Published in final edited form as:

Behav Ther. 2018 March ; 49(2): 212–224. doi:10.1016/j.beth.2017.07.011.

Using Retrieval Cues to Attenuate Return of Fear in Individuals With Public Speaking Anxiety

Ki Eun ShinMichelle G. Newman

Pennsylvania State University

Abstract

Even after successful exposure, relapse is not uncommon. Based on the retrieval model of fear extinction (e.g., Vervliet, Craske, & Hermans, 2013), return of fear can occur after exposure due to an elapse of time (spontaneous recovery) or change in context (contextual renewal). The use of external salient stimuli presented throughout extinction (i.e., retrieval cues [RCs]) has been suggested as a potential solution to this problem (Bouton, 2002). The current study examined whether RCs attenuated return of fear in individuals with public speaking anxiety. Sixty-five participants completed a brief exposure while presented with two RC stimuli aimed at a variety of senses (visual, tactile, olfactory, and auditory). Later, half the participants were tested for return of fear in a context different from the exposure context, and the other half in the same context. Half of each context group were presented with the same cues as in exposure, while the other half were not. Return of fear due to an elapse of time, change in context, and effects of RCs were evaluated on subjective, behavioral, and physiological measures of anxiety. Although contextual renewal was not observed, results supported effects of RCs in reducing spontaneous recovery on behavioral and physiological measures of anxiety. There was also evidence that participants who were reminded of feeling anxious during exposure by the RCs benefited more from using them at follow-up, whereas those who perceived the cues as comforting (safety signals) benefited less. Clinical implications of the findings are discussed.

Keywords

exposure; fear extinction; retrieval cues; return of fear; inhibitory learning

EXPOSURE INVOLVES REPEATEDLY CONFRONTING a feared object or situation in a safe environment with the goal of reducing fear reactions and/or learning to tolerate one's fear. It is considered a core component of treatments for anxiety-related disorders and has strong empirical support (Olatunji, Cisler, & Deacon, 2010). As a stand-alone intervention, exposure has comparable efficacy to cognitive therapy (Ougrin, 2011) and combined treatment packages such as cognitive-behavioral therapy (Deacon & Abramowitz, 2004). Nonetheless, there is still much room for improvement. Even after successful exposure,

Address correspondence to Ki Eun Shin, Department of Psychology, Pennsylvania State University, 378 Moore Building, University Park, PA 16802; kxs558@psu.edu.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

relapse is not uncommon, with some patients returning to their initial symptom levels (e.g., Rachman, 1989).

The Pavlovian fear-conditioning and extinction paradigm is one framework for understanding exposure mechanisms (e.g., Mineka & Zinbarg, 2006). According to an extinction-based model (Bouton, 2004), anxiety-related disorders involve an association between feared stimuli and aversive outcomes/threats. Exposure leads to the development of a secondary safety association via pairing the fear stimulus with lack of aversive outcome. This new association competes with the preexisting threat association rather than entirely replaces it (e.g., Bouton, 1993; Vervliet et al., 2013). Therefore, success of exposure depends on both the development of a safety association (i.e., inhibitory learning; Craske et al., 2008) and sustained inhibition of the threat association by the new safety association. A return of fear occurs if the threat association becomes more salient than the safety association for the extinguished stimulus.

Several factors can weaken inhibition of the threat association by the safety association. In both animals and humans, return of fear can occur due to an elapse of time following fear extinction (spontaneous recovery; Rescorla, 2004), confronting fear stimuli in a context different from the extinction context (contextual renewal; Bouton, 2004), and experiencing an unexpected aversive event that elicits fear after extinction (reinstatement; Rescorla & Heth, 1975). These phenomena have been observed in animals (e.g., Bouton, 2002), nonanxious humans (e.g., Milad, Orr, Pitman, & Rauch, 2005), and clinical analogues (e.g., Mystkowski, Craske, & Echiverri, 2002). This is in line with the retrieval model of extinction (Bouton, 1994, 2002), which posits that the safety association is bound to the extinction context. Change in contextual factors from the extinction context including time, physical environment, and internal emotional states can interfere with retrieval of the safety association.

Studies have examined various strategies that mitigate return of fear by enhancing retrieval of the safety association relative to the threat association. One such strategy is using retrieval cues (RCs), extra stimuli uniquely presented throughout fear extinction (e.g., Brooks & Bouton, 1993). In theory, RCs can take various forms, including sound (e.g., buzzer; Bouton, 1994), visual cues (e.g., "&" symbol; Dibbets, Havermans, & Arntz, 2008), or tactile cues (e.g., a uniquely textured pencil; Collins & Brandon, 2002). Through pairing with exposure, RCs can help individuals recover the safety association acquired during extinction. In rats, RCs have reduced return of conditioned appetitive responses caused by time lapse (Brooks & Bouton, 1993) or by change in physical context (Brooks & Bouton, 1994). RCs also reduced spontaneous recovery of other conditioned responses including taste aversion (Brooks, Palmatier, Garcia, & Johnson, 1999), alcohol tolerance (Brooks, Vaughn, Freeman, & Woods, 2004), and alcohol reactivity (Willcocks & McNally, 2014). Effects of these cues depended on their association with extinction and were distinguished from the effects of cues associated with fear acquisition.

RCs also have been tested in nonclinical human samples using computerized fear conditioning and extinction tasks (e.g., Dibbets et al., 2008; Vansteenwegen et al., 2006). Results were promising with RCs reducing return of a conditioned fear response caused by

change in context. As in animal research, effects of RCs were not confined to fear, generalizing to reactivity to alcohol cues in social drinkers (Collins & Brandon, 2002).

To our knowledge, there have been only three studies on the utility of RCs in analogue clinical samples: public speaking anxiety (Culver, Stoyanova, & Craske, 2011; Laborda et al., 2016) and spider phobia (Dibbets, Moor, & Voncken, 2013). Results indicated either lack of evidence (Dibbets et al., 2013; Laborda et al., 2016) or weak effects (Culver et al., 2011). One potential reason for these results is insufficient salience of RC stimuli. If the cues were not salient enough, participants may have failed to encode them as a part of the extinction context. In Culver et al. (2011) and Laborda et al. (2016), visual and tactile cues were used (e.g., unique textured pen, neon-colored clipboard), and Dibbets et al. (2013) used a bracelet with a distinct color and citronella odor. It is noteworthy that these studies used stimuli involving two sensory modalities. Nonetheless, it might have helped to recruit even more diverse sensory systems. Providing multiple sensory cues may have enhanced salience as well as increased the probability that one of the cues was successfully paired with exposure.

Another factor that could have led to weak or nonsignificant findings is attentional bias. Pathological, excessive anxiety has been linked to hypervigilance toward threatening stimuli (e.g., Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). During exposure, participants may have selectively focused on threat stimuli (e.g., spider for spider phobia) and failed to attend to RCs. In addition, some procedural factors may have weakened the effects of RCs. For instance, in Culver et al. (2011), two of the three cues were presented only prior to exposure. Because the cues were not presented throughout exposure, participants may not have associated them with fear extinction. Laborda et al. (2016) presented RCs during baseline and postexposure assessments as well as exposure, and Dibbets et al. (2013) gave participants bogus information about the RC to avoid participants' suspicion. These procedures could have led to an antecedent or extraneous association with RCs, interfering with forming the target association between cues and exposure. In addition, in Dibbets et al. (2013), participants wore the bracelet throughout the 1-week period between exposure and follow-up testing and were asked to remind themselves of their successful exposure whenever they saw the bracelet. Such approach may increase ecological validity given its applicability to actual practice (e.g., assigned homework to carry RCs). Nonetheless, wearing the cue outside the session may have led to additional associations with the cue, diluting the strength of the association with exposure.

