavioral and molecular aspects of latency, apparent stagnation, and spontaneous recovery. *Learning & Memory*, **10**, 16-25.

BJORK, R. A., & BJORK, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes, Vol. 2: From learning processes to cognitive processes* (pp. 35-67). Hillsdale, NJ: Erlbaum.

BJORK, R. A., & BJORK, E. L. (2006). Optimizing treatment and instruction: Implications of a new theory of disuse. In L.-G. Nilsson & N. Ohta (Eds.), *Memory and society: Psychological perspectives* (pp. 116-140). New York: Psychology Press.

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

BOUTON, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, **52**, 976-986.

BOUTON, M. E., & BOLLES, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning & Motivation*, **10**, 445-466.

CAIN, C. K., BLOUIN, A. M., & BARAD, M. (2003). Temporally massed CS presentations generate more fear extinction than spaced presentations. *Journal of Experimental Psychology: Animal Behavior Processes*, **29**, 323-333.

CHELONIS, J. J., CALTON, J. L., HART, J. A., & SCHACHTMAN, T. R. (1999). Attenuation of the renewal effect by extinction in multiple contexts. *Learning & Motivation*, **30**, 1-14.

CHILD ANXIETY NETWORK (2001). *Specialized programs*. Retrieved April 28, 2008, from www.childanxiety.net/Specialty_Programs .htm.

CRASKE, M. G., KIRCANSKI, K., ZELIKOWSKY, M., MYSTKOWSKI, J., CHOWDHURY, N., & BAKER, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research & Therapy*, **46**, 5-27.

DELGADO, M. R., OLSSON, A., & PHELPS, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, **73**, 39-48.

DENNISTON, J. C., CHANG, R. C., & MILLER, R. R. (2003). Massive extinction treatment attenuates the renewal effect. *Learning & Motivation*, **34**, 68-86.

DENNISTON, J. C., SAVASTANO, H. I., & MILLER, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Mahwah, NJ: Erlbaum.

DICKINSON, A., & BURKE, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgments. *Quarterly Journal of Experimental Psychology*, **49B**, 60-80.

DREW, M. R., YANG, C., OHYAMA, T., & BALSAM, P. D. (2004). Temporal specificity of extinction in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, **30**, 163-176.

GUNTHER, L. M., DENNISTON, J. C., & MILLER, R. R. (1998). Conducting exposure treatment in multiple contexts can prevent relapse. *Behaviour Research & Therapy*, **36**, 75-91.

HEIMBERG, R. G., & BECKER, R. E. (2002). *Cognitive–behavioral group therapy for social phobia: Basic mechanisms and clinical strategies*. New York: Guilford.

HERMANS, D., DIRIKX, T., VANSTEENWEGENIN, D., BAEYENS, F., VAN DEN BERGH, O., & EELEN, P. (2005). Reinstatement of fear responses in human aversive conditioning. *Behaviour Research & Therapy*, **43**, 533-551.

HOFMANN, S. G. (2008). Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. *Clinical Psychology Review*, **28**, 199-210.

HOWAT, M. G., & GRANT, D. A. (1958). Influence of intertrial interval during extinction on spontaneous recovery of conditioned eyelid responses. *Journal of Experimental Psychology*, **56**, 11-15.

KAUFMAN, M. A., & BOLLES, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, **18**, 318-320.

LANG, A. J., & CRASKE, M. G. (2000). Manipulations of exposure-based therapy to reduce return of fear: A replication. *Behaviour Research & Therapy*, **38**, 1-12.

LANG, A. J., CRASKE, M. G., & BJORK, R. A. (1999). Implications of a new theory of disuse for the treatment of emotional disorders. *Clinical Psychology: Science & Practice*, **6**, 80-94.

LEE, J. L. C., MILTON, A. L., & EVERITT, B. J. (2006). Reconsolidation and extinction of conditioned fear: Inhibition and potentiation. *Journal of Neuroscience*, **26**, 10051-10056.

LI, S. H., & WESTBROOK, R. F. (2008). Massed extinction trials produce better short-term but worse long-term loss of context conditioned fear responses than spaced trials. *Journal of Experimental Psychology: Animal Behavior Processes*, **34**, 336-351.

LOVIBOND, P. F., PRESTON, G. C., & MACKINTOSH, N. J. (1984). Context specificity of conditioning, extinction, and latent inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 360-375.

MACKINTOSH, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276-298.

MARTASIAN, P. J., & SMITH, N. F. (1993). A preliminary resolution of the retention of distributed vs massed response prevention in rats. *Psychological Reports*, **72**, 1367-1377.

