Conditioned fear extinction and reinstatement in a human fear-potentiated startle paradigm

Seth D. Norrholm,1,2,3 Tanja Jovanovic,1,2,3 Bram Vervliet,4,5 Karyn M. Myers,3 Michael Davis,1,3,6 Barbara O. Rothbaum,1 and Erica J. Duncan1,2,3

1 Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia 30322, USA; 2 Atlanta Veterans Affairs Medical Center, Mental Health Service Line 116A, Decatur, Georgia 30033, USA; 3 Center for Behavioral Neuroscience, Atlanta, Georgia 30302, USA; 4 Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium; 5 Social and Behavioral Sciences, University of Amsterdam, Roetersstraat, 1018 WB Amsterdam, The Netherlands; 6 Department of Psychology, Emory University, Atlanta, Georgia 30322, USA

The purpose of this study was to analyze fear extinction and reinstatement in humans using fear-potentiated startle. Participants were fear conditioned using a simple discrimination procedure with colored lights as the conditioned stimuli (CSs) and an airblast to the throat as the unconditioned stimulus (US). Participants were extinguished 24 h after fear conditioning. Upon presentation of unsigned USs after extinction, participants displayed significant fear reinstatement. In summary, these procedures produced robust fear-potentiated startle, significant CS+/CS− discrimination, within-session extinction, and significant reinstatement. This is the first demonstration of fear extinction and reinstatement in humans using startle measures.

A failure to inhibit fear is believed to underlie the pathophysiology of several anxiety disorders, including post-traumatic stress disorder (PTSD) (Cannistraro and Rauch 2003). One methodology for investigating fear inhibition is extinction, a form of learning in which the frequency and/or intensity of a conditioned response (CR), previously acquired through pairing a conditioned stimulus (CS; e.g., light) with an aversive unconditioned stimulus (US; e.g., shock), is reduced through the repeated presentation of the CS without the US (cf. Myers and Davis 2002). Evidence suggests that fear to a CS is not erased after extinction, but instead suppressed by a competing parallel inhibitory process (cf. Myers and Davis 2002), as demonstrated by the return of fear following extinction under specific conditions.

One example of fear return after extinction is reinstatement (Rescorla and Heth 1975; Bouton and Swartzentruber 1991). In a traditional reinstatement procedure, subjects undergo fear acquisition and extinction training, and are then presented with a small number of USs without the CS. In a later test session, the conditioned fear response will reappear upon presentation of the CS (Rescorla and Heth 1975; Bouton and Bolles 1979; Richard et al. 1999b; Myers and Davis 2002), provided that the test occurs in the context in which the unsigned USs were presented (Westbrook et al. 2002). Unsigned US presentations are believed to reinstate extinguished CRs through multiple mechanisms, including summation of fear conditioned to the context by the unsigned USs with subthreshold fear to the extinguished CS (Bouton and Bolles 1979; Bouton and King 1983). The context specificity of reinstatement has been demonstrated in animal and human studies (see Bouton 2004; LaBar and Phelps 2005).

Conditioned fear extinction and reinstatement have been widely investigated in animals by observing fear responses such as freezing, avoidance, or fear-potentiated startle (Falls et al. 1992; Gewirtz et al. 1997; Richardson et al. 1999a; Lu et al. 2001; Walker and Davis 2002; Westbrook et al. 2002; Ledgerwood et al. 2003, 2004; Chihwatl et al. 2005; Myers et al. 2006). In humans, reinstatement has been observed in verbal ratings of fear and US-expectancy (Hermans et al. 2005), re- action time task performance (Dritikl et al. 2004), and skin conductance (LaBar and Phelps 2005; Milad et al. 2005). These demonstrations are very important, as they provide a direct link between animal models of associative learning and potential applications in clinical settings. Clinically, the process of reinstatement may underlie the symptom exacerbation experienced by anxiety disorder patients as a result of re-exposure to trauma-related stimuli or negative life events (Steketee 1993; Wade et al. 1993).

The present study serves as a “translational bridge” and is the first to use fear-potentiated startle to examine extinction and reinstatement in humans. Fear-potentiated startle is the relative increase in the magnitude of the acoustic startle reflex when elicited in the presence of a CS previously paired with an aversive US. It is well-suited for translational studies of fear (Davis 1992; Grillon and Davis 1997; Ameli et al. 2001; Grillon and Baas 2003) as it provides an objective measure of fear responses, generates a non-zero baseline, offers cross-species generalization, and is mediated by a well-characterized neuronal system (Davis 1997). The establishment of a fear-potentiated startle paradigm for studying extinction and reinstatement in humans will provide a powerful tool for assessing the abnormalities in these processes associated with fear-related disorders. In addition, establishing such a paradigm will lay the groundwork for assessing the effects of psycho- pharmacological agents on these processes in people (e.g., D-cycloserine) (Walker et al. 2002; Ledgerwood et al. 2003, 2004, 2005).