Another limitation of prior studies is that they did not assess valence of RCs. There is preliminary evidence that RC effects may depend on their valence. In a study of nonanxious humans (Dibbets & Maes, 2011), cues rated as positively valenced (e.g., smiley face symbol) attenuated contextual renewal, but negatively valenced cues (e.g., hash sign) did not. Weak or nonsignificant effects in prior studies may have been due to perception of the cues as negatively or neutrally valenced. In addition, only one of the three studies (Dibbets et al., 2013) asked participants whether the cues helped them to retrieve the exposure memory. Given the proposed function of RCs, it is important to assess retrieval of the exposure memory directly.

Taken together, it remains unclear whether RCs can be effective and how best to use them to maximize their efficacy. The present study aimed to tackle these questions and extend the sparse literature on clinical utility of RCs with an analogue sample of people with public speaking anxiety. RCs were selected to activate multiple sensory systems (visual, tactile, olfactory, and auditory) to ensure their salience. In addition, cues were presented during exposure in a way intended to increase participants' engagement with and attention to them (e.g., holding a cue in hand). The cues were presented throughout exposure, and no cover story was provided on the purpose or nature of them to avoid creating any extraneous associations. We also assessed their salience, valence, and which exposure features the cues reminded participants of during the follow-up test. We predicted that RCs would attenuate spontaneous recovery and contextual renewal of fear.

Methods

PARTICIPANTS

Sixty-five public speaking anxious participants (50 females, 15 males) were recruited through mass screening of introductory psychology classes. Mean age was 19 years ($SD = 2.05$); 70.8% identified as Caucasian, 16.9% as Asian, 9.2% as African American, and 3.1% as biracial. Screening was based on a two-item measure previously used to select individuals with fears of public speaking (e.g., Tsao & Craske, 2000). Participants rated how anxious they would feel giving a formal speech in front of a live audience and how likely they would be to avoid a class that required an oral presentation on a 0–8 scale ranging from 0 (*none/never*) to 8 (*extremely/always*). Those scoring 6 or higher on both items were recruited. Anybody with a heart, respiratory, or neurological condition or with fragrance allergies was excluded.

Participants were randomly assigned to one of four groups: a group completing the follow-up test in a context different from the exposure context with RCs (DR), a group completing the follow-up test in a different context without RCs (DN), a group completing the follow-up test in the same context as the exposure context with RCs (SR), and a group completing the follow-up test in the same context without RCs (SN). Number of participants in each group was roughly equal: DR = 17, DN = 15, SR = 17, and SN = 16. There were no group differences in gender or ethnicity distribution, $\chi^2(3, N = 65) = .12, p = .99$; $\chi^2(9, N = 65) = 4.92, p = .89$. A one-way ANOVA also indicated no group differences in age, $F(3, 61) = .19, p = .90$.

MEASURES

Social Phobia Scale (SPS) and the Self-Statements During Public Speaking Scale (SSPS)—Participants completed the SPS (Mattick & Clarke, 1998) and the SSPS (Hofmann & DiBartolo, 2000) at baseline. The time frame was general (SPS: “Based on your experience, indicate the degree to which you feel the statement is true of you”; SSPS: “what you have typically felt”).

The SPS is a 20-item measure of anxiety about being scrutinized by others in a social setting. Items are rated on a 5-point scale from 0 (*not at all characteristic or true of me*) to 4

(*extremely characteristic or true of me*). It has shown high internal consistency ($\alpha = .87-.94$, $\alpha = .92-.93$ in the current sample), good retest reliability ($r = .66-.93$; Heimberg, Mueller, Holt, Hope, & Liebowitz, 1992; Mattick & Clarke, 1998), and convergent and discriminant validity. The cutoff score of 24 correctly classified 73% of a social phobia group (Heimberg et al., 1992).

The SSPS is a 10-item measure that assesses fearful thoughts during public speaking. It has two five-item subscales, Positive Self-Statements (SSPS-P) and Negative Self-Statements (SSPS-N). Participants rated the degree to which they agreed with each item on a 6-point scale from 0 (*do not agree at all*) to 5 (*agree extremely*). Both subscales had high internal consistency (SSPS-N: $\alpha = .87-.90$ in the current sample, SSPS-P: $\alpha = .80-.83$ in the current sample), and 3-month retest reliability (SSPS-N = .80, SSPS-P = .78; Hofmann & DiBartolo, 2000). Social phobia patients had a mean score of 12.3 on SSPS-N and 13.4 on SSPS-P, and score differences between social phobia and nonclinical samples were significant (Hofmann & DiBartolo, 2000).

Behavioral Avoidance Test (BAT)—Participants completed a BAT at pre- and post-exposure, and 1-week follow-up. They drew two topics from a bag and chose one to present. Topic options included capital punishment, gun control, legalizing marijuana, and TV violence. Chosen topics were removed to avoid having participants speak on the same topic in later BATs. Each speech occurred in front of a recording camera with an experimenter present. Participants were told that their speech would be video recorded and later evaluated for quality. Participants spoke for as long as they could up to 3 minutes. Speech duration was used as a behavioral measure of anxiety.

Subjective Units of Distress Scale (SUDS)—Subjective anxiety was measured with the 100-point SUDS (Wolpe, 1973) anchored by 0 (*no fear*), 25 (*mild fear*), 50 (*moderate fear*), 75 (*severe fear*), and 100 (*very severe fear*). Participants were prompted to provide their SUDS at the beginning and at each 1-minute interval during the BAT. Those who stopped before 3 minutes provided SUDS at the end of their speech. Two indices of subjective anxiety—mean SUDS (averaged across all ratings) and max SUDS (highest of the ratings)—were calculated.

Heart Rate (HR)—HR was also measured during each BAT, using an ambulatory monitoring device (Polar RS800CX; Polar Electro Oy, Kempele, Finland). The Polar monitor provides measurements comparable to an electrocardiogram (ECG; e.g., Terbizan, Dolezal, & Albano, 2002) and has been widely used in studies investigating physiological processes in anxiety (e.g., Alvares et al., 2013). It records HR at a sampling rate of 1,000 Hz. Participants wore an elastic belt around their chest and a wristwatch receiver for HR recording. Mean HR during the BAT was used as the physiological measure of anxiety.

STIMULI

Contexts—Two rooms were used as contextual stimuli. The rooms differed in size, shape, color of the walls, decoration, and furnishings. Room A was arranged as a meeting room, with white walls and watercolor paintings hung on the walls. There were beige sofa chairs

lined on the left and right sides of the room as well as a table, a desk, and a desktop computer. On the table, there were fake plants and flowers. Room B was smaller in size and used partially as a storage place. The color of the walls was black, and the room included a freezer, a mini refrigerator, several older televisions, and three black padded chairs. The two rooms were chosen to maximize distinctiveness between the contexts.

