Expanding the Intertrial Interval During Extinction: Response Cessation and Recovery

Alyssa J. Orinstein, **Gonzalo P. Urcelay**, and **Ralph R. Miller**
State University of New York at Binghamton

Abstract

We examined trial spacing during extinction following a human contingency learning task. Specifically, we assessed if an expanding retrieval practice schedule (Bjork & Bjork, 1992, 2006), in which the spacing between extinction trials was progressively increased, would result in faster immediate extinction and less recovery from extinction than uniformly spaced extinction trials. We used an ABB vs. ABA renewal design and observed that, whereas the expanding group extinguished faster during extinction treatment, the expanding and constant groups showed the same level of extinction with an immediate test in the extinction context (ABB) and the two groups showed equivalent ABA renewal at test in the training context. We conclude that the faster extinction observed in the expanding groups could be misleading in clinical treatment, if the therapist used the absence of fear during extinction as the basis for terminating treatment.

In Pavlovian conditioning, when a neutral stimulus is paired with an unconditioned stimulus (US), the initially neutral stimulus becomes a conditioned stimulus (CS) and comes to elicit a conditioned response (CR). Experimental extinction is the procedure in which the CS is repeatedly presented after paired training, but now in the absence of the US. Through this treatment, the CS progressively elicits a weaker CR. This is analogous to exposure-based therapies for the treatment of phobias and other anxiety-related disorders. Presumably, phobias involve (at least partially) learned associations between stimuli and dangerous events, and consequently can be viewed as a form of Pavlovian conditioning (e.g., Hofmann, 2008). In exposure-based therapy, a widely practiced form of cognitive-behavioral therapy, the patient is encouraged to confront feared objects or situations to reduce fear reactions that presumably result from these stimuli being associated with an aversive outcome (e.g., Foa, Cahill, & Pontoski, 2004). It is widely assumed that exposure therapy involves processes analogous to extinction: Repeatedly presenting the CS without the US after acquisition leads to a decrease in conditioned responding in a human differential fear-conditioning paradigm with visual stimuli as CSs and a loud aversive noise as the US (Vansteenwegen et al., 2005). During exposure the patient is presumably experiencing extinction in that the feared object encountered in the therapist's office does not result in the expected aversive event. Thus, research on the processes involved in experimental extinction can be of potential use in suggesting efficacious clinical treatments.

At one time, extinction was viewed as reflecting the unlearning of the CS–US association (e.g., Rescorla & Wagner, 1972). The central problem with this interpretation of extinction is that recovery from extinction should not occur because the CS–US memory should no longer exist. However, as will be described below, there are many manipulations devoid of further CS–US pairings that result in recovery from extinction. Newer theories (as well as some older theories,

Address correspondence to Ralph R. Miller, Department of Psychology, SUNY-Binghamton, Binghamton, NY 13902-6000; rmiller@binghamton.edu.

e.g., Hull, 1943; Pavlov, 1927) posit that extinction treatment results in the establishment of an inhibitory association between the CS and the US that coexists with the excitatory association. For example, Bouton (1993) proposed that after extinction treatment, the CS has two potentially available associations, one with the US (i.e., excitatory) and the other with the absence of the US (i.e., inhibitory). Presumably, after acquisition followed by extinction treatment, the CS is ambiguous because it has now two different meanings. This ambiguity is resolved at the time of testing by the presence or absence of the extinction context, the former of which is crucial for the expression of extinction learning. Thus, extinction learning is thought to be context-dependent.

There are many phenomena providing evidence that extinction learning is context-dependent. One such phenomenon is the renewal effect, which is observed when a change in the physical context between extinction treatment and test causes a restoration of conditioned responding to the extinguished CS. A critical characteristic of the renewal effect is that it occurs when testing is conducted in a context different from that in which extinction was conducted, even if that context is associatively neutral. The renewal effect has been observed in almost every conditioning preparation in which it has been investigated (for a review, see Bouton, 2002), which leads to two central conclusions. The first is that the extinction context modulates or "sets the occasion" for the retrieval of the CS–no US association (Bouton, 2004). In other words, the context primes the spatially local relationship of the CS with the absence of the US. The second conclusion is that extinction learning is more context-specific than is initial conditioning. This is apparent in renewal designs in which the CS is tested in a neutral context and recovery of responding is observed (ABC and AAB renewal, where the first letter denotes the context of acquisition, the second letter the context of extinction, and the third letter the context of test). If the original conditioning was more context dependent than the extinction learning, then the opposite result would be expected at least in the case of ABC renewal (i.e., no recovery of responding in the neutral context). However, because this is not the case, researchers have concluded that conditioning is less context-specific and consequently transfers better to a novel test context than extinction learning (but see Nelson, 2002).

A second phenomenon providing evidence that extinction is not unlearning but rather new, context-dependent learning, is spontaneous recovery. Spontaneous recovery is the recovery from extinction that occurs when a retention interval is interposed between extinction treatment and test. This effect was first observed by Pavlov (1927) and has been replicated by many others (e.g., Brooks & Bouton, 1993; Devenport, Hill, Wilson, & Ogden, 1997; Robbins, 1990). Spontaneous recovery can be considered a form of renewal that occurs when the CS is tested outside of the temporal context in which extinction occurred (Bouton, 2004). Thus, renewal and spontaneous recovery can both be viewed as due to a failure to retrieve memories of extinction outside of the extinction context, either because the physical context of extinction is not present (renewal) or the temporal context is not present (spontaneous recovery).

Although exposure therapy usually is highly effective in reducing fear and anxiety symptoms during therapy sessions (Foa et al., 2004), there is evidence that much of this reduction in fear is not permanent. The expression *return of fear* has been coined to denote this effect (see Rachman, 1979, 1989, for reviews). Return of fear generally occurs outside of the therapist's office, which can be seen as a form of renewal, and/or after a period of time since the cessation of treatment, which can be seen as a form of spontaneous recovery. For example, Mineka, Mystkowski, Hladek, and Rodriguez (1999) found that patients' arachnophobia returned when tested 1 week after treatment in a different context than the one where exposure therapy had occurred. Rodriguez, Craske, Mineka, and Hladek (1999) also observed a return of fear when they tested their patients in a different room with a different therapist. Additionally, Mystkowski, Craske, and Echiverri (2002) used indoor and outdoor contexts and found a return of fear (tested 1 week after exposure therapy) when participants were tested in the context other

than that in which exposure therapy had occurred. The results of these studies emphasize the similarity between the context dependency of experimental extinction and that of exposure therapy. Moreover, parallels between context-dependent memory phenomena (e.g., Smith, Glenberg, & Bjork, 1978) and return of fear provide a unique opportunity for both clinicians and behavioral psychologists to investigate ways of enhancing extinction and preventing recovery from extinction.

Various treatments that enhance extinction and alleviate recovery from it have been identified to date. Briefly, Gunther, Denniston, and Miller (1998) and Chelonis, Calton, Hart, and Schachtman (1999; but see Bouton, Garcia-Gutiérrez, Zilsky, & Moody, 2006) found that conducting extinction treatment of a CS in multiple contexts attenuates recovery from extinction. In an aversive conditioning preparation with rats, Gunther et al. found that extinguishing the CS in multiple contexts other than the training context alleviated the recovery from extinction that is typically observed when the stimulus is tested in a neutral context (i.e., ABC renewal). Another way to alleviate recovery from extinction is massive over-extinction, as demonstrated by Denniston, Chang, and Miller (2003). They observed that extensive extinction (800 trials), compared to moderate extinction (160 trials), attenuated renewal in both ABC and ABA designs. Rescorla (2000, 2006a; but see Urcelay, Lipatova, & Miller, 2008; Vervliet, Vansteenwegen, Hermans, & Eelen, 2007, for limitations on this procedure) observed that the co-presentation of another excitatory conditioned stimulus during extinction enhances extinction and attenuates spontaneous recovery compared with co-presentations with either a neutral stimulus or no other stimulus. Brooks and Bouton (1993, 1994) found that presenting a cue during test that was featured during extinction treatment reduced spontaneous recovery (1993) and renewal (1994), although the cue itself did not support any excitatory or inhibitory behavioral control. The spacing of extinction trials has also been suggested as a means of alleviating these recovery effects, with spaced extinction trials alleviating renewal and spontaneous recovery in rats (e.g., Urcelay, Wheeler, & Miller, 2008) and humans (Tsao & Craske, 2000). Thus, there are many ways in which an experimenter or clinician can try to alleviate recovery from extinction. New manipulations and improvements on the current methods continue to be examined, one being Bjork and Bjork's (1992, 2006) new theory of disuse.

The new theory of disuse (Bjork & Bjork, 1992, 2006) is a model devised to provide an account of certain fundamental phenomena of learning and memory. The following peculiarities of memory phenomena influenced the assumptions of the theory: (a) a great storage capacity exists that contrasts with a fallible retrieval process; (b) successful retrieval is dependent on the similarity between the acquisition and retrieval contexts; (c) the act of retrieval is a dynamic process that alters the subsequent state of the system; (d) in situations with ambiguous or conflicting information (e.g., two-phase treatment such as acquisition followed by extinction), the most recent information (extinction) wanes with increasing retention intervals, thereby allowing the earlier information (acquisition) to be more easily retrieved. The new theory of disuse heavily relies on the fact that learning and performance depend upon different (but related) aspects of a memory representation: retrieval strength, which refers to ease of retrieval, and storage strength, which refers to the strength of the memory independent of its retrievability.

