Title

Top-Down and Bottom-Up Fear Regulation: Experimental Combinations to Reduce the Return of Fear and an Examination of its Neural Correlates

Author

Sun, Michael

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Top-Down and Bottom-Up Fear Regulation: Experimental Combinations to Reduce the Return of Fear and an Examination of its Neural Correlates

A dissertation submitted in partial satisfaction of the

requirements for the degree

Doctor of Philosophy in Psychology

by

Michael Sun

2019

ABSTRACT OF THE DISSERTATION

Top-Down and Bottom-Up Fear Regulation: Experimental Combinations to Reduce the Return of Fear and an Examination of its Neural Correlates

by

Michael Sun

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2019

Professor Michelle Craske, Chair

The regulation of fear can be considered as driven from stimulus properties, considered bottom-up, and cognitive constructions, considered top-down. This dissertation contains three papers that investigated how these approaches may complement one another for the purposes of clinical translation to optimize long-term fear amelioration in treatments for fear-related disorders.

In Study 1, a fear-conditioning experiment was conducted manipulating the use of a low-cost, re-evaluative, and contingency-directed cognitive reappraisal against passive and active control conditions (i.e., react-as-normal and expressive suppression). The experiment examined how this strategy changed responses during extinction training and during a test of rapid reacquisition one-week later. In Study 2, the experiment was replicated twice. The first replaced

the test of rapid reacquisition with an induction of fear reinstatement. The second replaced the test of rapid reacquisition with an induction of context renewal. Results indicated that reappraisal led to faster reductions in threat expectancy to the CS- during extinction training relative to suppression. This was not observed when extinction training featured CSs overlaid atop visual contexts. Results also indicated that in one of three experiments, reappraisal, relative to reacting as normal, led to increases in CS+ valence after extinction training and reductions in the spontaneous recovery of skin conductance responses. Reappraisal, relative to suppression, also led to faster recovery of CS+ US expectancy after fear reinstatement. Suppression led to greater skin conductance responding to the CS- relative to reappraisal and reacting as normal during rapid reacquisition and greater skin conductance responding to the CS+ during context renewal relative to reacting as normal.

Study 3 examined how individuals who express above average spontaneous recovery in skin conductance responses differentially recruit neural activity in structures that putatively implement fear (amygdala, BNST, anterior insula), top-down regulation (dlPFC and vlPFC) and bottom-up regulation (vmPFC and sgACC) during a test of spontaneous recovery 24-48 hours after fear extinction training. Results suggested sparse evidence for fear over-generation in fear regions, and more evidence for misregulation, underregulation/disconnection, and competitive co-regulation in regions that implement fear regulation. Results are discussed in terms of the granular mechanisms underlying extinction and cognitive reappraisal, and implications for clinical application.

The dissertation of Michael Sun is approved.

Hakwan Lau

Jennifer Silvers

Annette Stanton

Michelle Craske, Committee Chair

University of California, Los Angeles

2019

Dedication

To Eli Olivia Sun, my daughter, and Stephanie Sun, my wife, my shining lights.

Table of Contents

List of Figures and Tables

Acknowledgements

This dissertation includes adaptations of the following publications: Sun, M., & Craske., M. G. (submitted). The Effects of Emotion Regulation on the Spontaneous Recovery of Conditional Fear and its Rapid Reacquisition; Sun, M., & Craske., M. G. (in preparation for publication). The Effects of Emotion Regulation on Reinstatement and Context-Renewal of Fear. For both publications, Michael Sun was the principal investigator and was responsible for study design, data collection, analysis, and write-up. Michelle Craske provided critical feedback regarding design, implementation, and the writing of the manuscripts.

The research in this dissertation was supported by various funding sources including a UCLA Graduate Research Mentorship Award (2015-2016) and grants from the National Institutes of Mental Health (2016-2017: T32-MH015750; 2017-2018: F31-MH111199).

I am incredibly grateful for the endless support of my family, friends, and mentors. My wife Stephanie, daughter Eli, father Danny, mother Hsiang-Chun, and sister Natalie, have supported me in innumerable ways. I have also had the privilege of benefiting from deep friendships with individuals like Will Tsai, Sisi Guo, Leslie Rith-Najarian, Timothy J. Williamson, and Lindsay K. Staples-Bradley. They made UCLA so memorable.

The inspirational minds that are responsible for my passion to this field are Drs. James J. Gross, Sylvia D. Kreibig, Marsha M. Linehan, Anita Lungu, Sharon Hsu, Anna S. Lau, Bruce F. Chorpita, Greg A. Miller, Hakwan C. Lau, and Vincent Taschereau-Dumouchel. Your intellectual support has not gone unappreciated. Of course, I would like to give heartfelt acknowledgement to my graduate advisor and committee chair Dr. Michelle G. Craske. Thank you for all of your patience and guidance throughout the years together. I will never forget what a force you are.

Vita

Education

2013-2014 University of California, Los Angeles, Los Angeles, CA
M.A., Clinical Psychology

2008-2011 University of Washington, Seattle, WA
B.A., Psychology; Minor: Japanese Language; *Cum Laude*

Positions and Employment

2018-2019 Minneapolis Veterans Affairs Health Care System, Minneapolis, MN
Psychology Intern, APA-Accredited Clinical Psychology Internship

Selected Academic Fellowships and Awards

2018 Dissertation Grant Award
Society for a Science of Clinical Psychology

2017 F. J. McGuigan Dissertation Award
American Psychological Foundation

2017 Ruth L. Kirschstein National Research Service Award F31-MH111199
National Institute of Mental Health

2016 Institutional Training Grant Predoctoral Fellowship T32-MH015750
National Institute of Mental Health

2015 Graduate Research Mentorship Fellowship
UCLA Graduate Division

2015 Graduate Summer Research Mentorship Fellowship
UCLA Graduate Division

2015 Graduate Research Fellowship Honorable Mention
National Science Foundation

2013 Distinguished University Fellowship
UCLA Graduate Division

Selected Publications

1. **Sun, M.,** Marquardt, C., Disner, S., Burton, P., Lissek, S., Sponheim, S. (invited, in preparation). Effects of post-traumatic stress on neural indices of emotion regulation as reflected by the MMPI-2-RF. *Personality Neuroscience Special Issue: Novel Investigations of the Connection between Quantitative Personality-Psychopathology Models and Neuroscience*

2. Rith-Najarian, L., Mesri, B., Park, A., **Sun, M.**, Chorpita, B., Chavira, D. (2018). Youth cognitive and behavioral therapy durability: A meta-analysis on long-term follow-up effects. *Behavior Therapy*. DOI: 10.1016/j.beth.2018.05.006

3. **Sun, M.**, Rith-Najarian, L., Williamson, T. J., Chorpita, B. (2018). Treatment features associated with youth cognitive behavioral therapy follow-up effects: meta-analysis. *Journal of Clinical Child and Adolescent Psychology.* DOI: 10.1080/15374416.2018.1443459

4. **Sun, M.**, Lau, A. S. (2018). Exploring cultural differences in expressive suppression and emotion recognition. *Journal of Cross-Cultural Psychology, 49*(4), 664-672. DOI: 10.1177/0022022118763749

5. **Sun, M.**, Vinograd, M., Miller, G., Craske, M. G. (2017). Research Domain Criteria (RDoC) and Emotion Regulation. In C. Essau, S. Leblanc, & T. Ollendick (Eds.), *Emotion Regulation and Psychopathology in Children and Adolescents*. New York, NY: Oxford University Press.

6. Guan, K., Levy, M.C., Kim, R. E., Brown, T. E., Reding, M. E. J., Rith-Najarian, L., **Sun, M.**, Lau, A.S., & Chorpita, B.F. (2017). Managing in-session "surprises:" Identifying provider responses to emergent life events during evidence-based treatment implementation. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(2), 164-176. DOI: 10.1007/s10488-015-0692-3

7. Lungu, A., **Sun, M.** (2016). Time for a change: College students' preference for technology-mediated versus face-to-face help for emotional distress. *Telemedicine and e-Health, 22*(12), 991-1000. DOI: 10.1089/tmj.2015.0214

8. Tsai, W., **Sun, M.**, Wang, S. W., Lau, A. S. (2016). Culture and emotion expression: Daily and trait interpersonal and intrapersonal functioning between Asian and European American college students. *Asian American Journal of Psychology, 7*(1), 52. DOI: 10.1037/aap0000043

General Introduction

He who has overcome his fears will truly be free. – Aristotle

Anxiety disorders are among the most prevalent and crippling of mental health conditions, afflicting at least 28.8% of the world population (in 12-month prevalence) (Kessler, Petukhova, Sampson, Zaslavsky, & Wittchen, 2012). They are the sixth leading global cause of time lost to disability (Baxter, Vos, Scott, Ferrari, & Whiteford, 2014), and they contribute significantly to worldwide economic burden (Chisholm et al., 2016; Hendriks et al., 2015). Contemporary exposure-based therapies, the gold-standard approach for alleviating these disorders, should be the big answer to this need. It was once thought of as such, as Rachman wrote in 1989, "contemporary exposure techniques are reasonably effective and reasonably durable" (Rachman, 1989). Unfortunately they have not done enough; 24-69% of individuals with anxiety do not respond and 21-64% experience relapse within eight years (Yonkers, Bruce, Dyck, & Keller, 2003). Answers to how exposure therapy can be enhanced will come from a broader examination of the emotion regulation literature.

Classically, fear-based disorders (e.g., specific phobia and panic disorder) are believed to develop through aversive classical conditioning (Bouton, Mineka, & Barlow, 2001; Mineka & Zinbarg, 2006). Through aversive learning, also known as fear acquisition, an excitatory association is learned through pairings of a neutral stimulus with an aversive outcome (i.e., an unconditional stimulus (US)). The neutral stimulus becomes a conditional stimulus (CS) that produces an excitatory conditional response (CR) in anticipation of the US (i.e., CS-US memory). A maladaptive, persistent, and overgeneralized CR is characteristic of fear-based disorders like anxiety (Vervliet, Craske, & Hermans, 2013).

Bottom-up Emotion Regulation through Conditional Learning

The strength of a fearful CR increases with learning, and this relationship is expressed mathematically with the Rescorla-Wagner model (Rescorla et al., 1972). This is expressed as $\Delta V_t = \alpha_x \bullet \beta(\lambda - V_{total})$, where ΔV_t is the change in CR strength on trial t, α_x is the salience of stimulus x, β is the associative value of the US, λ is the CS-US associative maximum, and V_{total} is the total strength of all stimuli present in the organism's sensory array. The Rescorla-Wagner model also models the extinction of fear. Fear extinction emphasizes learning through prediction error (i.e., CS prediction of US is incorrect) and occurs through repeated exposure to the CS in the absence of the US, diminishing the CR (Hermans, Craske, Mineka, & Lovibond, 2006). Because it is experience-based, it is considered conceptually bottom-up. According to Rescorla-Wagner, the prediction error needed for either fear acquisition or extinction learning requires a salient, intense, and aversive US, as ΔV_t is proportional to β no matter the direction. The more intense the US, the more likely its absence during extinction trials will elicit the prediction error that extinguishes the CR (Rescorla et al., 1972). Mismatch with expectancy provides the largest amount of new learning; hence, extinction should be augmented by initially heightened expectancies of the aversive outcome that contrast with the actual experience of no aversive outcomes. The extinction of conditional fear is a laboratory analogue of exposure therapy for anxiety disorders in humans (Craske & Mystkowski, 2006; Hermans et al., 2006).

One immediate target for the improvement of anxiety treatment is the mitigation of return of fear phenomena. As early as 1979, Rachman observed that anxious patients were experiencing relapse, and even despite observing complete extinction of the CR in his laboratory, there was a return of the response as soon as a week later (Rachman, Robinson, & Lopatka, 1987). This return of fear was evidence that the Rescorla-Wagner CS-US association was not extinguished.

Rather, the evidence suggested that a CS-noUS association must have arisen instead, acquired to compete with older CS-US associations to influence on the organism's behavioral output. Since then, evidence for this theory has accumulated, even at the level of the brain, where memory signatures coinciding with the CS-US representation has been observed in basolateral amygdala in rats (e.g., Anglada-Figueroa & Quirk, 2005; Barad, Gean, & Lutz, 2006; Herry, Trifilieff, Micheau, Lüthi, & Mons, 2006; Milad, Rosenbaum, & Simon, 2014) and in the amygdala more generally in humans (e.g., Agren et al., 2012; Barad et al., 2006; Milad et al., 2007). This idea is known as the *inhibitory learning* model (Bouton, 1993; Bouton & King, 1983). While there is mixed evidence suggesting that true extinction of fear may occur if conducted immediately following acquisition (Myers & Davis, 2007; Myers, Ressler, & Davis, 2006), this is not within the scope of the current work.

Return of fear phenomena are those phenomena that describe the conditions in which fearful CRs re-emerge after fear extinction (Bouton, 2002; Bouton et al., 2001; Hermans et al., 2006). They are the phenomena that putatively underlie anxious relapse. These include (1) *spontaneous recovery* (Baum, 1988; Napier, Macrae, & Kehoe, 1992), which occurs with the passage of time, (2) *rapid reacquisition* (Kehoe & Macrae, 1997; Rescorla & Heth, 1975; Rescorla et al., 1972), which involves re-pairing of the CS and US, (3) *reinstatement* (Pavlov, 1927; Rescorla & Heth, 1975; Rescorla et al., 1972), which involves an unsignaled US presentation, and (4) *context renewal* (Bouton, 1993), which involves CS presentation in a new context. Under these conditions, the original CS-US memories may prevail, causing a return of fear. In accordance with inhibitory learning theory, the return of conditional fear represents a failure to access or retrieve inhibitory memories (Bouton, 1993). Broadly they can be categorized as *context-free* or *context-based*.

Context-Free Return of Fear

The word *context* within the scope of this work refers primarily to stimuli surrounding the CS (i.e., the "background"). Spontaneous recovery and rapid reacquisition do not depend on contextual features and are thus classified as context-free. Spontaneous recovery is the most well studied of all return of fear phenomena, and is sometimes referred to in the literature simply as "return of fear". Despite the great success in treating anxious responding through exposure-based treatments such as flooding and desensitization at that time, it seemed simply a matter of time before fearful responding returned (Rachman, 1979). Spontaneous recovery is also studied when the aim is to examine "extinction recall". The two terms are almost completely conflated in the literature; spontaneous recovery refers to the response associated with a resurfaced CS-US memory, and extinction recall refers to the response suppressed due to a newly instated CS-noUS memory. Both spontaneous recovery and extinction recall are simply measured by the CR after some length of time (usually 1 week, but may be up to 3 months) after extinction training is completed.

Perhaps the most intuitive condition where the fearful CR returns is the effect of faster re-learning of the CS-US association. Rapid reacquisition (Kehoe & Macrae, 1997) is the concept where one can relearn associations that have once been forgotten faster than previously. In the laboratory, this is done by presenting CS-US pairings after extinction training. Rapid reacquisition is especially problematic given that for many fear-disorders, the feared catastrophe (i.e., the US) is likely to reoccur at some future point following successful treatment (e.g., social rejection for social phobia).

Context-Based Return of Fear

Reinstatement and context renewal are return of fear phenomena that depend on contextual features and are thus classified as context-based return of fear. Reinstatement is the return of fear when an unsignalled US is presented in the same context extinction had taken place. Reinstatement is problematic because its occurrence is not limited to encounters with the original US; an encounter with any sufficiently aversive stimulus will trigger reinstatement when the CS is re-encountered in its original context, making it a potentially common source of relapse.

The return of fear as a topic has become a rapidly increasing source of scientific inquiry as shown on the graph in Figure 1. As of May 19th, 2019, there are so far 4,010 publications containing the search term "return of fear". Figure 1 depicts the number of non-cumulative citations containing the term "Return of Fear" for each year, starting from the first article on the topic (Rachman, 1979). A three-year moving average, which provides the best fit of the current data, predicts around 477 ($y_{predicted}$ = 477, *SE* = 75.49) new publications containing the term by the end of 2019.

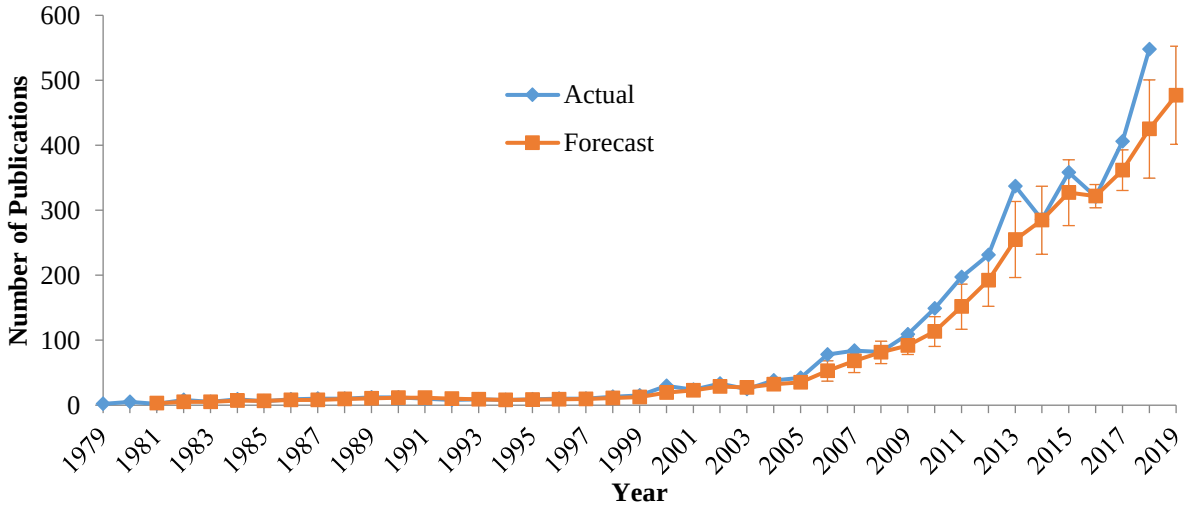


Figure 1: Year-by-year number of new publications containing the term "return of fear" along with predictions from a three-year moving average.

Top-Down Emotion Self-Regulation

Return of fear may be mitigated by incorporating insights from a relatively new school of thought, the emotion regulatory school, built on a foundation of literature by James J. Gross, and later infused by empirical findings in neurobiology by Kevin Ochsner. Emotion regulation is a central topic in contemporary psychological discourse, with research activity in individual health and well-being as well as developmental, social, and cultural psychology (Gross, 1998b, 2015a).

The Gross-Ochsner school of thought defines fearful responding based on a valuation-based model. It argues that stimulus evaluations (i.e., what is "bad for me/not bad for me/good for me") determine the emotional response. Top-down, or effortfully enacted, emotion regulation (ER) strategies can modify these evaluations with strategies such as cognitive reappraisal to reduce emotional responding. In the context of extinction, cognitive reappraisal could be used to re-evaluate the US as less aversive. For example, an electric shock (US) could be reappraised as not as intense as expected (e.g., a shock→ "bad for me" → respond in fear, to a shock → "not as bad for me" → respond less fearfully). Studies have shown that cognitive reappraisal of the stimulus is effective relative to analogous ER strategies focused on reducing the response to the stimulus, a strategy termed suppression (a shock→ "bad for me" → do not respond) (Gross, 1998b; Ochsner, Silvers, & Buhle, 2012). Accordingly, the valuation-based model suggests that changing the evaluation of the US through cognitive reappraisal should reduce fear CRs and attenuate the return of fear. However, this is counter to the Rescorla-Wagner premise that inhibitory learning is incumbent on experiential mismatch with expectancy – spurring a need for initially heightened expectancies of the aversive outcomes. As such, if and how cognitive approaches should be used in exposure indeed remains a central debate in the field (Vervliet et al., 2013).

An Integrative Computational Approach

Effortful regulatory strategies such as reappraisal come at a cost. Therefore, a new model termed the computational implementation model attempts to integrate the Rescorla-Wagner prediction error model with valuation accounts of emotional responding while accounting for such a cost (Etkin, Büchel, & Gross, 2015). In this model, resulting emotional responses are determined by experience and the cost of evaluative implementation. Cost relates to availability and believability of the evaluation (Etkin et al., 2015). Evaluations should be easy to understand, easy to believe, and quickly called upon, otherwise they are unlikely to regulate the emotional response. These propositions are denoted by the formula: $\Delta V_t = V_{t-1} + \rho\delta - C$. The terms are as follows: t denotes the timepoint, ΔV_t denotes the change in emotional response, V_{t-1} denotes the response made at the last timepoint, ρ denotes the learning rate of the individual, δ denotes prediction error, and C denotes the cost of an evaluative effort (Etkin et al., 2015). This integrated approach might inform us of an optimal combinatorial strategy for fear regulation. Leveraging the robust effects of extinction training with the flexibility of cognitive reappraisal, cognitive reappraisal may complement extinction training for potential enhancement so long as it is (1) re-evaluates the stimulus to be unfearful without (2) disrupting the CS-noUS contingency in a way that is (3) low in cost.

Neurobiological Support

Neurobiology supports the notion that cognitive reappraisal and extinction learning are amenable to complementation. CS-noUS memories are believed to be retained in the infralimbic cortex of the vmPFC-sgACC (Milad et al., 2007). When activated, the central amygdala, which is, responsible for enacting fear CRs, becomes inhibited (Duvarci & Pare, 2014; Milad & Quirk, 2012; Milad et al., 2007; Paré, Quirk, & Ledoux, 2004). Reappraisal is associated with

dorsolateral and ventrolateral prefrontal cortices (dlPFC and vlPFC: Beauregard, Lévesque, & Bourgouin, 2001; R. Kalisch et al., 2005; Raffael Kalisch, Wiech, Herrmann, & Dolan, 2006) which implements executive, organizational control of complex information (Dalley, Cardinal, & Robbins, 2004). Although they have few projections to amygdala (e.g., Buhle et al., 2014; Kohn et al., 2014; Vertes, 2006), they do project to medial areas including vmPFC-sgACC (Vertes, 2006). Cognitive reappraisal may be associated with increased vmPFC-sgACC activation, the same region associated with fear extinction to reduce central amygdala activation.

Overview of Studies

The overarching goal of this work was to test a theoretical complementary strategy combining top-down and bottom-up regulatory approaches and to examine its neurobiological plausibility at a time after extinction training had occurred. For such a complementary approach to have clinical utility, it would need to be robust against context-free and context-based return of fear phenomena. Therefore, two studies were conducted testing a combined approach to fear reduction. The first study was an experiment that examined an emotion regulatory combination in adults on context-free return of fear, assessed via spontaneous recovery and rapid reacquisition one-week later. The second study then replicated this procedure to examine the effects of emotion regulatory combinations on context-based return of fear phenomena, namely reinstatement, and context renewal. Finally, in a third study, I examined the neurobiological relationship between regions implicated in effortful emotion regulation and extinction recall. Results from these studies may serve to validate or discourage practices that are key in modern therapeutic approaches. They represent an important step in understanding the role of emotion regulation during inhibitory extinction learning – the foundational processes underlying anxiety treatment – across multiple units of analysis (i.e., self-report, neural and peripheral physiology).

The insights gained will also test the validity of the computational implementation model that, if valid, can inform how emotion regulation impacts emotional learning on an everyday basis in contexts outside of fear conditioning.

Study 1: The Effects of Emotion Regulation on the Spontaneous Recovery of Conditional Fear and its Rapid Reacquisition

Abstract

Valuation theories of emotion regulation posit that effortful strategies that manipulate self-relevant value can regulate responses toward feared stimuli. Learning theories emphasize maximizing prediction error correction in order to extinguish conditional responses to feared stimuli. In this study, we instructed participants to utilize a cognitive reappraisal strategy designed to complement prediction error correction by manipulating the stimulus value in a low-effort manner as participants underwent extinction training. This strategy was compared against active (suppression instruction) and passive (react-as-normal instruction) control groups. Results suggested that cognitive reappraisal with extinction training, caused reduced US expectancies by the end of extinction training. Combining emotion regulation strategies with extinction training did not benefit spontaneous recovery or rapid reacquisition, however, suppression exhibited a deleterious effect of increasing skin conductance responding to the CS- during rapid reacquisition.

Introduction

Exposure therapy is an effective treatment for fears and phobias and yet individuals remain vulnerable to a return of fear once treatment is over (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). Inhibitory retrieval models of Pavlovian fear extinction provide explanatory pathways for the return of fear, including spontaneous recovery and rapid reacquisition (Craske, Hermans, & Vervliet, 2018). Spontaneous recovery refers to the return of conditional fear proportional to the time elapsed since the end of extinction training (Rescorla, 2004). Spontaneous recovery implies that despite successful exposure treatment, fear will eventually reappear in the absence of repeated practice with the previously feared stimulus. Rapid reacquisition (Kehoe & Macrae, 1997) refers to relearning of fear associations after extinction training that is faster than original learning of fear associations. Rapid reacquisition is particularly applicable to clinical situations in which aversive fear conditioning is prone to re-occur following exposure treatment, as in the case of social rejection and the re-emergence of fears of social evaluation. The current study aimed to evaluate whether top-down, higher order cognitions that reconfigure the stimulus value of feared stimuli and associated outcomes augment extinction and attenuate spontaneous recovery and rapid reacquisition of fear.

Bottom-up, experience-based, prediction error correction that leads to new inhibitory association (i.e., CS-noUS association) is posited to be a critical mechanism underlying extinction and by translation exposure therapy (Craske et al., 2018). Prediction error correction implies that the perceiver makes predictions about threats, and the experience of error in that prediction (consciously or unconsciously) leads to new learning. The prediction error could involve either the likelihood of the threat (US) or its aversiveness, given the values assigned to both US probability and US salience in the Rescorla-Wagner model. Furthermore, the greater

the mismatch between the predicted outcome (US) and actual learning event (CS-noUS), the greater potential for prediction error correction.

Stimulus value has been studied with respect to fears and conditioning (Dirikx, Hermans, Vansteenwegen, Baeyens, & Eelen, 2004, 2007; Luck & Lipp, 2015) but there is no known work examining how cognitive strategies targeting stimulus valuations affect spontaneous recovery and rapid reacquisition. Cognitive reappraisal is an emotion regulation strategy that requires effortfully changing thoughts about a stimulus such that the value toward the self is changed. For a fearful stimulus such as a spider, the initial value toward the self is threat (i.e., *bad* for me). An example of a cognitive reappraisal would be to think of the spider as "not poisonous and will not hurt me" (i.e., *not bad* for me) or "having an important role to play in a healthy ecosystem" (i.e., *good* for me). This change of stimulus value regulates the unfolding emotional response. Indeed, cognitive reappraisal-based clinical interventions, such as cognitive restructuring, are commonly used in the clinical care of individuals with fear-based disorders (e.g., Clark, 1999; Resick, 2001).

Despite their wide clinical use, certain cognitive reappraisal strategies may interfere with bottom-up processes underlying extinction by mitigating prediction error correction. For example, cognitive reappraisal requires attentional effort, and by so doing may interfere with allocation of attentional resources towards the CS/US, which in turn impairs learning that the CS is no longer followed by the US (Mackintosh, 1975). In addition, cognitive reappraisal that lowers the perceived likelihood or aversiveness of the US may lessen the discrepancy between prediction and outcome, thereby reducing the potential for prediction error correction (Craske et al., 2014).

From the standpoint of emotion regulation, appraisals should change the relationship between a stimulus and the self (Gross, 2015b; Kross & Ayduk, 2011) but without substantial 'implementation cost', or the difference between the predicted outcome relative to the perceived resources necessary for implementation (Etkin et al., 2015). Costly implementation is hypothesized to be less effective (Etkin et al., 2015). By translation, difficulties generating examples of the importance of spiders to the ecosystem, or lack of confidence that such an appraisal would change emotional responses to spiders, would render a reappraisal strategy as costly and thereby less effective.

Therefore, to evaluate whether reappraisal strategies augment extinction learning, it is essential to select reappraisal strategies that change the relationship of a feared stimulus with the self without altering expectancies for the likelihood of US occurrence or its perceived aversiveness, since aversiveness likely contributes to US. At the same time, they should be strategies that are easily implementable with minimal cost. The aim of the current study was to assess whether such reappraisal strategies augment extinction training and weaken spontaneous recovery and rapid reacquisition one-week later. We hypothesized that the combination of reappraisal with extinction training would reduce conditional fear responses during extinction training and at tests of spontaneous recovery and rapid reacquisition compared to passive (i.e., react as normal) and active (i.e., suppress) control conditions.

Methods

Participants

Participants were 163 healthy adults over the age of 18, recruited from a participant subject pool at the University of California, Los Angeles. They received course credit for their participation. The average age was 21.76 years (SD = 5.78), 66.26% were female, and the

racial/ethnic breakdown consisted of 19.63% Caucasian, 39.88% Asian, 17.79% Hispanic/Latino, and 22.70% Other. Participants were excluded if they self-reported any of the following: diagnosed with a previous or current mental illness or psychiatric disorder, currently using psychoactive medication, uncorrected problems with vision or hearing, pregnant, had a serious medical condition, unable to speak or understand English, or recent caffeine intake that may affect measurement of peripheral physiology.

Design

The aim was to evaluate the impact of combining emotion regulatory strategies with extinction training upon extinction performance, spontaneous recovery and rapid reacquisition (see Figure 1). Participants were randomly-assigned to (1) Reappraisal, (2) Suppression (as an active control), or (3) React-as-Normal conditions (as a passive control).

Procedure

On Day 1, participants underwent Baseline, Habituation, Acquisition, and Extinction phases. In the Baseline phase, participants viewed a blank screen as physiology was recorded for 5 min. In the Habituation phase, participants were acclimated to the CS (neutral facial stimuli) through two six-second trials. During Acquisition, one (CS+) was paired with a 1-second 82dB scream sound as the US which co-terminated with the CS+. Scream sounds have been used successfully in previous fear conditioning studies (e.g., Culver, Vervliet, & Craske, 2015; Lau et al., 2008; Neumann & Waters, 2006). The other facial stimulus (CS-) was not paired with the US. Each CS+/- was presented for eight trials. CS+ and CS- images were counterbalanced across participants. CS trials were presented in random order, with no more than two consecutive trials of the same CS type.

According to randomization, one of three instructions appeared, consistent with instruction sets used in prior emotion regulation research (e.g., Gross, 1998a). In the Cognitive Reappraisal condition, participants were asked to think of themselves as a casting director who is evaluating a screaming actress trying out for a part in a scary movie with the following instruction: *"For the next couple of minutes, imagine that you are a Hollywood star hiring manager. You have just gotten the go-ahead on a big television horror-drama, and are looking to hire several lead actresses. You want to hire the actress with the most realistic scream possible. Please be as objective in your selection as you possibly can. Listen carefully and select your next lead actress."*

In the Suppression condition, participants were asked to force themselves to be unexpressive in order to prevent others from knowing how they feel as they listened to the scream. The instruction was as follows: "*For the next couple of minutes, we would like to see how well you can keep from showing any emotional response when you hear a scream. Try not to feel anything, and try not to have a physiological reaction. Also, see if you can act so that someone seeing the video with the sound off won't know that anything has happened. Try not to show any visible signs or feel anything before, during, or after the scream occurs. Try to look relaxed all the way through. See if you can fool the person who will be studying this video.*"

Finally, the React-as-Normal condition instructed participants to simply continue watching the computer screen. The exact instruction was as follows: "*For the next couple of minutes, you will be presented a series of faces and sounds. Simply attend to the computer screen as you would naturally do so. Do not distract yourself by thinking about other things.*"

All instructions were designed to be low in effort and to have no effect upon prediction error (i.e., no effects on either perceived US intensity or US expectation).