All subjects received the same experimental procedures during Acquisition. During Acquisition, subjects showed significant fear-potentiated startle and discrimination between lights A and B (BLOCK × TRIAL TYPE interaction, F16,264 = 9.43, P < 0.001, Fig. 1A). Fear acquisition was confirmed by US-expectancy responses of DANGER or SAFETY following A and B trials, respectively (F1,31) = 1147, P < 0.001; Fig. 1B).

There were no significant differences between subjects in the Control and Reinstatement groups regarding magnitude of fear-potentiated startle, rate of acquisition, discrimination be-
between the CS+ and CS−, or expectancy ratings during Acquisition.

All subjects received the same experimental procedures during Extinction (Blocks 1–5 on Day 2). While startle magnitude was significantly decreased across trial types during Extinction, there was a BLOCK × TRIAL TYPE interaction (Blocks 1–5; \( F_{1,43} = 16.49, P < 0.001 \); Fig. 2A). A marked decrement in Difference Scores on A trials indicated within-session extinction of fear-potentiated startle during Extinction (Blocks 1–5; \( F_{1,43} = 16.49, P < 0.001 \); Fig. 2B). The degree of within-session extinction did not differ between the Control and Reinstatement groups.

Subject expectancy ratings on the first presentation of each trial type during Extinction demonstrated retention of the Acquisition contingency (\( F_{1,41} = 189, P < 0.001 \)) (Fig. 2C). Ratings of DANGER on A trials were significantly extinguished during Extinction (Blocks 1–5; \( F_{1,40} = 130, P < 0.001 \)) (Fig. 2C). The Reinstatement group did not extinguish DANGER ratings on A trials as rapidly as the Control group (\( F_{1,39} = 5.56, P < 0.05 \)) (Fig. 3C,D); however, the terminal degree of within-session extinction did not differ between the two groups (Expectancy ratings for light A during Trial 20; \( F_{1,40} = 3.22, P > 0.05 \)).

At the conclusion of Extinction, one group received three unsignaled USs (Reinstatement, \( n = 22 \)) and a second group experienced an equivalent passage of time without the USs (Control, \( n = 22 \)). Twenty nonreinforced presentations of light A did not produce significant within-session extinction in all subjects and, as such, it was difficult to assess reinstatement in all subjects because of a ceiling effect. For this reason, we did an a posteriori assessment of reinstatement in a subset of subjects displaying at least 50% extinction (Difference Score to light A in Block 1 vs. Block 5). Subjects who met the 50% extinction criterion were referred to as “Extinguishers.” Extinguishers in the Control (\( n = 15 \)) and Reinstatement (\( n = 11 \)) groups displayed a mean within-session extinction decrement of 80% and 91%, respectively. (Difference Scores, Control: \( F_{1,14} = 10.54, P < 0.05 \); Reinstatement: \( F_{1,10} = 23.45, P = 0.001 \)) (Fig. 3, A and B, respectively). The degree of within-session extinction did not differ between Extinguishers in the Control and Reinstatement groups.

Extinguishers who received three unsignaled airblasts after Extinction (Reinstatement) showed a significant return of fear-potentiated startle (BLOCK × TRIAL TYPE interaction, \( F_{2,20} = 5.40, P = 0.011 \)) (Fig. 3B). Extinguishers who experienced an equivalent passage of time after Extinction (Control) did not show any reinstatement (Fig. 3A). In addition, Extinguishers in the Reinstatement group displayed increased startle responses to light A (and not to light B or NA) after the unsignaled USs compared with Extinguishers in the Control group (\( F_{1,25} = 4.21, P < 0.05 \)) (Fig. 3A, B).

In contrast to the incomplete extinction of fear-potentiated startle, US-expectancy ratings of DANGER on A trials were nearly completely extinguished by the end of Extinction in the Control and Reinstatement groups. This greater degree of extinction on US-expectancy allowed us to assess reinstatement in all subjects using this measure. In the Reinstatement group, US-expectancy ratings significantly increased on A trials following three unsignaled USs (Trial 20 vs. 21; \( F_{1,18} = 6.15, P < 0.05 \)) (Fig. 3D). As expected, expectancy ratings in the Control group did not change between Trials 20 and 21 (Fig. 3C). In addition, we observed similar results when we examined the US-expectancy data from Extinguishers only (data not shown).