Recently, this theory has received attention because it makes a prediction of potential practical importance regarding the spacing of trials that should enhance memory retention. The basic prediction is that progressively increasing the intertrial interval (ITI) should produce more enduring learning than when the ITI is held constant. This prediction has important implications not only for educational settings, but also for clinical practice. Because extinction may be characterized as new learning, when applied to extinction the theory predicts that progressively increasing ITIs during extinction treatment should result in more effective extinction (as

indexed by weaker renewal and spontaneous recovery) than extinction treatment with constant intervals (massed, as well as spaced). According to this view, the act of retrieval strengthens retrieval strength, which decays during ITIs. Hence, massing of extinction trials early in extinction should lead to a larger increase in retrieval strength of the extinction memory compared to spaced extinction trials, because of the loss of retrieval strength of the memory of extinction during the ITIs with spaced extinction trials. Thus, the theory predicts that when the cue is tested at a short retention interval (or during the extinction session), retrieval strength of the memory of extinction will be higher after massed extinction trials as compared to spaced extinction trials. However, spaced extinction trials result in greater increases in storage strength of the memory of extinction than do massed extinction trials because increments in storage strength are negatively related to current retrieval strength. Moreover, with spaced extinction trials, there is a greater loss of retrieval strength of the memory of extinction between extinction trials, which leads to greater increments in storage strength of the memory of extinction given successful retrieval. This accumulation of storage strength of the memory of extinction with spaced extinction trials slows the further loss of retrieval strength of the memory of extinction across retention intervals. The durability of the retrieval strength of the memory of extinction allows this memory to be more easily retrieved at test than the memory of reinforcement, thus preventing recovery from extinction over time.

One consequence of progressively expanding the ITI during extinction is that there would likely be little effect compared to uniformly massed extinction trials if participants were tested immediately after extinction treatment. This is because, when testing is conducted immediately after extinction, massed extinction trials result in a rapid loss of responding due to a high level of retrieval strength of the memory of extinction. However, the other consequence of expanding retrieval practice is that this schedule of extinction should lead to an alleviation of renewal and spontaneous recovery due to the higher levels of storage strength of the memory of extinction that result from the widely spaced trials at the end of the extinction treatment.

This unique prediction of the new theory of disuse has been tested in laboratory situations (Karpicke & Roediger, 2007) and in clinical populations of clients with arachnophobia (Rowe & Craske, 1998) and with acrophobia (Lang & Craske, 2000), but with mixed results. Rowe and Craske found that, although a massed trials group exhibited a faster decrease in fear responses, a generalization test (akin to a renewal test) immediately after the end of extinction treatment revealed more return of fear in the massed groups than the expanding group. Moreover, in a 1-month follow-up assessment, an increase in self-reported fear for the massed group and a decrease for the expanding group was observed. However, a replication of this study by Lang and Craske failed to find differences between groups, but in this replication extinction was apparently too extensive as none of the groups showed return of fear (a floor effect). In addition, high attrition rates in the massed groups raised concerns about the assumptions made about the data. Lang and Craske also mentioned that other variables such as environmental features, mood of patients, and precision of timing could have influenced their findings. As will be seen below, there are other confounds that impede an unambiguous interpretation of these results. The main objective of the present experiment was to test this prediction in a laboratory situation that avoided all these confounds.

Both Rowe and Craske (1998) and Lang and Craske (2000) compared an expanding schedule of extinction trials to massed extinction trials. This comparison, however, introduced a number of confounds with regard to the ITI and the intervals between extinction and test. For example, the mean ITI was appreciably shorter for the massed condition than the expanding condition. In a series of extinction experiments with nonhuman animals in a fear-conditioning preparation (Urcelay et al., 2008), we have recently observed that spacing extinction trials alleviates renewal and spontaneous recovery. It is possible that Rowe and Craske saw a difference between the expanding and massed groups simply because the mean ITI was larger for the

expanding group than for the massed group. Tsao and Craske (2000) sought to address this confound by comparing the effects of massed trials to uniform-spaced and expanding-spaced with the same mean ITI on fear reduction for self-reported public-speaking anxiety. One month after the last exposure therapy session, they found less return of fear in both spaced groups than the massed group; however, there was no difference in the amount of fear present between the uniform-spaced and expanding-spaced groups. Tsao and Craske speculated that this lack of differences may have been due to differences in treatment completion, arising from differential attrition between the uniform and expanding groups.

In the current experiment, expanding retrieval practice was contrasted with uniformly spaced extinction trials with the same mean ITI for both groups, and there was little attrition. Additionally, because Rescorla (2004) found that the interval between conditioning and extinction affected the magnitude of spontaneous recovery, timing of the different phases of treatment and testing was kept constant across groups in the present research. Specifically, the length of the interval between acquisition and the beginning of the extinction phase and the length of the interval between the end of the extinction phase and test were the same across all groups.

As previously mentioned, this experiment was conducted to test predictions concerning expanding extinction trials using a human contingency learning preparation, which gave us more experimental control than in a clinical setting over factors that have little to do with the learning mechanisms underlying this prediction, such as high attrition rates and lack of temporal control over stimuli presentations. In this task, participants were exposed on a computer screen to files containing information about foods (which served as CSs) that sometimes led to gastrointestinal problems (an aversive outcome which served as the US) in fictitious people who ate them in a particular restaurant (context). The name of the restaurant served as the context in which foods were eaten. In other words, participants learned the relationship between certain foods and negative consequences, which in some cases depended on contextual information.

Responses were recorded on each trial of treatment. Thus, this preparation allowed us to map the progression of extinction curves and assess the rate of extinction (response cessation), as well as determine if expanding retrieval practice alleviates ABA renewal. This is important because recent reviews have highlighted the importance of distinguishing between fear decreases during the session (within-session habituation) and fear decreases across sessions, which presumably is more related to retention of the extinction treatment (between-session habituation; Craske et al., 2008). They suggested that, although reported fear and physiological arousal may decline throughout a session, this does not guarantee that there will be significant long-term improvement, which is ultimately the goal of the treatment. We will return to this point later.

The Current Study

The human contingency task and many of the parameters used in this experiment were borrowed from studies by Rosas and Callejas-Aguilera (2006), in which they examined the effects of context switches on extinction in human contingency learning. Rosas and Callejas-Aguilera used a human predictive learning task in which participants had to estimate whether different foods that a hypothetical person ate at a specific restaurant would cause the person to have diarrhea. Participants did so by giving a contingency judgment concerning the relationship between the food and diarrhea. In their scenario, the different restaurants acted as the different contexts. Because Rosas and Callejas-Aguilera were successful in obtaining differences between contexts with their design, we used their task with a few minor changes that allowed us to better control temporal factors that are critical to assess the predictions of

the new theory of memory disuse (Bjork & Bjork, 1992, 2006). The main difference implemented in the current experiment was that the task was not self-paced. Instead, the trials were computer-paced to facilitate our manipulation of the intervals between extinction trials and maintain control over the duration of trials. We chose to use this task because it allowed us to assess, at a cognitive level, learned relationships between a predictor (foods) and the representation of an outcome (diarrhea) that is capable of evoking disgust reactions. Moreover, this task allowed us to further assess these relationships in particular contexts (restaurants). Disgust sensitivity has recently been strongly related to anxiety disorders (Teachman, 2006), and disgust reactions seem to share the same neurobiological basis with fear reactions (Zald, 2003).

Two contexts were used in order to assess whether extinction on an expanding extinction schedule alleviated renewal relative to both a group that experienced uniformly spaced extinction trials and a group for which the target cue was not extinguished. In addition to the target cue Q, which we trained and then extinguished with different ITIs, we used three nontarget (i.e., filler) cues, X, Y, and Z. These cues were trained under three different contingencies. X was never reinforced, Y was always reinforced, and Z was reinforced 50% of the time. The filler cues were included to prevent participants from adopting a global rule-based strategy that they could potentially generalize to all other cues. Additionally, during Phase 2, the number of fillers interposed between extinction trials of the target cue Q determined the length of the ITI between these trials. The current study was conducted with a 3×2 between-subjects factorial design in which type of extinction (Control vs. Constant vs. Expanding) and test context (ABB vs. ABA in which the letters denote the contexts of acquisition, extinction, and test, respectively) were factors. The intent was to test two main predictions derived from the new theory of memory disuse: (a) that expanding extinction trials results in faster extinction than uniformly spaced extinction trials (with the same mean ITI as the expanding extinction trials), and (b) that expanding extinction trials results in less renewal than uniformly spaced extinction trials. See Table 1 for the experimental design and predicted results.

Method

PARTICIPANTS

The participants were 240 female and male undergraduate students at the State University of New York at Binghamton, who participated in partial fulfillment of a course requirement. They were divided into six treatment groups, counterbalanced for gender to the extent possible. Our population was 63% female, had a mean of 19 years old, with a range from 18 to 27 years.

