8-2016

Emotion Regulation During Threat: Parsing the Time Course and Consequences of Safety Signal Processing

Kathryn R. Hefner
Yale University

Edelyn Verona
University of South Florida, everona@usf.edu

John J. Curtin
University of Wisconsin - Madison

HHS Public Access

Author manuscript

Psychophysiology. Author manuscript; available in PMC 2017 June 28.

Published in final edited form as:

Psychophysiology. 2016 August ; 53(8): 1193–1202. doi:10.1111/psyp.12660.

Emotion regulation during threat: Parsing the time course and consequences of safety signal processing

KATHRYN R. HEFNER[a,b,c], **EDELYN VERONA**[d], and **JOHN. J. CURTIN**[c]

[a]Veterans Health Administration Mental Illness Research and Clinical Centers (MIRECC), West Haven, Connecticut, USA

[b]Department of Psychiatry, School of Medicine, Yale University, West Haven, Connecticut, USA

[c]Department of Psychology, University of Wisconsin-Madison, Madison, Wisconsin, USA

[d]Psychology Department, University of South Florida, Tampa, Florida, USA

Abstract

Improved understanding of fear inhibition processes can inform the etiology and treatment of anxiety disorders. Safety signals can reduce fear to threat, but precise mechanisms remain unclear. Safety signals may acquire attentional salience and affective properties (e.g., relief) independent of the threat; alternatively, safety signals may only hold affective value in the presence of simultaneous threat. To clarify such mechanisms, an experimental paradigm assessed independent processing of threat and safety cues. Participants viewed a series of red and green words from two semantic categories. Shocks were administered following red words (cue+). No shocks followed green words (cue−). Words from one category were defined as safety signals (SS); no shocks were administered on cue+ trials. Words from the other (control) category did not provide information regarding shock administration. Threat (cue+ vs. cue−) and safety (SS+ vs. SS−) were fully crossed. Startle response and ERPs were recorded. Startle response was increased during cue+ versus cue−. Safety signals reduced startle response during cue+, but had no effect on startle response during cue−. ERP analyses (PD130 and P3) suggested that participants parsed threat and safety signal information in parallel. Motivated attention was not associated with safety signals in the absence of threat. Overall, these results confirm that fear can be reduced by safety signals. Furthermore, safety signals do not appear to hold inherent hedonic salience independent of their effect during threat. Instead, safety signals appear to enable participants to engage in effective top-down emotion regulatory processes.

Descriptors

Safety signals; Anxiety; Emotion regulation; Fear; Psychophysiology

Address correspondence to: Kathryn R. Hefner, Ph.D., MIRECC, VA Connecticut Healthcare System, Yale University School of Medicine, Department of Psychiatry, 950 Campbell Avenue, West Haven, CT 06516, USA. kathryn.hefner@yale.edu.

Effective emotion regulation is critical to adaptive functioning in clinical and nonclinical populations alike. In particular, improved downregulation of fear and anxiety responding has direct clinical applications to psychiatric disorders characterized by anxiety (e.g., generalized anxiety disorder, simple phobia, agoraphobia, posttraumatic stress disorder [PTSD]). Emotional responses can be modulated through several methods involving dynamic cognition-emotion interactions, and this is an active area of study with regard to the cognitive and affective dysregulation characteristic of various psychological disorders (e.g., depression, Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007; phobias, Larson et al., 2006). This work suggests that, just as emotional states can impair or facilitate behavioral responses on cognitive tasks (e.g., Gray, 2001), cognitive processes impact affect. For example, attention and reappraisal have been found to modulate subjective ratings of negatively valenced pictures (Ochsner, Bunge, Gross, & Gabrieli, 2002), autonomic responses to stressors (Butler et al., 2003; Gross, 2002), and activation in affect-relevant subcortical regions such as the amygdala (Ochsner et al., 2004; Pessoa, 2005). Further, areas of the prefrontal cortex (PFC) known to play an important role in cognitive processing, including the dorsolateral and medial PFC and the orbitofrontal cortex, are recruited to modulate fear responding (Ochsner et al., 2002; Ochsner & Gross, 2005). Indeed, PTSD is believed to be characterized by medial PFC hypoactivity, as top-down inhibition of limbic system hyperactivity is impaired in this disorder (Maier, 2015).