The primary aim of this study was to develop a protocol for future testing of psychopharmacological manipulations on human fear extinction and reinstatement (e.g., D-cycloserine) (see Walker et al. 2002; Ledgerwood et al. 2003, 2004, 2005). These procedures show promise as a psychophysiological tool for achieving this aim. With regard to fear-potentiated startle, we observed robust fear conditioning, significant CS+/CS− discrimination, and prolonged development of within-session extinction. The lack of complete extinction is ideal for future studies with pharmacological extinction-enhancing agents in that the probability of floor effects is reduced. Incomplete extinction did, however, prevent the analysis of reinstatement in some subjects. Thus, we observed reinstatement effects in a subset of subjects exhibiting at least 50% within-session extinction as measured by Difference Score. The protocol used in the present study may require slight modifications, depending on the goal of future investigations (e.g., modifying the time intervals between each phase to allow for drug administration or increasing the number of nonreinforced CS presentations during Extinction).

With regard to subject reports of US-expectancy, we observed successful fear conditioning, distinct retention of the CS+/CS− contingency, and significant within-session extinction. In contrast to the incomplete within-session extinction observed with fear-potentiated startle measures, the Control and Reinstatement groups displayed near complete extinction of US-expectancy. In addition, US-expectancy ratings were increased in the Reinstatement group after the unsignaled USs. The apparent dissociation between subject expectancy and startle measures is consistent with that reported by Hermans et al. (2005), who showed that fear ratings and psychophysiological measures are not always consistent. It is important to note that we have found no evidence in this or any other study to indicate that making a US expectancy rating can affect startle responses.
may be perceived as aversive and, as such, serve as a secondary US. However, our data indicate that potentiated startle responses to the CS− decrease during the course of the acquisition phase (Fig. 1C) even though these were followed by startle probes. The high level of fear-potentiated startle to the CS− in Block 1 more probably represented initial generalization between the CSs, often seen early in discrimination paradigms.

In conclusion, this is the first observation of conditioned fear extinction and reinstatement using fear-potentiated startle in humans. The development of a fear-potentiated startle paradigm for the analysis of conditioned fear extinction and reinstatement has significant clinical implications for understanding the underlying pathophysiology of fear-related psychiatric illnesses and for the screening of potential pharmacotherapies that may facilitate exposure therapy and prevent relapse.

Forty-five subjects (25 females/20 males with mean age of 29.4 ± 1.4 yr) participated after signing an informed consent form approved by the Emory University Institutional Review Board and the Atlanta VAMC Research and Development Committee. Inclusion requirements included corrected 20/20 vision without color blindness (assessed by an eye chart) and tone detection at 30 dB [A] SPL at frequencies ranging from 250 to 4000 Hz (assessed with a pure threshold audiometer, Grason-Stadler, Model GS1710). In addition, subjects were screened and excluded for current or past psychiatric illness and for current drug use by urine toxicology.

The eyblink component of the acoustic startle response was measured according to previously published methods (Jovanovic et al. 2005). The startle probe was a 108-dB [A], 40-msec burst of broadband noise with near instantaneous rise time (NA). The aversive stimulus (US) was a 250-msec airblast with 140 p.s.i. intensity directed to the larynx (Jovanovic et al. 2005). The conditioned stimuli were two different colored lights (A and B) matched for light transmission with color and serial position counterbalanced across subjects.

The Acquisition (Day 1) and Extinction/Reinstatement Testing (Day 2) phases were separated by 24 h. Each startle session began with a 1-min acclimation period consisting of 70-dB broadband noise, which continued throughout the session as background noise, followed by a habituation phase consisting of 6 NA presentations. Acquisition consisted of four blocks with 12 trials per block (three reinforced presentations of A [+], one nonreinforced presentation of A [−], four nonreinforced presentations of B [−], and four presentations of NA) for a total of 48 trials. As described above, 75% of the presentations of A were reinforced in an attempt to delay the onset of within-session extinction (LaBar et al. 1998; Haselgrove et al. 2004). The intertrial interval (ITI) was randomized between 9 and 22 sec.

Extinction and Reinstatement Testing (Day 2) consisted of six blocks with 12 trials per block (four trials each of A−, B−, and NA) for a total of 72 trials. None of the trials on Day 2 was reinforced with an airblast. The ITI was randomized between 9 and 22 sec.