Retrieval Cues—A puffer ball with a unique texture and bright neon green color and a diffuser generating white noise and emitting a peppermint scent were used as RCs. The ball was 2.5" × 2" in size, small enough to grab in one hand. It had soft, rubbery spikes on its surface, which created a distinct texture. A peppermint scent was presented using essential oil. In order to increase participants' engagement with the cues, participants were asked to grab the green ball in their dominant hand throughout exposure and the follow-up BAT. The diffuser was also placed near the participant (within 1.6 ft. or .5 m) to increase salience of the olfactory and auditory cues.

RC Manipulation Check—After the follow-up BAT, participants rated the pleasantness of RCs (green puffer ball, a diffuser generating a peppermint scent and white noise) on a 1 (*very unpleasant*) to 5 (*very pleasant*) Likert scale. Those presented with RCs at follow-up (RC conditions) were also asked if they recognized the cues from Session 1 in a yes/no format and what RCs reminded them of related to the Session 1 exposure in an open-ended format. Based on responses participants provided, we coded the responses into two themes: being reminded of anxiety during exposure, or perceiving the cues as safety signals.

PROCEDURE

The study procedure was similar to Culver et al.'s (2011). Session 1 included baseline and pre-exposure assessment, an exposure session, and post-exposure assessment. Session 2 was the 1-week follow-up assessment.

In Session 1, participants completed informed consent and the SPS and SSPS. Baseline anxiety was assessed using SUDS, and baseline HR was collected for 5 minutes while participants sat in a chair. Participants then received instructions and completed the pre-exposure BAT in a separate room. A 2-minute break followed the BAT.

Next, participants received instructions for exposure including a brief rationale for exposure therapy and fear extinction. Exposure consisted of giving four 5-minute speeches in front of a small audience and took place in the same room where participants completed the pre-exposure BAT. For each speech, participants drew three topics out of a bag and chose one to present. Topic options included same-sex marriage, ban on public speaking, abortion, euthanasia, animal testing, and nuclear power. The selected topic was removed from the bag to ensure different topics for each exposure speech. Speeches were given while standing in front of a three-person audience including the experimenter. Two other audience members were research personnel who were instructed to remain professional (e.g., do not laugh or make jokes) and to make constant eye contact with participants. In order to activate participants' fears sufficiently, each audience member had a clipboard, and made ratings for quality of the speech while participants gave their speech. Throughout exposure, all participants were presented with RCs. They were told to hold in their hand tightly the

uniquely textured green puffer ball and were presented continuously with a peppermint scent and white noise. The same stimuli were used as RCs at the 1-week follow-up BAT only for those in the DR and SR conditions.

Participants were prompted to provide SUDS ratings at baseline and at each 1-minute interval during exposure speeches. To minimize interference, they were asked to make their SUDS ratings quickly and continue with the speech. HR was recorded continuously during the speech. When participants stopped short of 5 minutes, they were encouraged to continue. There was a 2-minute break between each exposure speech. Participants were instructed to sit and relax for the first minute, and the last minute was used to select a subsequent speech topic.

After the last exposure speech, participants took a 2-minute break and returned to the room to complete the post-exposure BAT. SUDS ratings and HR were collected using the same procedure as in the pre-exposure BAT. Participants' second visit took place 6–8 days later.

In Session 2, baseline HR was collected during a 5-minute resting period. Participants then completed the follow-up BAT. Half the participants used the same room (e.g., Room A) as in previous BATs and exposure. The other half used a different room to create a context change. To control for potential carryover effects as well as effects of room types, half of the participants in each group (same context, different context) completed the first session in Room A, and the other half started with Room B. In addition, half of each context group were presented with RCs (green puffer ball, peppermint scent, and white noise) throughout the Session 2 BAT, and the other half were not. As before, SUDS ratings and HR were collected during the BAT. Following the BAT, participants completed the RC manipulation check. Participants were also debriefed and compensated with course credits for participation.

ANALYTIC STRATEGY

Power analyses were calculated for each a priori model to assess power of the current sample. Using the *simr* package (Green & MacLeod, 2016) in R, we ran Monte Carlo simulation, which is considered the gold standard for power analyses (Abraham & Russell, 2008; Arnold, Hogan, Colford, & Hubbard, 2011). Because Monte Carlo analyses involve a large number of randomly simulated data samples from specified populations, researchers can empirically evaluate sampling distributions for effects of interest. Our a priori effect-size target was Cohen's $d = 0.5$ (Thomas, 1997). We conducted 1,000 random simulations each for main effects of Time, Time $\times$ Context interaction, Time $\times$ RC interaction, and Time $\times$ Context $\times$ RC interaction. Results showed that we had 90.7% power, 95% CI [90.13, 90.94], to detect a main effect of Time; 83.7% power, 95% CI [82.01, 85.13], to detect a Time $\times$ Context interaction; 82.6% power, 95% CI [80.54, 84.42], to detect a Time $\times$ RC interaction; and 67.4% power, 95% CI [64.36, 70.35], to detect a Time $\times$ Context $\times$ RC interaction at a .05 significance level. Because the current sample was not fully powered to detect a three-way interaction of a medium effect size, results for the Time $\times$ Context $\times$ RC interaction should be interpreted with caution.

Using SPSS, we ran MANOVAs to examine differences between context and RC conditions on baseline symptom scores, pre-exposure BAT measures, and anxiety experienced during exposure. Main analyses used the R (version 3.3.1; R Core Team, 2015) packages lme4 (Bates, Machler, Bolker, & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2015). Multilevel modeling was used to account for the nesting of repeated measures within individuals. Level 1 predictor was time, and Level 2 predictors were context and RC conditions. Outcomes were BAT measures (pre-, post-exposure, and follow-up). Since both mean and max SUDS from the BAT showed a similar pattern of results, we report only results on max SUDS here, a commonly used metric in experimental exposure studies. Changes from pre- to post-exposure and post-exposure to follow-up were examined separately to test effects of exposure as well as return of fear and effects of RCs. Based on results from baseline models, the intercept was used as a random effect in all models. Fixed effects included intercept, time, context, RC condition, context $\times$ RC interaction, interaction between time and each condition variable, and three-way interactions between time, context, and RC. In the case of a significant interaction, we conducted simple slopes analyses to compare changes in each condition. For all models, Cohen's d and its confidence interval were calculated following the procedures in Feingold (2009, 2015).

Additional exploratory subgroup analyses were conducted for the RC manipulation check data. Using only data from the RC condition, we ran multilevel models to determine whether pleasantness of RCs and participant perceptions of the cues at follow-up (no memory of the cues from Session 1, being reminded of anxiety during exposure, or perceiving the cues as safety signals) affected change in BAT outcomes from post-exposure to follow-up. The three perception variables were dummy coded such that an endorsed perception was rated 1, and non-endorsed responses were rated 0. In multilevel models, the intercept was modeled as a random effect, and fixed effects included the intercept, time, a variable from the manipulation check, and interaction between time and the variable. Because these subgroup analyses are likely underpowered, the results need to be interpreted with caution.