The instructions were followed by Extinction Training, involving 24 trials of each CS without the US in random order, with no more than two consecutive trials of the same CS type.

One week later, participants underwent a test for Spontaneous Recovery, involving two trials of each CS without the US. This was followed by a test of Rapid Reacquisition, involving four CS+ trials paired with the US and four CS- trials, in random order, with no more than two consecutive trials of the same CS type.

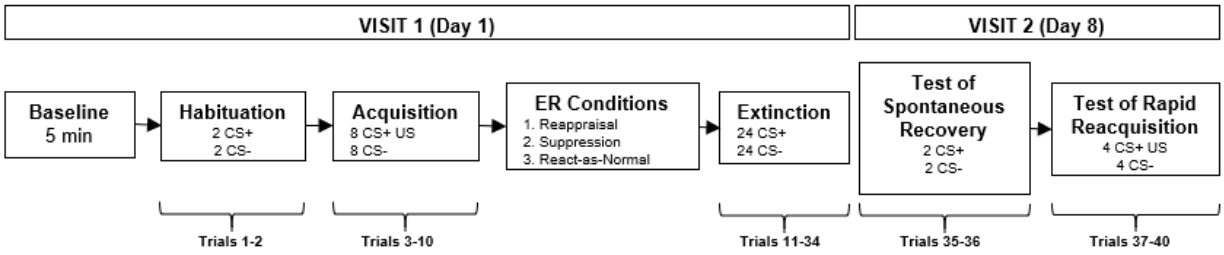


Figure 1: Experimental design.
Note. ER Conditions refers to the emotion-regulation instruction set shown to the participant prior to the onset of the Extinction phase.

Materials/Apparatus

Participants sat in front of a 21-inch monitor situated roughly 20-inches away from eye-level. Stimulus presentation was programmed and controlled using E-Prime 2.0 Professional (*version 2.0.10.353*; Psychology Software Tools, Pittsburgh, PA), installed on a personal computer running Windows 7. Two NimStim (Tottenham et al., 2009) images of Asian female faces with a neutral expression were chosen as conditional stimuli (CS) (counterbalanced between participants). To ensure attentional capture and sufficient salience, the height of the CSs matched the maximum height of the screen, with the width expanded to be proportional to the height.

A BioPac MP150 (Biopac System Inc., Goleta, CA, USA) collected peripheral physiology data. Specifically, BioPac amplifiers EDA100C measured skin conductance

responses (SCR) and EMG100C measured fear-potentiated startle blink through electromyography (EMG). Data were recorded from the BioPac MP150 using AcqKnowledge 4.2 software (Biopac System Inc., Goleta, CA, USA), and cleaned, inspected, and analyzed with ANSLab software (ANSLab v2.6, Wilhelm & Peyck, 2005).

Skin Conductance Responses. SCRs were recorded from two EL507 11 mm diameter Ag/AgCl electrodes placed on the distal phalanx of the index and middle fingers of the non-dominant hand (Bradley, Cuthbert, & Lang, 1990). Using an EDA100C amplifier and two LEAD110A electrode leads, SCR data were sampled at a rate of 2000 Hz and filtered using a finite impulse response (FIR) low pass filter with a frequency cutoff fixed at 2 Hz. SCR was calculated as a difference score between the maximum skin conductance value 1–6 s after CS onset minus the mean skin conductance value of the 2 seconds prior to CS onset. SCRs greater than zero were square root transformed to normalize the data (Levey, 1980). SCRs less than or equal to zero were coded as zero. To eliminate individual variability in SCR range, SCRs were *T*-score standardized using the formula shown below:

$$SCR_T = \frac{SCR_{Observed} - SCR_{mean}}{SCR_{SD}} \times 10 + 50$$

Self-Report. Online ratings of US expectancy were made via a BioPac TSD115 continuous sliding dial. Participants were asked to continually adjust how certain they were that the US will appear at the end of the trial ("how certain are you that you will hear a sound in the next few moments"). Participants received 3-second prompts to remind them of the expectancy dial at the beginning of each ITI and CS. The values ranged from 0 = "Certain no sound", 4.5 = "Uncertain", and 9 = "Certain sound". US expectancy was calculated as the mean rating 6.5–7 seconds after ITI, CS+ or CS− startle probe onset. Before and after every phase, valence, arousal,

and fearfulness ratings (from 1-low to 9-high) associated with each stimulus were assessed onscreen. Data were recorded via keypress after presentation of the scale on the computer screen.

Data Analysis

STATA 14.1 was used to perform all analyses. Preliminary examination of differential responding by Stimulus-type (CS+ vs. CS-) was examined using t-tests within the Acquisition phase (trials 3-10), Extinction phase (trials 11-34), and test of Rapid Reacquisition (trials 36-40).

Data for the Acquisition phase (trials 3-10) were then analyzed with a multilevel modelling framework (Bryk & Raudenbush, 1992) given its many advantages over an ANOVA framework (Kristjansson, Kircher, & Webb, 2007). The level 1 repeated measures were nested within individuals at level 2. Using the *mixed* command in STATA 14.1, best-fit prediction lines with stepwise-polynomial multilevel modelling for multi-trial outcomes (i.e., SCRs, US expectancies). These outcomes were predicted from Stimulus-type (0 = CS+, 1 = CS-), each of the polynomial components of the best-fit lines (e.g., instantaneous linear, quadratic component), and Stimulus x polynomial component interactions.

We did not believe that instructed emotion regulation would impact the different patterns of responding expected between the CS+ and CS- that normally results from aversive learning, just the properties of these patterns (e.g., speed). For the Extinction phase (trials 11-34), best-fit lines for each outcome were first estimated separately for CS+ and CS-. Predictors for Emotion Regulation terms (0 vs. 1 and 0 vs. 2, where 0 = React-as-Normal, 1 = Reappraisal, 2 = Suppression) and their interactions with polynomial slope components were then added to test the effect of Emotion Regulation condition on outcomes. The React-as-Normal condition was initially set as the reference group, and models were rerun to compare Reappraisal with Suppression conditions (i.e., 1 vs. 2). In all cases, intercepts refer to the response at the first trial

of the phase, the instantaneous linear slope refers to the initial change in responding. Further polynomial slopes refer to curvatures throughout the phase. To ensure that extinction was achieved and to examine if there were any differences in responding by the end of extinction due to Emotion Regulation, a two-way ANOVA was run estimating the final two trials of the Extinction phase by Stimulus-type, Emotion Regulation, and their interaction. Extinction was achieved if there was no significant Stimulus-type or Stimulus-type x Emotion Regulation interaction.

Simple-linear multilevel models were fitted to model CS+ and CS- responses separately to test Spontaneous Recovery (trials 34-35). SCR and US expectancy outcomes for this model were first predicted by period (0 = pre, 1 = post), then by Emotion Regulation and their interactions with period. The same approach was used to predict self-reported fear, arousal, and valence before and after the Extinction phase, test of Spontaneous Recovery, and test of Rapid Reacquisition.

Evidence for rapid reacquisition involved modelling the first two trials of the test of Rapid Reacquisition (trials 36-40) predicted by Stimulus-type, trial, and their interactions. Effects of Emotion Regulation were then tested by adding Emotion Regulation terms and their interactions with previously mentioned terms. To examine how much fear was retained by the end of the test of Rapid Reacquisition, a two-way ANOVA was run estimating the final two trials of the Extinction phase by Stimulus-type, Emotion Regulation, and their interaction.

Results

Acquisition

CS+ SCRs (B = 13.71, SE = 1.87, z = 7.33, p < .001, 95% CI=[10.04, 17.37]) and US expectancy ratings (B = 1.19, SE = 0.26, z = 4.54, p < .001, 95% CI=[0.67, 1.70]) exhibited

significantly increasing instantaneous linear components of the Acquisition phase. Significant Stimulus-type x phase interactions revealed greater changes to the CS+ than CS- for SCRs and US expectancy ratings (*p*s < .001). CS+ self-report ratings of fearfulness (*B* = 2.06, SE = .14, *z* = 14.36, *p* < .001, 95% CI = 1.78, 2.34]) and arousal (*B* = 1.42, SE = .14, *z* = 10.08, *p* < .001, 95% CI = [1.15, 1.70]) significantly increased from before to after Acquisition. CS+ self-report ratings of valence significantly decreased (*B* = -0.78, SE = 0.12, *z* = -6.54, *p* < .001, 95% CI = [-1.01, -0.54]). In all cases of self-report, significant Stimulus-type x phase interactions revealed greater changes to the CS+ than CS- (*p*s < .001).

Extinction

SCR. CS+ (*B* = -3.68, SE = 0.98, *z* = -3.74, *p* < .001, 95% CI = [-5.61, -1.75]) and CS- (*B* = -1.38, SE = 0.39, *z* = -3.56, *p* < .001, 95% CI = [-2.15, -0.62]) SCRs exhibited declining instantaneous linear slopes. Emotion Regulation did not significantly affect the slopes of the CS+ (*ps* > .148) or the CS- (*ps* > .263).

Two-way ANOVA revealed no significant differences in SCRs by Stimulus-type (*F*(1, 572) = 0.12, *p* = .733) at the end of the Extinction phase. There was also no significant effect of Emotion Regulation (*F*(2, 572) = 0.88, *p* = .416) or the Stimulus-type x Emotion Regulation interaction (*F*(2, 572) = 0.66, *p* = .515).

US Expectancy. CS+ (*B* = -1.73, SE = 0.19, *z* = -8.97, *p* < .001, 95% CI = [-2.11, -1.36]) and CS- (*B* = -1.13, SE = 0.17, *z* = -6.48, *p* < .001, 95% CI = [-1.47, -0.79]) US expectancies exhibited declining instantaneous linear slopes. Emotion Regulation did not significantly affect the slopes of the CS+ US expectancies (*ps* > .147).

Significant effects of Emotion Regulation were found on the slopes of CS- US expectancies. CS- US expectancies for the Suppression and React-as-Normal conditions were not

found to be different (*p*s > .580), however, the Reappraisal condition exhibited a declining instantaneous linear component that was significantly steeper relative to the Suppression condition (*B* = -1.02, SE = 0.44, *z* = 2.29, *p* = .022, 95% CI=[-1.89, -.15]) and marginally steeper relative to the React-as-Normal condition (*B* = -0.77, SE = 0.40, *z* = -1.91, *p* = .057, 95% CI=[-1.56, 0.02]). The Reappraisal condition also exhibited significantly different high-order polynomial components relative to the React-as-Normal condition (*p*s = .035 to .059), and similar patterns were evident between Reappraisal and Suppression conditions with high-order polynomial components that were marginally significantly different (*p*s = .067 to .087).

Two-way ANOVA revealed no significant differences in US expectancy by Stimulus-type (*p* = .614) at the end of the Extinction phase. There was also no significant effect of the Stimulus x Emotion Regulation interaction (*p* = .974). There was a significant main effect of Emotion Regulation (*F*(2, 556) = 3.03, *p* = .049), and pairwise comparisons revealed that US expectancies in the Reappraisal condition were significantly lower relative to the React-as-Normal condition (*B* = -0.76, SE = 0.32, *z* = -2.41, *p* = .016, 95% CI=[-1.39, -0.14), but not the Suppression condition (*p* = .117) by the end of the Extinction phase. Suppression and React-as-Normal conditions did not significantly differ (*p* = .593).

Self-Report. Extinction induced a significant reduction in self-reported CS+ arousal (*B* = -1.24, SE = 0.13, *z* = -9.23, *p* < .001, 95% CI = [-1.50, -0.97]) and CS+ fearfulness (*B* = -1.14, SE = 0.13, *z* = -8.75, *p* < .001, 95% CI = [-1.39, -0.88]), as well as a significant increase in self-reported CS+ valence (*B* = 0.87, SE = 0.12, *z* = 7.39, *p* < .001, 95% CI = [0.64, 1.10]). Stimulus-type x period interactions revealed that these changes were significantly different from changes in CS- (*p*s < .001). Emotion Regulation condition induced no significant differences in changes in fearfulness (*p*s > .064) or arousal (*p*s > .440). A significant Emotion Regulation x period

interaction revealed that Reappraisal induced a greater increase in CS+ valence relative to the React-as-Normal condition (B = 0.54, SE = 0.27, z = 1.98, p = .048, 95% CI = [0.01, 1.07]). No differences in self-reported CS+ valence change was found between Reappraisal and Suppression (ps = .134) or Suppression and React-as-Normal conditions (ps = .802), and no Emotion Regulation differences in self-reported CS- valence change was found (ps > .163).

Spontaneous Recovery

SCR. When tested one-week later, there was a significant spontaneous recovery of CS+ SCRs (B = 5.95, SE = 1.48, z = 4.02, p < .001, 95% CI = [3.05, 8.84]) and not CS- SCRs (p = .214). Emotion Regulation condition did not significantly affect SCR recovery of the CS+ (ps > .488) or the CS- (ps > .083).

US Expectancy. When tested one-week later, there was evidence of spontaneous recovery of US expectancy to the CS+ (B = 1.72, SE = 0.42, z = 4.14, p < .001, 95% CI = [0.91, 2.54]), but not to the CS- (p = .959). Spontaneous recovery in CS+ US expectancies did not exhibit significant differences due to Emotion Regulation one-week later (ps > .758). Interactions with CS- US expectancy revealed that recovery was significantly increased in the Reappraisal condition (B = 1.64, SE = 0.80, z = 2.06, p = .040, 95% CI = [0.08, 3.20]), and marginally increased in the Suppression condition (B = 1.56, SE = 0.90, z = 1.74, p = .082, 95% CI = [-0.20, 3.31]), relative to the React-as-Normal condition. Simple slopes analyses did not reveal significant recovery of CS- US expectancy in any Emotion Regulation condition (ps > .085).

Self-Report. One-week later, there was an increase in self-reported CS+ arousal (B = 0.25, SE = 0.11, z = 2.27, p = .023, 95% CI = [0.03, 0.47]), but no significant changes in self-reported CS+ valence (p = .347) or fearfulness (p = .792). Stimulus-type x period interactions revealed that these changes were not significantly different from the CS- (ps > .683). Emotion

Regulation condition induced no significant differences in changes in fearfulness (*p*s > .155), arousal (*p*s > .884), or valence (*p*s > .508).

Test of Rapid Reacquisition

SCR. CS+ SCRs exhibited significant rapid reacquisition (*B* = 3.24, SE = 1.27, *z* = 2.56, *p* = .011, 95% CI = [0.76, 5.73]), significantly differentiating itself from CS- SCRs (*p* = .030) in the first two trials of the Test of Rapid Reacquisition. There were no significant effects of Emotion Regulation condition (*p*s > .103). Two-way ANOVA of SCRs of the last two trials revealed no significant main effect of Emotion Regulation (*p* = .761), but there was a significant main effect of Stimulus-type ($F(444, 1) = 11.72$, $p < .001$), suggesting that the CS+ SCRs was consistently greater than CS- SCRs across Emotion Regulation conditions (*B* = 3.13, SE = 0.91, *t* = 3.42, *p* = .001, 95% CI = [1.33, 4.93]). Furthermore, there was a significant Stimulus-type x Emotion Regulation interaction ($F(444, 2) = 4.18$, $p = .016$). Follow-up pairwise comparisons within each Stimulus-type revealed that the Suppression condition exhibited significantly greater CS- SCRs than both the React-as-Normal condition (*B* = 3.70, SE = 1.61, *t* = 2.31, *p* = .022, 95% CI = [0.55, 6.86]) and the Reappraisal condition (*B* = 3.87, SE = 1.68, *t* = 2.31, *p* = .022, 95% CI = [0.57, 7.17]). No other pairwise comparisons within Stimulus-type were significant (*p*s > .153).

US Expectancy. CS+ US expectancy exhibited significant rapid reacquisition (*B* = 2.60, SE = 0.36, *z* = 7.16, *p* < .001, 95% CI = [1.89, 3.31]), significantly differentiating itself from CS- US expectancies (*p* < .001) in the first two trials of the Test of Rapid Reacquisition. Emotion Regulation was not found to significantly affect US expectancies during the Test of Rapid Reacquisition (*p*s > .053). Two-way ANOVA of SCRs of the last two trials revealed a significant main effect of Stimulus-type ($F(436, 1) = 265$, $p < .001$), suggesting that the CS+ SCRs was consistently greater than CS- SCRs across Emotion Regulation conditions (*B* = 4.58, SE = 0.28, *t*

= 16.28, $p < .001$, 95% CI = [4.02, 5.13]). There was no significant main effect of Emotion Regulation ($p = .745$) or the Emotion Regulation x Stimulus-type interaction ($p = .833$).

Self-Report. The test of Rapid Reacquisition led to a significant increase in self-reported CS+ fearfulness ($B = -0.89$, SE = 0.15, $z = 5.88$, $p < .001$, 95% CI = [0.59, 1.19]) and CS+ arousal ($B = 0.51$, SE = 0.13, $z = 3.87$, $p < .001$, 95% CI = [0.25, 0.77]). There was also a significant decrease in CS+ valence ($B = -0.79$, SE = 0.11, $z = 7.06$, $p < .001$, 95% CI = [-1.00, -0.57]). Stimulus-type x phase interactions revealed that these changes were significantly different from CS- ($ps < .001$). Emotion Regulation condition induced no significant differences in fearfulness ($ps > .108$), arousal ($ps > .175$), or valence ($ps > .226$).

Discussion

The present study evaluated the effects of combining emotion regulation strategies with extinction training upon initial fear attenuation and return of fear as measured by tests of spontaneous recovery and rapid reacquisition one-week later. We hypothesized that the addition of a cognitive reappraisal strategy designed to complement fear extinction training would outperform fear extinction combined with suppression or fear extinction alone. Our results supported the use of cognitive reappraisal in terms of skin conductance responses and expectancies for the US during extinction training and self-reported liking of the CS+ after extinction compared to no emotion regulation strategy. Tests one week later did not support additional benefits for combining cognitive reappraisal with extinction training for the reduction of spontaneous recovery or rapid reacquisition, but suppression with extinction training did result in unique drawbacks for rapid reacquisition.

A standard differential Pavlovian fear acquisition paradigm was successful to the degree that images (CS+) paired with aversive screams (US) were associated with increasingly stronger

expectations for the scream sound and elevated skin conductance responses to the onset of the CS+. Additionally, the CS+ was reported to be scarier, more arousing, and less likeable than the CS-.

In the initial trials of extinction training, SCRs and scream expectancies to the CS+ were observed to extinguish at the same rate, in all conditions. SCRs also ended at the same level by the end of extinction training for both the CS+ and CS- in all conditions. However, cognitive reappraisal sped up the reduction of CS- scream expectancies during extinction training, suggesting that the reappraisal instruction enabled a quicker recognition of the safety of the CS-, and scream expectancies were lower in general by the end of extinction training, compared to participants not instructed to use any emotion regulatory strategy. Furthermore, participants reported liking the CS+ more after extinction training relative to participants who had been reacting-as-normal during extinction training. The results therefore suggest that although cognitive reappraisal had an effect on the expectancy to threat during extinction training, it did not affect the sympathetic arousal elicited during extinction training, which SCRs are a sensitive index of (Critchley, Mathias, & Dolan, 2002).

Examining the return of fear, neither cognitive reappraisal nor suppression induced any changes in one-week spontaneous recovery of CS+ SCRs and scream expectancies. Rapid reacquisition is a potential pathway for the return of fear particularly for individuals subjected to re-exposure to aversive events as might occur in dangerous contexts. Yet, in contrast to predictions, combining cognitive reappraisal with extinction training did not change the rate of reacquisition measured using scream expectancy or SCRs, and did not change self-report ratings after the test of rapid reacquisition. However, at the end of the rapid reacquisition phase, suppression led to greater SCRs to the CS-. This suggests that not only was suppression

unhelpful in the short and long-term for the reduction or prevention of fear expression, it led to a deleterious side-effect of long-standing sympathetic arousal elicited by safe cues.

Emotion regulation effects on slopes of the CS+ relative to the CS- may have been precluded by sample size. Replication of these procedures with larger sample sizes may help uncover effects where no effect was found when comparing across conditions. Deriving reliable estimates of the magnitude and speed of these trajectories under specific conditions will be essential for understanding how to optimize the process of fear regulation.

Our results suggest that cognitive reappraisal in conjunction with extinction training can be beneficial by making the threat cue more likable, aiding in the recognition of safety early in extinction training, and abating fearful expectancies by the end of extinction training. Using cognitive reappraisal did not have any discernable costs, although it did not lessen the spontaneous recovery or rapid reacquisition of fear. Suppression, by contrast, appears detrimental to fear extinction in the long-term. These results can encourage clinical practice involving combining reappraisal strategies with extinction training for the reduction of fear during exposure therapy. However, the type of cognitive reappraisal may be critical. The reappraisal strategy used in this study was designed to complement Rescorla-Wagner assumptions, but is not reflective of typical cognitive restructuring techniques that are often used in therapy which are often focused on some combination of self-reassurance, relaxation, attentional orientation, thought challenging, experimenting with alternative thoughts and/or reality-testing. Experimental tests of such other reappraisal strategies that systematically addresses implementation cost, the degree of alteration of expectations, and the degree of the aversiveness or CSs or USs, must be considered toward steeling oneself against fear rapid

reacquisition. The procedures outlined here may lay the groundwork for a larger replication effort, and the eventual testing of such mechanisms.

Figures

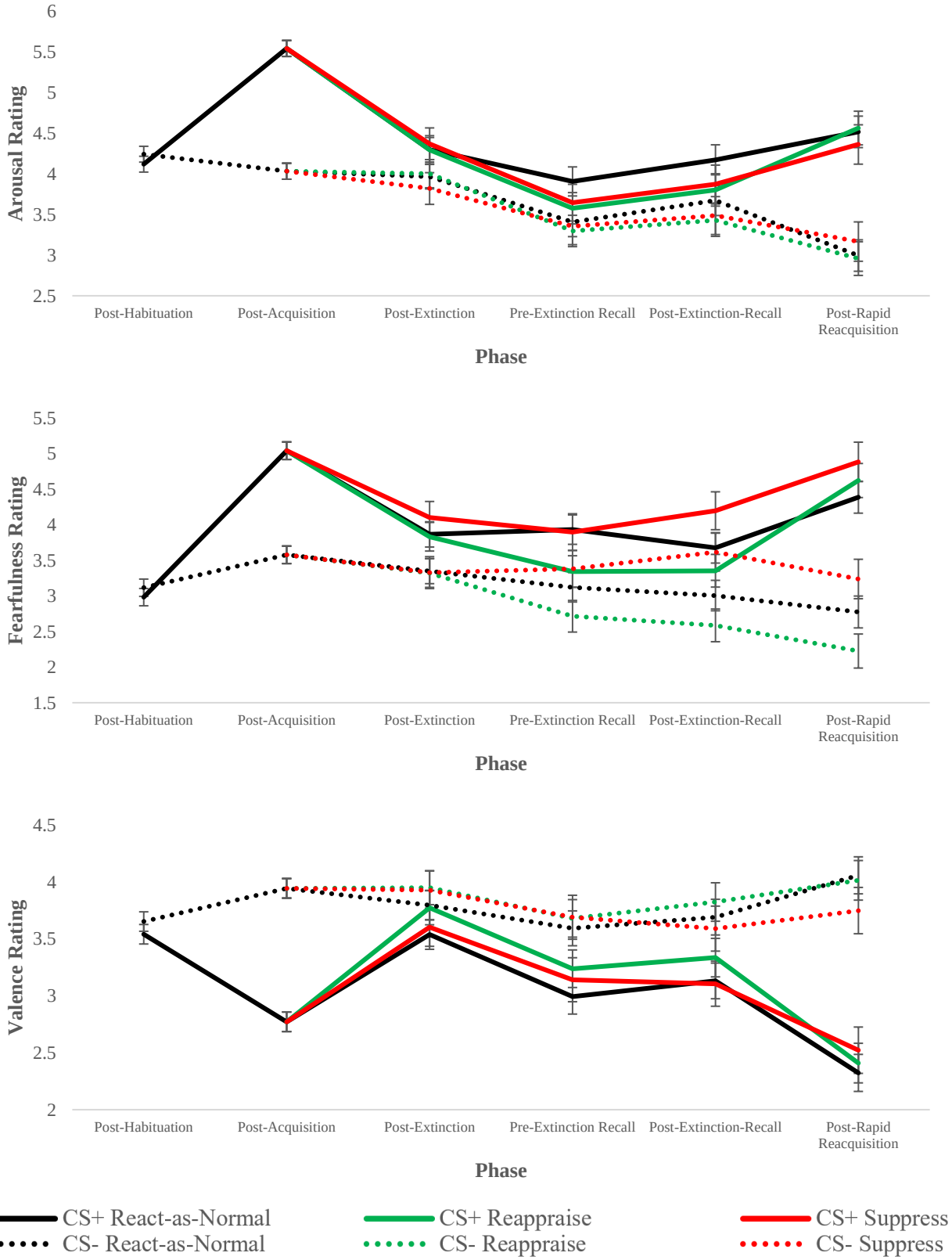


Figure 2: Self-report ratings per period.

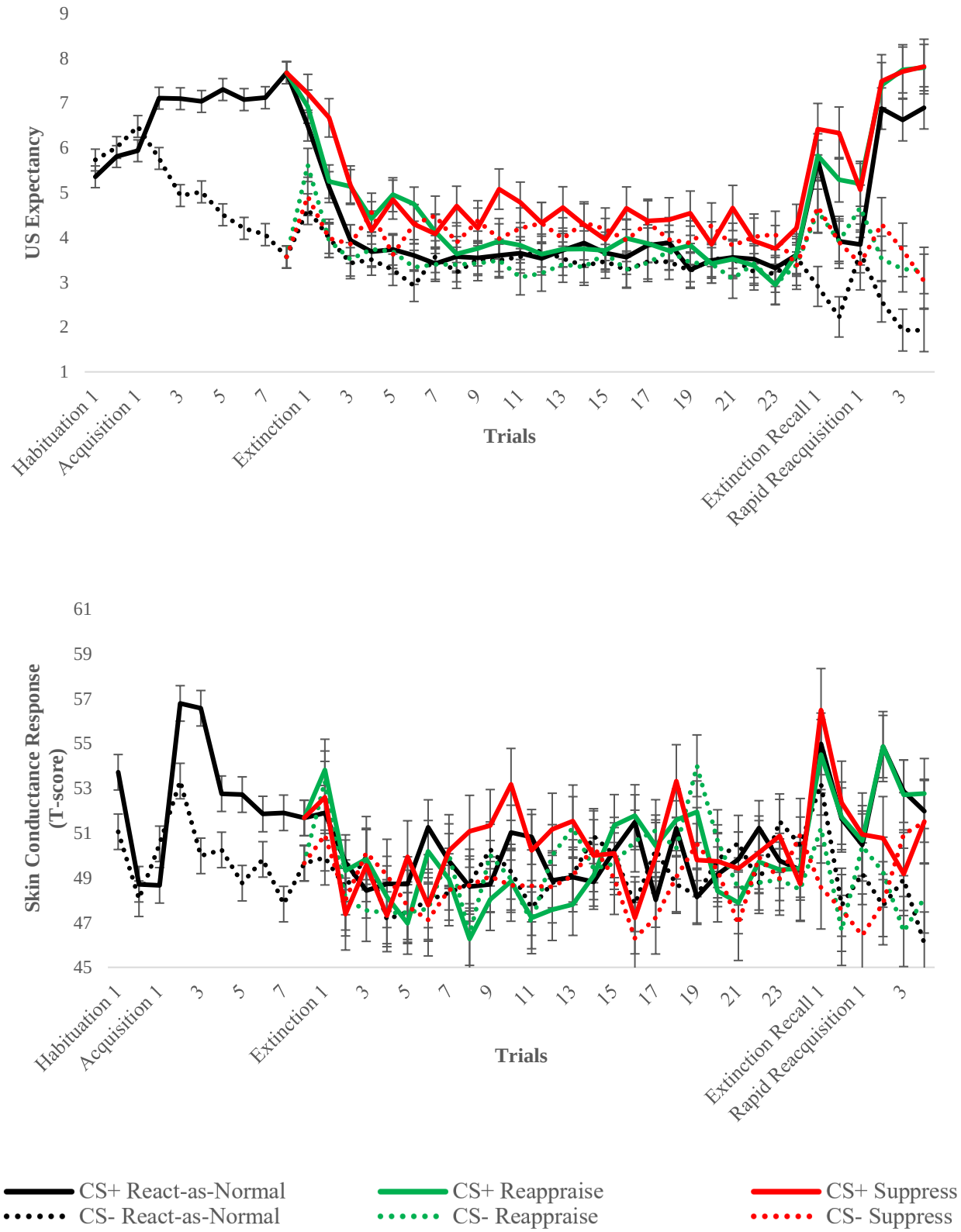


Figure 3: Continuous outcomes (US expectancy and skin conductance response) by trial.

Study 2: The Effects of Emotion Regulation on Reinstatement and Context-Renewal of Fear

Abstract

Across two studies, we examined the extent to which effortful strategies to regulate fear responses influence fear reinstatement and context renewal following extinction. We instructed participants to utilize a low-cost cognitive reappraisal strategy designed to complement prediction error correction as they underwent extinction training, and then tested spontaneous recovery, reinstatement, and context renewal one-week later. Cognitive reappraisal was compared against active (suppression instruction) and passive (react-as-normal instruction) control groups. Results suggested that cognitive reappraisal accelerated the extinction of CS- US expectancies relative to suppression, attenuated the spontaneous recovery of CS+ skin-conductance responses relative to react-as-normal, and accelerated the recovery of CS+ skin-conducted responses after fear reinstatement relative to suppression. However, no clear benefits were found in terms of context renewal. Suppression, on the other hand, was observed to produce deleterious long-term effects on context renewal relative to cognitive reappraisal, maintaining fear responses over the long term. Results are discussed in relation to clinical applications for fear-disorder treatment and relapse.

Introduction

Despite successful extinction of acquired fear, individuals who have been treated for fear-based disorders are susceptible to numerous return of fear phenomena. In the previous study (Sun & Craske, under review), we investigated the effect of instructing participants to implement a low-cost, conditional stimulus (CS)-directed cognitive reappraisal whilst undergoing fear extinction-training. Cognitive reappraisal was found to decrease US expectancies by the end of extinction training but did not attenuate spontaneous recovery (i.e., the return of fear due to the passage of time) or rapid reacquisition (i.e., the return of fear due to the repeated repairing of conditional and unconditional stimuli). The current study evaluates the effects of cognitive reappraisal upon reinstatement and context renewal of fear.