In and outside of the laboratory, emotion regulation may be achieved through selective attention, distraction, or through other top-down processes such as reappraisal, in which one might use attentional control to reevaluate the personal relevance of a stimulus or to process external cues (Ochsner et al., 2004). Importantly, contextual information can provide signals for emotion regulation, such as when a dark alley seems less threatening in the presence of a nearby police car. Such contextual cues, which convey security despite the presence of other apparent threats, have been referred to as “safety signals” (e.g., Maier, 2015; Rachman, 1984; Sartory, Master, & Rachman, 1989). Studies assessing emotional responses to safety signals during simultaneous threat presentation have the potential to inform mechanisms and processes involved in successful emotion regulation (i.e., downregulation of fear/anxiety responses). A large literature indicates that safety signals broaden while threat cues constrict attention; however, little is known about the psychophysiological underpinnings of this process (for review, see Friedman & Förster, 2010).

Improved understanding of the mechanisms involved in the effective regulation of such negative affective responses holds considerable potential to inform etiology and treatment of anxiety disorders. When used successfully, safety signals can be useful for regulating negative emotional reactions and coordinating behavioral responses to aversive situations. Thus, safety signals have relevant clinical applications for psychiatric populations characterized by emotional dysregulation and related functional impairment. For example, the ability to effectively identify cues that predict danger from those that are benign is profoundly disrupted in PTSD, in which afflicted patients overgeneralize fear responses to banal stimuli (Jovanovic, Kazama, Bachevalier, & Davis, 2012).

Furthermore, elucidating mechanisms by which safety signals may aid in successful emotion regulation, and to what extent different populations are able to benefit from them, may help

to clarify ongoing debates within clinical psychology. In particular, longstanding research and clinical attitudes suggest that safety signals (e.g., a pill bottle) are detrimental to certain anxiety treatments (e.g., exposure therapy) for the very reason that they reduce fear and anxiety, thereby preventing full extinction of fear (Foa & Kozak, 1985). In contrast, more recent research and commentary suggest the potential utility of safety signals, used judiciously, to effectively downregulate fear and anxiety under threat in both clinical (e.g., Carter, Hollon, Carson, & Shelton, 1995; Goldin & Gross, 2010; McKay, 2010) and nonclinical (e.g., Coan, Schaefer, & Davidson, 2006) populations. In particular, safety signals established via processes other than experiential learning (e.g., through cognitive behavioral therapy) may have utility in clinical populations. Thus, improved understanding of impact of safety signals established by instruction on behavior, fear responses to threat, and attention would help to clarify their affective, attentional, and motivational properties and speak to their clinical utility.

Animal Models of Learned Safety

Although relatively understudied to date in humans, safety signals and their effects on fear and anxiety responses have been well documented in animal models. In rodent models, safety signals can be established through conditioned inhibition that develops across a series of learning trials (e.g., Gewirtz, Falls, & Davis, 1997). Conditioned inhibition is a learning process in which an animal is trained to fear a conditioned stimulus (CS) due to its repeated pairing with an aversive unconditioned stimulus (US, e.g., electric shock). The conditioned inhibitor (i.e., the safety signal) is then presented along with the CS without the US, and the animal learns that the conditioned inhibitor prevents the feared stimulus and/or indicates safety, which is indicated by less freezing behavior. Substantial basic research with rodents has demonstrated conditioned inhibitors effectively decrease fear responding (e.g., Gewirtz et al., 1997). For example, the presence of a conditioned inhibitor (that predicted no shocks even in the presence of a threat) reduced physiological reactions associated with fear (fear-potentiated startle and corticosterone release) in response to the threat in rats (Campeau et al., 1997). Thus, learning processes such as conditioned inhibition can contribute to safety signals' ability to downregulate or otherwise inhibit fear in animals. Recent research suggests safety signals are processed in the basolateral amygdala (BLA) in primates (Genud-Gabai, Klavir, & Paz, 2013), and the posterior BLA and bed nucleus of the stria terminalis in rodents (Christianson et al., 2011).