Results

BASILINE SYMPTOMS AND ANXIETY

At baseline, participants reported symptoms of public speaking and social anxiety comparable to clinical levels on the SPS and SSPS (see Table 1 for descriptives collapsed across context groups). For symptom measures, a MANOVA indicated neither baseline differences in context group, $F(4, 57) = .58, p = .68, \eta_p^2 = .04, 95\% \text{ CI } [.001, .11]$; RC, $F(4, 57) = .07, p = .99, \eta_p^2 = .001, 95\% \text{ CI } [.00002, .002]$; nor context $\times$ RC interaction, $F(4, 57) = 1.51, p = .21, \eta_p^2 = .10, 95\% \text{ CI } [.002, .20]$. On pre-exposure max SUDS ratings, HR, and speech duration, there were also no main effects of context, $F(4, 53) = 1.43, p = .24, \eta_p^2 = .10, 95\% \text{ CI } [.002, .21]$; RC, $F(4, 53) = .21, p = .93, \eta_p^2 = .02, 95\% \text{ CI } [.001, .04]$; or context $\times$ RC interaction, $F(4, 53) = .94, p = .45, \eta_p^2 = .07, 95\% \text{ CI } [.002, .16]$. Group differences did not emerge in subjective and physiological levels of anxiety during exposure either for context, $F(3, 58) = 1.46, p = .24, \eta_p^2 = .07, 95\% \text{ CI } [.004, .18]$; RC, $F(3, 58) = .83,$

$p = .48$, $\eta_p^2 = .04$, 95% CI [.002, .13]; or context $\times$ RC, $F(3, 58) = .64$, $p = .59$, $\eta_p^2 = .03$, 95% CI [.001, .11].

EFFECTS OF EXPOSURE AT POST-TREATMENT

When pre- to post-exposure change was examined, a main effect of time showed that exposure was efficacious, leading to a decrease in anxiety on max SUDS,¹ $t(61) = -5.25$, $p < .001$, $d = -1.27$, 95% CI [-1.91, -.99]; speech duration, $t(61) = 5.18$, $p < .001$, $d = 1.15$, 95% CI [.73, 1.59]; and HR, $t(56) = -2.50$, $p = .02$, $d = -0.52$, 95% CI [-.87, -.21]. There were no significant Time $\times$ Context interactions across outcomes on max SUDS, $t(61) = .57$, $p = .57$, $d = 0.19$, 95% CI [-.14, 1.15]; speech duration, $t(61) = -.19$, $p = .85$, $d = -0.06$, 95% CI [-.66, .57]; and HR, $t(56) = -.31$, $p = .76$, $d = -0.09$, 95% CI [-.64, .41]. Similarly, three-way interactions between time, context, and RC were not significant on max SUDS, $t(61) = .12$, $p = .91$, $d = 0.06$, 95% CI [-1.62, .23]; speech duration, $t(61) = .15$, $p = .99$, $d = 0.07$, 95% CI [-.96, .78]; and HR, $t(59) = .46$, $p = .65$, $d = 0.26$, 95% CI [-.87, 1.35].

The Time $\times$ RC interaction was not significant on max SUDS, $t(61) = .05$, $p = .96$, $d = 0.06$, 95% CI [-.13, 1.18]; and speech duration, $t(61) = -.75$, $p = .46$, $d = -0.24$, 95% CI [-.85, .41]. However, there was an unexpected Time $\times$ RC interaction on HR, $t(58) = -2.33$, $p = .02$, $d = -0.72$, 95% CI [-1.27, -.14]. This was likely a spurious effect because the manipulation of RCs did not occur until follow-up testing. Based on simple slopes analyses, both RC, $p = -11.30$, $t(54) = -2.50$, $p = .02$, $d = -0.68$, 95% CI [-1.16, -.20]; and no-RC groups, $p = -27.15$, $t(54) = -5.33$, $p < .001$, $d = -1.45$, 95% CI [-2.01, -.87], showed significant reductions in HR, but the effect was greater in the no-RC group. Figure 1 presents the results collapsed across context groups. Nonetheless, an ANOVA showed no significant difference in post-exposure HR between RC and no-RC groups, $F(1, 57) = .26$, $p = .61$, $\eta_p^2 = .01$, 95% CI [.001, .09]. Therefore, despite a difference in change in HR, both groups ended up with similar levels of HR at post-exposure.

RETURN OF FEAR: EFFECTS OF CONTEXT AND RETRIEVAL CUES AT FOLLOW-UP

On max SUDS, the main effect of time for the post to follow-up change did not indicate a significant return of fear, $t(61) = 1.05$, $p = .30$, $d = 0.17$, 95% CI [-.11, .43]. However, speech duration and HR showed a significant post to follow-up return of fear, $t(61) = -5.25$, $p < .001$, $d = -1.27$, 95% CI [-1.69, -.77]; $t(56) = 4.23$, $p < .001$, $d = 0.92$, 95% CI [.48, 1.32]. Overall, results suggest that gains from exposure were maintained at 1-week follow-up on the subjective measure, but return of fear occurred on speech duration and HR.

Contextual renewal was not present from post to follow-up based on nonsignificant Time $\times$ Context interactions across outcome measures on max SUDS, $t(61) = -1.26$, $p = .21$, $d = -0.29$, 95% CI [-.70, .15]; speech duration, $t(61) = -.36$, $p = .72$, $d = -0.08$, 95% CI [-.50, .34]; and HR, $t(56) = 1.21$, $p = .23$, $d = 0.32$, 95% CI [-.18, .86].

¹As noted in the Methods section, we report results only on max SUDS, and not mean SUDS, for the sake of brevity. However, the pattern of results was the same for mean and max SUDS.

Despite no significant return of fear, there was a Time $\times$ RC interaction on the post to follow-up change on max SUDS, $t(61) = -2.21$, $p = .03$, $d = -0.52$, 95% CI $[-.95, -.10]$. Whereas the no-RC group showed reduction in anxiety, $p = -7.00$, $t(59) = -2.05$, $p = .04$, $d = -0.53$, 95% CI $[-1.05, -.10]$, the RC group did not, $p = 3.35$, $t(59) = 1.05$, $p = .30$, $d = 0.27$, 95% CI $[-.21, .72]$. However, comparison of max SUDS between the RC conditions at follow-up did not reveal a significant difference, $F(1, 61) = .83$, $p = .37$, $\eta_p^2 = .01$, 95% CI $[-.001, .12]$. Therefore, despite the difference in change on max SUDS, both groups showed similar ending levels of anxiety.

With respect to speech duration, which did show return of fear at follow-up, RCs significantly reduced the post to follow-up return of fear, $t(61) = -2.14$, $p = .04$, $d = -0.67$, 95% CI $[-.97, -.28]$. Based on simple slopes analysis, the no-RC group showed a significant post to follow-up increase in anxiety, $p = -30.33$, $t(59) = -3.99$, $p < .001$, $d = -1.04$, 95% CI $[-1.58, -.49]$. However, the RC group showed no such increase, $p = -10.77$, $t(59) = -1.51$, $p = .13$, $d = -0.39$, 95% CI $[-.85, .08]$. The Time $\times$ RC interaction was also significant for post to follow-up change on HR, $t(55) = 2.25$, $p = .03$, $d = 1.10$, 95% CI $[.44, 1.99]$. Based on simple slopes analyses, there was a significant post to follow-up return of fear in the no-RC group, $p = 19.61$, $t(54) = 3.77$, $p < .001$, $d = 1.03$, 95% CI $[.48, 1.57]$, but not in the RC group, $p = 3.69$, $t(54) = .77$, $p = .45$, $d = 0.21$, 95% CI $[-.26, .68]$. Figure 2 presents effects of RC on duration and HR collapsed across context groups. In sum, RCs prevented a return of fear on speech duration and HR.