APPARATUS AND STIMULI

All testing was done on six IBM-compatible personal computers. Inputs were made through a standard computer keyboard. All testing visuals were displayed on conventional monitor screens, using SuperLab Pro 2.0 (Cedrus). Stimuli Q, X, Y, and Z were the names of different foods. Stimulus Q was potatoes and stimuli X, Y, and Z were jello, pretzel, and bread, counterbalanced within groups. Contexts A and B were the names of two fictitious restaurants (The Argentinean Horn and The Ukrainian Tavern), counterbalanced within groups. On each trial, two types of screens were used: presentation and feedback. On the top of the presentation screen there was a sentence that read, "The person ate in the restaurant … [restaurant's name]." In the middle of the screen it said, "The person ate … [food]." On the bottom of the presentation screen, there was a sentence that read, "Press a number key on the keyboard to indicate the likelihood that this person will have…DIARRHEA." Following that sentence there was a 1 to 9 scale, with 1 labeled "Not Likely" and 9 labeled "Very Likely," to indicate the likelihood that the person would have diarrhea. On the feedback screen there was a sentence that read,

"The person ate in the restaurant …" followed by the name of the restaurant. The sentence, "This person had …" appeared below the restaurant name. The specific outcome was presented centered in the lower third of the screen. One of the two possible outcomes, "diarrhea" or "nothing," appeared in bold capital letters, "diarrhea" in red, and "nothing" in dark green. The name of the restaurant, The Argentinean Horn, was written in capital purple letters within an orange oval. The name of the restaurant, The Ukrainian Tavern, was written in capital purple letters within a grey rectangle. The names of the foods appeared in capital black letters in a yellow rectangular box. The rest of the text appeared in black font. The screen backgrounds, as well as the screen presented during the ITIs, were white.

PROCEDURE

Because the task was computer-paced and we expected participants to rate every cue on every trial, the participants were given a set of verbal instructions to ensure understanding of the task and to facilitate data collection. The verbal directions were read to all participants by the experimenter before participants were seated at their respective computers and were as follows:

> In this experiment, you will have to make quick decisions on whether or not certain foods cause diarrhea in a person who ate in a particular restaurant. The computer will first load the person's file and then tell you what restaurant the person ate in and what food the person ate. As you see the information, you will have to rate the likelihood that the person will have diarrhea, by pressing the number keys 1 through 9 on the keyboard, with 1 indicating that it is Not Likely that the person will have diarrhea and 9 indicating that it is Very Likely that the person will have diarrhea. You can use any of the number keys in between as well. You only have to press a key once for us to record your rating. The screen will look something like this [the experimenter provided a demonstration off screen and pointed at where the restaurant information and the food information goes], and it will be shown for a fixed amount of time. A few seconds after you see the information and give a rating, you will see another screen that will tell you whether or not the person actually had diarrhea. Remember, you will have to make your decisions as quickly as possible. I will now set each of you up on a computer. Be sure to carefully read the instructions before you begin the experiment.

The restaurant name and the food used in the demonstration screen with the verbal instructions were different from the information used during the actual experiment. At the beginning of the experiment, each participant was again presented with a set of instructions on the computer screen:

> Screen 1:
>
> Recent developments in food technology have led to chemical synthesis of food. This creates a great advantage because its cost is very low and it is easy to store and transport. This revolution in the food industry may solve hunger in third-world countries. Please press the spacebar to continue.
>
> Screen 2:
>
> However, it has been detected that some of the foods produce gastric problems in some people. For this reason we are interested in selecting a group of experts to identify the foods that lead to some type of illness and how severely it appears in each case. Please press the spacebar to continue.
>
> Screen 3:
>
> You are about to participate in a task in which you will be looking at the files of people that have ingested different foods in a specific restaurant. You will have to indicate

to what degree you think that a gastric problem will appear. To respond, you should press keys 1 through 9 (1=Not Likely; 9=Very Likely) to indicate whether or not the person will suffer any gastric problems. It is very important that you respond quickly so that those people who have problems can receive No Delay medical attention. Therefore, you will have only a few seconds to record your response. The computer will then tell you whether or not a gastric problem actually occurred. Your response will be random at the beginning, but do not worry; little by little you will become an expert. Please press the spacebar to continue.

Phase 1—Before Phase 1 began, the participants were shown a screen that had the sentence, "*Now you will analyze the files of people who ate at a restaurant called...*[restaurant's name]" (Context A). Above each cue presented, the restaurant name was displayed. All three groups received 8 trials each of Q+, X−, Y+ and Z+/− (where+indicates diarrhea and −indicates no diarrhea), in a fixed pseudo-random order. The order of the first presentation of each cue (the first 4 trials) was fully randomized (with Z being paired with diarrhea). The order of presentation for the next 28 trials (7 of each cue) was: Z−, X−, Q+, Y+, X−, Q+, Y+, Z+, Q+, Y+, Z−, X−, Y+, Z+, X−, Q+, Z−, X−, Q+, Y+, X−, Q+, Y+, Z+, Q+, Y+, Z−, X−. Participants rated each of the cues on each trial the cue was presented, by pressing the numeric key on the keyboard. Regardless of how quickly the participants responded, the cue remained on the screen for 5 seconds. The participants then received a 1.5-second feedback screen indicating the problem the person had (diarrhea or nothing). Following the feedback screen was an ITI of 1.5 seconds indicated by a screen with the sentence "*Loading file of...* [a randomly chosen person's name]." Full names were always different to create the impression that each file was drawn from a different person.

Phase 2—Before Phase 2 began, the participants were shown a screen that had the sentence, "*Now you will analyze the files of people who ate at a restaurant called...* [restaurant's name]" (Context B). Above each cue presented, the restaurant name was displayed. Both Expanding and Constant groups received extinction trials of Q. Critical here was that the ITI between extinction trials (in these two groups) was determined by the number of fillers between each nonreinforced presentation of the target cue Q. In Group Constant, each nonreinforced presentation of Q was always separated by 6 fillers. The participants in the Constant group received 6 nonreinforced trials of Q− and 10 trials each of X−, Y+, and Z+/−. Between each presentation of Q−, 6 filler cues were presented in a pseudo-random order. Thus, the cues were presented in the following order: Q−, Z+, X−, Y+, X−, Y+, Z−, Q−, Y+, Z+, X−, Z−, X−, Y +, Q−, X−, Y+, Z+, Y+, X−, Z−, Q−, X−, Z+, Y+, Z−, Y+, X−, Q−, Y+, X−, Z−, Y+, X−, Z +, Q−. In Group Expanding, each successive ITI was occupied by 0, 2, 4, 8, and 16 fillers, respectively. Thus, participants in the Expanding group also received 6 nonreinforced trials of Q− and 10 trials each of X−, Y+, and Z+/−. The difference in this group relative to the Constant group was that between each presentation of Q− an increasing number of filler cues (X−, Y+ and Z+/−) were presented in a pseudo-random order. The exact order of presentation of the cues was: Q−, Q−, Z−, Y+, Q−, Z+, X−, Y+, Z−, Q−, Y+, Z+, X−, Z−, X−, Y+, Z+, X−, Q−, Z−, Y+, X−, Y+, X−, Z+, Y+, X−, Y+, X−, Y+, X−, Z−, Y+, X−, Z+, Q−. Participants in the No Extinction group received 16 trials of X−, 10 trials of Y+ and 10 trials of Z+/−, in a fixed pseudo-random order, but Q was never presented during this phase. This group served as an extinction control that also controlled for retention, as the interval from the end of Phase 1 to the test session was the same for all groups. The order of cue presentations for this group was: X−, Z+, X−, Y+, X−, Y+, X−, Z−, X−, Z+, X−, Z−, X−, Y+, X−, Z+, Y+, X−, Y+, X−, Z−, Y +, X−, Z+, Y+, X−, Y+, Z−, X−, Y+, X−, Z−, X−, Y+, X−, Z+. The rest of the procedure for Phase 2 was the same as in Phase 1. We used six extinction trials because a pilot experiment (in which we observed a similar pattern as that observed in the conditions ABB in the present experiment – see below) determined that this number of trials would result in significant

decreases in the ratings to the target. Moreover, because prior research has suggested that massive extinction attenuates different forms of renewal (Denniston et al., 2003; Tamai & Nakajima, 2000), we did not want to administer an excessive number of extinction trials that would preclude us from observing a renewal effect. Typically, in these human contingency learning tasks, participants learn the contingencies in relatively few trials, even when a large number of cues are used (Larkin, Aitken, & Dickinson, 1998). Note that both extinction groups began and ended the extinction session with the same trial presentations, Q−, so that the interval between the last Q reinforced trial and the first Q extinction trial, and that between the last Q extinction trial and test of Q was the same. Moreover, because extinction might depend on the status of short-term memory rather than timing of events, both groups receiving extinction treatment experienced the same trial type (reinforced or nonreinforced) before experiencing an extinction trial of Q, so that any differences between groups could not be attributed to the effect of the trials immediately preceding the extinction trials. As in Phase 1, participants rated each cue presentation during each trial presentation.

Test—After completing Phases 1 and 2, the participants were asked to evaluate the potential of the individual cues to cause diarrhea. Before testing began, the participants were shown a screen that had the sentence, "*Now you will analyze the files of people who ate at a restaurant called...* [restaurant's name]" (Context A or B, depending on group assignment). In this test, the participants rated the potential of each food to cause diarrhea. Each participant was tested first on the target cue Q. After this, they rated cues X, Y, and Z, with the order randomized within groups. Above each cue was the name of the restaurant, which was Context A [restaurant's name] for Groups No Extinction–ABA, Constant-ABA, and Expanding-ABA, and Context B [restaurant's name] for Groups No Extinction–ABB, Constant-ABB, and Expanding-ABB. Each cue remained on the screen until the participant gave a rating for it. After completing this test, the program terminated with a thank you and debriefing screen.