Extinction memories acquired through repeated extinction training in one context may not generalize to new contexts. Context shifts may include a diverse array of stimuli such as "physical environments, reinforcer after-effects, drug states, emotions, and the passage of time" (Bouton & Swartzentruber, 1991, p. 1). This perspective on context subsumes the phenomena of spontaneous recovery (i.e., the passage of time), reinstatement (i.e., emotions), and context renewal (i.e., location). Spontaneous recovery, fear reinstatement, and context renewal are thought to be important underlying mechanisms for relapse after exposure therapy (Bouton, 2002). Spontaneous recovery is the return of fear as a function of time (Rescorla, 2004), which has been assessed in rats across time scales ranging from hours to weeks (G. J. Quirk, 2002), and humans at least 24 hours later (e.g., Zbozinek, Hermans, Prenoveau, Liao, & Craske, 2015). Fear reinstatement is the resurgence of conditional fear expression return of fear after unsignalled presentations of the unconditional stimulus (US) (Rescorla & Heth, 1975). "Context renewal" within the scope of this work refers primarily to location surrounding the CS. The effect of

context shifts leading to a renewal of fear is highly robust (Bouton & Bolles, 1979), and experimental designs have featured many different configurations, with contexts denoted by lettering (e.g., ABA vs. AAB design).

Cognitive reappraisal strategies, that are often used for self-regulation, have been demonstrated to affect neural (e.g., Goldin, McRae, Ramel, & Gross, 2008; Ochsner, Bunge, Gross, & Gabrieli, 2002) and peripheral (e.g., Gross, 1998) physiology, as well as self-reported emotional states (e.g., Gross & John, 2003). Cognitive reappraisal is often conceptualized as an effortfully implemented, "top-down" regulatory strategy, which contrasts with experientially-based "bottom-up" extinction-training. There is potential for productive strategic combinations of cognitive reappraisal and extinction training that may more effectively reduce or prevent return of conditional fear responses, relative to each strategy alone. Moreover, evidence for attenuation of the return of fear through cognitive reappraisal may encourage clinical interventions such as brief cognitive restructuring during exposure to reduce relapse after exposure therapy.

One potential problem of using cognitive reappraisal to reduce conditional fear responses is derived from the Rescorla-Wagner model (Rescorla & Wagner, 1972), in that cognitive reappraisal that interferes with prediction error normally induced by extinction training might impair extinction learning. For this reason, we designed a cognitive reappraisal instruction targeting US properties, as opposed to the prediction of US occurrence, which we posited would be less likely to interfere with, and instead complement, extinction learning. Across two studies, we assessed whether cognitive reappraisal augments the reduction of conditional fear responses during extinction training, and at tests of spontaneous recovery, fear reinstatement and context renewal one-week later. We hypothesized that, compared to passive (i.e., react as normal) and

active (i.e., suppress) control conditions, cognitive reappraisal would reduce conditional fear responses during extinction training, and reduce conditional fear responses at tests of spontaneous recovery (Study 1 and Study 2), reinstatement (Study 1), and context renewal (Study 2) one-week later.

Methods

Participants

For Studies 1 and 2, healthy adults over the age of 18 were recruited from a student subject pool at the University of California, Los Angeles, and received course credit for their participation. Participants were excluded if they were diagnosed with a previous or current mental health disorder (self-reported), if they were currently using psychoactive medication, if they had uncorrected problems with vision or hearing, if they were pregnant, if they had a serious medical condition, if they were unable to speak or understand English, or if they had recent caffeine intake that may affect online measurement of peripheral physiology. Sample characteristics for each study are described in their respective sections.

Design

For both studies, we evaluated the impact of combining emotion regulatory strategies with extinction training upon extinction performance, and spontaneous recovery. Participants were randomly-assigned to (1) Reappraisal, (2) Suppression (as an active control), or (3) React-as-Normal conditions (as a passive control), and the randomized regulatory instruction was presented after fear acquisition and before extinction-training. Spontaneous recovery was tested 1 week later. After spontaneous recovery, Study 1 investigated the impact of combining emotion regulatory strategies with extinction training upon reinstatement (Figure 1). Study 2 investigated

the impact of combining emotion regulatory strategies with extinction training upon context renewal (Figure 2).

Procedure

On Day 1, participants underwent Baseline, Habituation, Acquisition, and Extinction phases. In the Baseline phase, participants viewed a blank screen as physiology was recorded for 5 min. In the Habituation phase, participants were acclimated to the CS (neutral facial images) through two six-second trials. During Acquisition, the CS+ was paired with a 1-second 82dB scream sound as the US that co-terminated with the CS+. The CS- was not paired with the US. Each CS+/- was presented for eight trials. CS+ and CS- images were counterbalanced across participants. CS trials were presented in random order, with no more than two consecutive trials of the same CS type.

According to randomization, one of three instructions appeared after completion of the Acquisition phase. In the Cognitive Reappraisal condition, participants were asked to think of themselves as a casting director who is evaluating a screaming actress trying out for a part in a scary movie with the following instruction: *"For the next couple of minutes, imagine that you are a Hollywood star hiring manager. You have just gotten the go-ahead on a big television horror-drama, and are looking to hire several lead actresses. You want to hire the actress with the most realistic scream possible. Please be as objective in your selection as you possibly can. Listen carefully and select your next lead actress."*

In the Suppression condition, participants were asked to be unexpressive in order to prevent others from knowing how they feel as they listened to the scream. The instruction was as follows: "*For the next couple of minutes, we would like to see how well you can keep from showing any emotional response when you hear a scream. Try not to feel anything, and try not to*

have a physiological reaction. Also, see if you can act so that someone seeing the video with the sound off won't know that anything has happened. Try not to show any visible signs or feel anything before, during, or after the scream occurs. Try to look relaxed all the way through. See if you can fool the person who will be studying this video."

Finally, the React-as-Normal condition instructed participants to simply continue watching the computer screen. The exact instruction was as follows: "*For the next couple of minutes, you will be presented a series of faces and sounds. Simply attend to the computer screen as you would naturally do so. Do not distract yourself by thinking about other things.*"

All instructions were designed to be low in effort and to have no effect upon prediction error (i.e., no effects on either perceived US intensity or US frequency). This addresses a burgeoning concern about the cost of emotion regulatory implementation (Etkin et al., 2015; Ford & Troy, 2019). The instructions were followed by an Extinction phase, involving 24 trials of each CS without the US in random order, with no more than two consecutive trials of the same CS type. One week later, participants were tested for Spontaneous Recovery, which involved two trials of each CS without the US. Test of Spontaneous Recovery was followed by Reinstatement in Study 1 and context renewal in Study 2 (see details below).

Materials/Apparatus

Participants sat in front of a 21-inch monitor situated roughly 20-inches away from eye-level. Stimulus presentation was programmed and controlled using E-Prime 2.0 Professional (version 2.0.10.353; Psychology Software Tools, Pittsburgh, PA), installed on a personal computer running Windows 7. Two NimStim (Tottenham et al., 2009) images of Asian female faces wearing a neutral expression were chosen as conditional stimuli (CS) (counterbalanced between participants). To ensure attentional capture and sufficient salience, the height of the CSs

matched the maximum height of the screen, with the width expanded to be proportional to the height.

A BioPac MP150 (Biopac System Inc., Goleta, CA, USA) was used to collect peripheral physiology data. Specifically, the BioPac amplifiers EDA100C was used to measure skin conductance responses (SCR) and EMG100C was used to measure fear-potentiated startle blink through electromyography (EMG). Data were recorded from the BioPac MP150 using AcqKnowledge 4.2 software (Biopac System Inc., Goleta, CA, USA), and cleaned, inspected, and analyzed with ANSLab software (ANSLab v2.6, Wilhelm & Peyck, 2005).

Skin Conductance Responses. SCRs were recorded from two EL507 11 mm diameter Ag/AgCl electrodes placed on the distal phalanx of the index and middle fingers of the non-dominant hand. Using an EDA100C amplifier and two LEAD110A electrode leads, SCR data was sampled at a rate of 2000 Hz and filtered using a finite impulse response (FIR) low pass filter with a frequency cutoff fixed at 2 Hz. SCR was calculated as a difference score between the maximum skin conductance value 1–6 s after CS onset minus the mean skin conductance value of the 2 seconds prior to CS onset. SCRs greater than zero were square root transformed to normalize the data (Levey, 1980). SCRs less than or equal to zero were coded as zero. To eliminate individual variability in SCR range, SCRs were *T*-score standardized by subtracting each SCR by the mean and then divided by the standard deviation of the SCRs across all CSs across both days for each participant. This was then multiplied by 10 and added to 50. The formula for this conversion is as follows:

$$SCR_T = \frac{SCR_{Observed} - SCR_{mean}}{SCR_{SD}} \times 10 + 50$$

Self-Report. Online ratings of US expectancy were made via a BioPac TSD115 continuous sliding dial. Participants were asked to continually adjust how certain they were that

the US (scream sound) will appear at the end of the trial with the question “how certain are you that you will hear a sound in the next few moments”. Participants received 3-second prompts at the beginning of each ITI and CS reminding them to use the expectancy dial. The values range from 0 = “Certain no sound”, 4.5 = “Uncertain”, and 9 = “Certain sound”. US expectancy were calculated as the mean rating 6.5–7 seconds after ITI, CS+ or CS− startle probe onset. Before and after every phase, valence, arousal, and fearfulness ratings (from 1-low to 9-high) associated with each stimulus were assessed onscreen. Data were recorded via keypress after presentation of the scale on the computer screen.

Study 1. Reinstatement

Participant Characteristics

Participants were 112 individuals averaging 20.46 years of age (SD = 3.28), 76.84% of which were female. The racial/ethnic breakdown was 38.39% European American, 35.71% Asian, 15.18% Hispanic/Latino, and 10.72% Other.

Design

The aim of Study 1 was to evaluate the impact of combining emotion regulatory strategies with extinction training upon extinction performance, spontaneous recovery and fear reinstatement (see Figure 1). The test for Spontaneous Recovery was followed by Reinstatement, involving two unsignalled USs, and then four trials of each CS without the US as the Test of Reinstatement.

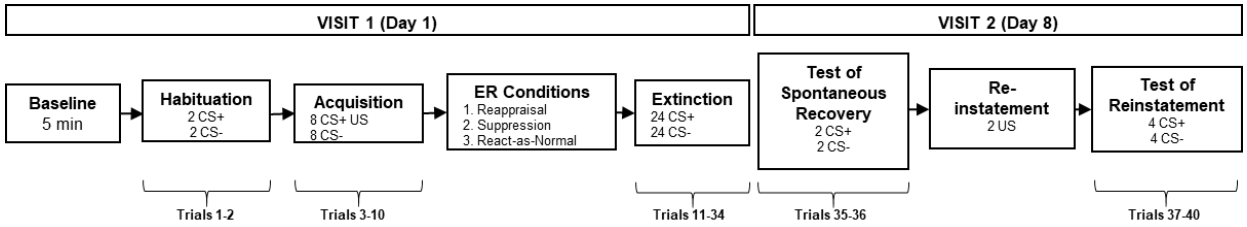


Figure 1: Experimental design of Study 1
Note: ER Conditions refers to the emotion-regulation instruction set shown to the participant prior to the onset of the Extinction phase.

Data Analysis

Statistical analyses were conducted from a multilevel modelling framework (Bryk & Raudenbush, 1992) in STATA 14.1, using the *mixed* command, setting repeated measures at level 1 nested within individuals at level 2. Examination of response separation by Stimulus type (CS+ vs. CS-) was conducted within the Acquisition phase (trials 3-10), fitting best-fit prediction lines with stepwise-polynomial regressors for multi-trial outcomes (i.e., SCRs and US expectancies). These outcomes were predicted from Stimulus-type (0 = CS+, 1 = CS-), each of the polynomial components of the best-fit lines (e.g., instantaneous linear, quadratic component), and Stimulus-type x polynomial component interactions.

Instructed emotion regulation was expected to affect the response to the CS+ and CS-, but not their essential aspect (e.g., the CS+ would not become the CS-, or vice versa). Therefore best-fit lines were estimated separately for the CS+ and CS- for the Extinction phase (trials 11-34), and the Test of Reinstatement (37-40). These lines were then predicted by Emotion Regulation condition terms (0 vs. 1 and 0 vs. 2, where 0 = React-as-Normal, 1 = Reappraisal, 2 = Suppression) and their interactions with polynomial slope components. Models were re-run to compare Reappraisal with Suppression conditions (i.e., 1 vs. 2). Intercepts refer to the response at the first trial of the phase, the instantaneous linear slope refers to the initial change in responding, and further polynomial slopes refers to curvatures in the change throughout the

phase. To ensure that extinction was achieved and to examine if there were any differences in responding by the end of extinction due to Emotion Regulation, a two-way ANOVA was run estimating the final two trials of the Extinction phase by Stimulus-type, Emotion Regulation, and their interaction. Extinction was achieved if there was no significant Stimulus-type or Stimulus-type x Emotion Regulation interaction.

Simple linear multilevel models were fitted to model SCR and US expectancy changes during the Test of Spontaneous Recovery (trials 34-35). Regressors for these models were Emotion Regulation, period (0 = pre, 1 = post), and their interactions. A similar approach was used to predict changes in self-reported fear, arousal, and valence before and after the Acquisition phase, Extinction phase, Test of Spontaneous Recovery, and the Test of Reinstatement, including Stimulus-type, Emotion Regulation, period, and their interactions as predictors.

Results

Acquisition

CS+ SCRs (B = 11.72, SE = 1.98, z = 5.92, p < .001, 95% CI = [7.84, 15.60]) and CS+ US expectancies (B = 0.75, SE = 0.33, z = 2.30, p = .022, 95% CI = [0.11, 1.39]) exhibited significantly increasing instantaneous linear components of the Acquisition learning curve. Significant Stimulus-type x phase interactions revealed greater changes to the CS+ than CS- for SCRs and US expectancies (ps < .001). After Acquisition, CS+ self-report ratings of fearfulness (B = 1.81, SE = .16, z = 11.62, p < .001, 95% CI = 1.51, 2.12]) and arousal (B = 1.37, SE = .16, z = 8.58, p < .001, 95% CI = [1.05, 1.68]) significantly increased, while CS+ valence significantly decreased (B = -1.12, SE = 0.15, z = -7.49, p < .001, 95% CI = [-1.42, -0.83]). In all cases of self-

report, Stimulus-type x phase interactions revealed that changes for CS+ were significantly different from changes for CS- (*p*s < .001).

Extinction

SCR. CS+ SCRs (*B* = -4.76, SE = 0.99, *z* = -4.79, *p* < .001, 95% CI = [-6.71, -2.81]) and CS- SCRs (*B* = -1.90, SE = 0.65, *z* = -2.91, *p* = .004, 95% CI = [-3.19, -0.62]) exhibited descending instantaneous linear reductions. Emotion Regulation condition terms were not found to affect the slopes of the best-fit lines to the CS+ (*p*s > .359) or the CS- (*p*s > .082). Two-way ANOVA revealed no significant differences in SCR responding by Stimulus-type at the end of the Extinction phase ($F(1, 358) = 2.94$, $p = .087$). There was also no significant effect of Emotion Regulation ($F(2, 358) = 0.72$, $p = .487$) or the Stimulus x Emotion Regulation interaction ($F(1, 358) = 2.94$, $p = .488$).

US Expectancy. CS+ US expectancies (*B* = -1.38, SE = 0.16, *z* = -8.65, *p* < .001, 95% CI = [-1.69, -1.07]) and CS- US expectancies (*B* = -0.70, SE = 0.15, *z* = -4.77, *p* < .001, 95% CI = [-0.98, -0.41]) exhibited significant instantaneous linear decreases in the Extinction phase. Emotion Regulation did not significantly affect the slopes of the CS+ US expectancy extinction curves (*p*s > .376). The CS- US expectancy extinction curves for the Reappraisal condition exhibited a significantly steeper instantaneous linear decrease relative to the Suppression condition (*B* = -0.86, SE = 0.39, *z* = -2.18, *p* = .029, 95% CI = [-1.63, -0.09]). No detectable differences were found in the slopes of the CS- extinction curves between Reappraisal and React-as-Normal conditions, or between Suppression and React-as-Normal conditions (*p*s > .164). A two-way ANOVA revealed no significant differences in US expectancy by Stimulus-type ($F(1, 362) = 0.06$, $p = .802$) at the end of the Extinction phase. There was also no significant

effect of Emotion Regulation ($F(2, 362) = 1.79$, $p = .168$) or the Stimulus x Emotion Regulation interaction ($F(2, 362) = 0.16$, $p = .856$).

Self-Report. Extinction induced a significant reduction in self-reported ratings of CS+ fearfulness ($B = -1.33$, SE = 0.16, $z = -8.55$, $p < .001$, 95% CI = [-1.64, -1.03]) and CS+ arousal ($B = -1.32$, SE = 0.16, $z = -8.24$, $p < .001$, 95% CI = [-1.64, -1.01]) as well as a significant increase in self-reported CS+ valence ($B = 0.83$, SE = 0.16, $z = 5.18$, $p < .001$, 95% CI = [0.52, 1.15]). Stimulus-type x phase interactions revealed that CS+ ratings were significantly different from changes in CS- ratings (ps < .001). Emotion Regulation did not affect Extinction-induced changes (ps = .092).

Tests of Spontaneous Recovery

SCR. When tested one-week later, there was marginal evidence of spontaneous recovery to the CS+ ($B = 2.11$, SE = 1.23, $z = 1.71$, $p = .087$, 95% CI = [-0.30, 4.53]) and CS- ($B = 1.98$, SE = 1.18, $z = 1.67$, $p = .094$, 95% CI = [-0.34, 4.30]) in SCRs. CS+ SCRs in the Reappraisal condition ($B = -6.11$, SE = 2.92, $z = -2.09$, $p = .037$, 95% CI = [-11.84, -0.37]), but not the Suppression condition ($p = .166$), was significantly reduced relative to the React-as-Normal condition. Simple slopes analyses revealed that while there was evidence of significant CS+ SCR spontaneous recovery in the React-as-Normal condition ($B = 5.39$, SE = 1.48, $z = 3.63$, $p < .001$, 95% CI = [2.48, 8.30]), there was not in the Reappraisal ($p = .821$) or Suppression conditions ($p = .459$).

Interactions with CS- SCR revealed that recovery was significantly reduced in the Suppression ($B = -6.84$, SE = 2.96, $z = -2.31$, $p = .021$, 95% CI = [-12.65, -1.03]), but not the Reappraisal condition ($p = .359$), relative to the React-as-Normal condition. Simple slopes analyses revealed that the React-as-Normal condition exhibited significant CS- SCR recovery (B

= 4.98, SE = 2.09, z = 2.39, p = .017, 95% CI = [0.89, 9.08]), whereas the Suppression (p = .327) and Reappraisal (p = .300) conditions did not.

US Expectancy. When tested one-week later, there was evidence of spontaneous recovery to the CS+ (B = 2.17, SE = 0.43, z = 5.04, p < .001, 95% CI = [1.33, 3.01]) and CS- (B = 0.84, SE = 0.35, z = 2.40, p = .016, 95% CI = [0.15, 1.52]). Spontaneous recovery in CS+ US expectancies (ps > .277) and CS- US expectancies (ps > .112) did not exhibit significant differences due to Emotion Regulation.

Self-Report. From the end of the Extinction phase to the Test of Spontaneous Recovery, we observed a significant reduction in self-reported ratings of CS+ arousal (B = -0.37, SE = 0.16, z = -2.27, p = .023, 95% CI = [-0.70, -0.05]) as well as a significant increase in self-reported CS+ valence (B = -0.28, SE = 0.13, z = -2.16, p = .031, 95% CI = [-0.54, -0.03]). CS+ fearfulness remained unchanged (p = .268). Stimulus-type x phase interactions did not reveal significant differences in these changes from changes in CS- ratings (ps > .129). Emotion Regulation did not affect these changes (ps = .433).

Test of Reinstatement

SCR. Throughout the Test of Reinstatement, CS- SCRs exhibited an instantaneous linear decline (B = -2.99, SE = 1.00, z = -2.99, p = .003, 95% CI = [-4.95, -1.03]) that slowed down over time as indicated with a significant positive quadratic component (B = 0.64, SE = 0.32, z = 1.99, p = .047, 95% CI = [0.01, 1.26]). The CS+ did not significantly change throughout the same period (p = .430). Emotion Regulation did not affect changes in the CS+ (ps > .068) or CS- (ps > .119).

US Expectancy. US expectancy ratings during the Test of Reinstatement exhibited a significantly higher intercept for the CS+ relative to the CS- (B = 1.04, SE = 0.26, z = 3.94, p

< .001, 95% CI = [0.52, 1.56]), evidencing reinstatement. CS+ US expectancies exhibited a significantly linear decrease (B = -0.53, SE = 0.09, z = -5.93, p < .001, 95% CI = [-0.70, -0.35]) throughout the Test of Reinstatement, while CS- US expectancies did not (p = .076). US expectancies in the Reappraisal condition exhibited a significantly steeper linear decrease relative to the Suppression condition for the CS+ (B = -0.51, SE = 0.23, z = -2.25, p = .025, 95% CI = [-0.95, -0.06]) and a marginally steeper linear decrease in CS- (B = -0.42, SE = 0.24, z = -1.77, p = .076, 95% CI = [-0.89, -0.04]). No differences were found comparing Reappraisal with React-as-Normal conditions (ps > .265) or Suppression with React-as-Normal conditions (ps > .118) in US expectancies throughout the Test of Reinstatement.

Self-Report. From before to after the Test of Reinstatement, there was a marginal decrease in CS+ fearfulness (B = -0.29, SE = 0.15, z = -1.88, p = .060, 95% CI = [-0.59, -0.01]), a significant increase in self-reported CS+ valence (B = 0.44, SE = 0.11, z = 3.93, p < .001, 95% CI = [0.22, 0.66]), and no significant change in CS+ arousal (p = .609). Stimulus-type x period interactions revealed that changes in CS+ valence were significantly different from the changes in CS- valence (p = .031) but not CS+ ratings were not significantly different from changes in CS- fearfulness (p = .254) or CS- arousal (p = .771). Emotion Regulation did not affect these changes (ps > .097).

Discussion

In this study of fear reinstatement, we evaluated the effects of combining emotion regulation strategies with extinction upon extinction training as well as spontaneous recovery and fear-reinstatement one-week later. We hypothesized that the addition of a cognitive reappraisal strategy designed to complement fear extinction training would outperform fear extinction

compared with suppression or fear extinction alone. Our results overall supported advantages to using cognitive reappraisal relative to suppression and reacting as normal.

Cognitive reappraisal induced a steeper decrease in US expectancy to the CS- during extinction relative to suppression. Conceivably, cognitive reappraisal led to increased certainty in the appraisal that the CS- is safe compared to suppressing emotion. One-week later, emotion regulation, compared to react-as-normal, affected the recovery of SCRs. Specifically, cognitive reappraisal attenuated spontaneous recovery of SCRs to the CS+, whereas suppression attenuated recovery of SCRs to the CS-. This set of findings could mean that cognitive reappraisal facilitated participants' memory that the CS+ was no longer threatening (as a result of extinction training), whereas suppression facilitated memory that the CS- never was threatening. However, the effect of suppression should be interpreted with caution as it was not hypothesized, and there is no known theoretical basis for why such an effect would be expected. In terms of fear reinstatement, cognitive reappraisal led to larger reductions of CS+ US expectancy compared to suppression. This suggests that while cognitive reappraisal did not attenuate fear reinstatement relative to extinction training alone (react as normal), it aided in recovery from reinstatement relative to suppression.

Study 2. Context Renewal

Participant Characteristics

The participants in Study 2 consisted of 132 individuals averaging 20.50 years of age (SD = 3.05), 79.66% of which were female. The racial/ethnic breakdown consisted of 34.17% European American, 42.50% Asian, 9.17% Hispanic/Latino, and 14.17% Other.

Design

The aim of Study 2 was to evaluate the impact of combining emotion regulatory strategies with extinction training upon extinction performance, spontaneous recovery and the context renewal of fear (see Figure 2). As with Study 2, participants were randomly assigned to (1) Reappraisal, (2) Suppression, or (3) React-as-Normal conditions. Context Renewal was tested using an ABA design. Contexts were represented as two counterbalanced background images, a living room setting (BG1) and an outdoor porch (BG2), presented 'behind' CS presentations. On the first day, Habituation involved trials of each CS and background image alone. Acquisition involved CSs superimposed on top of one background image (BG-A) and Extinction involved CSs superimposed on top of the other background image (BG-B), in counterbalanced order. One week later, the test of Spontaneous Recovery involved CSs re-presented superimposed on BG-B. The test of Spontaneous Recovery was followed by a Context Renewal phase, which involved testing CS superimposed on BG-A versus BG-B.

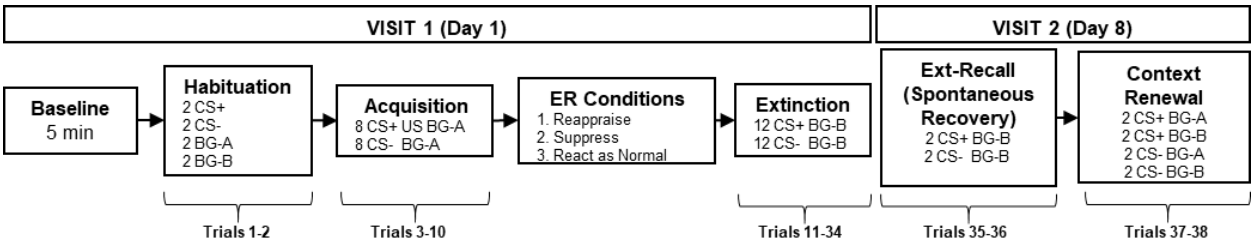


Figure 2: Experimental design of Study 2.
Note: ER Conditions refers to the emotion-regulation instruction set shown to the participant prior to the onset of the Extinction phase.

Data Analyses

Analyses up to the Test of Spontaneous Recovery were identical to Study 1 except for the Context Renewal phase. Context Renewal was assessed between the CS+/CS- trial superimposed on BG-B, and the CS+/CS- trial superimposed on BG-A. A first model examined the change in SCR and US expectancy from trials 36 to 37 with a Context Renewal condition variable (0 = CS- atop BGB on trial 36 and atop BGA on trial 37; 1 = CS+ atop BGB on trial 36 and atop BGA on

trial 37). A significant Context Renewal x trial interaction was tested to examine if there was a significant difference between CS+ and CS- responding when the context was changed. If so, CS+ and CS- were then modelled separately with simple linear multilevel modelling estimating responses predicted from trial (0 = trial 36; 1 = trial 37), Emotion Regulation terms (0 vs. 1 and 0 vs. 2, where 0 = React-as-Normal, 1 = Reappraisal, 2 = Suppression), and their interactions to test if Emotion Regulation affected the change in responding due to Context Renewal. Models were rerun to also compare Reappraisal with Suppression conditions (i.e., 1 vs. 2). Effects of Emotion Regulation on average response differences throughout the test of Context Renewal were then assessed with one-way ANOVA followed by pairwise comparisons. Simple linear multilevel modelling was used to predict changes in self-reported fear, arousal, and valence before and after the Context Renewal phase, including Stimulus-type and its interactions as predictors to compare the changes between CS+ and CS-.

Results

Acquisition

CS+ SCRs (B = 10.10, SE = 2.21, z = 4.58, p < .001, 95% CI=[5.78, 14.43]), and CS+ US expectancy ratings (B = 0.06 SE = 0.03, z = 2.21, p = .027, 95% CI=[0.01, 0.11]) exhibited significantly increasing instantaneous linear components of the Acquisition learning curve. Significant Stimulus-type x phase interactions revealed greater changes to the CS+ than CS- for SCRs and US expectancy ratings (ps < .001). After Acquisition, self-report ratings of CS+ fearfulness (B = 1.72, SE = 0.15, z = -8.27, p < .001, 95% CI = 1.42, 2.02]) and CS+ arousal (B = 1.26, SE = 0.15, z = 8.33, p < .001, 95% CI = [0.97, 1.56]) significantly increased. Self-report ratings of CS+ valence significantly decreased (B = -0.83, SE = 0.15, z = -5.61, p < .001, 95% CI

= [-1.13, -0.54]). In all cases, significant Stimulus-type x phase interactions revealed that the changes in CS+ were significantly different from changes in the CS- (*p*s < .001).

Extinction

SCR. CS+ (*B* = -0.61, SE = 0.23, *z* = -2.68, *p* = .007, 95% CI = [-1.05, -0.16]) and CS- (*B* = -1.41, SE = 0.41, *z* = -3.42, *p* = .001, 95% CI = [-2.22, -0.60]) SCRs exhibited decreasing instantaneous linear slopes at the beginning of the Extinction phase. Emotion Regulation did not significantly affect CS+ (*p*s > .343) or CS- (*p*s > .426) SCRs. Two-way ANOVA revealed no significant differences in SCR responding by Stimulus-type at the end of the Extinction phase ($F(1, 438) = 0.34$, $p = .561$). There was also no significant effect of Emotion Regulation ($F(2, 438) = 0.80$, $p = .451$) or the Stimulus x Emotion Regulation interaction ($F(2, 438) = 1.14$, $p = .321$).

US Expectancy. CS+ (*B* = -12.61, SE = 3.79, *z* = -3.33, *p* = .001, 95% CI = [-20.03, -5.19]) and CS- (*B* = -10.95, SE = 3.73, *z* = -2.93, *p* = .003, 95% CI = [-18.27, -3.63]) US expectancies exhibited significantly decreasing instantaneous linear slopes. However, no significant effects of Emotion Regulation were found on the slopes of CS+ (*ps* > .318) or CS- (*ps* > .098) US expectancies. Two-way ANOVAs revealed no significant differences in US expectancy by Stimulus-type at the end of the Extinction phase ($F(1, 438) = 1.18$, $p = .278$). There was also no significant effect of Emotion Regulation ($F(2, 438) = 0.48$, $p = .617$) or the Stimulus x Emotion Regulation interaction ($F(2, 438) = 0.00$, $p = .998$).

Self-Report. Extinction induced a significant reduction in self-reported ratings of CS+ fearfulness (*B* = -1.34, SE = 0.15, *z* = -8.77, *p* < .001, 95% CI = [-1.65, -1.04]) and arousal (*B* = -1.46, SE = 0.16, *z* = -8.98, *p* < .001, 95% CI = [-1.96, -1.33]), as well as a significant increase in self-reported CS+ valence (*B* = 0.89, SE = 0.15, *z* = 5.90, *p* < .001, 95% CI = [0.59, 1.18]).

Stimulus-type x phase interactions revealed that these changes were significantly different from the changes in the CS- (*p*s < .001). Emotion Regulation did not induce significant differences in changes in fearfulness (*p*s > .259), arousal (*p*s > .390), or valence (*p*s > .625).