Some animal researchers have suggested that one means by which safety signals reduce fear is through acquisition of their own affective quality (i.e., reinforcement, relief) independent of the threat stimulus (Dinsmoor, 2001). Admittedly, disentangling the fear-inhibiting versus rewarding qualities of safety signals, and the question of whether relief is processed as rewarding, are controversial (see Christianson et al., 2011, for review of animal literature). Dinsmoor and Sears (1973) suggested that the presence of a safety signal had a positive reinforcing effect in pigeons (increased lever presses to the safety signal even in the absence of threat) that was distinct from the negative reinforcing effect of terminating a threat signal. Rats have also been shown to suppress lever pressing for food in the presence of a danger cue, but increase lever pressing when a safety cue is presented (Walasek, Wesierska, & Zieliński, 1995), suggesting potential positive affective associations. In mice, learned safety

signals can become positive reinforcers and exert anxiolytic qualities (Rogan, Leon, Perez, & Kandel, 2005). However, another recent rodent study indicated that safety signals did not confer reinforcing properties or relief (Fernando, Urcelay, Mar, Dickinson, & Robbins, 2013), leaving this question unanswered.

Safety Signals in Humans: More Unanswered Questions

Despite the relevance of the results from the few existing learned safety paradigms and conditioned inhibition paradigms in animals, safety signal processing in humans has been drastically understudied, leaving many lingering questions (Kong, Monje, Hirsch, & Pollak, 2014). An important issue that has not yet been adequately addressed in human or animal research concerns the processing of stimulus associations not involving experiential learning. For example, fear responses to cued threat of electric shock can be established via instruction (Curtin, Lang, Patrick, & Stritzke, 1998; Curtin, Patrick, Lang, Cacioppo, & Birbaumer, 2001). Therefore, it is plausible that safety signals established by instruction might also serve to reduce fear. Indeed, many safety signals that humans typically encounter in the real world (e.g., a policeman on the corner walking down an otherwise deserted, dark city street; guard rails on a high balcony or cliff edge, etc.) have been established by processes other than experiential learning. Furthermore, recent research in humans indicates that experiential versus instructed learning is associated with greater uncertainty regarding US occurrences in panic disorder patients (Lissek et al., 2009), suggesting that these patients may stand to uniquely benefit from instructed safety. Moreover, effective safety signals established via instruction might bear substantial similarity to existing cognitive therapy techniques involving top-down regulation of emotional responses, such as reappraisal (Brennan, Beck, & Servatius, 2003). Safety signals established via instruction, reappraisal, or observation may play an important role in emotion regulation through top-down attentional control, though this has not been tested empirically. Therefore, improved understanding of the functional significance of safety signals established via instruction could lead to the development of new clinical tools for anxiety patients, despite historical admonitions against the use of safety signals in anxious populations (Foa & Kozak, 1985).

Although some animal researchers have postulated reinforcing and/or positive affective qualities of safety signals (Dinsmoor, 2001), human research to date has yielded equivocal results (Falls, Bakken, & Heldt, 1997; Falls & Davis, 1995; Josselyn, Falls, Gewirtz, Pistell, & Davis, 2005). Grillon and Ameli (1998) found that the affective startle response was significantly potentiated during presentations of a threat cue (signaling shock) versus a safety signal (signaling no shock). In their study, startle was also significantly reduced in the safety signal versus no-safety signal segments or intertrial intervals (ITIs) between cues, which the authors interpreted as signifying that threat cues and safety signals elicited opposing affective responses relative to no signal: fear and a positive affective state of relief, respectively (Grillon & Ameli, 1998). However, as the authors acknowledged, the no-safety signal context was not neutral (contextual fear was present); thus, it is possible that the safety signal modulated contextual fear in this paradigm. Even had this been a completely neutral context, interpretation of the results remains problematic because the attentional demands of the no-safety signal versus safe conditions differed substantially, as perceptual or cognitive load has been shown to modify affective reactions (Pessoa, Kastner, &

Ungerleider, 2002; Sadeh & Verona, 2012). Thus, improved clarity regarding whether safety signals elicit positive affective responses in humans may be achieved through more precise methodological control, such as matching of stimulus properties and cognitive load.