DATA ANALYSIS

Participants who gave a rating to stimulus X (which was never reinforced) that was equal or higher than that to stimulus Y (which was always reinforced) during testing were excluded from all analyses. We based this criterion on the assumption that if these participants did not learn the basic contingencies of the filler cues, their ratings of the target stimulus Q would be unreliable at best. Based on this selection criterion, 175 participants were used in the final test analyses: 30 in the No Extinction–ABB group, 30 in the Constant-ABB group, 30 in the Expanding-ABB group, 27 in the No Extinction–ABA group, 29 in the Constant-ABA group, and 29 in the Expanding-ABA group. The data were subjected to mixed or between-subjects analyses of variance (ANOVAs), depending on the analysis. The sources of significant interactions were determined with planned comparisons using the overall error term of the ANOVA. We report effect sizes calculated using the algorithm provided by Myers and Well (2003). Power analyses were performed with the use of G-Power (Erdfelder, Faul, & Buchner, 1996). To correct for the possibility of Type I errors, we used the rather "pessimistic" Bonferroni adjustment (Abdi, 2007).

Results

The left side of Fig. 1 depicts the ratings for all four stimuli (Q, X, Y, and Z) during acquisition. Fig. 1 shows smooth asymptotic acquisition for stimuli Q and Y, which were always reinforced; low ratings for stimulus X, which was never reinforced; and intermediate ratings for stimulus Z, consistent with its 50% reinforcement schedule. The figure (and analysis) represent data from 94 participants who responded on all acquisition trials. The acquisition data were analyzed using a 4 (stimuli)× 8 (trials) within-subjects ANOVA. This analysis revealed main effects of stimulus, $F(3, 279)=129.03$, $p<.01$, Cohen's $f=2.02$, $\eta^2 = 1.00$; and trial, $F(7, 651)=84.52$, $p<.$

01, Cohen's f=2.49, η^2 = 1.00; as well as a Stimulus×Trial interaction, $F(21, 1953)$=20.03, p<.01, Cohen's f=2.06, η^2 =1.00.

The right side of Fig. 1 shows the extinction curves of stimulus Q for the Constant and Expanding groups collapsed across the ABA vs. ABB conditions (which did not appreciably differ) for 107 participants from Conditions Constant and Expanding who responded during all extinction trials. There are no extinction data for the No Extinction groups because these groups did not receive any Q presentations during the extinction phase. As can be seen in the figure, extinction proceeded much faster in the Expanding groups than in the Constant groups. However, at the end of extinction training, the two groups were rating Q similarly. These observations were confirmed with a 2 (Group: constant vs. expanding)×6 (Trials: 1–6) mixed ANOVA. This analysis detected main effects of group, $F(1, 105)$ = 13.45, p <.01, Cohen's f=0.34, η^2 =1.00; trial, $F(5, 525)$=10.24, p<.01, Cohen's f=0.65, η^2 =1.00; as well as an interaction, $F(5, 525)$=6.78, p<.01, Cohen's f=0.51, η^2 =1.00. To ascertain the source of this interaction, we conducted planned comparisons. Based on the pattern of the interaction and a priori expectations, we decided to conduct 5 planned comparisons. Thus, after the Bonferroni adjustment, the alpha level was p=.01. In the Constant and Expanding groups, within-groups comparisons indicated that the ratings for Trial 6 were lower than for Trial 1, $F(1, 105)$=8.68, p<.01, Cohen's f=0.26, η^2 =.99, and $F(1, 105)$=11.77, p<.01, Cohen's f=0.31, η^2 =.99, respectively, which indicates that extinction occurred in both groups. Further between-groups planned comparisons showed that ratings in the Expanding groups were lower than those in the Constant groups on Trial 2, $F(1, 105)$=34.75, p<.01, Cohen's f=0.56, η^2 =1.00; Trial 3, $F(1, 105)$=16.82, p<.01, Cohen's f=0.38, η^2 =1.00; and Trial 4, $F(1, 105)$=6.83, p=.01, Cohen's f=0.23, η^2 =.99. Thus, the interaction seems to be driven by different levels of ratings in the Constant and Expanding groups on Trials 2, 3, and 4. Overall, these analyses confirm one of the predictions made by the new theory of memory disuse: expanding the ITI should result in a faster decrement in ratings relative to a group that always receives the same ITI. Of course Trials 2, 3, and 4 had more recent presentations of Q in the Expanding condition than in the Constant condition, which qualifies the conclusion that can be drawn from these differences.

Fig. 2 shows the test data for the target stimulus Q. The Constant-ABB and the Expanding-ABB groups, which were tested in the same context in which extinction occurred, both showed extinction relative to their respective controls (i.e., no extinction of Q). However, the Constant-ABA and Expanding-ABA groups both showed recovery from extinction when testing was conducted in the training context, suggesting the presence of an ABA renewal effect. The Q test data were analyzed using a 3 (Group: Control vs. Constant vs. Expanding) ×2 (Test: ABB vs. ABA) between-subjects ANOVA. This analysis detected main effects of Group, $F(2, 169)$ =6.74, p<.01, Cohen's f=0.25, η^2 =1.00, and Test, $F(1, 169)$=17.44, p<.01, Cohen's f=0.30, η^2 =1.00, as well as a Group×Test interaction, $F(2, 169)$=3.81, p<.05, Cohen's f=0.18, η^2 =.99. Planned comparisons were conducted to verify the basic extinction effect in the Constant and Expanding groups. The alpha level was, after the Bonferroni adjustment, equal to .01. These comparisons revealed that immediately after extinction (without a context change, condition ABB), extinction was evident in both Constant and Expanding groups relative to Group Control–ABB, $F(1, 169)$=17.24, p<.01, Cohen's f=0.52, η^2 =.99; and $F(1, 169)$=14.43, p<.01, Cohen's f =0.48, η^2 =.99, respectively. The comparison between Constant-ABB and Expanding-ABB was not significant, $F(1, 169)$=0.12, p=.72, suggesting that these two groups showed similar levels of extinction. This is not consistent with the prediction that an expanding series of extinction trials will result in more extinction (at least on an immediate test) than the same number of uniformly spaced extinction trials. A planned comparison between Constant-ABB and Constant-ABA was significant, $F(1, 169)$=15.87, p<.01, Cohen's f=0.51, η^2 =.99, indicating that there was recovery from extinction in the Constant condition when testing was conducted in the training context (i.e., renewal). Similarly, a planned comparison between Expanding-ABB and Expanding-ABA proved significant, $F(1, 169)$=9.55, p<.01, Cohen's

f=0.38, η² =.99, indicating that there was recovery from extinction in the Expanding condition. These planned comparisons revealed that the Constant-ABB and the Expanding-ABB groups showed the same amount of extinction when tested in the extinction context and that both the Constant-ABA and Expanding-ABA groups showed recovery from extinction when tested in the acquisition context (i.e., renewal).

Participants in the No Extinction groups received an extra 6 trials of stimulus X in place of the 6 extinction trials of stimulus Q received by the participants in the Extinction groups. Additionally, participants in the ABB condition were tested on stimuli X, Y, and Z in Context B (the extinction context), whereas participants in the ABA groups were tested on stimuli X, Y, and Z in Context A (the acquisition context). Preliminary statistical analyses revealed that neither the number of presentations of stimulus X, nor the test context affected the ratings for stimuli X, Y, and Z at test. Therefore, all participants were pooled for the analysis of the mean ratings of stimuli X, Y, and Z, which were 1.58, 7.98, and 5.14, respectively (see Fig. 3). The X, Y, and Z test data were analyzed using a 1-way within-subject ANOVA with stimulus (X vs. Y vs. Z) as a factor. The ANOVA was significant, $F(2, 348)=390.54$, $p<.01$, Cohen's *f*=2.10, η² =1.00, indicating that there were different levels of responding to the three stimuli. In accordance with the overall reinforcement contingencies for these three cues, stimulus X, which was never reinforced, received the lowest ratings; stimulus Y, which was always reinforced, received the highest ratings; and stimulus Z, which was partially reinforced (50%), received intermediate ratings (see Fig. 3). Based on these data, we conclude that the Q data reflect extinction processes and not some other nonspecific interference processes. If the latter were the case, the X, Y, and Z data would also have reflected such interference.

Overall, the test results of this experiment suggest that expanding retrieval practice does not attenuate renewal relative to constant controls matched for mean ITI. However, the extinction data clearly show that there were differences in the progression of decrement in rating between the two groups, with faster rating decrement occurring in the Expanding group.

Discussion

This experiment used an aversive human contingency learning task in which foods (cues) were paired with diarrhea (outcome) in certain restaurants (context) to determine whether progressively increasing the ITI between successive extinction trials improves extinction memory and attenuates recovery from extinction when the test is conducted in the acquisition context (ABA renewal). In this experiment, expanding retrieval practice was contrasted with uniformly spaced extinction trials, with the mean ITIs being equal, in contrast to previous reports (e.g., Lang & Craske, 2000; Rowe & Craske, 1998) that compared expanding retrieval practice with massed extinction trials. Furthermore, the intervals between the beginning and end of the different phases were kept constant for all groups. Thus, our results cannot be interpreted in terms of differences in the training-to-extinction or extinction-to-test intervals (Rescorla, 2004).