Tests of Spontaneous Recovery

SCR. When tested one-week later in the same context as extinction (BG-B), there was significant spontaneous recovery in CS+ (*B* = 6.50, SE = 1.87, *z* = 3.47, *p* = .001, 95% CI = [2.83, 10.18]) but not CS- (*p* =.062) SCRs. There were no significant effects of Emotion Regulation condition upon CS+ (*p*s > .132) or CS- (*p*s > .716) SCR recovery.

US Expectancy. There was a significant spontaneous recovery in the CS+ US expectancy (*B* = 0.79, SE = 0.37, *z* = 2.15, *p* = .031, 95% CI = [0.07, 1.50]) but not CS- US expectancy (*p* = .446). Emotion Regulation condition did not significantly affect CS+ (*p*s > .548) or CS- (*p*s > .163) recovery of US expectancy ratings.

Self-Report. From the end of the Extinction phase to the Test of Spontaneous Recovery, we observed a significant decrease in self-reported CS+ valence (*B* = -0.49, SE = 0.13, *z* = -3.79, *p* < .001, 95% CI = [-.74, -0.24]), with no significant corresponding change in self-reported CS+ arousal (*p* = .066) or fearfulness (*p* = .195). Stimulus-type x phase interactions revealed that these changes were not significantly different from changes in the CS- (*p*s > .099). Emotion Regulation did not induce significant differences in changes in fearfulness (*p*s > .701), arousal (*p*s > .419), or valence (*p*s > .755).

Context Renewal

SCR. Significant Context Renewal was found for CS+ SCRs (*B* = 5.99, *SE* = 2.00, *z* = 3.00, *p* = .003, 95% CI = [2.08, 9.91]) and CS- SCRs (*B* = 3.95, *SE* = 1.46, *z* = 2.70, *p* = .007, 95% CI = [1.08, 6.81]), but their relative changes were not significantly different from one

another (p = 0.407) No differences were observed in SCR changes across Emotion Regulation conditions to either the CS+ (ps > .516) or CS- (ps > .588). A one-way ANOVA revealed a significant effect of Emotion Regulation condition on SCRs to the CS+ superimposed on BG-A (F(2, 177) = 3.40, p = .036) such that Suppression exhibited significantly elevated SCRs relative to both Reappraisal (B = 5.55, SE = 2.44, t = 2.28, p = .024, 95% CI = [0.74, 10.36]) and React-as-Normal (B = 5.38, SE = 2.51, t = 2.14, p = .034, 95% CI = [0.42, 10.34]). SCRs to the CS+ superimposed on BG-B (F(2, 177) = 0.10, p =.909), to the CS- superimposed on BG-A (F(2, 175) = 1.65, p = .195), and to the CS- superimposed on BG-B (F(2, 178) = 0.41, p = .663) were not significantly affected by Emotion Regulation condition during the test of Context Renewal.

US Expectancy. Significant Context Renewal was found for CS+ US expectancies (B = 1.81, SE = 0.39, z = 4.70, p < .001, 95% CI = [1.06, 2.57]) and CS- US expectancies (B = 0.60, SE = 0.29, z = 2.09, p = .037, 95% CI = [0.04, 1.17]), and increase in CS+ US expectancy was significantly greater than the CS- US expectancy (B = 2.22, SE = 0.48, z = 2.55, p = .011, 95% CI = [0.28, 2.16). No differences in US expectancy change was observed across Emotion Regulation conditions to either the CS+ (ps > .674) or CS- (ps > .202). A one-way ANOVA did not reveal any significant differences in US expectancy by Emotion Regulation condition to the CS+ superimposed on BG-A (F(2, 177) = 1.22, p = .298), to the CS+ superimposed on BG-B (F(2, 177) = 0.15, p = .863), to the CS- superimposed on BG-A (F(2, 175) = 0.06, p = .944), or to the CS- superimposed on BG-B (F(2, 178) = 0.99, p = .375).

Self-Report. Pre- to post-Context Renewal phase, we observed a significant decrease in self-reported CS+ fearfulness (B = -0.37, SE = 0.15, z = -2.56, p = .010, 95% CI = [-0.66, -0.09]) and CS+ arousal (B = -0.31, SE = 0.12, z = -2.60, p = .009, 95% CI = [-0.54, -0.08]). There was no significant change in CS+ valence (p = .282). Stimulus-type x phase interactions revealed that

these changes were not significantly different from changes in the CS- (*p*s > .469). Emotion Regulation did not induce significant differences in changes in fearfulness (*p*s > .173), arousal (*p*s > .486), or valence (*p*s > .580).

Discussion

In Study 2, we evaluated the effects of combining emotion regulation strategies with extinction upon extinction training, spontaneous recovery and context renewal one-week later. As with Study 1, we hypothesized that the addition of a cognitive reappraisal would outperform fear extinction compared with suppression or fear extinction alone.

When re-presenting conditional stimuli superimposed on the original context in which fear was acquired (that differed from the extinction context), context renewal was clearly indexed in US expectancy, tracking expected associative properties, whereas it was not as clearly indexed in SCR despite featuring significant resurgence.

In contrast to hypotheses, cognitive reappraisal had no observable effects upon extinction training, spontaneous recovery, or context renewal, relative to extinction alone (i.e., react as normal). However, context renewal of the SCR to the CS+ was significantly greater when extinction was combined with suppression relative to extinction training alone (i.e., react as normal). These results could mean advising against suppression during extinction, as doing so increases the autonomic preparation for aversive outcomes when extinguished stimuli are re-encountered in original fear-conditioning contexts.

Our findings with regard to context renewal are confounded by tests of spontaneous recovery that always occurred first. In other words, pre-exposure to the CSs superimposed on BG-B occurred immediately prior to the test of context renewal. Consequently, due to the counterbalancing of stimulus presentations in the test of Context Renewal, half of the

participants received two pre-exposure trials to the CS+ superimposed on BG-B, while the other half received three pre-exposure trials to the CS+ superimposed on BG-B. The Rescorla-Wagner model (Rescorla & Wagner, 1972) would suggest that these pre-exposures should not affect the conditional fear strength elicited by BG-A, but that they would act as additional extinction trials that further strengthen the inhibitory conditioning of the CSs, potentially mitigating context renewal.

Conclusion

The current studies investigated whether low-cost emotion regulatory strategies augmented the effects of extinction training and extinction learning, as indexed by tests of spontaneous recovery, reinstatement, and context renewal. Overall, we observed benefits of cognitive reappraisal relative to suppression and react-as-normal in terms of extinction training and spontaneous recovery when the background context was unchanged (Study 1) and detriments from suppression relative to using cognitive reappraisal or react-as-normal in terms of context renewal (Study 2).

Cognitive reappraisal was found to strengthen reductions in US expectancy to the CS- during extinction relative to suppression, and to attenuate spontaneous recovery of SCR to the CS+ relative to react-as-normal in Study 1; these benefits from cognitive reappraisal were not found in Study 2. In Study 1, fear acquisition, extinction, and spontaneous recovery occurred on a blank white background. In Study 2, fear acquisition and spontaneous recovery occurred within the context of one visual background, while extinction occurred in another. Thus, one interpretation of the discrepant findings across the two studies is that the benefits of cognitive reappraisal upon extinction and spontaneous recovery are only apparent when context does not vary. Alternatively, a shift in contexts may outweigh the effects of instructed emotion

regulation. Although formal replication and randomization to study designs is necessary to support this interpretation, if true, this result may imply that cognitive reappraisal is of little value when administering exposure therapy in settings other than the context where fear was acquired. Unfortunately, exposure therapy is rarely, if ever, conducted in contexts where fear was acquired. Moreover, the benefits from reappraisal during extinction were only found in comparison to conditions in which participants are attempting to suppress their emotions. Whereas some individuals undergoing exposure therapy may engage in suppression efforts naturalistically, many may not and thus the benefits for cognitive reappraisal would be even more restricted.

There was no evidence to suggest that cognitive reappraisal during extinction training attenuated either reinstatement or context renewal. However, Study 1 demonstrated that cognitive reappraisal during extinction training led to a significantly faster CS+ US expectancy recovery (i.e., reduction) from reinstatement relative to suppression. Furthermore, Study 2 demonstrated deleterious effects of suppressing during extinction training upon context renewal, since suppression led to increased CS+ SCRs to context renewal relative to cognitive reappraisal and react-as-normal. Together, the findings suggest that simple and low-cost cognitive reappraisal-based interventions can have modest benefits for the recovery of fear from reinstatement (to the CS+) compared to suppression when tested in the same context as fear acquisition. We presume that this is because the appraised value of the US is changed after the reappraisal instruction whereas the value remained unchanged after the suppression instructions, and because suppression requires active regulatory effort that impedes attentional processing and implicit prediction error correction that occurs during extinction alone. Furthermore, suppression exhibited deleterious effects upon context renewal relative to cognitive reappraisal and react-as-

normal. This may be due to at least two reasons. First, the requirement of active regulatory effort without cognitive change may additively affect context renewal to CS+ SCRs and consequently prolongs the recovery of fear. Second, suppression may reallocate attention from environmental stimuli to the self (e.g., Ellis & Ashbrook, 1989; Goldin et al., 2008; Richards & Gross, 2000) potentially drawing attention away from the CS+ thereby mitigating the development of CS+ inhibitory associations during extinction training. This may then lead to an enhanced return of conditional fear responding when the CS+ and the fear-inducing context from the acquisition phase are presented together.

There are several strengths of these studies that are worth highlighting. First, we recruited sample sizes that were large relative to the extant literature on reinstatement and context renewal. Second, we used multiple modalities of measurement, namely, self-report ratings of fear, arousal, and valence, skin-conductance responding, and continuous US expectancy. Third, all units of analyses evidenced acquisition and extinction. Finally, our studies involved a direct experimental manipulation of emotion regulation that tied together two models, those being Pavlovian associative prediction error and emotion regulation.

Several limitations are worth considering for informing future work. First, the novel results obtained for these studies were not consistent across measures. Definitive conclusions about the psychological mechanisms involved in modulating fear will require evidence from multiple indexes. Second, our studies did not include a manipulation check to ascertain whether participants followed through on the administered emotion regulation instructions. Finally, the emotion regulatory effects upon the CS- present some interpretational difficulties. Our hypotheses centered on emotion regulatory effects that would reduce fear to the CS+ without

affecting fear to the CS-. Yet, we observed effects upon the CS-, which may suggest that emotion regulation changes the way the CS- is valued. From a valuation perspective (Etkin, Büchel, & Gross, 2015), the value set to the CS+ following fear acquisition is clear: it is "bad for me" because it is linked with an aversive scream sound. The CS- presumably was valued either as "irrelevant to me" following fear acquisition because it was not linked with any outcome, or it was valued as "better for me than the CS+". During extinction, cognitive reappraisal may have contributed to the CS- value being shifted more towards that of "good for me" or "better for me than the CS+" values. We also unexpectedly observed that suppression led to significantly reduced CS- SCRs relative to react-as-normal when spontaneous recovery was tested one-week later. Although reappraisal responses did not significantly differ from react-as-normal responses in the same way suppression did, simple slopes revealed that reacting-as-normal led to significant CS- SCR recovery, while recovery was absent for individuals who either reappraised and suppressed during extinction-training. It could be the case that mere regulatory effort can have transient downregulatory effects on spontaneous recovery, as it engenders some certainty as to how to respond when faced with the CS-.

Overall, we found that in the short-term, cognitive reappraisal reduced threat expectancies to the CS- during extinction relative to suppression. In the long-term, cognitive reappraisal reduced spontaneous recovery of threat-relevant (in terms of SCR) relative to reacting-as-normal. These benefits unfortunately appear to be sensitive to contexts, as they were not apparent once the contexts were changed, thus limiting the utility of cognitive reappraisal for the reduction of fear in treatment. Moreover, the evidence suggests that the use of cognitive reappraisal led to a faster recovery from reinstatement relative to suppression, but whether this effect is also sensitive to changes in visual contexts remains to be seen. Cognitive reappraisal as

designed in this study did not exhibit any substantive costs, as there was no evidence of increases in fear at tests of reinstatement or context renewal relative to control conditions. Although this may help allay fears that reappraisal disrupts prediction error and mitigates extinction learning, it must be noted that the reappraisal effect induced in these studies were designed explicitly to not interfere with extinction learning. The results herein therefore cannot be taken to broadly pertain to typical cognitive restructuring strategies used in therapy. However, the therapeutic implementation of cognitive reappraisal strategy may do well to follow the guidelines used here to complement extinction learning by minimizing cognitive cost, maintaining attention to the CSs, changing the appraisal value of the US to "good for me", while not changing the expected frequency of the US. Suppression did not evidence any substantive benefits in threat responses to the CS+ in the short or long term, and in fact engendered a prolonging of context renewal, making its use during exposure therapy a potential risk factor for the return of fear.

Figures

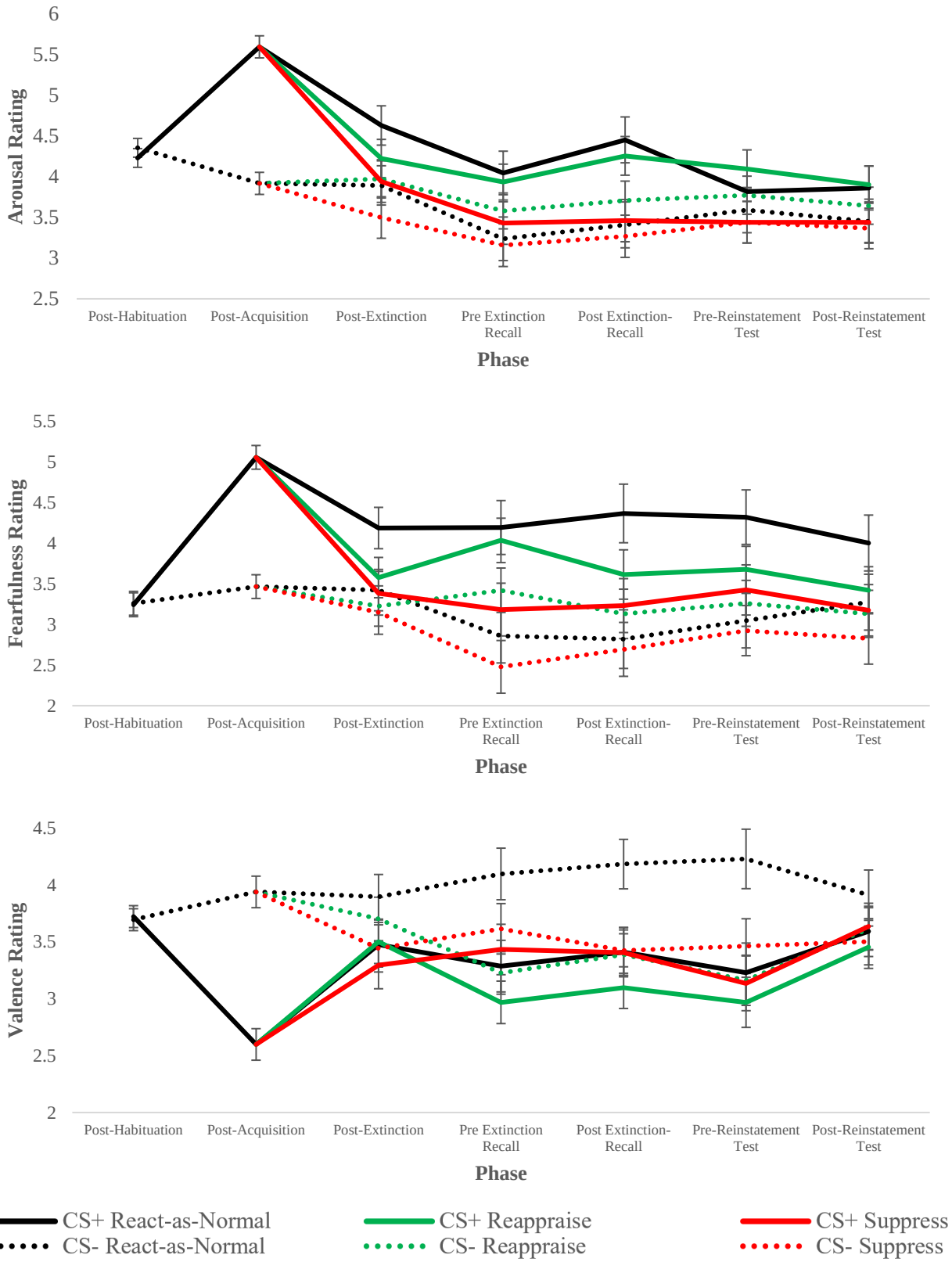


Figure 3. Study 1 self-report ratings per period.

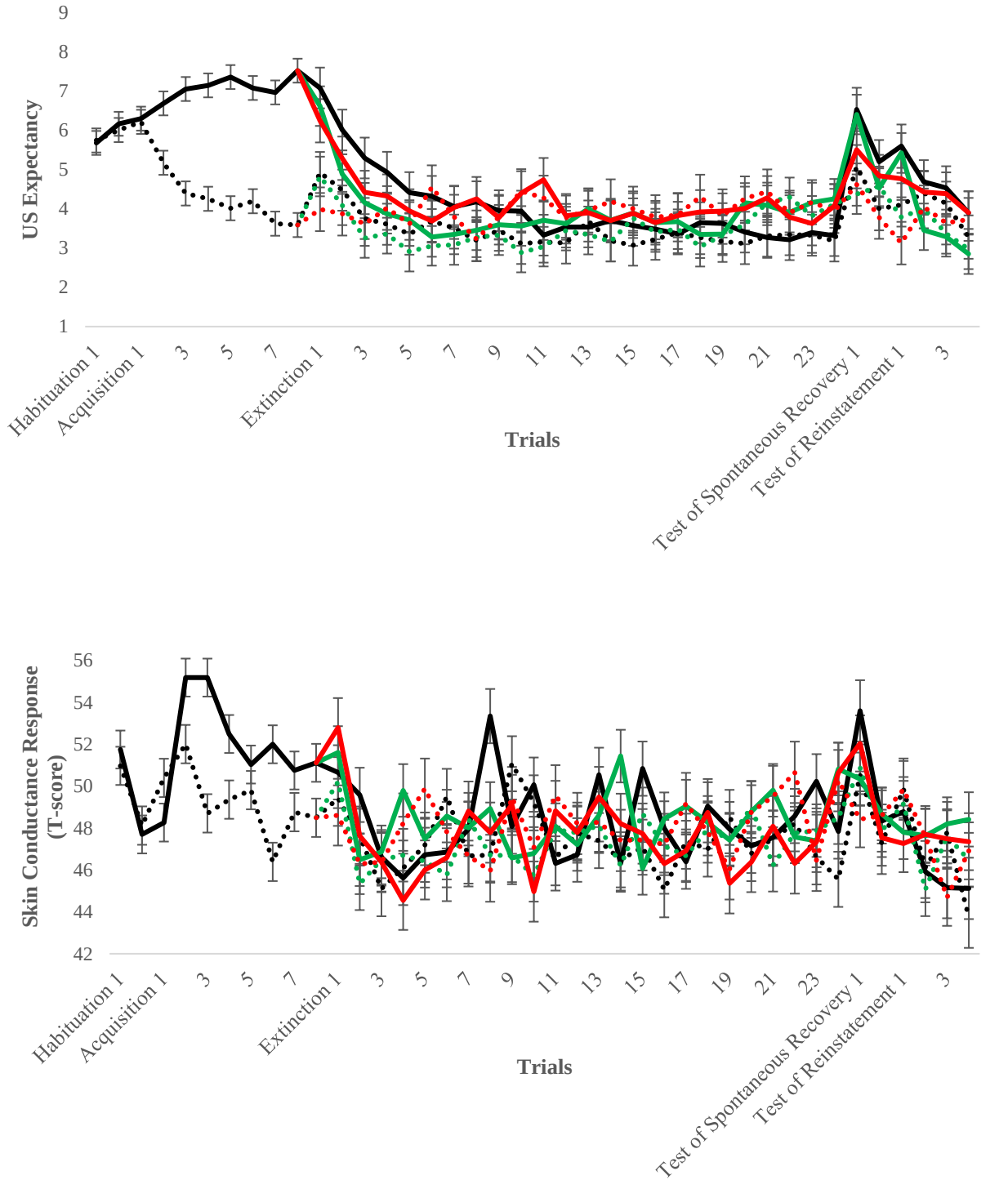


Figure 4. Study 1 continuous outcomes (US expectancy and skin conductance response) by trial.

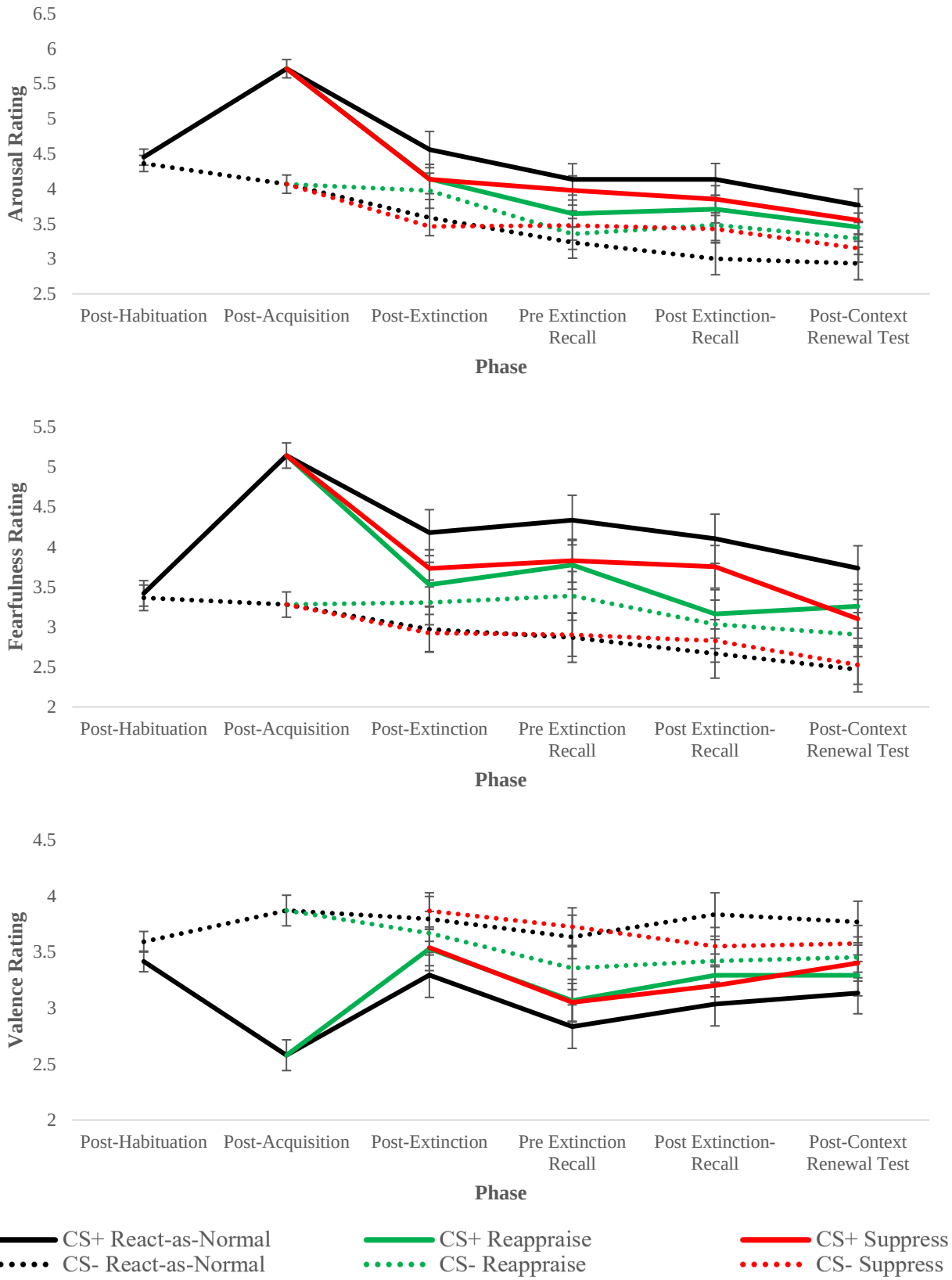


Figure 5. Study 2 self-report ratings per period.

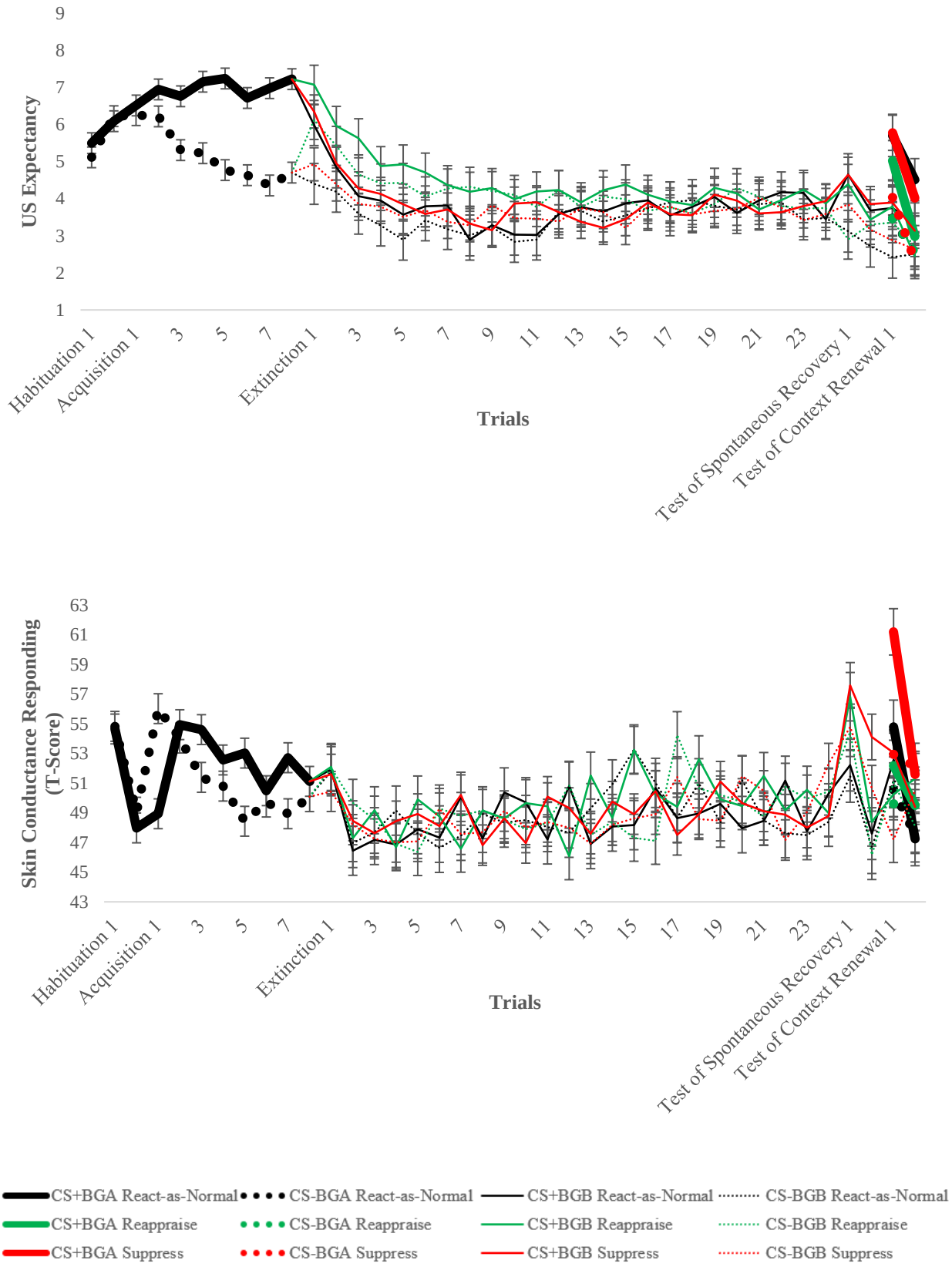


Figure 6. Study 2 continuous outcomes (US expectancy and skin conductance response) by trial.

Study 3: Neural Correlates of Fear Spontaneous Recovery and its Regulation

Abstract

The spontaneous recovery of fear is a common phenomenon that describes the return of a previously extinguished fear association sometime after fear extinction. It purportedly underlies relapse of fear-based disorders (e.g., anxiety, phobia, and post-traumatic stress) among individuals in remission. Fear spontaneous recovery may be conceptualized as a problem of fear over-generation or fear dysregulation. Fear dysregulation can be characterized as misregulation, underregulation/disconnection, or competitive co-regulation. In this fMRI study, we compared activation in neural threat circuitry between groups of individuals with high and low levels of fear spontaneous recovery (defined using skin conductance responses to the CS+ relative to the CS-) during the extinction recall phase of a two-day fear conditioning paradigm. We found sparse evidence for fear over-generation concurrently with evidence for misregulation, underregulation/disconnection, and competitive co-regulation. Differences in how threat and safe stimuli are processed and related to, and their clinical implications are discussed.

Introduction

Relative to healthy controls, individuals with fear-based disorders have been shown to exhibit hyperactivation in the neural areas associated with fear responding, and hypoactivation in areas of the ventromedial prefrontal cortex associated with fear regulation (Etkin & Wager, 2007; Lissek et al., 2014; Milad & Quirk, 2012; Milad et al., 2007; Zelikowsky et al., 2013). It is possible that the use of extinction-complementing cognitive reappraisal and the activation of CS-noUS memories through extinction training might co-activate vmPFC-sgACC to reduce fear and minimize its return.

Neurobiology supports the notion that cognitive reappraisal and extinction learning are complementary processes. CS-noUS memories are believed to be retained in the infralimbic cortex of the vmPFC-sgACC (Milad et al., 2007). This area has glutamatergic projections to the strip of intercalated cells between basolateral and central amygdala (Hurley, Herbert, Moga, & Saper, 1991; Likhtik, Popa, Apergis-Schoute, Fidacaro, & Paré, 2008; Vertes, 2004). When these cells are activated, GABA-ergic projections inhibit the central amygdala responsible for enacting fear CRs (Duvarci & Pare, 2014; Milad & Quirk, 2012; Milad et al., 2007; Paré et al., 2004). Reappraisal has been shown repeatedly to be associated with dorsolateral and ventrolateral prefrontal cortices (dlPFC and vlPFC; (Beauregard, Lévesque, & Bourgouin, 2001; R. Kalisch et al., 2005; Raffael Kalisch, Wiech, Herrmann, & Dolan, 2006) believed to reflect the function of executive, organizational control of complex information (Dalley et al., 2004). These areas have few projections to amygdala (e.g., (Buhle et al., 2014; Kohn et al., 2014; Vertes, 2006) but do project to medial areas including vmPFC-sgACC (Vertes, 2006). Cognitive reappraisal may be associated with increased vmPFC-sgACC activation, the same region associated with fear extinction to reduce central amygdala activation.