For reinforced trials (A+), the “A” light was illuminated for a total of 4995 msec. A startle probe (40 msec) was administered 4000 msec after light onset. The airblast US (250 msec) was then presented 500 msec after the startle probe. The light terminated 205 msec after offset of the airblast. For nonreinforced trials (A− or B−), the light was illuminated for a total of 4245 msec. Again, a startle probe (40 msec) was administered 4000 msec after light onset. The light terminated 205 msec after the startle probe.

A three-button response keypad (SuperLab, Cedrus Corp.) was used in the startle sessions in coordination with the EMG startle-response monitoring system (SR-LAB, San Diego Instruments) to collect trial by trial ratings of US-expectancy similar to previously published methods (Jovanovic et al. 2005). In short,
participants received verbal instructions prior to the Acquisition and Extinction phases. Prior to Acquisition, subjects were first instructed on proper use of the response pad. For each presentation of A or B, subjects pressed a button marked “+” if they expected a light to be followed by the US (DANGER), a button marked “−” if they did not expect a light to be followed by the US (SAFETY), and a button marked “0” if they were uncertain of what to expect. Subjects were also explicitly instructed that they would see two different colored lights, and that one of these lights would be followed by an airblast some of the time (partial reinforcement), while the other light would never be followed by an airblast. Subjects were also explicitly instructed to remember what they had learned during Acquisition. These instructions were included to enhance subject retention of the CS/US contingency between the Acquisition and Extinction phases.

At the conclusion of Extinction, a Reinstatement group (n = 22) received three unsignaled USs, while a Control group (n = 22) experienced an equivalent passage of time without the USs. The Acquisition and Extinction phases for these groups were identical. In the Reinstatement group, the time between the last extinction trial of Block 5 and the first reinstating US was 19 sec. The Extinction and Reinstatement data from one subject were not included due to a technical error. Following the unsignaled USs or the equivalent passage of time, all subjects were presented with a block of 12 trials (four presentations each of A or B, and NA) to assess reinstatement. The time between the reinstating USs (or an equivalent time period) and the reinstatement block was 18 sec. Reinstatement was assessed immediately following the Extinction phase in a manner consistent with previous human reinstatement studies (e.g., Dirikx et al. 2004) and in an effort to assess reinstatement independent of spontaneous recovery.

Raw startle amplitude for each trial was calculated by the EMG startle recording software. Digital signals were full-wave rectified and smoothed by an averaging routine that calculated a rolling average of 10 data points. The mean startle magnitude for each trial type (A, B, or NA) in each block was used to calculate the Difference Score using the following formula:

\[
\text{Difference Score} = \frac{\text{Mean startle magnitude to startle probe in the presence of light A or B}}{\text{Mean startle magnitude to startle probe alone (NA)}}
\]

In order to assess potentiation of startle to the CS+ (light A) on Days 1 and 2, we used a repeated measures analysis of variance (ANOVA) with TRIAL TYPE (A or NA) and BLOCK as within-subjects variables and startle magnitude as a dependent variable. To assess discrimination between the CS+ and CS− on Days 1 and 2, we also used repeated measures ANOVA with TRIAL TYPE (A or B) and BLOCK as within-subjects variables and Difference Score as a dependent variable. To assess the extinction of potentiated startle on Day 2, we examined the within-session decrement in mean Difference Score per block on A− trials. A one-way repeated measures ANOVA was used for this analysis with BLOCK as a within-subjects variable and Difference Score on A− trials as a dependent variable. To assess reinstatement of fear-potentiated startle on Day 2, we compared the mean Difference Scores with light A during Blocks 5 and 6 prior to and immediately following the unsignaled USs.

Acknowledgments
We thank Jennifer Fennell for her technical assistance. This work was supported by the Mental Health Service, Atlanta VAMC; the STC Program, the Center for Behavioral Neuroscience of the National Science Foundation under Agreement No. IBN-9876754 (Venture grant, E.J.D.); the American Psychiatric Association/GlaxoSmithKline (E.J.D.), National Institute of Mental Health Grants 1R24MH067314-01A1 (B.O.R.), and R37 MH47840 (M.D.), and the Woodruff Foundation, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine.

References
Human fear extinction

Received August 2, 2006; accepted in revised form September 27, 2006.

Learning & Memory 685

www.learnmem.org
Conditioned fear extinction and reinstatement in a human fear-potentiated startle paradigm


Learn. Mem. 2006, 13:
Access the most recent version at doi:10.1101/lm.393906