Cognitively, our aversive-like task presumably engaged the same neurobiological substrates as those found to be hyperactive in patients with anxiety disorders (Phelps, 2006). That is, brain imaging studies have detected bilateral amygdala activation to aversive stimuli of multiple sensory modalities (i.e., olfactory, gustatory, visual, and auditory; Zald, 2003). Moreover, amygdala activation has been observed after instructed fear, in which participants, instead of experiencing the aversive stimulus, are told that a neutral cue is associated with an aversive stimulus (Phelps et al., 2001). This is the neural correlate of the verbal conditioning experiments conducted in the last century (Cook & Harris, 1937). The outcome (i.e., diarrhea) used in this task has a negative valence and produces a disgust reaction that is particularly significant for some forms of anxiety disorders like OCD (Teachman, 2006). Moreover, in a recent study with

a nonclinical population of children, measures of disgust sensitivity were positively correlated with symptoms of specific phobias, social phobia, agoraphobia, and OCD, which further strengthen at a behavioral level the parallels at a neurobiological level described above (Muris, van der Heiden, & Rassin, 2008). The experiment showed that during extinction treatment, ratings decreased in the Expanding group faster than the Constant group. This pattern provides evidence supporting the first prediction derived from Bjork and Bjork's (1992, 2006) new theory of disuse. That is, expanding the spacing of extinction trials presumably results in faster decrements in response or rating potential than do uniformly spaced extinction trials. If only the data from during extinction treatment were considered (e.g., Kushner et al., 2007, see below), the benefits of expanding retrieval practice would be clear (i.e., response cessation was achieved in fewer trials). However, the observation of faster decreases in ratings with fewer extinction trials given expanding trials is confounded by time since the immediately prior extinction trial. That is, rate of extinction is an ambiguous measure with which to compare uniformly spaced extinction trials and expanding extinction trials given that early expanding trials are massed. Moreover, the test results showed that expanding extinction trials did not prevent renewal relative to uniformly spaced trials. Thus, it seems that expanding the ITI during extinction may have some short-term benefits for extinction that can be observed as extinction treatment progresses but not afterwards. Centrally, the test data collected in this experiment do not support the second prediction derived from Bjork and Bjork's (1992, 2006) new theory of disuse. That is, renewal was not alleviated with an expanding ITI between trials relative to uniformly spaced trials.

This failure to observe alleviated renewal from extinction with expanding ITIs during extinction does not necessarily mean that Bjork and Bjork's (1992, 2006) assumptions concerning expanding retrieval practice are incorrect. As the new theory of disuse does not provide formal parameters for determining optimal trial spacing, the spacing and number of the extinction trials in our experiments were chosen on the basis of pilot data that we had previously collected to assure we would observe renewal. All that we can conclude from our failure to see attenuation of recovery from extinction using the current expanding trial spacing parameters is that improvement as a result of expanding extinction trials minimally is at least parameter-dependent. This fact is actually implicit in Bjork and Bjork's description of the process. For expanding retrieval practice to be successful, extinction trials need to be spaced so that they are far enough apart that storage strength increases, but not so far apart that there is not enough retrieval strength to remember the previous extinction trial. In their original formulation, Bjork and Bjork suggested that retrieval had to be successful for the expanding series to result in better performance. However, Storm, Bjork, Bjork, and Nestojko (2006) recently found that successful retrieval of one memory is not necessary to induce forgetting of a competing memory. That is, simply attempting to retrieve information, regardless of whether retrieval was successful or not, may be sufficient for retrieval-induced forgetting of the competing memory to occur. In the present experiment, participants had to rate the cues on each trial; minimally, this should at least have served as a retrieval attempt. In fact, our extinction curves suggest that such retrieval was successful; otherwise performance would have been the same during extinction treatment in both the Expanding and Constant groups.

In a clinical setting, the optimal trial spacing for each individual client would probably vary greatly, thus making it extremely hard for a therapist to know how to space the trials in exposure therapy for a particular client. Consistent with this concern, Lang and Craske (2000) discussed the problem of identifying the optimal interval between extinction trials after failing to find that expanding retrieval practice resulted in better immediate fear reduction and less return of fear than massed extinction trials. They stated that the delay between exposure trials may promote forgetting of aspects of the conditional stimulus, thereby increasing the generalization of extinction learning within treatment. However, they noted that the delay should not be long enough to allow for complete return of fear. These variables relate to the amount of exposure

and initial level of fear, which presumably would be different for each individual. Indeed, a recent report suggests that the conditions under which expanding retrieval practice provides a benefit for long-term retention are very selective (Hays & Bjork, 2007).

Although expanding retrieval practice is likely to result in faster fear reduction during treatment, it does not necessarily alleviate recovery from extinction. Similarly, when examining expanding retrieval practice during acquisition, Balota, Duchek, and Logan (2007) found that participants in an expanding condition were at a higher level of retrieval success during a longer retention interval than participants in a uniformly spaced condition. But, no such benefit translated to the final recall test. Balota et al. proposed that this may be due to the fact that the number of spaced trials administered during acquisition correlates with performance at test. That being said, participants in an expanding condition typically receive at least one less spaced trial than participants in a uniformly spaced condition, because in the expanding condition the second trial immediately follows the first trial. This might mitigate results on a final recall test. Similarly, Karpicke and Roediger (2007) reported that delaying the initial retrieval attempt, but not expanding the interval between repeated trials, is necessary for promoting long-term retention. In other words, they observed that an expanding retrieval practice schedule confounds the interval between the last study trial and the first retrieval (test) trial (which is analogous to our first extinction trial) and that this latter interval was more important than the rest of the schedule. Because in the present experiment the delay between the last acquisition trial and the first extinction trial was the same in both extinction conditions, our failure to find a long-term difference is consistent with this view. However, more research needs to be done before any final conclusion can be made regarding to the efficacy of expanding retrieval practice.

Another model that explains part of the present data set is Wagner's (1981) SOP model. This is a real-time model in which memory representations can be activated in one of three memory states. When a stimulus is physically presented, its representation moves from the inactive state (I) to the active state A1. The representation then decays from A1 into a less active state A2, and subsequently into the inactive state. To explain extinction, the model proposes that when the cue is presented, it activates a representation of the outcome into state A2 (because the outcome is remembered, but not experienced) and therefore an inhibitory association between the cue representation in A1 and the outcome in A2 is formed. If extinction is conducted with massed trials, fewer elements of the cue will be activated into A1 because some elements will still be in A2 (from the previous trial), which should result in less inhibitory learning (i.e., less long-term extinction). Moreover, this model accounts for the rapid decrease in responding that we observed during extinction with massed trials. With massed early trials fewer cue elements should be available in state I to be activated into A1, and this should activate fewer outcome elements into A2, which is the primary cause of conditioned responding. Thus, the model predicts the rapid response cessation that we observed early in the extinction session. However, it also predicts that long-term extinction with constant ITIs, with a mean ITI equal to that of the expanding ITI, should be better than with the expanding series, which is contrary to what we observed when testing was conducted in the extinction context. A recent variant of SOP (replaced elements; Wagner, 2003; Wagner & Brandon, 2001) predicts the context dependency of extinction, for which the original SOP cannot account. It poses that some elements will acquire an inhibitory association with the outcome but this association will be context-dependent, which should result in recovery from extinction if testing is conducted outside of the extinction context. In summary, this model describes well the faster loss of responding during extinction treatment with increasing ITIs, but it also predicts that this treatment should result in less extinction (and more recovery), which we did not observe.

Another issue relevant to the present research that needs to be addressed is the notion that extinction can be assessed by different metrics and that the conclusions reached by each metric

should be convergent (Drew, Yang, Ohyama, & Balsam, 2004). One way to assess extinction is by examining the rate of extinction (response cessation) during the extinction session, and a second way of assessing extinction is by looking at transfer of extinction learning from extinction treatment to test. Others have distinguished these two metrics by referring to them as extinction training (response cessation) and extinction retention (transfer; Davis, Ressler, Rothbaum, & Richardson, 2006). Others have differentiated these processes as within-session habituation and between-session habituation (Craske et al., 2008: this distinction parallels different mechanisms of habituation in Wagner's [1981] SOP model). Because of economic and time constrains, most clinicians tend to rely on assessments or reports obtained during the therapeutic session (response cessation). These assessments rely on the assumption that performance during extinction training is a precise predictor of extinction learning. However, experimental variables applied during training have long been known to have two different kinds of effects (Estes, 1955; Guthrie, 1952; Hull, 1942; Skinner, 1938; Tolman, 1932). First, the experimental variables can have permanent effects, which translate into what is considered learning by most definitions. In other words, the experimental manipulation results in relatively permanent changes in conditioned responding. Second, the experimental variables can have temporary effects, in which performance is either increased or reduced during training but there is no evidence of a change in behavior on a subsequent test. Thus, only certain kinds of changes in performance should be considered "learning effects" (Schmidt & Bjork, 1992). Behavioral differences during training could be the result of performance factors, learning itself, or a combination of both. To determine if and how much extinction learning actually occurred, it is necessary to implement a subsequent test in which there is a sharp division between treatment and test to ensure that the results are due to enduring learning and not just the result of performance variables that affect behavior during extinction treatment but do not perseverate. The data observed in this experiment speak to this distinction because we observed faster decrease in ratings during extinction treatment with an expanding retrieval practice schedule, but this faster decrease was not reflected as alleviated recovery in a renewal tests. Thus, the faster decrease in ratings observed during extinction treatment should not be taken as indicative of enhanced extinction because this difference in performance did not transfer to subsequent tests.

Drew et al. (2004). discussed the differences between cessation of responding during nonreinforcement and the transfer of extinction learning to a subsequent test. They explained this disparity by assuming that excitation and extinction are independent processes with different generalization gradients. Moreover, Drew et al. suggested that cessation of responding is modulated by the degree to which excitation generalizes from the excitatory training context to the extinction context. In this case, manipulations that increase the difference between acquisition and extinction training will result in faster decreases in responding. This is evident in the present experiment when one compares ratings of the target stimulus at the end of training (conducted in context A) and the first trial of extinction (conducted in context B; see Fig. 1). Note that on the first extinction trial, participants rated the target cue before they received any feedback; thus, these lower ratings cannot be explained in terms of extinction. Critically, the significantly lower ratings during the second trial (in Condition Expanding, which received the second trial immediately after the first one) potentially reflects the operation of extinction processes. It is interesting to note that studies with nonhuman animals (i.e., crabs) have suggested that the molecular machinery that is responsible for extinction starts to operate within less than a minute after CS-offset (Pedreira, Pérez-Cuesta, & Maldonado, 2004; Pérez-Cuesta, Hepp, Pedreira, & Maldonado, 2007). However, performance during the test phase in a novel context is dependent in part on both the extent to which extinction learning generalizes from the extinction context to the novel test context and the extent that conditioning generalizes to that test context. Because excitation has a broader generalization gradient than extinction (Bouton, 1993; but see Rescorla, 2006b), recovery from extinction is likely to occur when testing occurs outside of the extinction context.