Several studies provide support for the above assertions. Neurobiologically, vmPFC thickness has been correlated with both cognitive reappraisal (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007) and extinction recall (Milad et al., 2005). Furthermore, these cross-sectional studies suggest that reappraisal was correlated with vmPFC activation (Johnstone et al., 2007) while suppression was not (Welborn et al., 2009). One study found that affect labelling, a similar linguistic emotion-regulatory strategy, increased activity in vlPFC and dlPFC while it decreased activity in amygdala. This process was found to be mediated by vmPFC activation (Lieberman et al., 2007).

This study aimed to improve the understanding of the network underlying fear and fear regulation during fear recall using the perspective of the extended process model of emotion regulation (Gross, 2015b). This model, which delineates the differences between emotion generation and regulation, along with its implications for dysregulation (Sun, Vinograd, Miller, & Craske, 2017), can help to better understand the profile of activation relationships during the regulation of spontaneous recovery. Maladaptive expression of fear may result from a confluence of fear overgeneration, fear misregulation (regulatory effort that counterproductively increases fear), fear underregulation or disconnection (inefficacious or insufficient fear regulatory effort), or competitive co-regulation of fear (where one regulatory area is increasing fear while, another area is simultaneously decreasing it).

Regions Implementing Fear Spontaneous Recovery

Although the central amygdala is thought to enact fear CRs, other networked regions are also active to implement various properties of fear and fear salience. Along with bilateral amygdala, we focused on bilateral anterior insula, an area that is frequently implicated in the reportable experience of negative emotion, and the bilateral bed nuclei of the stria terminalis

(BNST), part of the "extended amygdala" with strong neural interconnections between amygdala and insula.

The anterior insula is densely packed with spindle neurons, which allow for the rapid communication across the brain, implicating interoceptive awareness of negative emotional feelings like fear (Phan, Wager, Taylor, & Liberzon, 2002; Vilares, Howard, Fernandes, Gottfried, & Kording, 2012). This interoception is thought to feed back to give rise to fear salience the body, and this area appears to play a role in anxiety disorders (Paulus & Stein, 2006), and general emotion dysregulation (Thayer & Lane, 2000). Right anterior insula in particular may regulate the interaction between the salience of attended goal-stimuli and the salience of fearful arousal created to maintain focus upon the relevant part of the environment. This salience regulation in turn regulates the vigilance required to complete challenging and fatiguing tasks and over-regulation may induce the hypervigilance components of anxiety (Eckert et al., 2009).

The BNST consists of a band of fibers running along the surface of the thalamus. It serves as a major output pathway of the amygdala. In general, where the amygdala responds to immediate, predictable, and proximal fearful stimuli, the BNST responds to sustained, distant, and unpredictable anxiety responses. The unique roles that BNST appears to play is that it mediates sustained responses to contextual, diffuse, and unpredictable threats (Sullivan et al., 2004; Waddell, Morris, & Bouton, 2006). It also appears to mediate hypervigilance and arousal (Davis, Walker, Miles, & Grillon, 2010), increased sensitization to the environment (Davis & Walker, 2014), stress-enhanced learning (Bangasser & Shors, 2008), and a myriad of stress and anxiety-related behaviors (Kim et al., 2013). It may act as a relay site within the hypo-pituitary-adrenal axis to regulate the response to acute stress (Somerville, Whalen, & Kelley, 2010). As

such, the BNST is an area of burgeoning focus in the anxiety literature (Avery, Clauss, & Blackford, 2015). It not only connects to amygdala and anterior insula, it connects to frontal regions as well in dorsomedial prefrontal cortex, and vmPFC via the sgACC.

Top-Down Emotion Regulatory Regions

Dorsolateral prefrontal cortex (dlPFC), largely composed of spatially selective neurons, is one of the most recently derived parts of the human brain (Olson & Luciana, 2008). It is heavily implicated in working memory, with controversy over lateralized specialization for verbal versus visuospatial working memory (Barbey, Koenigs, & Grafman, 2013; Smith, Jonides, & Koeppe, 1996). It is also involved in many executive functions such as motor planning, organization, and regulation, as well as executive subfunctions of sensory input, retention in short-term memory, and motor signaling. Given that dlPFC is required for the comparing two items in memory (Goldman-Rakic, 1994), it may be an essential structure for implementing cognitive reappraisal (Gross, 2015a). DlPFC may be involved in threat-induced anxiety, as those who rated themselves as behaviorally inhibited show greater tonic (resting) activity in right-posterior dlPFC and dlPFC activity correlated with individual experiences of vigilance and uncertainty (Shackman, McMenamin, Maxwell, Greischar, & Davidson, 2009).

Ventrolateral prefrontal cortex (vlPFC) has distinct and well known functional lateralizations. Right vlPFC is a critical substrate of control (Levy & Wagner, 2011), engaged to stop or override motor responses (Aron, Robbins, & Poldrack, 2004). It is also thought to govern reflexive reorienting (Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002). VlPFC is the end point of the ventral pathway that brings in information an executively processes stimulus characteristics (Lee, Blumenfeld, & D'Esposito, 2013). Left vlPFC, which holds Broca's area, is linked with language production, and more recently with the cognitive control of memory. It has

been implicated in cognitive control processes that guide access to relevant information from semantic memory (Croxson, 2005; Petrides & Pandya, 2002b, 2002a). It is responsible for controlled retrieval, activating goal-relevant knowledge in a top-down manner, and post-retrieval selection, resolving competition between simultaneously active representations. Stimulus representations can be retrieved in multiple forms, and left vlPFC is a selector (Fletcher, Shallice, & Dolan, 2000; Moss et al., 2005).

Bottom-Up Emotion Regulatory Regions

The vmPFC, while understood to store CS-noUS memories to regulate fear CRs, is more broadly implicated in the processing of risk and fear and plays a role in the inhibition of emotional responses and in decision making and self-control. It does so by integrating environmental information and the goal prioritization of the frontal regions. Importantly, vmPFC demarcation is not universally agreed upon, and within the fear conditioning and emotion regulation literature, this area has been described as including some or all of Brodmann areas 10-14, 25, and 32. These areas are connected to and receive input from ventral tegmentum, amygdala, temporal lobe, olfactory system, and the dorsomedial thalamus, and sends signals to temporal lobe, amygdala, lateral hypothalamus, hippocampus, cingulate cortex, and other regions of the prefrontal cortex. It appears necessary in extinction training (Madsen, Guerin, & Kim, 2017), and plays a role in general memory consolidation (Nieuwenhuis & Takashima, 2011) including the type needed for extinction learning (Gregory J. Quirk, Russo, Barron, & Lebron, 2000). Clinically, patients with larger vmPFCs tend to have lower CRs to extinguished CS+s, suggesting stronger and more enacted extinction memories (Milad et al., 2005). The confusion over what vmPFC does may be obfuscated by the fact that within the large general area lies subareas with different molecular and cellular make-ups. Areas of the vmPFC near the frontal

surface may not have direct connections with limbic areas, but may communicate more directly with top-down prefrontal areas. The cortical overlap between vmPFC and subgenual anterior cingulate cortex (sgACC) is rich in serotonin transporters, and influences amygdala and insula with direct afferents and efferents. Given the expanse of brain matter that is ostensibly vmPFC, I attempted to attain more specificity on the nature of limbic regulation by examining smaller vmPFC areas, a frontal vmPFC area and a vmPFC-sgACC area.

Methods

Participants

275 college-aged participants from the greater Los Angeles and Evanston communities were recruited through fliers and online advertisements of an ongoing longitudinal trial examining neurobiological changes in emotional development. They averaged 19.56 years of age (SD = 3.43), 66.06% of whom were female, 0.37% was transgender (n = 1), and the racial/ethnic breakdown consisted of 33.58% Non-Hispanic White, 28.41% Asian, 19.93% Hispanic/Latino, and 18.98% Other. We excluded individuals based on the presence of metal in the body, a history of diagnosed severe psychiatric or neurological disorders, and/or current use of psychoactive medications.

Procedures

While in the fMRI scanner, participants first underwent a differential Pavlovian Fear Learning Task programmed in E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA USA) and presented to participants using a mirror and projector system. This slow event-related fMRI paradigm has been widely used in prior studies of healthy and anxious participants (Milad et al., 2009; Milad et al., 2007). It consisted of four phases: habituation, fear acquisition, fear extinction (all conducted on day 1) and extinction recall (conducted on day 2: 24-48hrs later). During

habituation participants viewed each of three CS images to reduce the novelty of these stimuli for the subsequent phases. During acquisition, two of the images served as CS+ stimuli, which was followed by a US 62.5% of the time, and one served as a CS- stimulus. CS images were office or conference rooms (context) with different colored lights (red/yellow/blue), and color order and context images were counterbalanced across participants. During each trial, participants first viewed the context image (3 seconds), followed by the context and CS (6 seconds). There were 8 trials of each CS+ (16 trials total) and 16 trials of the CS-. Five (out of eight) of the CS+ trials of each type were followed immediately by a 500-millisecond shock (consisting of 10 2-millisecond shocks at 20Hz) applied to the left bicep serving as the US. Inter-trial intervals varied from 12-18sec (mean 15sec) and included a jitter of 125 milliseconds per trial to reduce slice timing bias. US shocks were delivered using a DS7a constant current high voltage stimulator (Digitimer Ltd, England) at Los Angeles and using a STMISOC constant voltage stimulator (Biopac Systems Inc., Goleta, CA, USA) at Evanston. Current levels were determined for each participant during a 'work-up' procedure conducted on the Day 1 scanning session. In this procedure, participants were presented with shocks of increasing intensity and were asked to rate each on a pain scale of 1-10 (1= 'not at all painful', 10 = 'most pain imaginable'). Participants were informed we aimed to reach a level of shock that was 'uncomfortable but not painful' and 'took some effort to tolerate' (a rating of 5-6 that they were willing to tolerate for the experiment).

During extinction, participants viewed 16 trials of one of the CS+'s with no shocks (the extinguished CS, CS+E) and 16 trials of the CS- image. During extinction recall (Day 2), participants viewed 8 trials of the extinguished CS+ (CS+E), the unextinguished CS+ (CS+U) and 16 trials of the CS-.

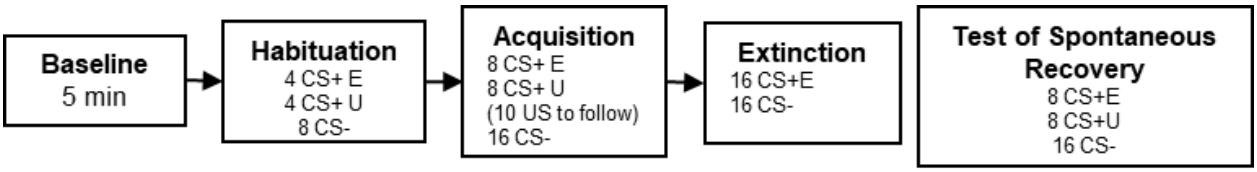


Figure 1: Study design.

Materials/Apparatus

FMRI. Data were acquired on Prisma 3.0 Tesla whole-body scanners using 64-channel head coils (Siemens Medical Systems, Iselin, New Jersey) at the UCLA Ahmanson-Lovelace Brain Mapping Center and the Northwestern University Center for Translational Imaging. High resolution structural images (T1-weighted) were acquired using a magnetized prepared rapid acquisition gradient echo (MPRAGE) sequence containing 0.8mm isotropic voxels, TR/TE/flip angle=2300ms/2.99ms/7°, FOV= 256mm^2, 208 slices. Blood oxygenation level-dependent (BOLD, T2*-weighted) functional images were acquired parallel to the AC-PC line using Siemens AutoAlign function, containing 2mm isotropic voxels, TR/TE/flip angle=2000ms/25ms/80°, FOV = 208mm^2, 64 slices, 380 volumes (per task phase).

Skin Conductance Responses. A BioPac MP150 (Biopac System Inc., Goleta, CA, USA) with an EDA100C amplifier was used to record the skin conductance response (SCR) during all phases of the task. Data was collected using AcqKnowledge 4.2 software (Biopac System Inc., Goleta, CA, USA) and then cleaned, inspected, and analyzed in ANSLab software (ANSLab v2.5, Wilhelm & Peyck, 2005).

Data Analysis

FMRI Analysis. Raw dicom files taken by the fMRI scanner were converted to NIFTI format using dcm2nii (MRIcroN, http://www.cabiatl.com/mricro/mricron/dcm2nii.html). Data was processed and analyzed using FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl).

Structural data was corrected for spatial intensity variations (bias field correction) using FAST (FMRIB's Automated Segmentation Tool) (Zhang, Brady, & Smith, 2001) and brain extraction was performed using optiBET (optimized brain extraction) (Lutkenhoff et al., 2014). Functional data was first assessed for outlier volumes (75^{th} percentile +1.5 times interquartile range) based on framewise displacement (average of rotation and translation parameter differences, using weighted scaling (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) as implemented in the fslmotionoutliers function (FSL). Runs with >10% outliers were not included in group analyses. Outlier volumes were subsequently censored in first level analyses by including a regressor with a single time point corresponding to each outlying volume. Functional data was brain extracted using BET (Brain Extraction Tool, FSL) (Smith, 2002) and bias field corrected using N4BiasFieldCorrection, run twice (ANTS registration suite Tustison et al., 2010).

FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00. Registration to high resolution structural space images was carried out using FLIRT (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson, Jenkinson, & Smith, 2007a, 2007b). The following pre-statistics processing was applied: motion correction using MCFLIRT (Jenkinson et al., 2002), slice-timing correction using Fourier-space time-series phase-shifting, spatial smoothing using a Gaussian kernel of FWHM 4.0mm, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and high-pass temporal filtering (0.01Hz) to remove low frequency artifacts.

First-level analyses included regressors of interest (detailed below), temporal derivatives, and six motion regressors. Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). Regressors consisted of the first four images of the CS+U/CS+E/CS-, the last four images of the CS+U/CS+E/CS-, and all other CS+U/CS+E/CS- images on the second day. Contrasts for spontaneous recovery was defined as the first four CS+E versus the first four CS- images. For the purposes of the present analyses, activity to the CS+U was ignored, and the CS+E will be referred to hereinafter as simply the CS+.

Region of interest (ROI) parameter estimates from bilateral amygdala, bilateral anterior insula, bilateral BNST, frontal vmPFC, vmPFC-sgACC, bilateral dlPFC, and bilateral vlPFC were extracted from the aversive conditioning fMRI task (see Figures 11-17). The bilateral amygdala ROI, bilateral anterior insula, and bilateral vmPFC-sgACC were defined through masks from the Harvard-Oxford MNI probabilistic atlas. ROIs for frontal vmPFC (-2, 56, -14), left dlPFC (-36, 44, 22), right dlPFC (34, 44, 32), left vlPFC (-48, 16, 6), and right vlPFC (34, 44, 32) were defined as 5mm spheres centered on coordinates from the most recent meta-analysis on fear conditioning by Fullana and colleagues (2016). The bilateral BNST mask was created from a 7T gradient spin echo (GRASE) MRI image (Avery et al., 2015).

SCR Analysis. SCRs greater than zero were square root transformed to normalize the data (Levey, 1980). SCRs less than or equal to zero were coded as zero. To eliminate individual variability in SCR range, SCRs were *T*-score standardized using the following formula:

$$SCR_T = \frac{SCR_{Observed} - SCR_{mean}}{SCR_{SD}} \times 10 + 50$$

Spontaneous recovery was operationalized through the SCR of the first four CS+ trials averaged minus the first four CS- trials averaged.

Preliminary Analyses

193 participants exhibited interpretable SCRs, and 3 of these participants were removed due to problems with fMRI data, leaving a final *n* of 190. SCRs to the CS+ and CS- were estimated with best-fit curvilinear models to check for fear acquisition, fear extinction, and the return of fear. As expected (see Figure 10), fear conditioning differentiated the CS+ from the CS- in SCRs during the Acquisition phase (Stimulus-type x Trial: B = -6.40, SE = 1.04, $p < .001$) and SCRs to the CS+ were extinguished by the end of the Extinction phase (95% confidence intervals between CS+ and CS- begin overlapping by trial 27). SCR spontaneous recovery during the Test of Spontaneous Recovery was not significantly induced on average (M = -0.40, SD = 7.12, $t(189) = -0.67$, $p = .749$). Seventy-three (38.42%) participants did exhibit spontaneous recovery and 52 (27.37%) exhibited higher-than-average spontaneous recovery (M = 8.73, SD = 5.04, $t(51) = 12.50$, $p < .001$). Descriptive statistics are provided on Table 1, comparing the group of individuals exhibiting higher-than-average SCR spontaneous recovery (hereinafter the high SCR spontaneous recovery group) from others (hereinafter the low SCR spontaneous recovery group) using t-tests and chi-square tests of equivalence.

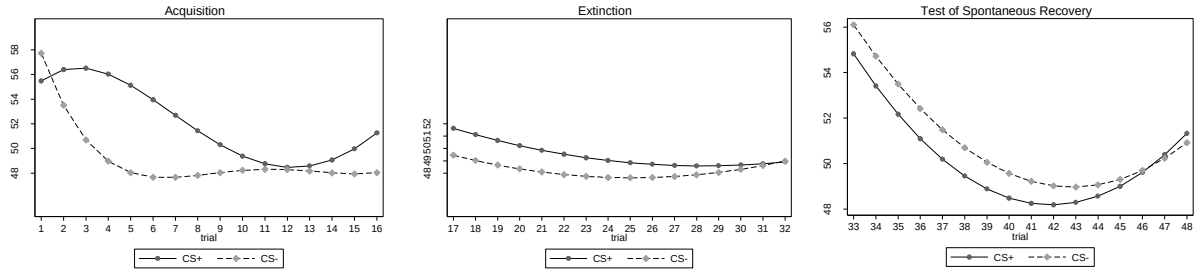


Figure 2: Predicted skin conductance responding during Acquisition, Extinction, and Test of Spontaneous Recovery phases.

Table 1
Individuals Exhibiting Low and High SCR Spontaneous Recovery

Demographic Characteristic	Low SCR Spontaneous Recovery (n = 138)	High SCR Spontaneous Recovery (n = 52)	t / χ^2	p
Gender			$\chi^2(1) = 5.31$	.021*
Male	37 (26.81%)	23 (44.23%)		
Female	101 (73.19%)	29 (55.77%)		
Mean Age (SD)	19.20 (0.50)	19.13 (0.61)	$t = 0.82$	.208
Race/ethnicity			$\chi^2(6) = 9.40$	.153
White, non-Hispanic	46 (33.33%)	15 (28.85%)		
Hispanic White	32 (23.19%)	10 (19.23%)		
Asian	33 (23.91%)	23 (44.23%)		
Black	14 (10.14%)	2 (3.85%)		
Native	2 (1.45%)	0 (0.00%)		
Multiracial	10 (7.25%)	2 (3.85%)		
None endorsed by choice	1 (0.72%)	0 (0.00%)		

Path Modelling Specification Procedures

Path modelling was conducted with the STATA 14.1 *sem* command to address whether top-down or bottom-up control regions predicted activity in regions that implement fear spontaneous recovery. This provided estimates of direct and indirect regulatory paths, as well as total effects from the activity of control ROIs on limbic spontaneous recovery. The totality of the regulatory activations may lead to complementary mediation, competitive mediation, indirect-only mediation, direct-only mediation, or non-mediation (Zhao, Lynch Jr, & Chen, 2010) between top-down and bottom-up regulatory mechanisms. Additionally, path modelling provides estimates of covariances between region categories (i.e., spontaneous recovery regions, top-down regulatory regions, bottom-up regulatory regions). SEM disattenuates parameter estimates from measurement error (Ullman & Bentler, 2003), allowing a more accurate assessment of the relationship between the fear response and its putative neural mechanisms. Estimates were computed using maximum-likelihood.

To understand the emotion regulation-spontaneous recovery activation network, a saturated configural path model was first fit for the sample, and then re-fit with a grouping variable, clustering individuals by low and high SCR spontaneous recovery (0 = average to lower than average SCR spontaneous recovery; 1 = higher than average SCR spontaneous recovery). This path model included top-down regulatory ROI activity during exposure to CS+/CS- as exogenous variables (8 variables), bottom-up regulatory ROI activity during exposure to CS+/CS- as intermediary variables (4 variables), and spontaneous recovery activity in each limbic ROI as endogenous variables (6 variables). The intermediary variables are predicted by the exogenous variables and the endogenous variables are predicted by both exogenous and intermediary variables. This model had 189 parameters estimated per group (104 directed paths, 28 exogenous covariances from 8 exogenous means and their variances, and 21 error covariances from 10 intercepts and their error variances).

Fit indices are not available for just-identified configural models, yet such models may not fit well, so we explored the effects of testing group invariance by constraining classes of paths and examining the fit after imposing constraints relative to the configural model with all paths free-to-vary. This involved testing models after constraining (a.) the directed paths from top-down to bottom-up regulatory activity, (b.) the directed paths from top-down regulatory activity to spontaneous recovery, (c.) the directed paths from bottom-up regulatory activity to spontaneous recovery, (d.) the bottom-up regulatory covariances, (e.) the spontaneous recovery covariances, (f.) the top-down regulatory covariances, (g.) the top-down regulatory activity means, (h.) the bottom-up regulatory activity intercepts, and (i.) the spontaneous recovery intercepts. Goodness of fit was determined using the chi-square test of the difference between the constrained and the configural model, the Tucker-Lewis Index (TLI), and Root Mean Square

Error of Approximation (RMSEA). Population level goodness-of-fit was estimated with the 90% Confidence Interval of the RMSEA. Adequate fit was defined as a non-significant chi-square difference, a TLI above .90 and a RMSEA under .08. Path classes were freed if constraining them led to poor model fit. Variance explained in each model was quantified with the Coefficient of Determination (CD). A final model was chosen by constraining all classes of paths that maintained good fit after being constrained.

Paths that remained unconstrained across groups were tested for group differences with a univariate Wald test. Direct paths for the high SCR spontaneous recovery group estimated from regulatory areas to spontaneous recovery were interpreted from an extended process model viewpoint of emotion dysregulation (Gross, 2015b). These direct paths, when significantly different from the respective paths of the comparison group, were interpreted as emotion misregulation for significantly positive paths that were greater than comparison paths, emotion underregulation for significantly positive paths of that were also significantly smaller in magnitude, and emotion non-regulation/regulatory-disconnection for paths that were no longer significant when compared with comparison paths. 95% confidence intervals were inspected within path classes to explore how paths compared to one another.

Results

Tests of Invariance

The configural model, after including the grouping variable, increased the CD from .706 to .805. Direct paths from top-down to bottom-up regulatory activity ($\chi^2(32) = 39.32$, $p = .175$, TLI = .922, RMSEA = .049, 90% CI = (.00, .10)), mean ($\chi^2(8) = 5.22$, $p = .734$, TLI = 1.118, RMSEA = .00, 90% CI = (.00, .09)) and variance ($\chi^2(8) = 7.32$, $p = .502$, TLI = 1.029, RMSEA = .00, 90% CI = (.00, .11)) estimates of top-down regulatory activity, intercept estimates of

bottom-up regulatory activity ($\chi^2(4) = 4.82$, $p = .306$, TLI = .931, RMSEA = .046, 90% CI = (.00, .168)), and variance estimates of spontaneous recovery ($\chi^2(6) = 4.75$, $p = .576$, TLI = 1.070, RMSEA = .00, 90% CI = (.00, .12)) could be constrained without sacrificing model fit. When concurrently constrained, the final model fit well ($\chi^2(58) = 59.01$, $p = .438$, TLI = .994, RMSEA = .014, 90% CI = (.00, .07)) and increased the CD to .814. All tests of invariance and their model fit statistics are shown on Table 2.

Table 2

Tests of invariance after constraining categories of paths to be equal

Model	χ^2	df	*p* for χ^2	TLI	RMSEA	90% CI	CD
Saturated model without grouping variable	0	0	.	1.00	.00	(.00, .00)	.706
Configural model with grouping variable	0	0	.	1.00	.00	(.00, .00)	.805
Equal Paths (top-down → bottom-up)	**39.322**	**32**	**.175**	**.922**	**.049**	**(.00, .095)**	**.804**
Equal Paths (top-down → spontaneous recovery)	72.700	48	.012	.826	.074	(.035, .107)	.774
Equal Paths (bottom-up → spontaneous recovery)	34.499	24	.076	.852	.068	(.00,.115)	.797
Equal Disturbance Variances (bottom-up)	14.329	4	.006	.125	.165	(.078, .261)	.826
Equal Disturbance Variances/Covariances (bottom-up)	19.519	10	.034	.678	.100	(.027, .166)	.839
Equal Disturbance Variances (spontaneous recovery)	**4.752**	**6**	**.576**	**1.070**	**.00**	**(.00, .117)**	**.813**
Equal Disturbance Variances/Covariances (spontaneous recovery)	29.425	21	.104	.864	.065	(.00,.116)	.821
Equal Exogenous Variances (top-down)	**7.321**	**8**	**.502**	**1.029**	**.00**	**(.00, .114)**	**.810**
Equal Exogenous Variances/Covariances (top-down)	52.332	36	.038	.846	.069	(.017, .108)	.826
Equal Means (top-down)	**5.220**	**8**	**.734**	**1.118**	**.00**	**(.00, .088)**	**.806**
Equal Intercepts (bottom-up)	**4.819**	**4**	**.306**	**.931**	**.046**	**(.00, .168)**	**.805**
Equal Intercepts (spontaneous recovery)	10.271	6	.114	.759	.087	(.00, .174)	.804
Final Model[1]	59.010	58	.438	.994	.014	(.00, .065)	.814

Note. Rows in bold signify path categories that may be constrained without significantly sacrificing model fit.

[1]The final model consists of the configural model while constraining all paths in bold.

Means and Intercepts

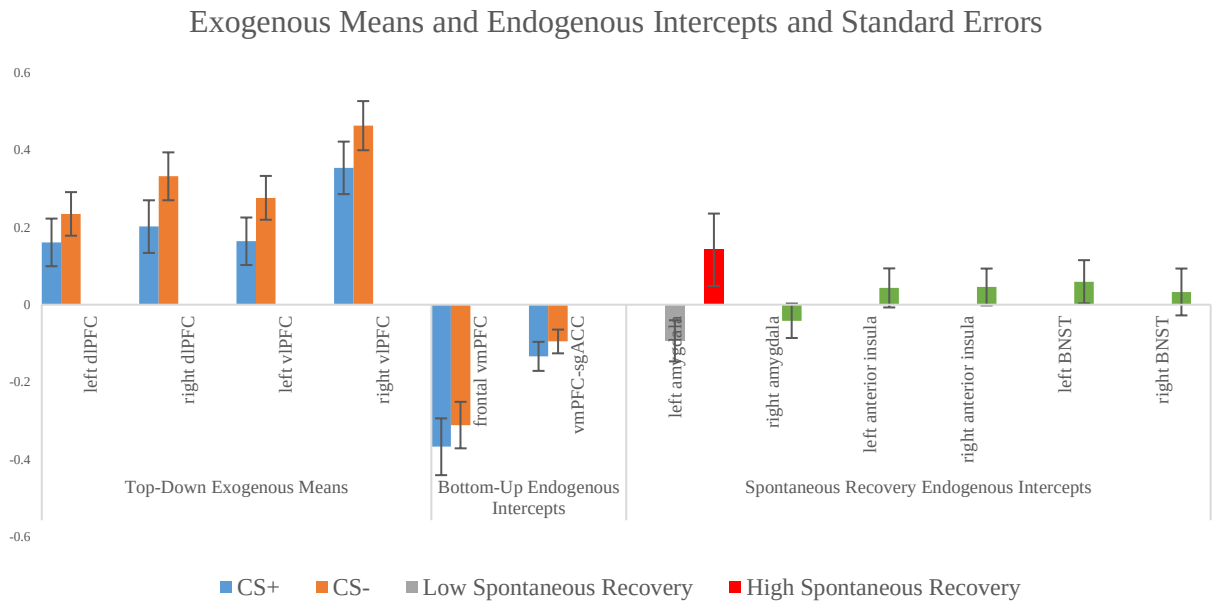


Figure 3: Region-of-interest activity at the test of spontaneous recovery
Note. dlPFC: dorsolateral prefrontal cortex, vlPFC: ventrolateral prefrontal cortex, vmPFC: ventromedial prefrontal cortex, sgACC: subgenual anterior cingulate, BNST: bed nucleus of the stria terminalis. Bars are differentiated between the Low and the High Spontaneous Recovery groups if significantly different.

Across the entire sample, mean activity of top-down regulatory areas were all significantly active to the CS+ and CS- (*p*s > .003). Moreover, inspecting the 95% confidence intervals revealed that mean right vlPFC activity to the CS- (95% CI = [0.16, 0.24]) was significantly greater than mean left vlPFC (95% CI = [0.04, 0.28]) as well as bilateral dlPFC (right: 95% CI = [0.07, 0.34]); left: 95% CI = [0.04, 0.28]) activity to the CS+. The variability of top-down regulatory activity were not observed to be significantly different across activations.

Intercepts of bottom-up regulatory activity evidenced significant frontal vmPFC and vmPFC-sgACC deactivations to the CS+ and CS- (*p*s < .002). Confidence intervals revealing that frontal vmPFC evidenced significantly greater deactivations to the CS+ (95% CI = [-0.21, -

0.06]) and CS- (95% CI = [-0.16, -0.03]) relative to vmPFC-sgACC deactivations to CS+ (95% CI = [-0.51, -0.22]) and CS- (95% CI = [-0.43, -0.19]). Confidence intervals of bottom-up regulatory activity revealed that frontal vmPFC activity to the CS+ (95% CI = [0.61, 0.92]) and the CS- (95% CI = [0.41, 0.61]) was significantly more variable than vmPFC-sgACC activity to the CS+ (95% CI = [0.16, 0.24]) and CS- (95% CI = [0.11, 0.16]).

Neural spontaneous recovery intercepts were not significant for any region (*p*s > .232), which is consistent with the sample's absence of significant expressed spontaneous recovery in SCRs. Individuals who did exhibit higher levels of SCR spontaneous recovery exhibited relatively higher left amygdalar spontaneous recovery intercept than the rest of the sample (Wald $\chi^2 = 4.59$, $p = .032$), although the estimated activity was not significant from zero ($B = 0.14$, SE = 0.09, $z = 1.51$, $p = 0.13$, 95% CI = [-0.04, 0.33]). Confidence intervals of the variability of limbic ROI spontaneous recovery intercepts revealed that the left BNST spontaneous recovery intercept (95% CI = [0.41, 0.61]) was significantly more variable than the right amygdalar spontaneous recovery intercept (95% CI = [0.19, 0.28]), and that variability of the right BNST spontaneous recovery intercept (95% CI = [0.34, 0.52]) was significantly greater than that of bilateral amygdalar spontaneous recovery (left amygdala: 95% CI = [0.21, 0.32]).