Three-way interactions between time, context, and RC for the post to follow-up change were not significant across outcome measures on max SUDS, $t(61) = .12$, $p = .91$, $d = 0.06$, 95% CI $[-.81, .97]$; speech duration, $t(61) = .35$, $p = .73$, $d = 0.15$, 95% CI $[-.71, 1.02]$; and HR, $t(61) = .46$, $p = .65$, $d = .26$, 95% CI $[-.87, 1.38]$.

SUBJECTIVE RESPONSES TO RETRIEVAL CUES

Pleasantness of Cues—Participants perceived the puffer ball ($M = 3.81$, $SD = .71$) and the diffuser-generating peppermint scent and white noise ($M = 3.75$, $SD = .90$) as somewhat more pleasant than neutral. RC and no-RC conditions did not differ in pleasantness ratings, $F(2, 56) = 1.15$, $p = .32$, $\eta_p^2 = .04$, 95% CI $[.002, .14]$. In the RC condition, perceived valence of the cues did not significantly affect post-exposure to follow-up change, ball, on max SUDS, $t(30) = -.70$, $p = .49$, $d = -0.15$, 95% CI $[-.51, .11]$; duration, $t(30) = 1.26$, $p = .22$, $d = 0.23$, 95% CI $[-.16, .39]$; HR, $t(27) = -.99$, $p = .33$, $d = -0.33$, 95% CI $[-.74, .24]$; scent and noise on max SUDS, $t(30) = -1.19$, $p = .25$, $d = -0.21$, 95% CI $[-.62, .22]$; duration, $t(30) = 1.26$, $p = .23$, $d = 0.29$, 95% CI $[-.11, .56]$; and HR, $t(26) = .87$, $p = .39$, $d = 0.24$, 95% CI $[-.24, .92]$.

Memory and Perceived Effects of RCs—At follow-up BAT, 9 out of 34 participants in the RC condition reported that they did not recognize that the same cues were present during exposure. However, explicit memory of the cues did not significantly affect the post-exposure to follow-up change on max SUDS, $t(31) = .16$, $p = .87$, $d = 0.06$, 95% CI $[-.56, .58]$; speech duration, $t(31) = -.66$, $p = .51$, $d = -0.21$, 95% CI $[-.80, .38]$; and HR, $t(27) = -.88$, $p = .39$, $d = -0.38$, 95% CI $[-1.18, .45]$.

Of the remaining 25 participants in the RC condition who recognized the cues from Session 1, 10 participants indicated being reminded of feeling anxious during the exposure (e.g., “I was reminded of how anxious I felt”) to an open-ended question on what they remembered from the RCs. Another 10 said that the RCs made them feel more comfortable and relaxed. Five did not specify what they were reminded of by the cues. Compared with the other 24 participants in the RC condition, the 10 individuals who reported being reminded of feeling anxious by the cues showed *less* return of fear from post-exposure to follow-up at a marginal level on max SUDS, $t(31) = -2.05$, $p = .05$, $d = -0.60$, 95% CI $[-1.02, -.01]$; duration, $t(31) = 1.90$, $p = .07$, $d = 0.48$, 95% CI $[-.01, .91]$; and HR, $t(29) = -1.95$, $p = .06$, $d = -0.89$, 95% CI $[-1.71, .02]$. However, contrary to their report, the 10 participants who *perceived* the cues as safety signals did not show less return of fear from post-exposure to follow-up at a significant level compared with the other 24 participants on max SUDS, $t(31) = 1.50$, $p = .15$, $d = 0.54$, 95% CI $[-.11, .86]$, and HR, $t(28) = -.12$, $p = .91$, $d = -0.06$, 95% CI $[-.82, .86]$. Despite nonsignificance, the effect size for max SUDS was medium, suggesting potentially *greater* return of fear in participants who perceived the cues as safety signals. The safety perception of the cues led to greater return of fear for speech duration at a marginal level, $t(31) = -1.91$, $p = .07$, $d = -.50$, 95% CI $[-.98, .02]$.

Discussion

The current study examined whether RCs attenuated return of fear after exposure in individuals with public speaking anxiety. Exposure was effective at reducing anxiety. Only spontaneous recovery, but not contextual renewal, was evidenced. RCs reduced spontaneous recovery on behavioral and physiological measures of anxiety, the only measures that showed a return of fear over time. On subjective measures of anxiety, return of fear was not evidenced, and anxiety level at follow-up was not significantly different between those presented with RCs and those who were not. Perceived valence or explicit memory of RCs did not significantly influence the cues' effects. However, there was preliminary evidence that participants for whom the cues were a reminder of their anxiety during exposure showed better outcomes than others, whereas those who perceived the cues as safety signals demonstrated greater return of fear than others.

Relative to prior studies that examined effects of RCs in clinical analogues (Culver et al., 2011; Dibbets et al., 2013; Laborda et al., 2016), the current study found stronger support for effects of RCs in attenuating return of fear. For example, RCs significantly reduced return of fear on the behavioral measure of anxiety (speech duration) with a small to large effect size based on the confidence interval. This result is encouraging because behavioral change and its maintenance are an important treatment goal in exposure therapy for anxiety disorders. RCs also significantly reduced return of fear on the physiological measure (HR) with a medium to large effect size. However, it is important to note a spurious effect in which exposure led to greater pre- to post reduction in HR in the no-RC group than the RC group. This result might be due to the fact that physiological data often does not align with behavioral or self-report data in experimental studies on exposure (e.g., Emmelkamp & Felten, 1985). It is also possible that the act of squeezing a ball during the speech increased physiological reactivity for the RC group relative to the no-RC group. Given the spurious effect, the finding on HR might be due to regression to the mean, with the no-RC group

having more room to show return of fear than the RC group. This interpretation, however, is contradicted by the nonsignificant difference in mean HR at post-exposure between the RC and no-RC groups. Thus, the interpretation that RCs attenuated spontaneous recovery of fear at a physiological level may be warranted.

The current study may have found stronger support for RCs than prior studies due to strategies to increase cue salience. For instance, cues were used to activate more diverse sensory systems (four vs. two) than in previous studies (Culver et al., 2011; Dibbets et al., 2013; Laborda et al., 2016). Also, we included an olfactory cue (in addition to visual, auditory, and tactile cues) which has been shown to elicit emotional memories more effectively than auditory or visual cues at both subjective (Chu & Downes, 2002) and neural levels (Herz, Eliassen, Beland, & Souza, 2004). Only one prior study (Dibbets et al., 2013) used an olfactory cue and did not find significant effects. However, this might be due to procedural factors that could have limited cue potency (e.g., presenting bogus information about the cues). We also instructed participants to engage with the cues (e.g., pressing a ball in hand) to enhance their salience. These strategies might have facilitated pairing of the cues with exposure. Evidence for this idea is that although 26% of our retrieval group did not remember the cues, in the one other study that assessed RC salience, 78% did not remember the cues at follow-up (Dibbets et al., 2013). It is also possible that our stronger effects were due to a larger sample size than prior studies. Compared with the 65 participants in our study, Culver et al. (2011) had a maximum of 40 participants (across three studies and ranging from 18 to 40), Dibbets et al. (2013) had 33 participants, and Laborda et al. (2016) had 38 participants.

The hypothesized effects of RCs were supported only on behavioral and physiological measures of anxiety, but not subjective measures. However, it is important to note that unlike speech duration and HR, SUDS ratings did not show return of fear from post-exposure to follow-up. Thus, at a subjective level, participants retrieved the safety association with the feared situation without help of RCs. Such desynchrony between different components of fear responses (subjective, physiological, and behavioral) is a well-documented phenomenon (e.g., Rachman, 1976). Given the proposed function of RCs in facilitating retrieval of the safety association, the absence of subjective return of fear might have reduced sensitivity to detect effects of RCs.