The difference between performance during extinction and evidence of extinction on a subsequent test can have a large impact in a clinical setting. Therapists ordinarily try to structure their sessions with patients to ensure maximal reduction in responding during therapy. For example, in sessions with phobic patients, the goal of the therapist is often to reduce the fear as much as possible during exposure therapy. Specifically, in a recent study with OCD individuals, Kushner et al. (2007) saw that D-cycloserine (DCS; relative to a placebo) augmented exposure therapy as evidenced by a decrease in obsession-related fear during four sessions, but differences disappeared with further sessions. Specifically, they used the Subjective Unit of Distress Scale after subjects had ranked the 10 most disturbing obsession-related stimuli (e.g., touch a toilet seat, which evoked a disgust reaction). Moreover, they administered the Yale-Brown Obsessive Compulsive Scale and observed that these scores did not differ between DCS and placebo groups after the fourth session, at the end of treatment, or 3 months later. The risk is lack of fear exhibited by the patient during treatment may lead the therapist to conclude that the patient needs no further therapy. However, the absence of fear could reflect a short-term performance effect, rather than appreciable learning of the new extinction contingency. The extinction curves and subsequent test data in this experiment speak to this problem. If the amount of learning was based on the level of apparent extinction during the extinction trials, the participants in the Expanding group would appear to have learned the extinction contingency after the first three trials. However, when assessed in a different context or after a retention interval, responding was the same in the Expanding and Constant conditions. Thus, there was relatively little transfer of learning from extinction training to test. This is consistent with recent data (reviewed in Craske et al., 2008) suggesting very little relation between fear reduction observed during the session and that seen outside the session (indicative of a long term effect of extinction treatment). In the present experiment we saw significant differences in the speed of reduction during the extinction session, but no differences in a subsequent renewal test.

With respect to the present data, if a therapist had used the absence of fear during a session as an index for treatment termination, treatment would have been concluded after Trial 3 for the Expanding group. However, when tested after a contextual shift, there would likely be a higher level of fear during test than during extinction training if an expanding retrieval practice schedule was used. The extinction data from the current experiment provide a strong argument against using the absence of fear during exposure therapy as an index of extinction of fear. Rowe and Craske (1998) found similar results in their experiments with arachnophobia. Patients who experienced massed exposure trials displayed greater reductions in self-reported anxiety and self-reported physiological symptoms during exposure therapy compared to patients who received expanding exposure trials. Additionally, when tested immediately after extinction in the same context, patients who received massed exposure trials exhibited significantly less fear than the patients who received expanding exposure trials. Thus, if the data during extinction and the immediate post-extinction test were taken as indicative of extinction of fear, Rowe and Craske would have concluded that massed exposure trials were better at reducing fear than expanding exposure trials. However, results from a generalization test to a novel spider, as well as results from a follow-up assessment 1 month later, showed that the patients who received expanding exposure trials had relatively little return of fear, whereas the patients who received massed exposure trials had significant return of fear. Thus, their results are in accordance with the current experiment, at least in indicating that performance during extinction training is not necessarily indicative of learning of the extinction contingency. The present results suggest that expanding the spacing of extinction trials will result in a faster loss of responding than uniformly spaced extinction trials. However, the expanding schedule did not alleviate recovery from extinction when the physical context was changed (i.e., renewal).

Acknowledgments

References

Abdi, H. Bonferroni and Sidak corrections for multiple comparisons. In: Salkind, NJ., editor. Encyclopedia of measurement and statistics. Thousand Oaks, CA: Sage; 2007. p. 103-107.

Balota, DA.; Duchek, JM.; Logan, JM. Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extant literature. In: Nairne, JS., editor. The foundations of remembering: Essays in honor of Henry L. Roediger, III. New York: Psychology Press; 2007. p. 83-105.

Bjork, RA.; Bjork, EL. A new theory of disuse and an old theory of stimulus fluctuation. In: Healy, AF.; Kosslyn, SM.; Shiffrin, RM., editors. Essays in honor of William K. Estes, Vol. 2: From learning processes to cognitive processes. Erlbaum; Hillsdale, NJ: 1992. p. 35-67.

Bjork, RA.; Bjork, EL. Optimizing treatment and instruction: Implications of a new theory of disuse. In: Nilsson, L.; Ohta, N., editors. Memory and society: Psychological perspectives. New York: Psychology Press; 2006. p. 116-140.

Bouton ME. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. Psychological Bulletin 1993;114:80–99. [PubMed: 8346330]

Bouton ME. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. Biological Psychiatry 2002;52:976–986. [PubMed: 12437938]

Bouton ME. Context and behavioral processes in extinction. Learning and Memory 2004;11:485–494. [PubMed: 15466298]

Bouton ME, García-Gutiérrez A, Zilski J, Moody EW. Extinction in multiple contexts does not necessarily make extinction less vulnerable to relapse. Behaviour Research and Therapy 2006;44:983–994. [PubMed: 16198302]

Brooks DC, Bouton ME. A retrieval cue for extinction attenuates spontaneous recovery. Journal of Experimental Psychology: Animal Behavior Processes 1993;19:77–89. [PubMed: 8418218]

Brooks DC, Bouton ME. A retrieval cue for extinction attenuates response recovery (renewal) caused by a return to the conditioning context. Journal of Experimental Psychology: Animal Behavior Processes 1994;20:366–379.

Chelonis JJ, Calton JL, Hart JA, Schachtman TR. Attenuation of the renewal effect by extinction in multiple contexts. Learning and Motivation 1999;30:1–14.

Cook SW, Harris RE. The verbal conditioning of the galvanic skin response. Journal of Experimental Psychology 1937;21:202–210.

Craske MG, Kircanski K, Zelikowsky M, Mystkowski J, Chowdhury N, Baker A. Optimizing inhibitory learning during exposure therapy. Behaviour Research and Therapy 2008;46:5–27. [PubMed: 18005936]

Davis M, Ressler K, Rothbaum BO, Richardson R. Effects of D-cycloserine on extinction: Translation from preclinical to clinical work. Biological Psychiatry 2006;60:369–375. [PubMed: 16919524]

Denniston JC, Chang RC, Miller RR. Massive extinction treatment attenuates the renewal effect. Learning and Motivation 2003;34:68–86.

Devenport L, Hill T, Wilson M, Ogden E. Tracking and averaging in variable environments: A transition rule. Journal of Experimental Psychology: Animal Behavior Processes 1997;23:450–460.

Drew MR, Yang C, Ohyama T, Balsam PD. Temporal specificity of extinction in autoshaping. Journal of Experimental Psychology: Animal Behavior Processes 2004;30:163–176. [PubMed: 15279508]

Erdfelder E, Faul F, Buchner A. GPOWER: A general power analysis program. Behavior Research Methods, Instruments & Computers 1996;28:1–11.

Estes WK. Statistical theory of distributional phenomena in learning. Psychological Review 1955;62:369–377. [PubMed: 13254976]

Foa EB, Cahill SP, Pontoski K. Factors that enhance treatment outcome of cognitive-behavioral therapy for anxiety disorders. CNS Spectrums 2004;9:6–17.

Gunther LM, Denniston JC, Miller RR. Conducting exposure treatment in multiple contexts can prevent relapse. Behaviour Research and Therapy 1998;36:75–91. [PubMed: 9613018]

Guthrie, ER. The psychology of learning. Rev.. Oxford, England: Harper; 1952.

Hays, MJ.; Bjork, RA. Expanding-interval retrieval practice and the Goldilocks principle. Poster presented at the annual convention of the American Psychological Association; San Francisco, CA. 2007 Augustt.

Hofmann SG. Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. Clinical Psychology Review 2008;28:200–211.

Hull CL. Conditioning: Outline of a systematic theory of learning. Yearbook of the National Society for the Study of Education 1942;41:61–95.

Hull, CL. Principles of behavior: An introduction to behavior theory. Oxford, England: Appleton-Century; 1943.

Karpicke JD, Roediger HL. Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. Journal of Experimental Psychology: Learning, Memory, and Cognition 2007;33:704–719.

Kushner MG, Kim SW, Donahue C, Thuras P, Adson D, Kotlyar M, et al. D-cycloserine augmented exposure therapy for obsessive-compulsive disorder. Biological Psychiatry 2007;62:835–838. [PubMed: 17588545]

Lang AJ, Craske MG. Manipulations of exposure-based therapy to reduce return of fear: A replication. Behaviour Research and Therapy 2000;38:1–12. [PubMed: 10645020]

Larkin MJW, Aitken MRF, Dickinson A. Retrospective revaluation of causal judgments under positive and negative contingencies. Journal of Experimental Psychology: Learning, Memory, and Cognition 1998;24:1331–1352.