Structural Covariances

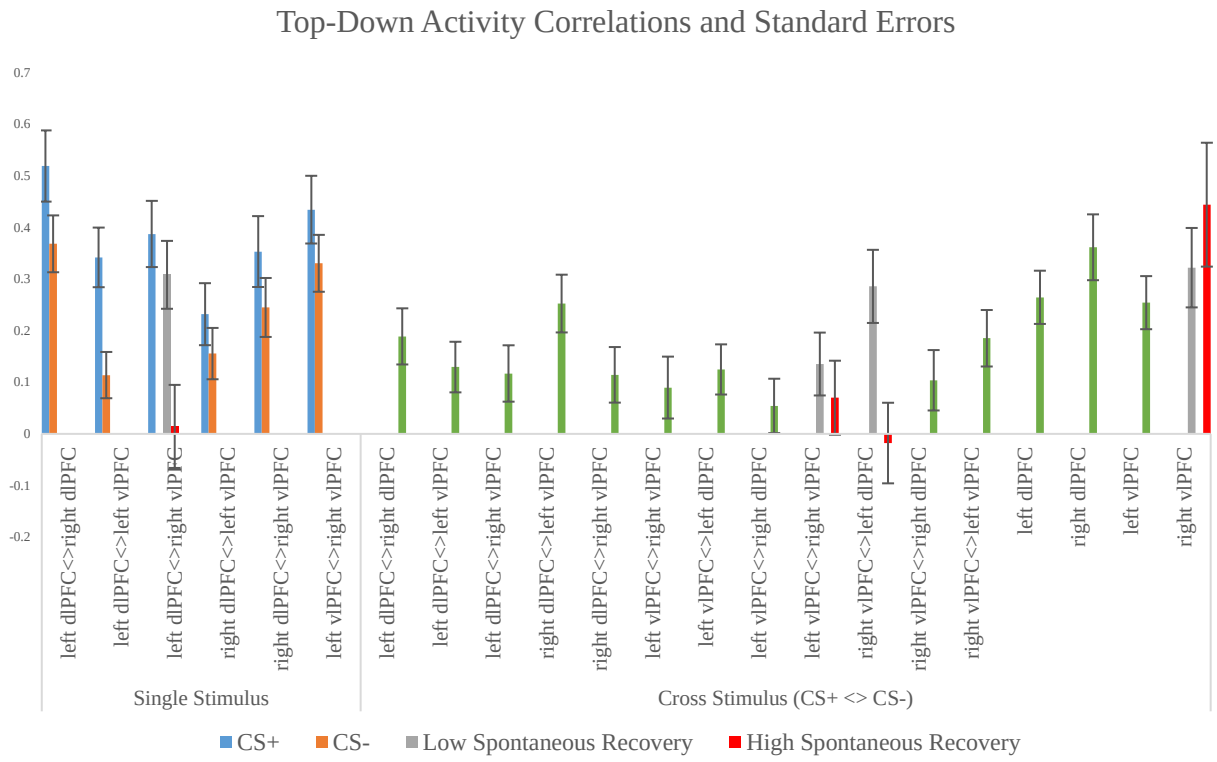


Figure 4: Top-down region-of-interest activity correlations at the test of spontaneous recovery
Note. dlPFC: dorsolateral prefrontal cortex, vlPFC: ventrolateral prefrontal cortex. Bars are differentiated between the Low and the High Spontaneous Recovery groups if significantly different. Low vs. High Spontaneous Recovery bars for left dlPFC<>right vlPFC activity are for the CS-.

Between Top-Down Regulatory Regions. For the entire sample, all top-down regulatory activity to the CS+ ($ps < .001$) and CS- ($ps < .011$) were significantly positively correlated. Cross-region correlations of activity between CS+ and CS- were also positive and significant ($ps < .034$) with exceptions of correlations between right dlPFC to the CS+ and right vlPFC to the CS- ($p = .135$), between right dlPFC to the CS- right vlPFC to the CS+ ($p = .076$), between right dlPFC activity to the CS- and left vlPFC to the CS+ ($p = .303$), and between right vlPFC to the CS- and left vlPFC to the CS+ ($p = .063$).

Individuals with higher than average SCR spontaneous recovery exhibited correlations between left dlPFC to the CS- and right vlPFC to the CS+ (Wald $\chi^2 = 4.19$, $p = .041$), between left dlPFC to the CS- and right vlPFC to the CS- (Wald $\chi^2 = 4.86$, $p = .028$), and between left vlPFC to the CSE and right vlpfc to the CS- (Wald $\chi^2 = 6.06$, $p = .014$), that significantly differed from those who did not exhibit higher than average SCR spontaneous recovery such that these correlations were no longer significant (*p*s > .328). At the same time, right vlPFC activity between the CS+ and CS- was significantly increased (low group: $B = 0.32$, SE = 0.08, $z = 4.18$, $p < 0.001$, 95% CI = [0.17, 0.47]; high group: $B = 0.44$, SE = 0.12, $z = 3.70$, $p < 0.001$, 95% CI = [0.21, 0.68]) for individuals with higher than average SCR spontaneous recovery (Wald $\chi^2 = 6.56$, $p = .010$).

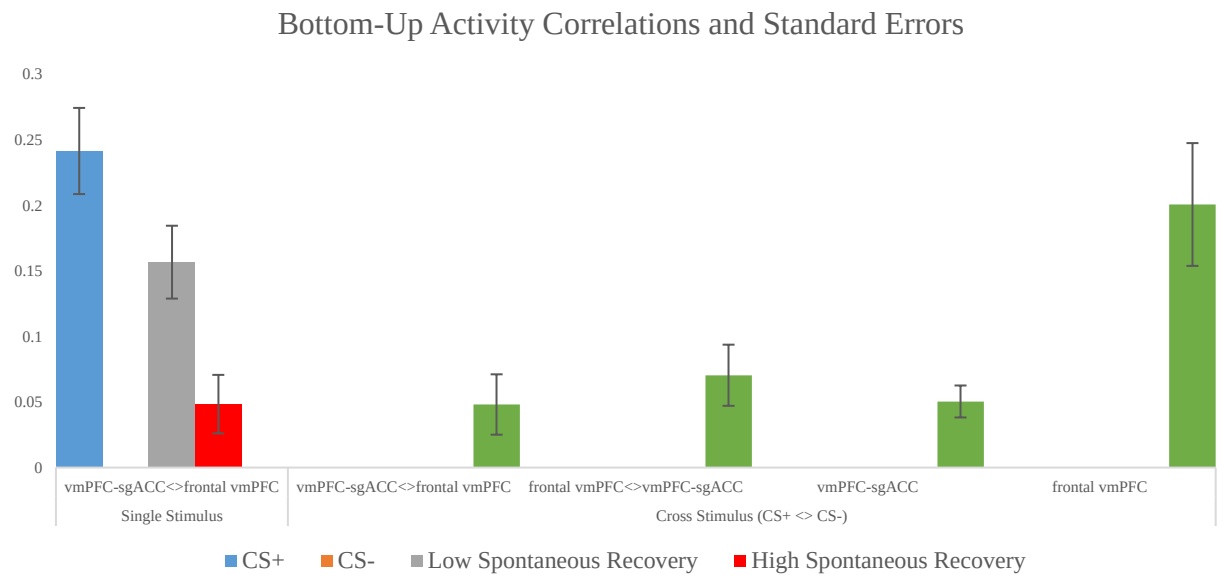


Figure 5: Bottom-up region-of-interest activity correlations at the test of spontaneous recovery
Note. vmPFC: ventromedial prefrontal cortex, sgACC: subgenual anterior cingulate. Low vs. High Spontaneous Recovery bars for vmPFC-sgACC<>frontal vmPFC activity are for the CS-.

Between Bottom-Up Regulatory Regions. All bottom-up regulatory activity to the CS+ (*p* < .001), CS- (*p* < .001), and between the CS+ and CS- (*p*s < .037) were positively correlated.

Confidence intervals revealed that the correlation between frontal vmPFC and vmPFC-sgACC activity to the CS+ (95% CI = [0.18, 0.31]) was significantly greater than all other correlations between bottom-up regulatory activity (95% CI upper bound < 0.17) except between frontal vmPFC activity to the CS+ and CS- (95% CI = [0.11, 0.29]). Frontal vmPFC activity to the CS+ and CS- was in turn significantly greater than correlations between vmPFC-sgACC activity to the CS+ and CS- (95% CI = [0.03, 0.07]), and between vmPFC-sgACC activity to the CS+ and frontal vmPFC activity to the CS- (95% CI = [0.003, 0.09]).

Individuals with higher than average SCR spontaneous recovery exhibited a significantly reduced correlation (Wald $\chi^2 = 7.94$, $p = .005$) between vmPFC-sgACC activity and frontal vmPFC activity to the CS- (low group: $B = 0.16$, SE = 0.03, $z = 5.64$, $p < 0.001$, 95% CI = [0.10, 0.21]; high group: $B = 0.05$, SE = 0.02, $z = 2.17$, $p = 0.03$, 95% CI = [0.005, 0.09]).

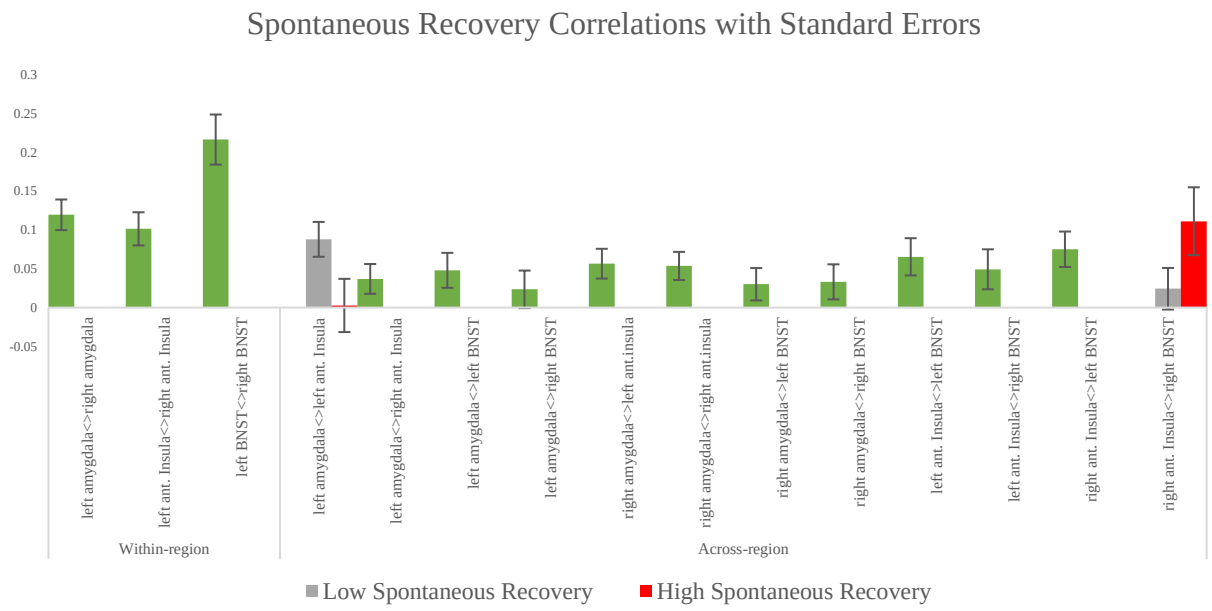


Figure 6: Spontaneous recovery region-of-interest activity correlations.
Note. Ant: anterior. Bars are differentiated between the Low and the High Spontaneous Recovery groups if significantly different.

Between Regions Implementing Spontaneous Recovery. All implementations of neural spontaneous recovery types evidenced positive correlations (e.g., between proximal fear, sustained fear, and interoception; *p*s < .033), although significant correlations were not evidenced between every region. Those exceptions were for correlations between left amygdala and right anterior insula ($p = .054$), left amygdala and right BNST ($p = .333$), right amygdala and left BNST ($p = .149$), right amygdala and right BNST ($p = .143$), and left anterior insula and right BNST ($p = .056$). The correlation between left and right BNST spontaneous recovery (95% CI = [0.15, 0.28]) was significantly greater than all other bivariate correlations of spontaneous recovery (95% CI lower bound < 0.14) except for the correlation between left and right amygdalar spontaneous recovery (95% CI = [0.08, 0.16]). The correlation between left and right amygdalar spontaneous recovery was in turn significantly greater than correlations between left amygdala and right anterior insula (95% CI = [-0.001, 0.07]), left amygdala and right BNST (95% CI = [-0.02, 0.07]), right amygdala and left BNST (95% CI = [-0.01, 0.07]) and, right amygdala and right BNST (95% CI = [-0.01, 0.08]).

Individuals with higher than average SCR spontaneous recovery exhibited a positive relationship between right anterior insular and right BNST spontaneous recovery ($B = 0.11$, SE = 0.04, $z = 2.54$, $p = 0.011$, 95% CI = [0.03, 0.20]), which was statistically different from (Wald $\chi^2 = 5.01$, $p = .025$) and not observed significantly in others ($p = .366$). However, the relationship between left amygdalar and left anterior insular spontaneous recovery was reduced to non-significance ($p = .937$), which was also significantly different from others (Wald $\chi^2 = 4.46$, $p = .034$).

Direct Paths

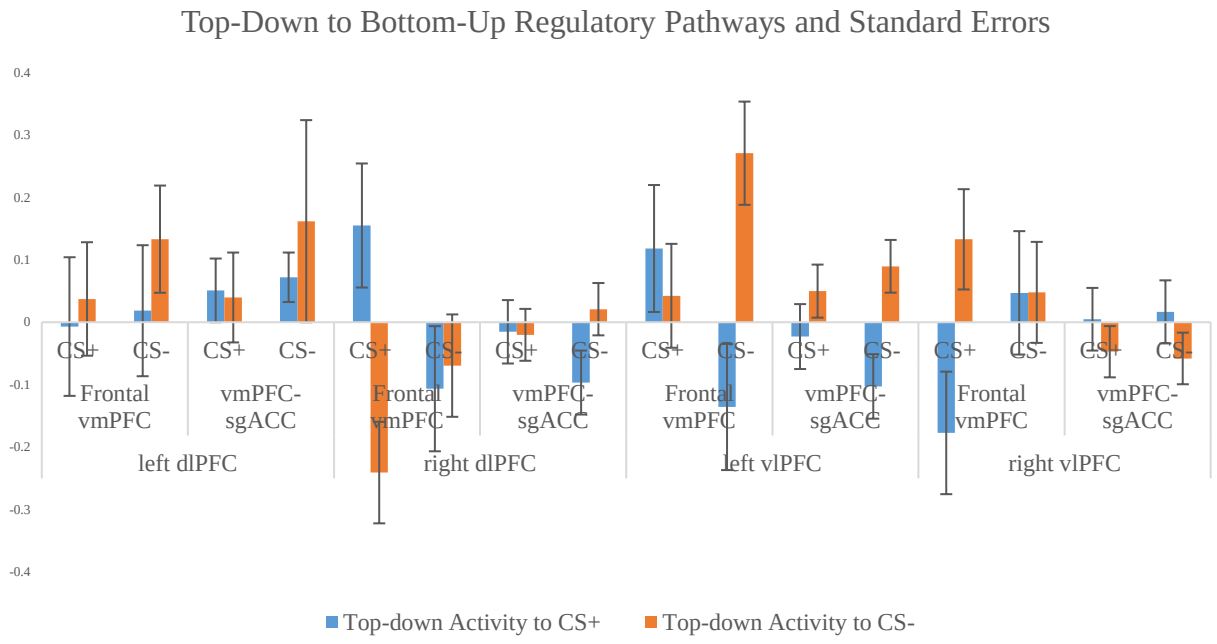


Figure 7: Direct paths from top-down to bottom-up activity in regions-of-interest. *Note.* dlPFC: dorsolateral prefrontal cortex, vlPFC: ventrolateral prefrontal cortex, vmPFC: ventromedial prefrontal cortex, sgACC: subgenual anterior cingulate.

Top-Down to Bottom-Up Regulatory Activity. Greater left vlPFC activity to the CS- predicted increased frontal vmPFC activity to the CS- (B = 0.27, SE = .08, z = 3.27, p = .001, 95% CI = [0.11, 0.43]). Greater right dlPFC activity to the CS- predicted reduced frontal vmPFC activity to the CS+ (B = -0.24, SE = .10, z = -2.40, p = .016, 95% CI = [-0.44, -0.04]).

Greater left dlPFC (B = -0.16, SE = .04, z = 3.69, p < .001, 95% CI = [0.08, 0.25]) and left vlPFC (B = 0.09, SE = .04, z = 2.12, p = .034, 95% CI = [0.01, -0.17]) activity to the CS- predicted increased vmPFC-sgACC activity to the CS-. Greater right dlPFC (B = -0.10, SE = .04, z = -2.33, p = .020, 95% CI = [-0.18, -0.02]) and left vlPFC (B = -0.10, SE = .04, z = -2.42, p = .016, 95% CI = [-0.19, -0.020]) activity to the CS+ predicted reduced vmPFC-sgACC activity

to the CS-. Activity in top-down regulatory regions did not significantly predict any other activity in bottom-up regulatory regions (*p*s > .071).

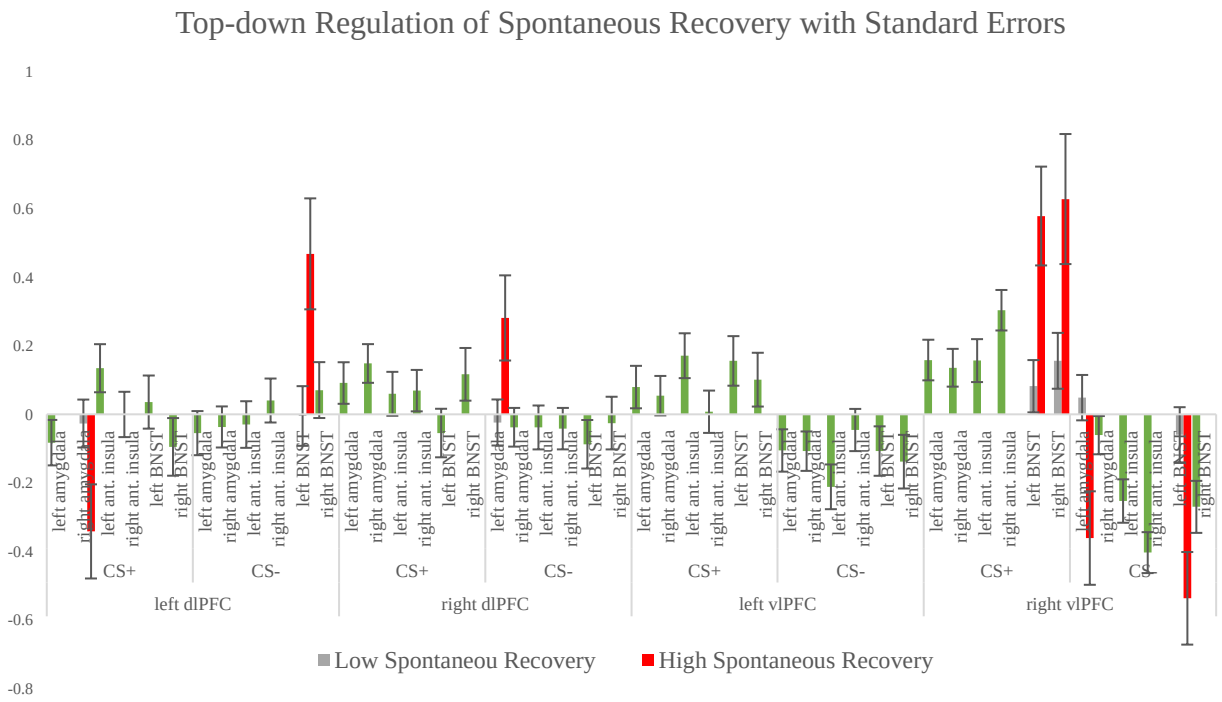


Figure 8: Direct paths from top-down to spontaneous recovery activity in regions-of-interest. *Note.* dlPFC: dorsolateral prefrontal cortex, vlPFC: ventrolateral prefrontal cortex, ant: anterior. Bars are differentiated between the Low and the High Spontaneous Recovery groups if significantly different.

Top-Down Regulatory Activity to Spontaneous Recovery. Across the entire sample, left dlPFC activity did not predict spontaneous recovery (*p*s > .213). Right dlPFC activity to the CS+ predicted significantly greater spontaneous recovery in right amygdala (B = 0.15, SE = .06, z = 2.63, p = .009, 95% CI = [0.04, 0.26]), but right dlPFC activity to the CS- did not significantly predict spontaneous recovery (*p*s > .218). Left vlPFC activity to the CS+ predicted significantly greater spontaneous recovery in left anterior insula (B = 0.17, SE = .07, z = 2.61, p = .009, 95% CI = [0.04, 0.30]) and left BNST (B = 0.16, SE = .07, z = 2.16, p = .031, 95% CI = [0.01, 0.30]), while left vlPFC activity to the CS- predicted decreased spontaneous recovery in

left anterior insula (B = -0.21, SE = .07, z = -3.24, p = .001, 95% CI = [-0.34, -0.08]). Right vlPFC activity to the CS+ predicted increased spontaneous recovery in bilateral amygdala (left: B = 0.16, SE = .06, z = 2.67, p = .008, 95% CI = [0.04, 0.27]; right: B = 0.14, SE = .06, z = 2.46, p = .014, 95% CI = [0.03, 0.24]), bilateral anterior insular spontaneous recovery (left: B = 0.16, SE = .06, z = 2.50, p = .012, 95% CI = [0.03, 0.28]; right: B = 0.17, SE = .07, z = 2.61, p = .009, 95% CI = [0.04, 0.30]), and right BNST spontaneous recovery (B = 0.20, SE = .08, z = 2.61, p = .009, 95% CI = [0.05, 0.34]). It also predicted marginally increased left BNST spontaneous recovery (B = -0.15, SE = .07, z = -2.09, p = .037, 95% CI = [-0.28, -0.01]). Right vlPFC activity to the CS- predicted decreased spontaneous recovery in bilateral anterior insula (left: B = -0.25, SE = .06, z = -3.99, p < .001, 95% CI = [-0.38, -0.13]; right: B = -0.40, SE = .06, z = -6.75, p < .001, 95% CI = [-0.52, -0.29]) and bilateral BNST (left: B = -0.15, SE = .07, z = -2.09, p = .037, 95% CI = [-0.28, -0.01]; right: B = -0.27, SE = .08, z = -3.56, p < .001, 95% CI = [-0.42, -0.12]).

For individuals with higher than average SCR spontaneous recovery, greater right dlPFC activity to the CS- (Wald χ^2 = 4.48, p = .034; B = 0.28, SE = 0.12, z = 2.26, p = .024, 95% CI = [0.04, 0.53]) and insufficient right vlPFC activity to the CS- (Wald χ^2 = 7.00, p = .008; B = -0.36, SE = 0.14, z = -2.65, p = .008, 95% CI = [-0.63, -0.09]) predicted increased left amygdalar spontaneous recovery. Insufficient left dlPFC activity to the CS+ predicted greater right amygdalar spontaneous recovery (Wald χ^2 = 4.32, p = .038; B = -0.34, SE = 0.14, z = -2.49, p = .013, 95% CI = [-0.61, -0.07]). Greater right vlPFC activity to the CS+ (Wald χ^2 = 7.00, p = .008; B = 0.58, SE = 0.14, z = 4.01, p < .001, 95% CI = [0.30, 0.86]) and left dlPFC activity to the CS- (Wald χ^2 = 5.00, p = .025; B = 0.47, SE = 0.16, z = 2.89, p = .004, 95% CI = [0.15, 0.79]) predicted greater left BNST spontaneous recovery. Insufficient right vlPFC activity to the

CS- predicted greater left BNST spontaneous recovery (Wald χ^2 = 7.07, p = .008; B = -0.54, SE = 0.14, z = -3.97, p < .001, 95% CI = [-0.80, -0.27]). Finally, greater right vlPFC activity to the CS+ predicted greater right BNST spontaneous recovery (Wald χ^2 = 4.95, p = .026; B = 0.63, SE = 0.19, z = 3.31, p = .001, 95% CI = [0.26, 1.00]). All of these regulatory influences were not significantly present in individuals without higher than average SCR spontaneous recovery (ps > .055).

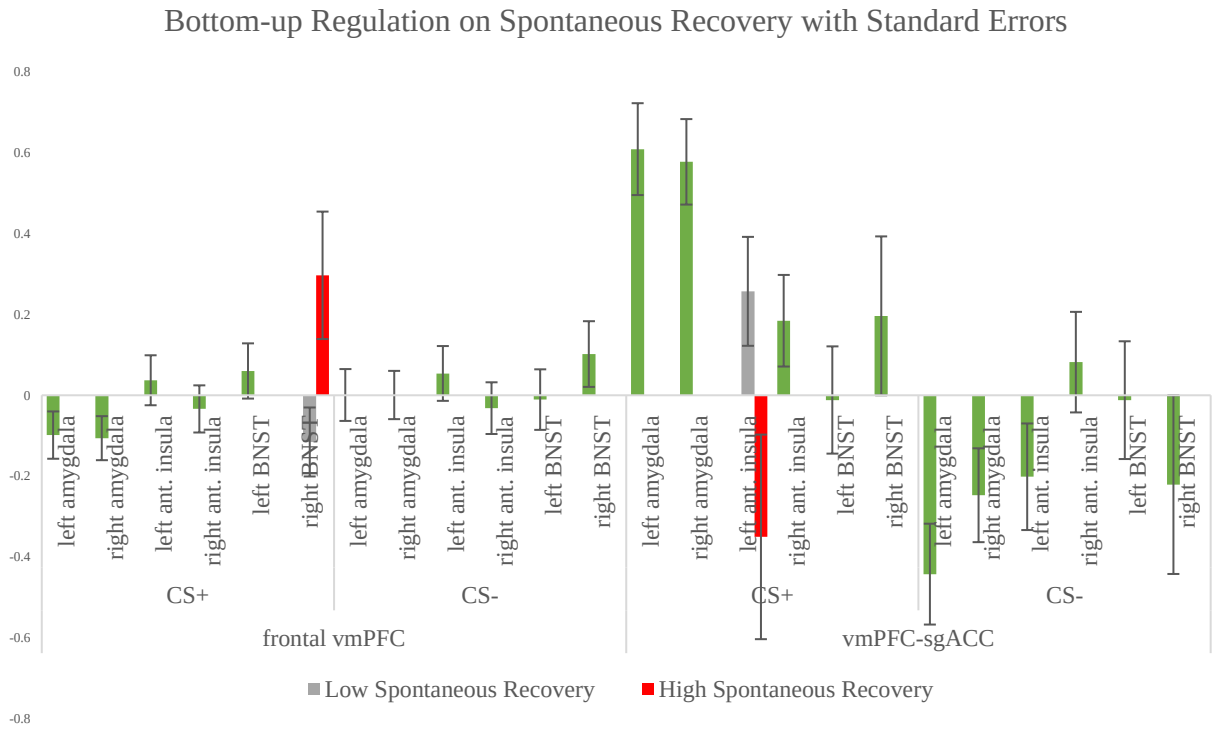


Figure 9: Direct paths from bottom-up to spontaneous recovery activity in regions-of-interest. *Note.* vmPFC: ventromedial prefrontal cortex, sgACC: subgenual anterior cingulate, ant: anterior. Bars are differentiated between the Low and the High Spontaneous Recovery groups if significantly different.

Bottom-up Regulatory Activity to Spontaneous Recovery. Across the entire sample, frontal vmPFC activity did not significantly affect spontaneous recovery (ps > .051). VmPFC-sgACC activity to the CS+ predicted greater spontaneous recovery in bilateral amygdala (left: B = 0.61, SE = 0.11, z = 5.36, p < .001, 95% CI = [0.39, 0.83]; right: B = 0.58, SE = 0.11, z = 5.46,

$p < .001$, 95% CI = [0.37, 0.78]). Insufficient vmPFC-sgACC activity to the CS- also predicted increased spontaneous recovery in bilateral amygdala (left: $B = -0.44$, SE = 0.12, $z = -3.55$, $p < .001$, 95% CI = [-0.69, -0.20]; right: $B = -0.25$, SE = 0.12, $z = -2.13$, $p = .033$, 95% CI = [-0.47, -0.02]). There was no evidence of significant paths relating bottom-up activity with spontaneous recovery in bilateral anterior insula (ps > .104) or bilateral BNST (ps > .162).

Individuals with higher than average SCR spontaneous recovery exhibited a marginally significant path such that greater frontal vmPFC activity to the CS+ predicted greater spontaneous recovery in right BNST ($B = 0.30$, SE = 0.16, $z = 1.88$, $p = .06$, 95% CI = [-0.01, 0.61]), which was significantly greater relative to others (Wald $\chi^2 = 5.04$, $p = .025$; low SCR spontaneous recovery: $B = -0.12$, SE = 0.09, $z = -1.35$, $p = .176$, 95% CI = [-0.28, 0.05]). They also exhibited a significantly different (Wald $\chi^2 = 4.65$, $p = .031$), inverted, and non-significant relationship from vmPFC-sgACC activity to the CS+ to left anterior insular spontaneous recovery (high spontaneous recovery: $B = -0.35$, SE = 0.25, $z = -1.38$, $p = .166$, 95% CI = [-0.85, 0.15]; low spontaneous recovery: $B = 0.25$, SE = 0.13, $z = 1.91$, $p = .056$, 95% CI = [-0.01, 0.52]).

Mediating Paths

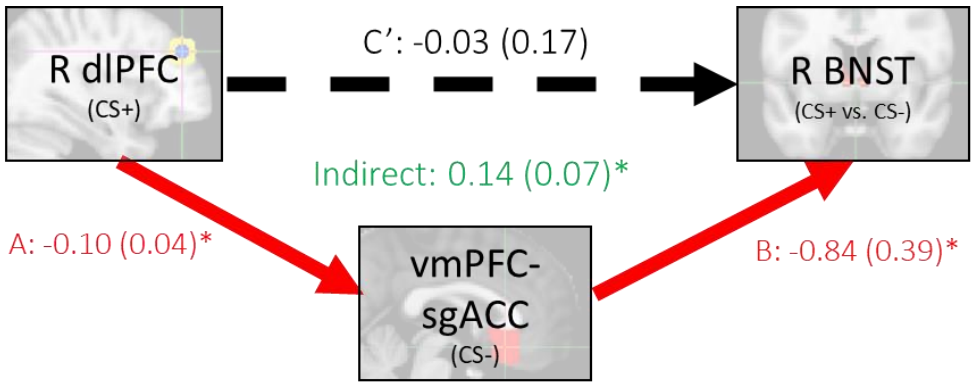


Figure 10: Significant indirect pathway for individuals exhibiting high levels of SCR spontaneous recovery.