Some researchers have suggested that subjective fear might be elicited through distinct mechanisms from nonconscious, physiological, and behavioral expressions of fear (e.g., LeDoux, 2014). Associative learning involved in fear conditioning/extinction does not presume conscious awareness. For instance, people showed conditioned autonomic fear responses despite lacking awareness of being conditioned (e.g., Bechara, Tranel, Damasio, & Adolphs, 1995) or when conditioned threats were presented subliminally (e.g., Olsson & Phelps, 2004). The amygdala, a brain area implicated in fear conditioning, is also activated in response to threats even without conscious awareness of threats (Liddell et al., 2005). Thus, subjective, physiological, and behavioral components of fear might be amenable to different interventions. For instance, inhibitory learning that occurs at an implicit level might be sufficient to change physiological and behavioral responses to a feared stimulus. However, reduction in subjective fear might require modification of explicit threat

expectancies. Similarly, RCs might reduce return of physiological or behavioral fear responses even when cues are not explicitly linked to exposure, whereas such strategy (e.g., explaining the cues as a reminder of how participants overcame their fear during exposure) might be helpful for preventing return of subjective fear.

The current study extended prior studies by investigating how perceptions of RCs affected outcomes. In the present sample, perceived valence of RCs did not affect return of fear. This is inconsistent with a prior finding from a nonanxious sample that more positively valenced cues attenuated return of fear better than negatively valenced cues (Dibbets & Maes, 2011). The nonsignificant result might be due to a restricted range of scores. About 80% of the sample rated the cues as neutral to pleasant (three to four out of zero to five). On the one hand, this speaks to successful standardization of stimuli. On the other hand, such restricted range would have limited statistical power to detect an effect of valence compared with the previous approach of preselecting distinctly positively and negatively valenced stimuli (Dibbets & Maes, 2011).

Explicit memory of the presence of RCs during exposure also did not affect return of fear. This is consistent with the finding that RCs attenuated return of fear only at physiological and behavioral levels. If attenuation occurred through implicit associative learning between cues and fear extinction, explicit recognition of the cues may not have been essential.

One attribute of RCs that did affect outcomes was whether the cues reminded participants of their anxiety during exposure. Being more reminded of anxiety led to less return of fear at a marginal level across outcome measures. This is contrary to the theorized function of RCs, to be more associated with extinction than the experience of anxiety. Several explanations can be considered. First, even when RCs were successfully associated with inhibitory learning at an implicit level, participants may have recalled feeling anxious at a conscious level because the experience of anxiety was more salient. In this regard, being reminded of anxiety by RCs might be a proxy for effective pairing of the cues with fear extinction. Second, based on the inhibitory learning model, which posits that successful exposure does not depend on fear reduction, other factors such as sustained fear arousal may have been important (e.g., Craske et al., 2008). Sustained arousal may increase salience of the threat stimulus during exposure, providing more opportunities for new learning to occur. Participants who recalled feeling anxious by RCs might have experienced sustained anxiety during exposure, which could have facilitated extinction learning. Also consistent with this model, the cues may have reminded these individuals of tolerating their anxiety. Thus, RCs may have helped participants retrieve a sense of distress tolerance efficacy. Unfortunately, no participants spontaneously mentioned this, but future studies might ask about a distress tolerance association more directly.

A subset of participants presented with RCs at follow-up perceived them as safety signals (e.g., relaxing). However, those participants benefited less from RCs than others who did not perceive the cues as safety signals. By definition, safety signals should lessen fear responding to any threat-related stimulus. Therefore, the cues did not function in the way we traditionally view safety signals despite participants' report. As noted above, an association with distress tolerance may be most important for RCs to function as safety signals. The

current findings also indicate that therapists may need to employ direct behavioral testing rather than rely on clients' reports when assessing whether RCs became safety signals. In addition, regardless of actual behavioral outcomes, it still might be important to address clients' subjective perception of the cues as safety signals because such perception may induce psychological dependency on the cues and decrease clients' self-efficacy in coping with their anxiety.

The current findings have implications for using RCs in clinical practice. First, it may be advantageous to use multimodal sensory stimuli to increase cue salience. However, there is a potential trade-off of increased client burden. In theory, after completing exposure, clients should have RCs available for times when they confront extinguished feared stimuli. If multiple sensory cues are used, it can be cumbersome or infeasible to carry them around. One solution might be using a single cue designed to engage different senses simultaneously (e.g., a neon ball with a unique texture and peppermint scent). Such cue might serve as a combination of several single-modal cues. In addition, mentally rehearsing the exposure experience and context has been shown to reduce return of fear in spider phobia (Mystkowski, Craske, Echiverri, & Labus, 2006). Mental rehearsal might be easier to implement than physical RCs. One way to extend the present findings is to test alternate forms of RCs (e.g., a single multimodal cue vs. multiple unimodal cues; abstract vs. physical) to identify types of cues that are both effective and practical.

Another issue to consider is the risk that people develop dependence on RCs as safety signals. For nonanxious controls, RCs functioned as safety signals when response to unextinguished fear stimuli was tested in the exposure context (Dibbets et al., 2008). Also, positively valenced, but not negatively valenced, cues became safety signals (Dibbets & Maes, 2011). When RCs became safety signals, their effects generalized to nonextinguished fear stimuli. However, since the goal of treatment is to help clients cope with their anxiety better and increase their self-efficacy in doing so, relying on safety signals can be counter-therapeutic. Therefore, it may be most beneficial to restrict the usage of RCs to a short-term period following exposure and encourage clients to cope with their anxiety independently in the long term.

The present study had several limitations. First, manipulation of contexts did not elicit renewal of fear. Although the two contexts were chosen to maximize distinctiveness, and similar manipulations led to contextual renewal in a previous study (Culver et al., 2011), more extensive variations of contextual stimuli may be needed. For instance, some prior studies included variation of the experimenter or BAT materials (e.g., Mystkowski et al., 2006).

Another limitation is the use of relatively brief exposure. It remains an empirical question whether the results generalize to a treatment-seeking sample undergoing regular-length exposure therapy. Nonetheless, previous studies showed that longer one-time exposure can be as effective as multiple-exposure sessions for circumscribed fears (e.g., Ost, 1989). Therefore, results from the current study merit replication with a longer single-exposure or multiple-exposure sessions.

It is important to note that even small effects can be meaningful for a low-cost, low-burden intervention like retrieval cues (e.g., Glass, McGaw, & Smith, 1981). However, inspection of our confidence intervals suggests lack of power in the current sample to detect small effects reliably. Despite being larger than prior studies of RCs in clinical analogues, our sample size was still small. Thus, some of our null findings may be due to insufficient power to detect small effects. Our power analysis also indicated less than ideal power to detect a three-way interaction of a medium effect size. Therefore, findings on three-way interactions and from exploratory subgroup analyses should be interpreted with caution. At the same time, it is noteworthy that all three-way interactions had small effect sizes ($< .3$) and were nonsignificant. In the current study, the effects of RCs were of medium to large sizes, but replication is needed, using a larger sample with sufficient power for detecting small effects.