Mineka S, Mystkowski JL, Hladek D, Rodriguez BI. The effects of changing contexts on return of fear following exposure therapy for spider fear. Journal of Consulting and Clinical Psychology 1999;67:599–604. [PubMed: 10450633]

Muris P, van der Heiden S, Rassin E. Disgust sensitivity and psychopathological symptoms in non-clinical children. Journal of Behavior Therapy and Experimental Psychiatry 2008;39:133–146. [PubMed: 17433251]

Myers, JL.; Well, AD. Research design and statistical analysis. 2. Mahwah, NJ: Lawrence Erlbaum; 2003.

Mystkowski JL, Craske MG, Echiverri AM. Treatment context and return of fear in spider phobia. Behavior Therapy 2002;33:399–416.

Nelson JB. Context specificity of excitation and inhibition in ambiguous stimuli. Learning and Motivation 2002;33:284–310.

Pavlov, I. Conditioned reflexes. New York: Oxford University Press; 1927.

Pedreira ME, Pérez-Cuesta LM, Maldonado H. Mismatch between what is expected and what actually occurs triggers memory reconsolidation or extinction. Learning and Memory 2004;11:579–585. [PubMed: 15466312]

Pérez-Cuesta LM, Hepp Y, Pedreira ME, Maldonado H. Memory is not extinguished along with CS presentation but within a few seconds after CS-offset. Learning and Memory 2007;14:101–108. [PubMed: 17272655]

Phelps EA. Emotion and cognition: Insights from studies of the human amygdala. Annual Review of Psychology 2006;57:27–53.

Phelps EA, O'Connor KJ, Gatenby JC, Gore JC, Grillon C, Davis M. Activation of the left amygdala to a cognitive representation of fear. Nature Neuroscience 2001;4:437–441.

Rachman SJ. The return of fear. Behaviour Research and Therapy 1979;17:164–166. [PubMed: 426747]

Rachman SJ. The return of fear: Review and prospect. Clinical Psychology Review 1989;9:147–168.

Rescorla RA. Extinction can be enhanced by a concurrent excitor. Journal of Experimental Psychology: Animal Behavior Processes 2000;26:251–260. [PubMed: 10913990]

Rescorla RA. Spontaneous recovery varies inversely with the training-extinction interval. Learning and Behavior 2004;32:401–408. [PubMed: 15825882]

Rescorla RA. Deepened extinction from compound stimulus presentation. Journal of Experimental Psychology: Animal Behavior Processes 2006a;32:135–144. [PubMed: 16634656]

Rescorla RA. Stimulus generalization of excitation and inhibition. Quarterly Journal of Experimental Psychology 2006b;59:53–67.

Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. Classical conditioning II: Current research and theory. New York: Appleton-Century-Crofts; 1972. p. 64-99.

Robbins SJ. Mechanisms underlying spontaneous recovery in autoshaping. Journal of Experimental Psychology: Animal Behavior Processes 1990;16:235–249.

Rodriguez BI, Craske MG, Mineka S, Hladek D. Context-specificity of relapse: Effects of therapist and environmental context on return of fear. Behaviour Research and Therapy 1999;37:845–862. [PubMed: 10458048]

Rosas JM, Callejas-Aguilera JE. Context switch effects on acquisition and extinction in human predictive learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 2006;32:461–474.

Rowe MK, Craske MG. Effects of an expanding-spaced vs. massed exposure schedule on fear reduction and return of fear. Behaviour Research and Therapy 1998;36:701–717. [PubMed: 9682526]

Schmidt RA, Bjork RA. New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science 1992;3:207–217.

Skinner, BF. The behavior of organisms: An experimental analysis. Appleton-Century; Oxford, England: 1938.

Smith SM, Glenberg A, Bjork RA. Environmental context and human memory. Memory and Cognition 1978;6:342–353.

Storm BC, Bjork EL, Bjork RA, Nestojko JF. Is retrieval success a necessary condition for retrieval-induced forgetting? Psychonomic Bulletin and Review 2006;13:1023–1027. [PubMed: 17484429]

Tamai N, Nakajima S. Renewal of formerly conditioned fear in rats after extensive extinction training. International Journal of Comparative Psychology 2000;13:137–146.

Teachman BA. Pathological disgust: In the thoughts, not the eye, of the beholder. Anxiety, Stress, and Coping 2006;19:335–351.

Tolman, EC. Purposive behavior in animals and men. London: Century/Random House; 1932.

Tsao JCI, Craske MG. Timing of treatment and return of fear: Effects of massed, uniform-, and expanding-spaced exposure schedules. Behavior Therapy 2000;31:479–497.

Urcelay, GP.; Lipatova, O.; Miller, RR. Generalization decrement constrains extinction in the presence of an excitor. 2008. Manuscript submitted for publication

Urcelay, GP.; Wheeler, DS.; Miller, RR. Spacing extinction trials alleviates renewal and spontaneous recovery. 2008. Manuscript submitted for publication

Vansteenwegen D, Hermans D, Vervliet B, Francken G, Beckers T, Baeyens F, et al. Return of fear in a human differential conditioning paradigm caused by a return to the original acquistion context. Behaviour Research and Therapy 2005;43:323–336. [PubMed: 15680929]

Vervliet B, Vansteenwegen D, Hermans D, Eelen P. Concurrent excitors limit the extinction of conditioned fear in humans. Behaviour Research and Therapy 2007;45:375–383. [PubMed: 16597433]

Wagner, AR. SOP: A model of automatic memory processing in animal behavior. In: Spear, NE.; Miller, RR., editors. Information processing in animals: Memory mechanisms. Hillsdale, NJ: Lawrence Erlbaum; 1981. p. 5-47.

Wagner AR. Context-sensitive elemental theory. Quarterly Journal of Experimental Psychology 2003;56B:7–29. [PubMed: 12623534]

Wagner, AR.; Brandon, SE. A componential theory of Pavlovian conditioning. In: Mowrer, RR.; Klein, SB., editors. Handbook of contemporary learning theories. Mahwah, NJ: Lawrence Elbaum; 2001. p. 23-64.

Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. Brain Research Reviews 2003;41:88–123. [PubMed: 12505650]
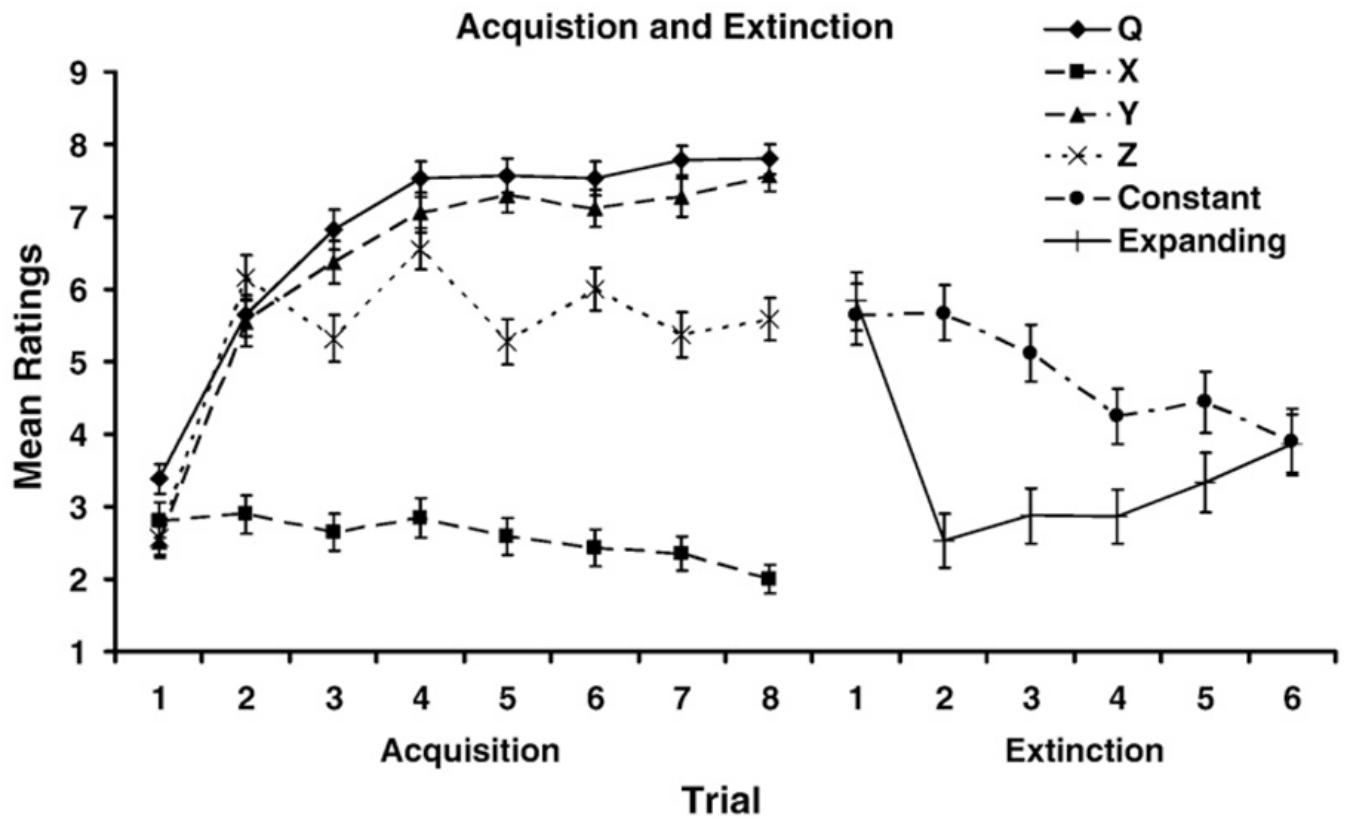
FIGURE 1.
Mean ratings for each stimulus (Q, X, Y, and Z) on the eight acquisition trials and mean ratings on each extinction trial in the Constant and Expanding groups. Error bars denote standard errors of means.
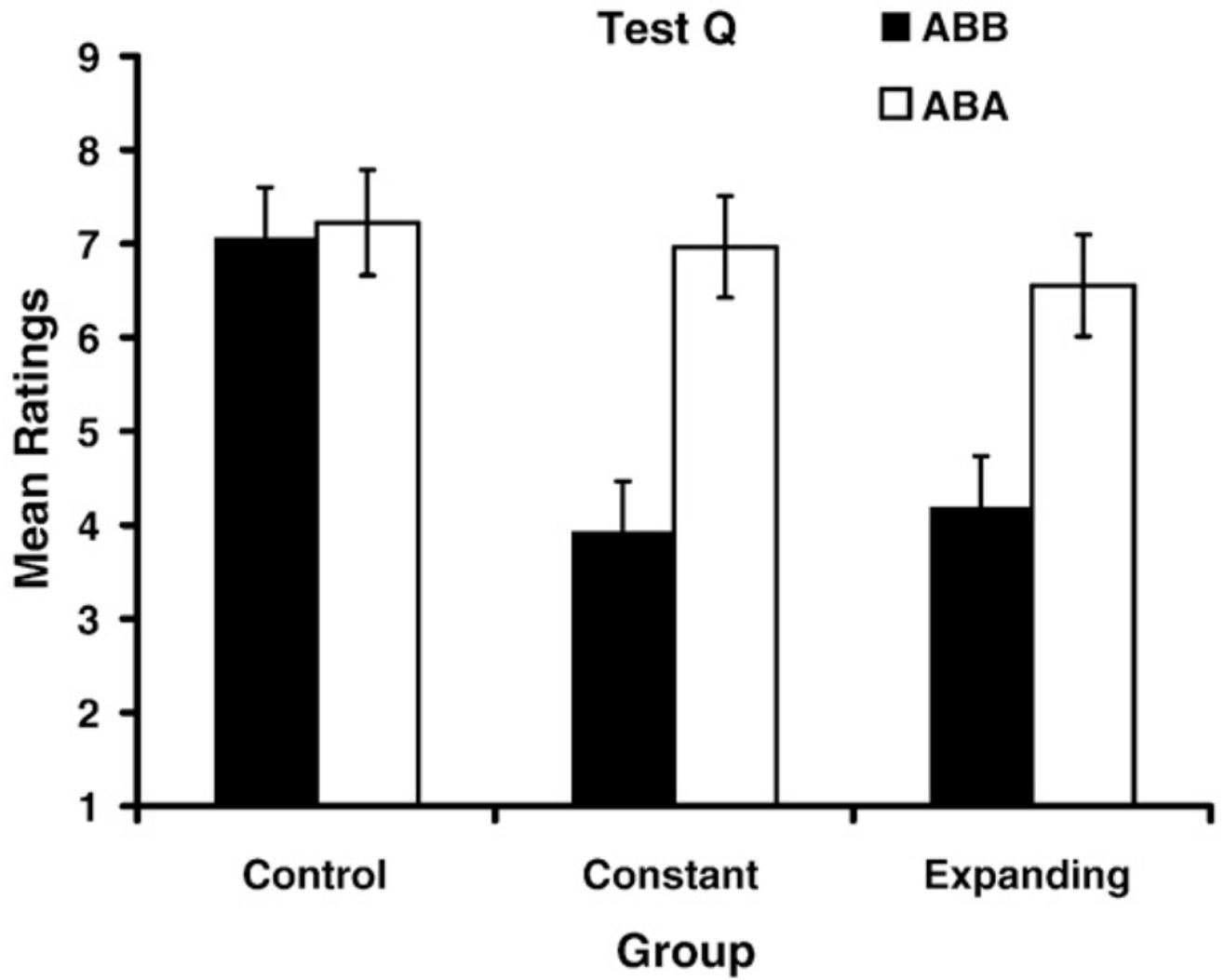
FIGURE 2.
Mean ratings for stimulus Q during test. Error bars denote standard errors of means.
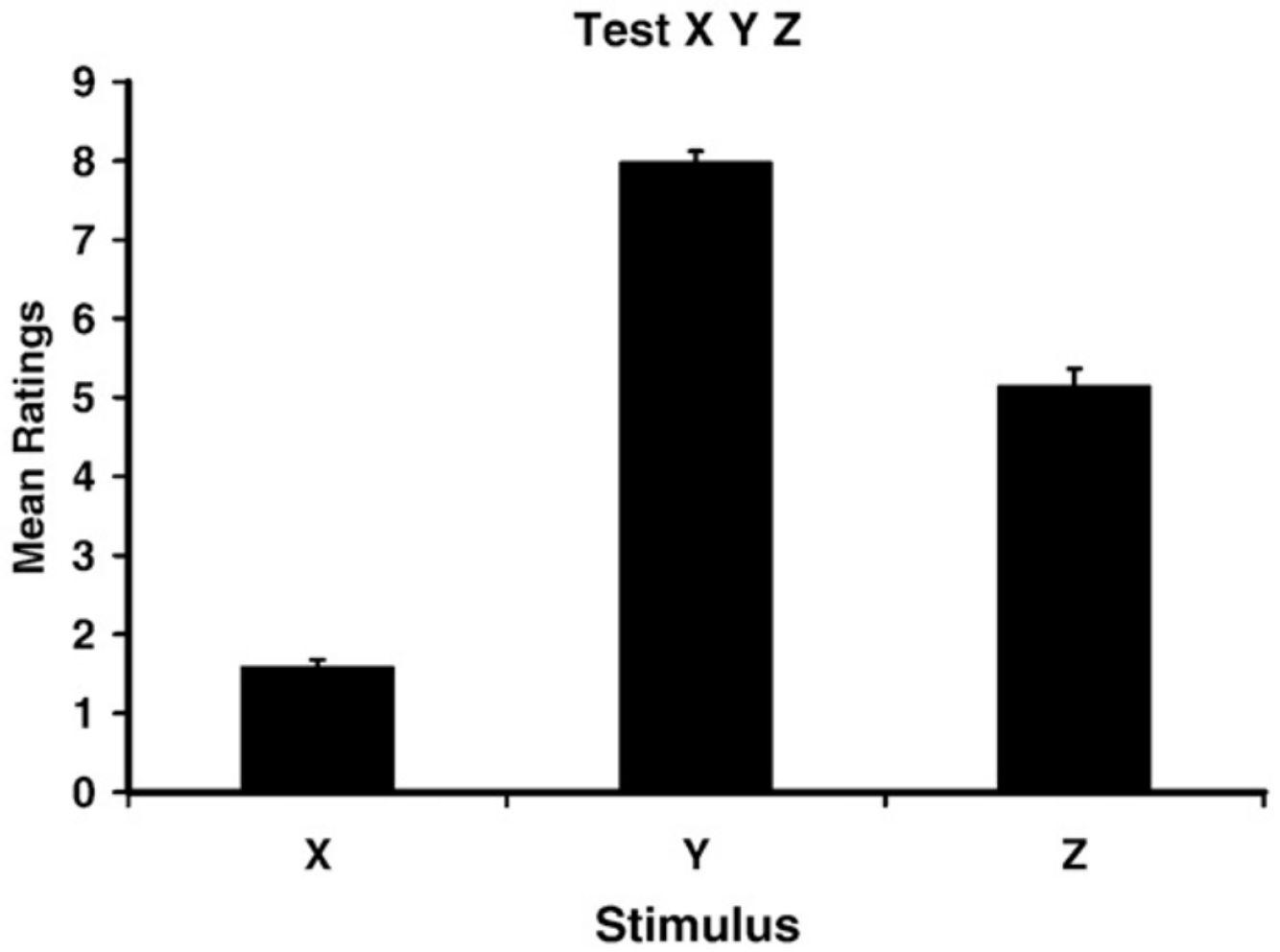
FIGURE 3.
Mean ratings for filler cues X, Y, and Z during test. Error bars denote standard errors of means.

Table 1

Design of the experiment

Group	Phase 1 (Ctx A)	Phase 2 (Ctx B)	Test Context	Test Q	Test X	Test Y	Test Z
No Extinction - ABB	8Q+	16X−, 10Y+, 10Z+/−	Ctx B	CR	cr	CR	Cr
Constant - ABB	8X−	6Q−, 10X−, 10Y+, 10Z+/−	Ctx B	cr	cr	CR	Cr
Expanding - ABB	8Y+	6Q−, 10X−, 10Y+, 10Z+/−	Ctx B	cr	cr	CR	Cr
No Extinction - ABA	8Z+/−	16X−, 10Y+, 10Z+/−	Ctx A	CR	cr	CR	Cr
Constant - ABA		6Q−, 10X−, 10Y+, 10Z+/−	Ctx A	CR	cr	CR	Cr
Expanding - ABA		6Q−, 10X−, 10Y+, 10Z+/−	Ctx A	Cr	cr	CR	Cr

Note. Q represents the target stimulus, which was potatoes; X, Y, and Z represents the filler cues, which were jello, pretzel, and bread, counterbalanced within groups; + denotes that the food cues were followed by diarrhea; − denotes that food cues were not followed by diarrhea; +/− denotes that foods were followed by diarrhea 50 % of the time. Contexts A and B were different names of restaurants *The Argentinean Horn* and *The Ukrainian Tavern,* counterbalanced within groups. The case of CR indicates the magnitude of the anticipated response based on the new theory of disuse (Bjork & Bjork, 1992, 2006). Commas separate interspersed trials.