Note. dlPFC: dorsolateral prefrontal cortex, vmPFC: ventromedial prefrontal cortex, sgACC: subgenual anterior cingulate, BNST: bed nucleus of the stria terminalis, CS: conditional stimulus. Reported as unstandardized B (Standard Error) * $p < .05$

There was no evidence of indirect effects in the general sample (*p*s > .157) or within the low SCR spontaneous recovery group (*p*s > .088) to suggest that effects on spontaneous recovery by activity in top-down regulatory regions were mediated by bottom-up regulatory regions. Individuals with higher than average SCR spontaneous recovery exhibited a significant indirect pathway from right dlPFC activity to the CS+ to increased right BNST spontaneous recovery through vmPFC-sgACC activity to the CS- (*B* = 0.14, SE = 0.07, *z* = 2.10, *p* = .036, 95% CI = [0.01, 0.27]). There was no significant total effect between right dlPFC to the CS+ and right BNST spontaneous recovery (*p* = .485), suggesting indirect-only mediation.

Discussion

In this study we modelled fear spontaneous recovery by examining the activity of the neural network of regions implementing emotion generation and emotion regulation during a test of fear spontaneous recovery 24-48 hours after fear had been acquired. Our specified model of the interconnections between emotion and emotion regulatory regions suggests a 10.8% increase

in the variance explained over a naïve model with no assumptions and no differentiation between individuals who express high levels of autonomic SCR spontaneous recovery and individuals who do not. The model suggests evidence for robust pathways of emotion regulation. It suggests that abnormally expressed autonomic fear spontaneous recovery is not due to differences in the activity or the variability of activity of top-down emotion regulatory regions (i.e., dlPFC and vlPFC), differences in the activity of bottom-up emotion regulatory regions (i.e., frontal vmPFC and vmPFC-sgACC), or differences in the relations between top-down and bottom-up emotion regulatory activity. Furthermore, there were no differences found in the variance of neural spontaneous recovery activity in amygdala, BNST, or anterior insula. In our sample, participants were not instructed to regulate their fears, yet spontaneous recovery was not evidenced either in SCRs or neural activity. However, as we detail below, we found evidence for differential emotion generation, as well as regulation, misregulation, and regulatory conflict in individuals who did express higher than average SCRs at a test of fear spontaneous recovery. A disproportionate amount of those individuals were males. This model is likely to replicate for future samples, as it demonstrates good fit (.065) at the upper bound of the RMSEA 90% confidence interval.

Emotion Generation

There was no evidence of neural spontaneous recovery. Our data does suggest that the variability of amygdalar spontaneous recovery was less than that of BNST, suggesting that the spontaneous recovery of proximal fear responding is more stable on average than the spontaneous recovery of sustained fear responding. Intercorrelations between the spontaneous recovery areas suggest that greater return of proximal fear tends to coincide with greater return of sustained fear and greater interoception of the CS+ compared to the CS-. However, the

intercorrelations between amygdalar spontaneous recovery and BNST spontaneous recovery are sparse, limited only to between left amygdala and left BNST. This, coupled with the observation that the correlation between left amygdala and left BNST is the smallest in magnitude of the significant spontaneous recovery activations (B = .048, p = .03, 95% CI = [.004, .09]), suggests that the spontaneous recovery of proximal fear and the spontaneous recovery of sustained fear are relatively independent.

Emotion Regulation

Top-down regulatory areas were all significantly active to the CS+ and CS- relative to their implicit baselines. The greatest amount of activity, perhaps surprisingly, was generated from the right vlPFC toward the CS-, suggesting a strong general effort toward motor control when the safe cue was presented. This area was more active than left vlPFC or bilateral dlPFC activity toward the CS+, which are relevant for cognitive control, syllogistic reasoning, and stress regulation, all of which are putatively important aspects of regulation when presented with a threat. There was significant intercorrelation among top-down regulatory activities to the CS+ and CS-, with exceptions between right dlPFC and bilateral vlPFC. These areas coactivated toward either CS+ or CS-, but coactivated more sparsely otherwise, only correlating between right dlPFC to the CS+ left vlPFC to the CS-.

Bottom-up regulatory areas of frontal vmPFC and vmPFC-sgACC were deactivated to the CS+ and CS- relative to implicit baselines when all top-down regulatory regions are inactive. Frontal vmPFC was more deactivated but more variable relative to the vmPFC-sgACC, These deactivations may suggest that on average, these areas are ceding the regulatory processing of the conditional stimuli to top-down areas, and activity relevant for extinction processing to the CS+ and CS- is more stable than activity relevant for the goal prioritization of CS+ and CS-.

Frontal vmPFC activity correlated with vmPFC-sgACC activity, suggesting that goal prioritization co-occurs in proportion with extinction processing to organize the perceptual import of the CS+ and CS-.

Top-Down Influence on Bottom-Up Regulatory Activity.

Goal Prioritization. Greater frontal vmPFC activity to the CS+ was associated with increases in left vlPFC activity to the CS- and with decreases in right dlPFC activity to the CS-. This suggests that cognitive control of the CS- increases the goal prioritization of the CS+, while stress regulation effort devoted to the CS- impedes the goal prioritization of the CS+.

Extinction Processing. There were no significant top-down influences on vmPFC-sgACC activity to the CS+, suggesting that activation from top-down regions do not influence the extinction processing of the CS+. Activity in bilateral dlPFC and left vlPFC was associated with vmPFC-sgACC activity to the CS- suggesting that reducing stress regulation and cognitive control processing of the CS+ and increasing syllogistic reasoning and cognitive control of the CS- would predict a reduction of CS- extinction processing.

Top-Down and Bottom-Up Influences on Spontaneous Recovery.

Proximal Fear. Right dlPFC, right vlPFC, and vmPFC-sgACC activity to the CS+ predicted significantly greater spontaneous recovery in amygdala, suggesting that greater stress regulation, motor control, and extinction processing to the CS+ signals the spontaneous recovery of proximal fear. Reductions in proximal fear spontaneous recovery was associated with greater vmPFC-sgACC activity to the CS-. This may suggest that attention paid to the safety value of the CS- during extinction training may play an important role in the reduction of the spontaneous recovery of proximal fear.

Sustained Fear. Left vlPFC activity to the CS+ predicted significantly greater spontaneous recovery in left BNST, suggesting that the implementation of cognitive control on the CS+ may increase the spontaneous recovery of sustained fear.

Interoception. Bilateral vlPFC activity to the CS+ predicted significantly greater spontaneous recovery in anterior insula, suggesting that cognitive and motor control processing to the CS+ predicted increases in CS+ interoception relative to CS- interoception. Right vlPFC activity to the CS- predicted a decreased difference between threat and safety interoception, suggesting that a motoric action plan downregulates internal, potentially fearful dialogues and the attentiveness to potentially fearful internal feeling states.

Emotion Dysregulation

Emotion Overgeneration. Our data suggests that high SCR spontaneous recovery coincides with an elevated level of baseline left amygdalar spontaneous recovery, which dovetails with the understanding that amygdala activity has a proximal relationship with skin conductance levels (Öhman & Soares, 1994). This suggests that the stage for the return of proximal fear is set, although the amount of spontaneous recovery estimated was not significant from zero without influence from regulatory areas. Taken together, these data suggest that high autonomic spontaneous recovery may ultimately have more to do with fear dysregulation than fear generation.

Competitive Co-regulation. Individuals with high SCR spontaneous recovery exhibited an increased correlation of right vlPFC activity between CS+ and CS-, suggesting that for these individuals, the implementation of motoric regulation is applied across safe and threat stimuli. This is particularly important as right vlPFC activity, implementing motor control, was more associated with all estimates of spontaneous recovery: proximal fear, sustained fear, and

interoception. However, right vlPFC activity to the CS+ and CS-, while more positively correlated with each other, also predicted spontaneous recovery in opposite directions. This suggest that implementation of a similar motoric plan across stimuli may induce competitive co-regulation. For such individuals there is perhaps a greater need for them to disassociate the motor strategy used to approach the two stimuli.

Emotion Misregulation. Individuals with high SCR spontaneous recovery uniquely exhibited greater spontaneous recovery of proximal fear with greater right dlPFC activity to the CS-, suggesting that their stress regulation attempts toward safe cues are misregulatory and counterproductive. They also exhibited greater spontaneous recovery of sustained fear with greater right vlPFC activity to the CS+ and greater left dlPFC activity to the CS-, suggesting that their motor contingencies directed to threats and their syllogistic reasoning employed on safe cues serve to sustain fear. There was also marginal indication that greater spontaneous recovery of sustained fear was associated with greater frontal vmPFC activity to the CS+, suggesting that threat cues are perceptually mis-prioritized, although it is not clear through our current results whether it is over- or under-prioritized (and may in fact be person-specific).

Individuals with high SCR spontaneous recovery also exhibited indirect stress misregulation toward threat cues, leading to increased spontaneous recovery of sustained threat. Indirect associations increasing spontaneous recovery in right BNST were found attributed to right dlPFC activity to the CS+ through decreased vmPFC-sgACC activity to the CS-. Stress regulation of the CS+ was not associated with spontaneous recovery of sustained threat but it was associated with decreased extinction processing of the CS-. Since increased extinction processing of the CS- decreases the spontaneous recovery of sustained threat, reducing it through

increased stress regulation of the CS+ serves to indirectly increase spontaneous recovery of sustained threat.

Underregulation and Disconnectivity. Individuals with high SCR spontaneous recovery showed disconnects in the relationship between left dlPFC activity to the CS- and right vlPFC to the CS+, between left dlPFC activity to the CS- and right vlPFC activity to the CS-, and between left vlPFC to the CS+ and right vlPFC to the CS-. All of this is to suggest that such individuals exhibit more widespread discordance in the top-down organization and conceptualization of how threat and safe stimuli relate after extinction training. These individuals also exhibited a reduction in the correlation between frontal vmPFC and vmPFC-sgACC activity to the CS-, suggesting an impairment in the coordination between goal prioritization and extinction processing of the safe cue.

Increased spontaneous recovery of proximal fear was predicted by insufficient left dlPFC activity to the CS+ and right vlPFC activity to the CS-, suggesting that for individuals that express high SCR spontaneous recovery, there needs to be improvements in their syllogistic reasoning about threats, and they need to implement a motoric plan when safe.

In addition, increased spontaneous recovery of sustained fear was also predicted by insufficient right vlPFC activity to the CS-, again highlighting the importance of motoric regulation in the face of safety.

Individuals with high SCR spontaneous recovery also exhibited disconnection between left anterior insular spontaneous recovery and vmPFC-sgACC activity to the CS+. VmPFC-sgACC activity to the CS+ marginally (p = .056) increased left anterior insular spontaneous recovery for individuals with low SCR spontaneous recovery. This may suggest that for most individuals, increased extinction processing of the CS+ (or at least the need to activate extinction

processing) was associated with greater interoceptive processing of threat, while this was not the case for individuals that express high SCR spontaneous recovery.

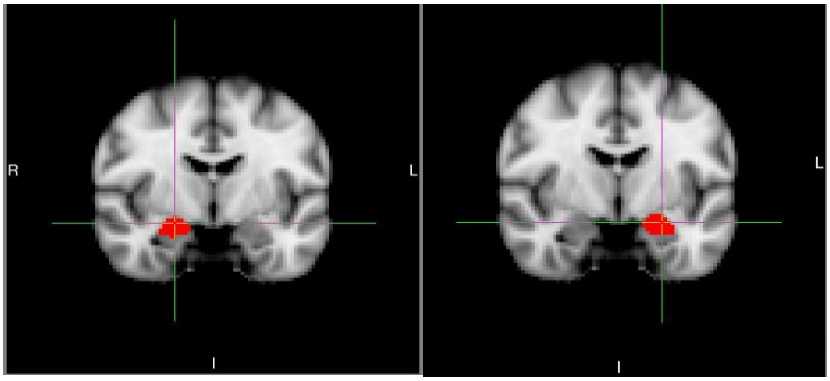
Conclusion

We found that expressed fear spontaneous recovery through SCR, an autonomic index, may be a result of abnormal proximal fear generation, but it is more likely accounted for by a mosaic of fear dysregulation patterns, characterized by the competitive co-regulation of fear interoception and sustained fear, general misregulation of proximal and sustained fear, and emotion underregulation or disconnection of regulatory activity on proximal, sustained, and interoceptive fear. Unexpectedly, we found that regulatory processing of the CS+ threat cue usually served to increase fear spontaneous recovery, calling attention to the fact that when participants are not instructed about *how* to think or act on threats, they may likely perform counterproductively. We also found that regulatory activity toward the CS- safe cue was the most direct pathway to downregulate fear spontaneous recovery, especially through bottom-up activity for the downregulation of the spontaneous recovery of proximal fear and through right vlPFC for the downregulation of fear interoception.

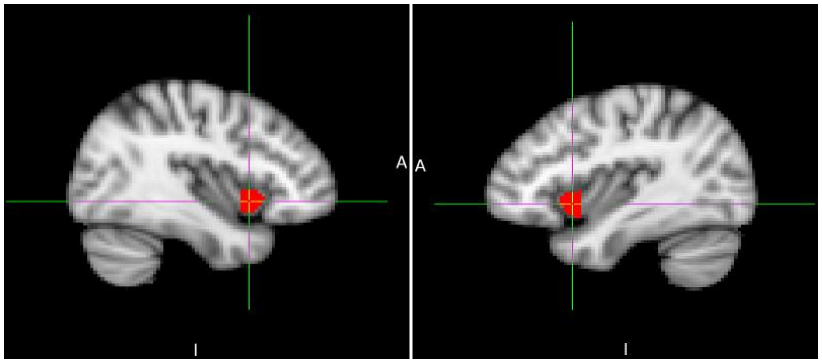
The evidence presented here may implicate clinical conceptualization and clinical practice of treating fear and its spontaneous recovery. The model of fear and fear regulation presented is a reflection of activity soon after extinction-training, which may be similar to the activity profile of an individual that has completed exposure therapy. The profile of dysregulation, or parts of it, may reflect in behavior as predictive indicators for fear-disorder relapse. Therefore, a clinician should be wary of verbal indications of disorganized or misregulating conceptualizations of threat or safe stimuli, inadequate syllogistic reasoning around threats, and perhaps most importantly, motoric plans that are overgeneralized and fail to

discriminate threat from safety, or motoric plans that fail to address safety at all. The evidence also suggests that top-down approaches to safe cues are an underappreciated aspect of spontaneous recovery reduction or prevention in contemporary clinical psychology. Instead, individuals at risk for fear relapse, may benefit most from cognitive didactics and behavioral practices that promote adaptive motor control strategies and contingencies that regulate right vlPFC responses to threat and safety cues.

Figures



Figures 11a and 11b: (a) Right and (b) left amygdala regions-of-interest defined by the Harvard-Oxford MNI probabilistic atlas.



Figures 12a and 12b: (a) Right and (b) left anterior insular cortex regions-of-interest defined by the Harvard-Oxford MNI probabilistic atlas.

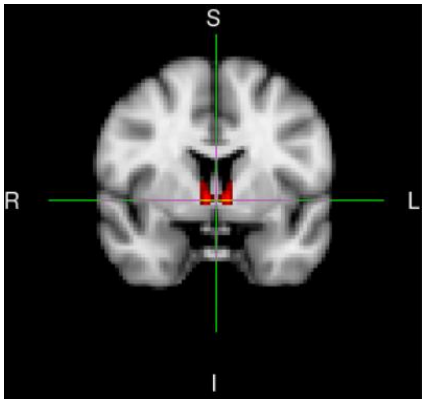
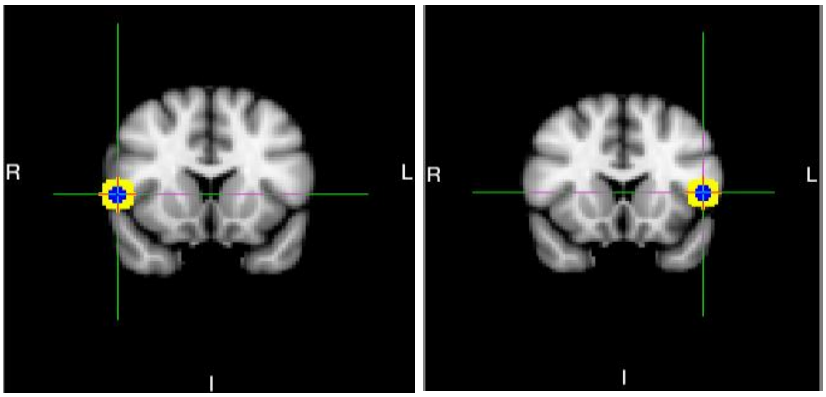
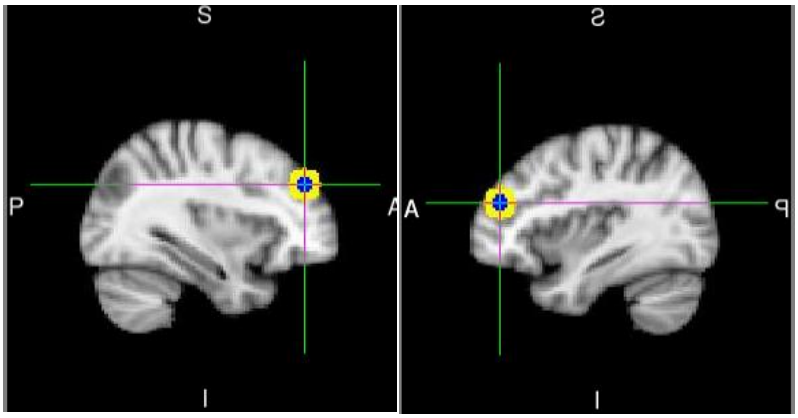


Figure 13: Right and left bed nuclei of the stria terminalis regions-of-interest defined by Avery et al, 2015.



Figures 14a and 14b: (a) Right and (b) left ventrolateral prefrontal cortex 5mm region-of-interest spheres in blue centered on coordinates found in Fullana et al. 2016.



Figures 15a and 15b: (a) Right and (b) left dorsolateral prefrontal cortex 5mm region-of-interest spheres in blue centered on coordinates found in Fullana et al. 2016.

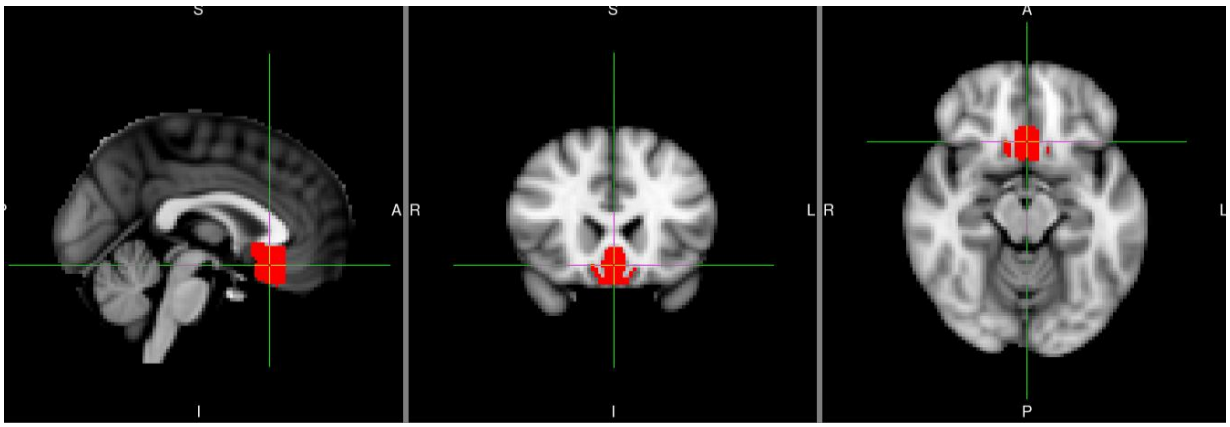


Figure 16: Ventromedial prefrontal-subgenual anterior cingulate cortices (vmPFC-sgACC) defined by the Harvard-Oxford MNI probabilistic atlas.

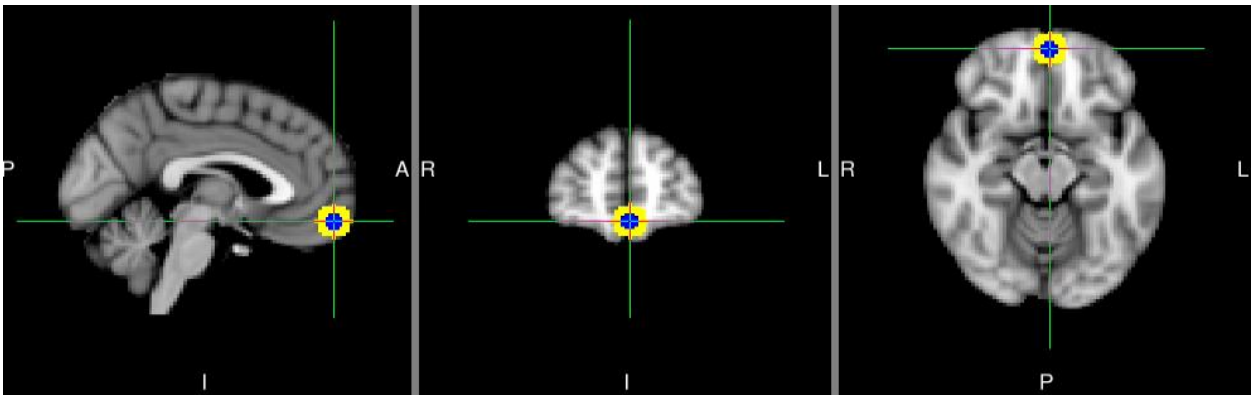


Figure 17: Ventromedial prefrontal cortex (vmPFC) 5mm region-of-interest spheres in blue defined by Fullana et al. 2016

General Discussion

"To perceive the world differently, we must be willing to change our belief system, let the past slip away, expand our sense of now, and dissolve the fear in our minds." – William James

The three studies in this dissertation assessed aspects of emotion regulation and emotion generation on the return of fear. Return of fear was examined through paradigms of rapid reacquisition (Study 1), fear reinstatement and context renewal (Study 2), and spontaneous recovery (Studies 1 and 2). We also attempted to elucidate the neural underpinnings of the return of fear through spontaneous recovery (Study 3). This series of studies sets the stage for a program of future investigations that bridge the neural, peripheral physiological, and self-report phenomenology of conditional fear learning and the return of fear.

Within the framework of fear learning, we configured the cognitive reappraisal instruction featured in Studies 1 and 2 to be directed at the US-self relationship. Specifically, it asked participants to rethink the scream sound as syntonic to the goals of one's role as a Hollywood movie director. Salience of the CS or the US was intentionally not targeted in order to preserve the prediction error required for reinforcing CS-noUS extinction memories. This study also importantly minimized the cost of cognitive reappraisal, in accordance with the computation implementation model of emotion regulation, by instructing the participant to take on the role of a movie director, presumably an easy thing to imagine especially for participants drawn from the southern Los Angeles area. Also, this cognitive reappraisal was not a distraction induction, as the instruction entailed paying close attention to the stimuli. One would presume that distraction would lead to a worsened fear trajectory, given that learning of the contingencies that lead to extinction would be disrupted.

The results of the experimental studies testing the effects of our low-cost cognitive reappraisal modestly supports its inclusion alongside extinction training for short-term fear amelioration, as it more quickly reduced CS- threat expectancies in two of the three experiments. This result was not observed in the experiment testing for the effects of context renewal. This may be because the test of spontaneous recovery in that experiment featured CS presentations with visual contexts, stimuli that is sufficiently distracting to mitigate the benefits of cognitive reappraisal. When translating these observations toward clinical application, two things are worth noting. The first is that cognitive reappraisal consistently did not evidence significant effects on fear responding to the CS+, so cognitive reappraisal does not directly enhance the extinction of conditional fear. The second is that the enhancements in the understanding that safe stimuli are safe, how one might interpret the faster reduction in threat expectancy to the CS-, are rather quite fragile. Even if they can be replicated in laboratory settings, given the sheer number of ecological distractors inevitable during exposure therapy in the clinic or elsewhere exposure may be indicated, we would be unlikely to see any short-term benefits of implementing cognitive reappraisal alongside exposure therapy.

Cognitive reappraisal also showed very limited benefit for long-term fear amelioration, showing efficacy for the reduction of SCR spontaneous recovery in one out of three tests, and showing no evidence of efficacy in the reduction or prevention of rapid reacquisition, fear reinstatement, or context renewal. Taken together, this can be interpreted in a positive light in that cognitive reappraisal, at least the one designed in this set of studies, did not result in any deleterious effects, which might have been predicted by learning theory, purporting that cognitive reappraisal can adversely impact prediction error, leading to mitigated extinction learning (Vervliet et al., 2013). Rather, deleterious effects were observed consistently for the use

of emotion suppression in combination with extinction training, which dovetails with the literature frequently describing suppression as costly (e.g., Richards & Gross, 2000, 2000; Roberts, Levenson, & Gross, 2008; Srivastava, Tamir, McGonigal, John, & Gross, 2009), ironic (e.g., Ben-Naim, Hirschberger, Ein-Dor, & Mikulincer, 2013; Burns et al., 2008; Burns, Quartana, & Bruehl, 2007; Butler, Young, & Randall, 2010; Dalgleish, Yiend, Schweizer, & Dunn, 2009; Quartana & Burns, 2007, 2010; Quartana, Yoon, & Burns, 2007), exacerbatory (e.g., Campbell-Sills, Barlow, Brown, & Hofmann, 2006; Dennis, 2007; Langner, Epel, Matthews, Moskowitz, & Adler, 2012), and deleterious in general (Chapman, Fiscella, Kawachi, Duberstein, & Muennig, 2013; Moore, Zoellner, & Mollenholt, 2008).

Despite the emotion regulation inductions, the pattern of data largely adhered to learning curves typical of the Rescorla-Wagner model. A single cognitive reappraisal strategy in the context of fear conditioning can take the form of re-evaluating the CS, US, the self, or any relation or set of relations between these. It is not clear if any one strategy is superior to another, and it remains to be seen how each form of reappraisal corresponds to the variables in the Rescorla-Wagner model. Given the reductions in fear due to the cognitive reappraisal instruction focused here, which was directed at the US-self relationship, this instruction may have reduced β, the associative value of the US, λ is the CS-US associative maximum, or both. A hypothesis whereby a US-self reappraisal is superior to a CS-targeted or US-targeted reappraisal, especially ones that may theoretically increase α_x, β, or λ, should be explicitly tested. It will be worth understanding what types of cognitive reappraisals are useful and what types may potentially be detrimental to build a theoretical interface ontology between learning and evaluative terminology.

It may be that the cognitive reappraisal instruction used in Studies 1 and 2 lies within classes of multiple cognitive reappraisal strategies (McRae, Ciesielski, & Gross, 2012). One criticism may be that the instructions used conflates a self-role reappraisal with a distancing reappraisal, and the effects may be mixed between the two types amongst participants. How a self-role reappraisal affects the fear response in this context may depend on the inherent value that the role of "movie director" may take, as this could be a simple neutral role shift, or a shift that takes on a positive valence as it mentally shifts one from a position of low-power in an undergraduate student to a position of high power. Other than sample differences in trait variables, such as personality or emotion regulatory habits, this may be underlying the mismatch in the effects of cognitive reappraisal on spontaneous recovery between Study 1 and Study 2. Future work should strive to better understand the role of cognitive effort on attention, as increased general cognitive effort may be driving enhanced learning of CS-noUS contingencies.

The reasons why we observed unreliable efficacy of cognitive reappraisal on spontaneous recovery may be better understood from the neural perspective garnered from Study 3. Although in Studies 1 and 2, we designed a cognitive reappraisal that was hypothesized to target the US, we did not analyze how US activity predicted spontaneous recovery levels, so it is not known if this is the optimal pathway. However, it should be noted that our cognitive reappraisal strategy also reappraised the CSs as "photos of actresses" and it reappraised the CS-US relationship as "an actress's scream". Cognitive mechanisms of emotion regulation such as cognitive reappraisal are said to be "top-down" and implemented by high-order frontal brain structures such as the dlPFC and vlPFC. Learning mechanisms of emotion regulation such as extinction learning are said to be "bottom-up" and implemented by more ventral regions such as vmPFC and sgACC. Although Studies 1 and 2 suggest that although the regulation of fear may effectively result from

a simple low-cost cognitive reappraisal instruction, cognitive reappraisal may not be used if uninstructed and the results of Study 3 dovetails with this with potential explanation as to why.

Without instruction, the data from Study 3 suggests that activity in right dlPFC and right vlPFC (along with vmPFC-sgACC) to the CS+ *increased* neural activity in amygdala at a test of spontaneous recovery. For individuals who expressed levels of autonomic spontaneous recovery that was higher on average, there was evidence that a regulatory pathway through right dlPFC toward the CS+ indirectly upregulated sustained fear spontaneous recovery by decreasing vmPFC-sgACC activity to the CS-. Taken together, this discourages the CS+ targeting of cognitive reappraisal.

The data from Study 3 also suggests that neural processing of the CS- in service of reducing fear spontaneous recovery may be underappreciated. Reductions in amygdalar spontaneous recovery was associated with greater vmPFC-sgACC activity to the CS-. Furthermore, we observed that individuals expressing higher than average autonomic spontaneous recovery exhibited hypercorrelated activity between the CS+ and CS- in the right vlPFC, which is posited to be involved with the planning of goal-directed motoric planning. This suggests that there was a more similar motor response to the CS+ and CS- planned within those individuals relative to individuals who exhibited a healthier spontaneous recovery profile. Right vlPFC activity to the CS+ and CS- competitively co-regulate amygdalar spontaneous recovery, as right vlPFC activity to the CS+ increased amygdalar spontaneous recovery, whereas right vlPFC activity to the CS- decreased amygdalar spontaneous recovery. A natural hypothesis that emerges from this would be that perhaps the most effective cognitive reappraisals for individuals who exhibit high levels of spontaneous recovery disassociates motoric plans between the CS+ and CS- and implements a motoric plan specifically directed at the CS-. To extend our

Hollywood director example, we may instruct participants “log in your mind instances when you do not hear a scream, because that should go against your decision to hire her.”

The analysis of Study 3 uses structural equation modelling, which implies causal effective connectivity relationships between regions. However, the nature of the estimated associations in this model cannot be deemed causal due to the study design. We observed that many of the direction of effects from regulatory regions to fear implementation regions as modelled were in the positive direction, which suggests that regulatory areas typically increase emotional activity when participants are uninstructed. Notably, the results reported are also general associations between emotional and regulatory neural activity with stimulus presentation within a relatively long time-span. This strategy is similar to analytical strategies employed in other studies (Fullana et al., 2016; Milad et al., 2005), but it shares the weakness that it is unable to tease apart the temporal dynamics of these relationships. Although the hypotheses center on relationships such that stimulus properties would produce activation in fear regions followed by influences by activation in regulatory regions, alternative interpretations are that our estimates may represent how fear region activations produce regulatory region activations, or how stimulus presentations produce activation in both fear and regulatory regions simultaneously. Future work proposing different SEMs, or using such methods as psychophysiological interaction (PPI) analyses, can help tease out these explanations by supplying evidence as to whether a relationship between the BOLD activity between regulatory and fear regions differ between the CS-types, and if interregional connectivity is interactive versus independent. Beyond the static modelling of PPI and SEM, dynamic causal modelling (DCM; Friston, Harrison, & Penny, 2003) may provide evidence for a more exact model of information flow through the brain between

regulatory and emotion regions than the one assumed in Study 3 using systematic model comparison.