The current sample was a clinical analogue and relatively homogeneous with most participants identifying as female and Caucasian. The gender imbalance is consistent with a higher prevalence rate of social anxiety disorder in females than males (e.g., Xu et al., 2012), but the findings await replication in treatment-seeking and more diverse samples. In addition, we did not include a control group for whom RCs were not present during exposure. Concurrent presentation of external stimuli throughout extinction may lessen extinction efficacy regardless of whether the stimuli are inhibitory or excitatory (e.g., Lovibond, Davis, & O'Flaherty, 2000). Although exposure led to a significant reduction in anxiety across measures in the current study, it would be worthwhile for future studies to evaluate how the presence of RCs affects exposure efficacy. Future research should also investigate boundary conditions (e.g., reminding of anxiety during exposure) for effectiveness of RCs to optimize their use in clinical practice.

In sum, the current study found support for effects of RCs in reducing spontaneous recovery of fear after exposure in individuals with public speaking anxiety. Effects were specifically observed in behavioral and physiological measures of anxiety, but not in subjective measures. There was also preliminary evidence that participants who were reminded of feeling anxious during exposure by the cues benefited more from using them, whereas those who perceived the cues as safety signals ironically benefited less. These findings indicate the importance of ensuring salience of the RCs for maximizing their efficacy and suggest potential distinction in effects of RCs for implicit and explicit processes of anxiety.

References

- Abraham WT, & Russell DW (2008). Statistical power analysis in psychological research. *Social and Personality Psychology Compass*, 2, 283–301. 10.1111/j.1751-9004.2007.00052.x
- Alvares GA, Quintana DS, Kemp AH, Van Zwieten A, Balleine BW, Hickie IB, & Guastella AJ (2013). Reduced heart rate variability in social anxiety disorder: Associations with gender and symptom severity. *PLoS ONE*, 8, e70468 10.1371/journal.pone.0070468 [PubMed: 23936207]
- Arnold BF, Hogan DR, Colford JM Jr., & Hubbard AE (2011). Simulation methods to estimate design power: An overview for applied research. *BMC Medical Research Methodology*, 11, 94 10.1186/1471-2288-11-94 [PubMed: 21689447]
- Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, & van IJzendoorn MH (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, 133, 1–24. 10.1037/0033-2909.133.L1 [PubMed: 17201568]

- Bates D, Machler M, Bolker B, & Walker S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. 10.18637/jss.v067.i01
- Bechara A, Tranel D, Damasio H, & Adolphs R (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, 269, 1115–1116. 10.1126/science.7652558 [PubMed: 7652558]
- Bouton ME (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114, 80–97. 10.1037/0033-2909.114.1.80 [PubMed: 8346330]
- Bouton ME (1994). Context, ambiguity, and classical conditioning. *Current Directions in Psychological Science*, 3, 49–53. 10.1111/1467-8721.ep10769943
- Bouton ME (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52, 976–986. 10.1016/S0006-3223(02)01546-9 [PubMed: 12437938]
- Bouton ME (2004). Context and behavioral processes in extinction. *Learning and Memory*, 11, 485–494. 10.1101/lm.78804 [PubMed: 15466298]
- Brooks DC, & Bouton ME (1993). A retrieval cue for extinction attenuates spontaneous recovery. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 77–90. 10.1037/0097-7403.19.1.77
- Brooks DC, & Bouton ME (1994). A retrieval cue for extinction attenuates response recovery (renewal) caused by a return to the conditioning context. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 366–381. 10.1037/0097-7403.20.4.366
- Brooks DC, Palmatier MI, Garcia EO, & Johnson JL (1999). An extinction cue reduces spontaneous recovery of a conditioned taste aversion. *Animal Learning and Behavior*, 27, 77–88. 10.3758/BF03199433
- Brooks DC, Vaughn JM, Freeman AJ, & Woods AM (2004). An extinction cue reduces spontaneous recovery of ataxic ethanol tolerance in rats. *Psychopharmacology*, 176, 256–265. 10.1007/s00213-004-1882-y [PubMed: 15164156]
- Chu S, & Downes JJ (2002). Proust nose best: Odors are better cues of autobiographical memory. *Memory and Cognition*, 30, 511–518. 10.3758/BF03194952 [PubMed: 12184552]
- Collins BN, & Brandon TH (2002). Effects of extinction context and retrieval cues on alcohol cue reactivity among nonalcoholic drinkers. *Journal of Consulting and Clinical Psychology*, 70, 390–397. 10.1037/0022-006X.70.2.390
- Craske MG, Kircanski K, Zelikowsky M, Mystkowski J, Chowdhury N, & Baker A (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46, 5–27. <http://dx.doi.org/10.1016/j.brat.2007.10.003> [PubMed: 18005936]
- Culver NC, Stoyanova M, & Craske MG (2011). Clinical relevance of retrieval cues for attenuating context renewal of fear. *Journal of Anxiety Disorders*, 25, 284–292. 10.1016/j.janxdis.2010.10.002
- Deacon BJ, & Abramowitz JS (2004). Cognitive and behavioral treatments for anxiety disorders: A review of meta-analytic findings. *Journal of Clinical Psychology*, 60, 429–441. 10.1002/jclp.10255
- Dibbets P, Havermans R, & Arntz A (2008). All we need is a cue to remember: The effect of an extinction cue on renewal. *Behaviour Research and Therapy*, 46, 1070–1077. 10.1016/j.brat.2008.05.007 [PubMed: 18675398]
- Dibbets P, & Maes JH (2011). The effect of an extinction cue on ABA-renewal: Does valence matter? *Learning and Motivation*, 42, 133–144. 10.1016/j.lmot.2010.12.003
- Dibbets P, Moor C, & Voncken MJ (2013). The effect of a retrieval cue on the return of spider fear. *Journal of Behavior Therapy and Experimental Psychiatry*, 44, 361–367. 10.1016/j.jobtep.2013.03.005
- Emmelkamp PM, & Felten M (1985). The process of exposure in vivo: Cognitive and physiological changes during treatment of acrophobia. *Behaviour Research and Therapy*, 23, 219–223. 10.1016/0005-7967(85)90034-8 [PubMed: 2860891]
- Feingold A (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, 14, 43–53. 10.1037/a0014699 [PubMed: 19271847]