MARTASIAN, P. J., SMITH, N. F., NEILL, S. A., & RIEG, T. S. (1992). Retention of massed vs distributed response-prevention treatments in rats and a revised training procedure. *Psychological Reports*, **70**, 339-355.

MILLER, R. R., & MATZEL, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 51-92). Orlando, FL: Academic Press.

MOODY, E., SUNSAY, C., & BOUTON, M. (2006). Priming and trial spacing in extinction: Effects on extinction performance, spontaneous recovery, and reinstatement in appetitive conditioning. *Quarterly Journal of Experimental Psychology*, **59**, 809-829.

MORRIS, R. W., FURLONG, T. M., & WESTBROOK, R. F. (2005). Recent exposure to a dangerous context impairs extinction and reinstates lost fear reactions. *Journal of Experimental Psychology: Animal Behavior Processes*, **31**, 40-55.

MYERS, J. L., & WELL, A. D. (2003). *Research design and statistical analysis* (2nd ed.). Mahwah, NJ: Erlbaum.

MYSTKOWSKI, J. L., CRASKE, M. G., & ECHIVERRI, A. M. (2002). Treatment context and return of fear in spider phobia. *Behavior Therapy*, **33**, 399-416.

NAPIER, R. M., MACRAE, M., & KEHOE, E. J. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 182-192.

NEUMANN, D. L. (2006). The effects of physical context changes and multiple extinction contexts on two forms of renewal in a conditioned suppression task with humans. *Learning & Motivation*, **37**, 149-175.

PAVLOV, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press. (Original work published 1926)

PEARCE, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, **94**, 61-73.

PEARCE, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, **101**, 587-607.

PEARCE, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior*, **30**, 73-95.

PEARCE, J. M., & HALL, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review*, **87**, 532-552.

PEDREIRA, M. E., PÉREZ-CUESTA, L. M., & MALDONADO, H. (2004). Mismatch between what is expected and what actually occurs triggers memory reconsolidation or extinction. *Learning & Memory*, **11**, 579-585.

RACHMAN, S. J. (1989). The return of fear: Review and prospect. *Clinical Psychology Review*, **9**, 147-168.

RESCORLA, R. A. (2000). Extinction can be enhanced by a concurrent excitor. *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 251-260.

RESCORLA, R. A. (2001). *Experimental extinction*. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 119-154). Mahwah, NJ: Erlbaum.

RESCORLA, R. A. (2006). Deepened extinction from compound stimulus presentation. *Journal of Experimental Psychology: Animal Behavior Processes*, **32**, 135-144.

RESCORLA, R. A., & DURLACH, P. J. (1987). The role of context in intertrial interval effects in autoshaping. *Quarterly Journal of Experimental Psychology*, **39B**, 35-48.

RESCORLA, R. A., & HETH, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, **1**, 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 research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

ROWE, M. K., & CRASKE, M. G. (1998a). Effects of an expanding-spaced vs massed exposure schedule on fear reduction and return of fear. *Behaviour Research & Therapy*, **36**, 701-717.

ROWE, M. K., & CRASKE, M. G. (1998b). Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behaviour Research & Therapy*, **36**, 719-734.

SANGHA, S., SCHEIBENSTOCK, A., MORROW, R., & LUKOWIAK, K. (2003). Extinction requires new RNA and protein synthesis and the soma of the cell right pedal dorsal 1 in *Lymnaea stagnalis. Journal of Neuroscience*, **23**, 9842-9851.

STOUT, S. C., & MILLER, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, **114**, 759-783.

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., & AYRES, J. J. B. (2004). Use of the ABA fear renewal paradigm to assess the effects of extinction with co-present fear inhibitors or excitors: Implications for theories of extinction and for treating human fears and phobias. *Learning & Motivation*, **35**, 22-52.

TSAO, J. C. I., & CRASKE, M. G. (2000). Timing of treatment and return of fear: Effects of massed, uniform-, and expanding-spaced exposure schedules. *Behavior Therapy*, **31**, 479-497.

URCELAY, G. P., LIPATOVA, O., & MILLER, R. R. (2008). *Generalization decrement constrains extinction in the presence of an excitor.* Manuscript submitted for publication.

VERVLIET, B., VANSTEENWEGEN, D., HERMANS, D., & EELEN, P. (2007). Concurrent excitors limit the extinction of conditioned fear in humans. *Behaviour Research & Therapy*, **45**, 375-383.

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

WESTBROOK, R. F., SMITH, F. J., & CHARNOCK, D. J. (1985). The extinction of an aversion: Role of the interval between non-reinforced presentations of the averted stimulus. *Quarterly Journal of Experimental Psychology*, **37B**, 255-273.