Studies 1, 2, and 3, broadly recorded indices of conditional fear autonomically through skin conductance responding, reportably through the continuous assessment of US expectancy, and intermittently through self-reports of fear, valence, and arousal. Many more methodologies are available and were not assessed, such as reaction time recording to threatening stimuli, electroencephalogram, and fMRI. This type of broad application of methodologies to index conditional fear might be required to uncover unobvious benefits and detriments and to better specify the dynamics of using clinical interventions that modulate simultaneous combinations of psychological (e.g., cognitive restructuring, mindfulness, behavioral activation). However, such an approach should first be viewed as exploratory, requiring follow-up observations to uncover the granular elements and processes that underlie unobvious changes in outcomes such as conditional fear.

The current work provides the foundation for a program of research that aims to elucidate the neural, peripheral physiological, behavioral, and reportable links that manifest the return of fear and its interaction with attempts at self-regulation. Future work can utilize the experimental manipulations in Studies 1 and 2 in the fMRI scanner to examine how different emotion regulatory strategies direct impact the neural network of activations in the frontal regulatory-emotion generation system outlined in Study 3. Further work can re-examine this activation network in the contexts of return of fear phenomena other than spontaneous recovery, namely, rapid reacquisition, fear reinstatement, and context renewal. Furthermore, future work that follows the themes of (1) clarifying the costs of enacting cognitive reappraisal, (2) clarifying the effects of specific classes of reappraisal on fear, (3) clarifying the effects of reappraisal

properties such as cost, value, and target, and (4) mapping the constituent elements underlying reappraisal and suppression to the formulas of Rescorla-Wagner and the computational implementation model will be important next steps for a field of optimizing fear regulation. Such a research trajectory can elucidate combinable mechanisms that accelerates fear extinction and reduce or prevent the return of fear.

Clinical Research and Translation

Zooming out to a clinical translational model for fear extinction as outlined by the model described by Craske, Hermans, & Vervliet, (2018), these studies are steps forward in the elucidation of cognitive-emotional processes that moderate fear extinction (see Figure 1). These studies also advance a nomological net connecting inhibitory regulation with the disconfirmation of expectancies, and the attention to feared stimuli components of optimizing exposure therapy (see Figure 2).

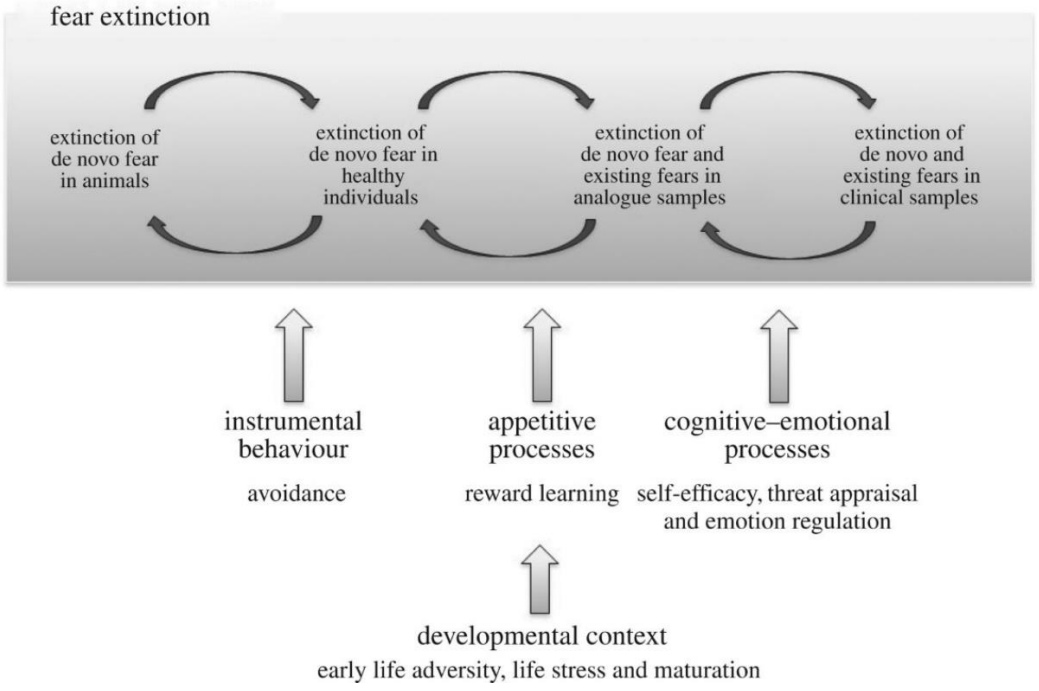


Figure 1: Translational model for fear extinction adapted from Craske et al., 2018

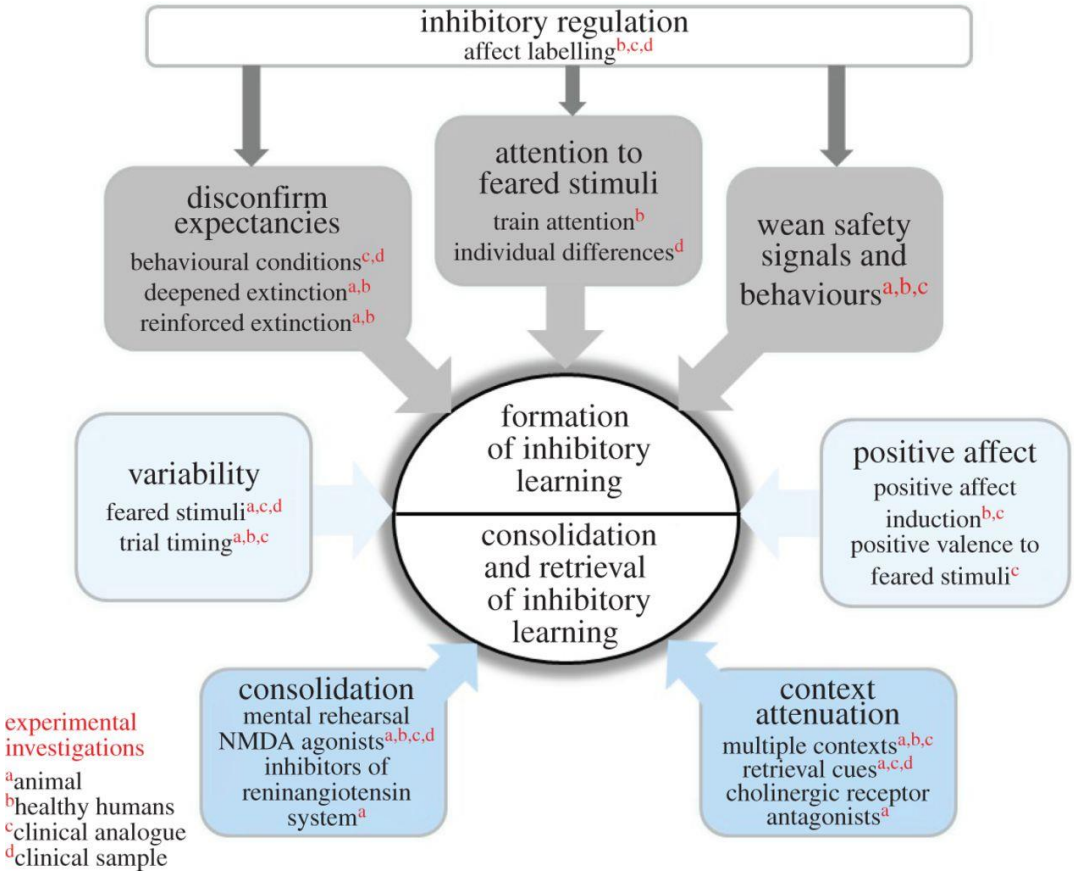


Figure 2: Extinction-derived strategies for optimizing exposure therapy adapted from Craske et al., 2018

It is unclear whether inducing purer constituents underlying cognitive reappraisal instructions of Study 1 would improve the fear trajectory or translate well clinically. However, it is important to know how these constituents work, for example, to know what cognitive reappraisal classes are appropriate to combine for crafting clinical interventions. This is especially in light of the fact that cognitive restructuring in a clinical setting is unlikely to consist solely on one, or even centered on one strategy. A clinician treating a dog phobic, over the course of a single visit, may say "most dogs will not bite you" (a CS-not bad for me reappraisal), "you are overestimating the probability of these bites" (CS-US contingency reappraisal), "and you are overreacting to the consequences of a bite" (US-not as bad for me reappraisal), "so what would you say to your friend if your friend was afraid of dogs?" (distancing and self-role

reappraisal). Furthermore, examining how cognitive reappraisals work as interventions involving a dyadic relationship must take relational aspects into account, such as communicative delivery (e.g., tone and prosody) from the interventionist, and individual differences of the client such as their communicative receptiveness and cognitive rigidity.

Beyond cognitive reappraisal and extinction training, other combinations targeted at the fear response may play a role, and should be induced and tested. This might include such factors as physiological change through relaxation, breathing exercises, or biofeedback. The role of all these regulatory factors speak to the important and dynamic role that clinicians and clinical technologies such as apps or devices to aid in mental health play in the intervention. That is, addressing fear amelioration appropriately will require, the appropriate "recipe" or emotion regulatory strategies that ensures the modulation of appropriate psychophysiological outcomes should be prepared and induced when appropriate.

References

Agren, T., Engman, J., Frick, A., Björkstrand, J., Larsson, E.-M., Furmark, T., & Fredrikson, M. (2012). Disruption of reconsolidation erases a fear memory trace in the human amygdala. *Science*, *337*(6101), 1550–1552.

Anglada-Figueroa, D., & Quirk, G. J. (2005). Lesions of the basal amygdala block expression of conditioned fear but not extinction. *Journal of Neuroscience*, *25*(42), 9680–9685.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177. https://doi.org/10.1016/j.tics.2004.02.010

Avery, S. N., Clauss, J. A., & Blackford, J. U. (2015). The Human BNST: Functional Role in Anxiety and Addiction. *Neuropsychopharmacology*, *41*, 126.

Bangasser, D. A., & Shors, T. J. (2008). The Bed Nucleus of the Stria Terminalis Modulates Learning after Stress in Masculinized But Not Cycling Females. *Journal of Neuroscience*, *28*(25), 6383–6387. https://doi.org/10.1523/JNEUROSCI.0831-08.2008

Barad, M., Gean, P.-W., & Lutz, B. (2006). The role of the amygdala in the extinction of conditioned fear. *Biological Psychiatry*, *60*(4), 322–328.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205.

Baum, M. (1988). Spontaneous recovery from the effects of flooding (exposure) in animals. *Behaviour Research and Therapy*, *26*(2), 185–186.

Baxter, A. J., Vos, T., Scott, K. M., Ferrari, A. J., & Whiteford, H. A. (2014). The global burden of anxiety disorders in 2010. *Psychological Medicine*, *44*(11), 2363–2374.

Beauregard, M., Lévesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *The Journal of Neuroscience*, *21*(18), 6993–7000.

Ben-Naim, S., Hirschberger, G., Ein-Dor, T., & Mikulincer, M. (2013). An experimental study of emotion regulation during relationship conflict interactions: The moderating role of attachment orientations. *Emotion*, *13*(3), 506.

Biopac System Inc. (n.d.). Goleta, CA, USA: Biopac System Inc.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986.

Bouton, M. E., & Bolles, R. C. (1979). Role of conditioned contextual stimuli in reinstatement of extinguished fear. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*(4), 368.

Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 248.

Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review*, *108*(1), 4.

Bouton, M. E., & Swartzentruber, D. (1991). Sources of relapse after extinction in Pavlovian and instrumental learning. *Clinical Psychology Review*, *11*(2), 123–140.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., … Ochsner, K. N. (2014). Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990.

Burns, J. W., Quartana, P., Gilliam, W., Gray, E., Matsuura, J., Nappi, C., … Lofland, K. (2008). Effects of anger suppression on pain severity and pain behaviors among chronic pain patients: Evaluation of an ironic process model. *Health Psychology*, *27*(5), 645.

Burns, J. W., Quartana, P. J., & Bruehl, S. (2007). Anger management style moderates effects of emotion suppression during initial stress on pain and cardiovascular responses during subsequent pain-induction. *Annals of Behavioral Medicine*, *34*(2), 154–165.

Butler, E. A., Young, V. J., & Randall, A. K. (2010). Suppressing to please, eating to cope: The effect of overweight women's emotion suppression on romantic relationships and eating. *Journal of Social and Clinical Psychology*, *29*(6), 599–623.

Campbell-Sills, L., Barlow, D. H., Brown, T. A., & Hofmann, S. G. (2006). Acceptability and suppression of negative emotion in anxiety and mood disorders. *Emotion*, *6*(4), 587.

Chapman, B. P., Fiscella, K., Kawachi, I., Duberstein, P., & Muennig, P. (2013). Emotion suppression and mortality risk over a 12-year follow-up. *Journal of Psychosomatic Research*, *75*(4), 381–385.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Clark, D. M. (1999). Anxiety disorders: Why they persist and how to treat them. *Behaviour Research and Therapy*, *37*(1), S5.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1742), 20170025. https://doi.org/10.1098/rstb.2017.0025

Craske, M. G., & Mystkowski, J. L. (2006). *Exposure Therapy and Extinction: Clinical Studies.* Retrieved from http://psycnet.apa.org/index.cfm?fa=main.doiLanding&uid=2006-09216-011

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron*, *33*(4), 653–663.

Croxson, P. L. (2005). Quantitative Investigation of Connections of the Prefrontal Cortex in the Human and Macaque using Probabilistic Diffusion Tractography. *Journal of Neuroscience*, *25*(39), 8854–8866. https://doi.org/10.1523/JNEUROSCI.1311-05.2005

Culver, N. C., Vervliet, B., & Craske, M. G. (2015). Compound extinction: using the Rescorla–Wagner model to maximize exposure therapy effects for anxiety disorders. *Clinical Psychological Science*, *3*(3), 335–348.

Dalgleish, T., Yiend, J., Schweizer, S., & Dunn, B. D. (2009). Ironic effects of emotion suppression when recounting distressing memories. *Emotion*, *9*(5), 744.

Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews*, *28*(7), 771–784. https://doi.org/10.1016/j.neubiorev.2004.09.006

Davis, M., & Walker, D. L. (2014). Role of bed nucleus of the stria terminalis and amygdala AMPA receptors in the development and expression of context conditioning and sensitization of startle by prior shock. *Brain Structure and Function*, *219*(6), 1969–1982. https://doi.org/10.1007/s00429-013-0616-5

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology*, *35*(1), 105.

Dennis, T. A. (2007). Interactions between emotion regulation strategies and affective style: Implications for trait anxiety versus depressed mood. *Motivation and Emotion*, *31*(3), 200–207.

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2004). Reinstatement of extinguished conditioned responses and negative stimulus valence as a pathway to return of fear in humans. *Learning & Memory*, *11*(5), 549–554.

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2007). Reinstatement of conditioned responses in human differential fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(3), 237–251.

Duvarci, S., & Pare, D. (2014). Amygdala Microcircuits Controlling Learned Fear. *Neuron*, *82*(5), 966–980. https://doi.org/10.1016/j.neuron.2014.04.042

Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2009). At the heart of the ventral attention system: The right anterior insula. *Human Brain Mapping*, *30*(8), 2530–2541. https://doi.org/10.1002/hbm.20688

Ellis, H. C., & Ashbrook, P. W. (1989). The" state" of mood and memory research: A selective review. *Journal of Social Behavior and Personality*, *4*(2), 1.

E-Prime 2.0 Professional version 2.0.10.353. (n.d.). Pittsburgh, PA: Psychology Software Tools.

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*(11), 693–700.

Etkin, A., & Wager, T. D. (2007). Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia. *American Journal of Psychiatry*, *164*(10), 1476–1488. https://doi.org/10.1176/appi.ajp.2007.07030504

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the Response Space"—An Account of Left Prefrontal Activation at Encoding. *NeuroImage*, *12*(4), 404–417. https://doi.org/10.1006/nimg.2000.0633

Ford, B. Q., & Troy, A. S. (2019). Reappraisal Reconsidered: A Closer Look at the Costs of an Acclaimed Emotion-Regulation Strategy. *Current Directions in Psychological Science*, 0963721419827526.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*(4), 1273–1302.

Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Àvila-Parcet, A., & Radua, J. (2016). Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry*, *21*(4), 500–508. https://doi.org/10.1038/mp.2015.88

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*(6), 577–586.

Goldman-Rakic, P. S. (1994). Working memory dysfunction in schizophrenia. *The Journal of Neuropsychiatry and Clinical Neurosciences*.

Gross, J. J. (1998a). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224.

Gross, J. J. (1998b). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(3), 271–299. http://dx.doi.org/10.1037/1089-2680.2.3.271

Gross, J. J. (2015a). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*(1), 1–26.

Gross, J. J. (2015b). The Extended Process Model of Emotion Regulation: Elaborations, Applications, and Future Directions. *Psychological Inquiry*, *26*(1), 130–137.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362. http://dx.doi.org/10.1037/0022-3514.85.2.348

Hendriks, S. M., Spijker, J., Licht, C. M., Hardeveld, F., de Graaf, R., Batelaan, N. M., … Beekman, A. T. (2015). Long-term work disability and absenteeism in anxiety and depressive disorders. *Journal of Affective Disorders*, *178*, 121–130.

Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biological Psychiatry*, *60*(4), 361–368.

Herry, C., Trifilieff, P., Micheau, J., Lüthi, A., & Mons, N. (2006). Extinction of auditory fear conditioning requires MAPK/ERK activation in the basolateral amygdala. *European Journal of Neuroscience*, *24*(1), 261–269.

Hurley, K. M., Herbert, H., Moga, M. M., & Saper, C. B. (1991). Efferent projections of the infralimbic cortex of the rat. *Journal of Comparative Neurology*, *308*(2), 249–276.

Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *The Journal of Neuroscience*, *27*(33), 8877–8884.

Kalisch, R., Wiech, K., Critchley, H. D., Seymour, B., O'Doherty, J. P., Oakley, D. A., … Dolan, R. J. (2005). Anxiety Reduction through Detachment: Subjective, Physiological, and Neural Effects. *Journal of Cognitive Neuroscience*, *17*(6), 874–883. https://doi.org/10.1162/0898929054021184

Kalisch, Raffael, Wiech, K., Herrmann, K., & Dolan, R. J. (2006). Neural correlates of self-distraction from anxiety and a process model of cognitive emotion regulation. *Journal of Cognitive Neuroscience*, *18*(8), 1266–1276.

Kehoe, E. J., & Macrae, M. (1997). Savings in animal learning: Implications for relapse and maintenance after therapy. *Behavior Therapy*, *28*(1), 141–155.

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States: Anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Kim, S.-Y., Adhikari, A., Lee, S. Y., Marshel, J. H., Kim, C. K., Mallory, C. S., … Deisseroth, K. (2013). Diverging neural pathways assemble a behavioural state from separable features in anxiety. *Nature*, *496*(7444), 219–223. https://doi.org/10.1038/nature12018

Kohn, N., Eickhoff, S. B., Scheller, M., Laird, A. R., Fox, P. T., & Habel, U. (2014). Neural network of cognitive emotion regulation—an ALE meta-analysis and MACM analysis. *Neuroimage*, *87*, 345–355.

Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: An introduction to growth curve modeling. *Psychophysiology*, *44*(5), 728–736.

Kross, E., & Ayduk, O. (2011). Making meaning out of negative experiences by self-distancing. *Current Directions in Psychological Science*, *20*(3), 187–191.

Langner, C. A., Epel, E. S., Matthews, K. A., Moskowitz, J. T., & Adler, N. E. (2012). Social hierarchy and depression: the role of emotion suppression. *The Journal of Psychology*, *146*(4), 417–436.

Lau, J. Y., Lissek, S., Nelson, E. E., Lee, Y., Roberson-Nay, R., Poeth, K., … Pine, D. S. (2008). Fear conditioning in adolescents with anxiety disorders: results from a novel experimental paradigm. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(1), 94–102.

Lee, T. G., Blumenfeld, R. S., & D'Esposito, M. (2013). Disruption of Dorsolateral But Not Ventrolateral Prefrontal Cortex Improves Unconscious Perceptual Memories. *Journal of Neuroscience*, *33*(32), 13233–13237. https://doi.org/10.1523/JNEUROSCI.5652-12.2013

Levey, A. B. (1980). Measurement units in psychophysiology. *Techniques in Psychophysiology*, 597–628.

Levy, B. J., & Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating: Cognitive control and right ventrolateral PFC. *Annals of the New York Academy of Sciences*, *1224*(1), 40–62. https://doi.org/10.1111/j.1749-6632.2011.05958.x

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science, 18*(5), 421–428.

Likhtik, E., Popa, D., Apergis-Schoute, J., Fidacaro, G. A., & Paré, D. (2008). Amygdala intercalated neurons are required for expression of fear extinction. *Nature*, *454*(7204), 642–645. https://doi.org/10.1038/nature07167

Lissek, S., Bradford, D. E., Alvarez, R. P., Burton, P., Espensen-Sturges, T., Reynolds, R. C., & Grillon, C. (2014). Neural substrates of classically conditioned fear-generalization in humans: a parametric fMRI study. *Social Cognitive and Affective Neuroscience*, *9*(8), 1134–1142. https://doi.org/10.1093/scan/nst096

Luck, C. C., & Lipp, O. V. (2015). A potential pathway to the relapse of fear? Conditioned negative stimulus evaluation (but not physiological responding) resists instructed extinction. *Behaviour Research and Therapy*, *66*, 18–31.

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276.

Madsen, H. B., Guerin, A. A., & Kim, J. H. (2017). Investigating the role of dopamine receptor- and parvalbumin-expressing cells in extinction of conditioned fear. *Neurobiology of Learning and Memory*, *145*, 7–17. https://doi.org/10.1016/j.nlm.2017.08.009

McRae, K., Ciesielski, B., & Gross, J. J. (2012). Unpacking cognitive reappraisal: goals, tactics, and outcomes. *Emotion*, *12*(2), 250.

Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10706–10711.

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Annual Review of Psychology*, *63*, 129–151.

Milad, M. R., Rosenbaum, B. L., & Simon, N. M. (2014). Neuroscience of fear extinction: implications for assessment and treatment of fear-based and anxiety related disorders. *Behaviour Research and Therapy*, *62*, 17–23.

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry*, *62*(5), 446–454.

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*(1), 10.

Moore, S. A., Zoellner, L. A., & Mollenholt, N. (2008). Are expressive suppression and cognitive reappraisal associated with stress-related symptoms? *Behaviour Research and Therapy*, *46*(9), 993–1000.

Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., & Tyler, L. K. (2005). Selecting Among Competing Alternatives: Selection and Retrieval in the Left Inferior Frontal Gyrus. *Cerebral Cortex*, *15*(11), 1723–1735. https://doi.org/10.1093/cercor/bhi049

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry*, *12*(2), 120–150.

Myers, K. M., Ressler, K. J., & Davis, M. (2006). Different mechanisms of fear extinction dependent on length of time since fear acquisition. *Learning & Memory*, *13*(2), 216–223.

Napier, R. M., Macrae, M., & Kehoe, E. J. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(2), 182.

Neumann, D. L., & Waters, A. M. (2006). The use of an unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biological Psychology*, *73*(2), 175–185.

Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research*, *218*(2), 325–334. https://doi.org/10.1016/j.bbr.2010.12.009

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking Feelings: An fMRI Study of the Cognitive Regulation of Emotion. *Journal of Cognitive Neuroscience*, *14*(8), 1215–1229. https://doi.org/10.1162/089892902760807212

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*(1), E1–E24.

Öhman, A., & Soares, J. J. (1994). " Unconscious anxiety": phobic responses to masked stimuli. *Journal of Abnormal Psychology*, *103*(2), 231.

Olson, E. A., & Luciana, M. (2008). The development of prefrontal cortex functions in adolescence: Theoretical models and a possible dissociation of dorsal versus ventral subregions. *Handbook of Developmental Cognitive Neuroscience*, 575–590.

Paré, D., Quirk, G. J., & Ledoux, J. E. (2004). New Vistas on Amygdala Networks in Conditioned Fear. *Journal of Neurophysiology*, *92*(1), 1–9. https://doi.org/10.1152/jn.00153.2004

Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological Psychiatry*, *60*(4), 383–387.

Pavlov, I. P. (1927). Conditioned reflexes. *An Investigation of the Physiological Activity of the Cerebral Cortex*.

Petrides, M., & Pandya, D. N. (2002a). Association pathways of the prefrontal cortex and functional observations. *Principles of Frontal Lobe Function*, *1*, 31–50.

Petrides, M., & Pandya, D. N. (2002b). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI. *NeuroImage*, *16*(2), 331–348. https://doi.org/10.1006/nimg.2002.1087

Psychology Software Tools, I. (2012). E-Prime 2.0. *Available*.

Quartana, P. J., & Burns, J. W. (2007). Painful consequences of anger suppression. *Emotion*, *7*(2), 400.

Quartana, P. J., & Burns, J. W. (2010). Emotion suppression affects cardiovascular responses to initial and subsequent laboratory stressors. *British Journal of Health Psychology*, *15*(3), 511–528.

Quartana, P. J., Yoon, K. L., & Burns, J. W. (2007). Anger suppression, ironic processes and pain. *Journal of Behavioral Medicine*, *30*(6), 455–469.

Quirk, G. J. (2002). Memory for Extinction of Conditioned Fear Is Long-lasting and Persists Following Spontaneous Recovery. *Learning & Memory*, *9*(6), 402–407. https://doi.org/10.1101/lm.49602

Quirk, Gregory J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *Journal of Neuroscience*, *20*(16), 6225–6231.

Rachman. (1979). The return of fear. *Behaviour Research and Therapy*, *17*(2), 164–166.

Rachman, S., Robinson, S., & Lopatka, C. (1987). Is incomplete fear-reduction followed by a return of fear? *Behaviour Research and Therapy*, *25*(1), 67–69.

Rachman, Stanley. (1989). The return of fear: Review and prospect. *Clinical Psychology Review*, *9*(2), 147–168.

Rescorla, R. A. (2004). Spontaneous recovery. *Learning & Memory*, *11*(5), 501–509.

Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*(1), 88.

Rescorla, R. A., Wagner, A. R., & others. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Resick, P. A. (2001). Cognitive therapy for posttraumatic stress disorder. *Journal of Cognitive Psychotherapy*, *15*(4), 321.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*(3), 410.

Roberts, N. A., Levenson, R. W., & Gross, J. J. (2008). Cardiovascular costs of emotion suppression cross ethnic lines. *International Journal of Psychophysiology*, *70*(1), 82–87.

Shackman, A. J., McMenamin, B. W., Maxwell, J. S., Greischar, L. L., & Davidson, R. J. (2009). Right dorsolateral prefrontal cortical activity and behavioral inhibition. *Psychological Science*, *20*(12), 1500–1506.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, *6*(1), 11–20.

Somerville, L. H., Whalen, P. J., & Kelley, W. M. (2010). Human Bed Nucleus of the Stria Terminalis Indexes Hypervigilant Threat Monitoring. *Biological Psychiatry*, *68*(5), 416–424. https://doi.org/10.1016/j.biopsych.2010.04.002

Srivastava, S., Tamir, M., McGonigal, K. M., John, O. P., & Gross, J. J. (2009). The social costs of emotional suppression: a prospective study of the transition to college. *Journal of Personality and Social Psychology*, *96*(4), 883.

Sullivan, G. M., Apergis, J., Bush, D. E. A., Johnson, L. R., Hou, M., & Ledoux, J. E. (2004). Lesions in the bed nucleus of the stria terminalis disrupt corticosterone and freezing

responses elicited by a contextual but not by a specific cue-conditioned fear stimulus. *Neuroscience*, *128*(1), 7–14. https://doi.org/10.1016/j.neuroscience.2004.06.015

Sun, M., & Craske, M. G. (in preparation). *The effects of emotion regulation on the spontaneous recovery of conditional fear and its rapid reacquisition*.

Sun, M., Vinograd, M., Miller, G. A., & Craske, M. G. (2017). Research Domain Criteria (RDoC) and Emotion Regulation. *Emotion Regulation and Psychopathology in Children and Adolescents*, 79.

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*(3), 201–216.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Ullman, J. B., & Bentler, P. M. (2003). *Structural equation modeling*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/0471264385.wei0224/full

Vertes, R. P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse*, *51*(1), 32–58. https://doi.org/10.1002/syn.10279

Vertes, R. P. (2006). Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neuroscience*, *142*(1), 1–20. https://doi.org/10.1016/j.neuroscience.2006.06.027

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: state of the art. *Annual Review of Clinical Psychology*, *9*, 215–248.

Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A., & Kording, K. P. (2012). Differential Representations of Prior and Likelihood Uncertainty in the Human Brain. *Current Biology*, *22*(18), 1641–1648. https://doi.org/10.1016/j.cub.2012.07.010

Waddell, J., Morris, R. W., & Bouton, M. E. (2006). Effects of bed nucleus of the stria terminalis lesions on conditioned anxiety: Aversive conditioning with long-duration conditional stimuli and reinstatement of extinguished fear. *Behavioral Neuroscience*, *120*(2), 324–336. https://doi.org/10.1037/0735-7044.120.2.324

Welborn, B. L., Papademetris, X., Reis, D. L., Rajeevan, N., Bloise, S. M., & Gray, J. R. (2009). Variation in orbitofrontal cortex volume: relation to sex, emotion regulation and affect. *Social Cognitive and Affective Neuroscience*, nsp028.

Wilhelm, F. H., & Peyck, P. (2005). *ANSLab: Autonomic nervous system laboratory (Version 2.5)*.

Yonkers, K. A., Bruce, S. E., Dyck, I. R., & Keller, M. B. (2003). Chronicity, relapse, and illness—course of panic disorder, social phobia, and generalized anxiety disorder: findings in men and women from 8 years of follow-up. *Depression and Anxiety*, *17*(3), 173–179.

Zbozinek, T. D., Hermans, D., Prenoveau, J. M., Liao, B., & Craske, M. G. (2015). Post-extinction conditional stimulus valence predicts reinstatement fear: Relevance for long-term outcomes of exposure therapy. *Cognition and Emotion*, *29*(4), 654–667. https://doi.org/10.1080/02699931.2014.930421

Zelikowsky, M., Bissiere, S., Hast, T. A., Bennett, R. Z., Abdipranoto, A., Vissel, B., & Fanselow, M. S. (2013). Prefrontal microcircuit underlies contextual learning after

hippocampal loss. *Proceedings of the National Academy of Sciences*, *110*(24), 9938–9943. https://doi.org/10.1073/pnas.1301691110

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206.