- Feingold A (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting and Clinical Psychology*, 83, 157. 10.1037/a0037721
- Glass GV, McGaw B, & Smith ML (1981). *Meta-analysis in social research*. Thousand Oaks, CA: SAGE.
- Green P, & MacLeod CJ (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493–498. 10.1111/2041-210x.12504
- Heimberg RG, Mueller GP, Holt CS, Hope DA, & Liebowitz MR (1992). Assessment of anxiety in social interaction and being observed by others: The Social Interaction Anxiety Scale and the Social Phobia Scale. *Behavior Therapy*, 23, 53–73. 10.1016/S0005-7894(05)80308-9
- Herz RS, Eliassen J, Beland S, & Souza T (2004). Neuroimaging evidence for the emotional potency of odor-evoked memory. *Neuropsychologia*, 42, 371–378. 10.1016/j.neuropsychologia.2003.08.009 [PubMed: 14670575]
- Hofmann SG, & DiBartolo PM (2000). An instrument to assess self-statements during public speaking: Scale development and preliminary psychometric properties. *Behavior Therapy*, 31, 499–515. 10.1016/S0005-7894(00)80027-1 [PubMed: 16763666]
- Kuznetsova A, Brockhoff PB, & Christensen RHB (2015). lmerTest: Tests in linear mixed effects models. R package version 2.0–29. Retrieved from <http://CRAN.Rproject.org/package=lmerTest>.
- Laborda MA, Schofield CA, Johnson EM, Schubert JR, George-Denn D, Coles ME, & Miller RR (2016). The extinction and return of fear of public speaking. *Behavior Modification*, 40(6), 901–921. 10.1177/0145445516645766 [PubMed: 27118054]
- LeDoux JE (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences*, 111, 2871–2878. 10.1073/pnas.1400335111
- Liddell BJ, Brown KJ, Kemp AH, Barton MJ, Das P, Peduto A, ... Williams LM (2005). A direct brainstem-amygdala-cortical “alarm” system for subliminal signals of fear. *Neuroimage*, 24, 235–243. 10.1016/j.neuroimage.2004.08.016 [PubMed: 15588615]
- Lovibond PF, Davis NR, & O’Flaherty AS (2000). Protection from extinction in human fear conditioning. *Behaviour Research and Therapy*, 38, 967–983. 10.1016/S0005-7967(99)00121-7 [PubMed: 11004736]
- Mattick RP, & Clarke JC (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36, 455–470. 10.1016/S0005-7967(97)10031-6 [PubMed: 9670605]
- Milad MR, Orr SP, Pitman RK, & Rauch SL (2005). Context modulation of memory for fear extinction in humans. *Psychophysiology*, 42, 456–464. 10.1111/j.1469-8986.2005.00302.x [PubMed: 16008774]
- Mineka S, & Zinbarg R (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It’s not what you thought it was. *American Psychologist*, 61, 10–26. 10.1037/0003-066X.61.L10 [PubMed: 16435973]
- Mystkowski JL, Craske MG, & Echiverri AM (2002). Treatment context and return of fear in spider phobia. *Behavior Therapy*, 33, 399–416. 10.1016/S0005-7894(02)80035-1
- Mystkowski JL, Craske MG, Echiverri AM, & Labus JS (2006). Mental reinstatement of context and return of fear in spider-fearful participants. *Behavior Therapy*, 37, 49–60. <http://dx.doi.org/10.1016/j.beth.2005.04.001> [PubMed: 16942960]
- Olatunji BO, Cisler JM, & Deacon BJ (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics of North America*, 33, 557–577. 10.1016/j.psc.2010.04.002 [PubMed: 20599133]
- Olsson A, & Phelps EA (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychological Science*, 15, 822–828. 10.1111/j.0956-7976.2004.00762.x [PubMed: 15563327]
- Ost LG (1989). One-session treatment for specific phobias. *Behaviour Research and Therapy*, 27, 1–7. 10.1016/0005-7967(89)90113-7 [PubMed: 2914000]
- Ougrin D (2011). Efficacy of exposure versus cognitive therapy in anxiety disorders: Systematic review and meta-analysis. *BMC Psychiatry*, 11, 200. 10.1186/1471-244X-11-200 [PubMed: 22185596]

- R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Rachman S (1976). The passing of the two-stage theory of fear and avoidance: Fresh possibilities. *Behaviour Research and Therapy*, 14, 125–131. 10.1016/0005-7967(76)90066-8 [PubMed: 7228]
- Rachman S (1989). The return of fear: Review and prospect. *Clinical Psychology Review*, 9, 147–168. 10.1016/0272-7358(89)90025-1
- Rescorla RA (2004). Spontaneous recovery. *Learning and Memory*, 11, 501–509. 10.1101/lm.77504 [PubMed: 15466300]
- Rescorla RA, & Heth CD (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 1, 88. 10.1037/0097-7403.1.1.88
- Terbizan DJ, Dolezal BA, & Albano C (2002). Validity of seven commercially available heart rate monitors. *Measurement in Physical Education and Exercise Science*, 6, 243–247.
- Thomas L (1997). Retrospective power analysis. *Conservation Biology*, 11, 276–280. 10.1046/j.1523-1739.1997.96102.x
- Tsao JCI, & Craske MG (2000). Timing of treatment and return of fear: Effects of massed, uniform-, and expanding-spaced exposure schedules. *Behavior Therapy*, 31, 479–497. 10.1016/S0005-7894(00)80026-X
- Vansteenwegen D, Vervliet B, Hermans D, Beckers T, Baeyens F, & Eelen P (2006). Stronger renewal in human fear conditioning when tested with an acquisition retrieval cue than with an extinction retrieval cue. *Behaviour Research and Therapy*, 44, 1717–1725. 10.1016/j.brat.2005.10.014 [PubMed: 16457776]
- Vervliet B, Craske MG, & Hermans D (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology*, 9, 215–248. 10.1146/annurev-clinpsy-050212-185542
- Willcocks AL, & McNally GP (2014). An extinction retrieval cue attenuates renewal but not reacquisition of alcohol seeking. *Behavioral Neuroscience*, 128, 83. 10.1037/a0035595 [PubMed: 24512068]
- Wolpe J (1973). *The practice of behavior therapy*. Oxford, UK: Pergamon Press.
- Xu Y, Schneier F, Heimberg RG, Princiisvalle K, Liebowitz MR, Wang S, & Blanco C (2012). Gender differences in social anxiety disorder: Results from the national epidemiologic sample on alcohol and related conditions. *Journal of Anxiety Disorders*, 26, 12–19. 10.1016/j.janxdis.2011.08.006 [PubMed: 21903358]

FIGURE 1.
Mean values and standard errors of the means for pre- and post-exposure Behavioral Avoidance Test heart rate by retrieval cue condition collapsed across context conditions.

FIGURE 2.

Mean values and standard errors of the means for post-exposure and follow-up Behavioral Avoidance Test speech duration and heart rate by retrieval cue condition collapsed across context conditions.

Table 1

Means and Standard Deviations for Outcomes at Pre-, Post-exposure, and 1-Week Follow-Up by Retrieval Cue Condition

		Pre-exposure	Post-exposure	Follow-up
Mean SUDS	RC	39.40 (18.33)	19.88 (16.31)	20.43 (12.71)
	No RC	40.46 (16.93)	25.87 (20.61)	22.69 (19.14)
Max SUDS	RC	47.94 (20.82)	24.65 (17.63)	25.15 (14.59)
	No RC	48.55 (18.08)	28.68 (21.93)	26.77 (21.39)
Speech duration	RC	92.85 (38.76)	143.65 (44.58)	131.06 (47.07)
	No RC	97.84 (51.98)	136.03 (43.14)	108.97 (49.94)
Heart rate	RC	103.44 (17.53)	90.53 (16.14)	98.47 (18.60)
	No RC	106.21 (26.54)	88.52 (12.51)	105.39 (18.30)
SPS	RC	24.88 (13.83)	—	21.85 (14.52)
	No RC	21.17 (11.77)	—	20.65 (12.05)
SSPS-P	RC	12.65 (4.99)	—	12.38 (4.55)
	No RC	11.77 (5.72)	—	11.13 (5.47)
SSPS-N	RC	8.24 (6.31)	—	6.59 (5.78)
	No RC	8.37 (5.26)	—	7.26 (6.00)

Note. RC = presence of RCs at the follow-up testing; No RC = absence of RCs at the follow-up testing; SPS = Social Phobia Scale; SSPS-P = Self-Statements During Public Speaking Scale-Positive Self-Statements Subscale; SSPS-N = Self-Statements During Public Speaking Scale-Negative Self-Statements Subscale.