Another important remaining question involves how safety signals are processed in the context of simultaneous threat; that is, whether safety signals are processed configurally (i.e., threat and safety information are processed as a compound stimulus) as opposed to elementally (i.e., threat and safety are processed as two separate stimuli that are synthesized to determine threat status), and to what extent this impacts emotional response. Elucidating attentional processing of safety signals could uncover whether safety and threat information are processed sequentially, which may allow top-down regulation of fear responses. Sequential processing of safety signal and threat information would bear similarity to emotion regulation strategies such as cognitive reappraisal where one can focus on reinterpreting the situation (e.g., “It is safe here despite the appearance of danger”; Brennan et al., 2003).

Research addressing this question in humans is rare. Grillon and Ameli (2001) found that, although a safety signal reduced startle responses to a cue signaling shock, the safety signal did not transfer its “safe” property to a new threat signal, suggesting that the safety signal was only effective in conjunction with the threat in which its safety was established. The authors interpreted this finding to indicate that participants used a configural approach in conditioning to a safety signal (i.e., the composite stimulus array of threat cue and safety signal was categorized as a unitary cue signaling safety rather than as individual stimuli; Grillon & Ameli, 2001). Indeed, others have suggested that, unlike animals, humans tend to perceive compound stimuli as a unique, single entity (configurally) rather than as an array of separate parts (i.e., elementally; Williams, Sagness, & McPhee, 1994). This hypothesis could be evaluated by assessing the time course of attention while safety status is determined (e.g., by measuring ERPs during cognitive processing of concurrent threat and safety information). Such data could inform the nature of early attentional processing of the two sources of information (i.e., whether threat and safety cues are processed configurally), and could also be used to evaluate their impact on later emotional and behavioral responses to threat. Thus, due to the critical role that emotional regulatory processes play in adaptive functioning for clinical and nonclinical populations alike, it is important to fully characterize the time course of attentional processing of, and resulting emotional responses to, safety signals both in the presence (to assess downregulation of fear and impact on behavior) and absence (to assess whether safety signals garner independent positive affective qualities) of threat.

The Current Study

To date, precise mechanisms for the operation of safety signals in the downregulation of fear and anxiety in humans have not been fully elucidated. Thus, our goals for the present study were:

1. To experimentally establish that safely signals established by instruction can impact emotional responses to otherwise threatening stimuli, and affect subsequent behavior.
2. To preliminarily test whether safety signals acquire positive affective properties that subsequently attenuate fear responses, as has been previously suggested (Christianson et al., 2011; Dinsmoor, 2001; Dinsmoor & Sears, 1973; Grillon & Ameli, 1998), or whether safety signals are affectively meaningful solely in the context of simultaneous threat.
3. To determine whether safety signals acquire motivational salience, garnering attention early in cognitive processing that is independent of the threat cue, allowing top-down regulation of fear responses. This possibility may depend on whether safety signals are processed configurally in combination with the threat cues (i.e., safety and threat cue information are combined such that the four conditions are reduced to simply threat and no-threat) or elementally (i.e., threat and safety signal information is processed independently; Grillon & Ameli, 2001).

We used a novel paradigm to evaluate the effects of safety signals established via instruction on fear responses to threat in humans. We also aimed to characterize the time course of the associated attentional processing of this information. In addition, we examined behavioral responses (reaction time to determine threat status). The paradigm presented words (animals or body parts). Each trial included the presence or absence of a threat cue (indicated by ink color of word) that was fully crossed with the presence or absence of a safety signal (semantic category of word), within the same stimulus. The use of a safety signal (semantic categorization) that required effortful cognitive processing in order to determine safety status was chosen due to our interest in emotion-cognition interactions. Such an approach more closely mimics emotion regulation strategies in complex real life situations, which are likely to unfold over time as different sources of information are synthesized to guide emotional response and behavior.

We measured affective response and attention to threat and safety signals with well-validated psychophysiological indices including fear-potentiated startle response and ERP, respectively. Although the startle response may be a more sensitive index of negative affect (e.g., fear-potentiated startle; Kaye, Bradford, & Curtin, in press), startle response inhibition can also be used to document possible positive affective response if safety signals acquire hedonic qualities. The ability to assess the time course of processing the two sources of information afforded by this design can speak to the question of whether threat and safety information are processed configurally (Grillon & Ameli, 2001; Williams et al., 1994) or, rather, elementally (Jovanovic et al., 2005). This was accomplished by examining differential ERP amplitude by condition to determine attentional effects of these complex cues. We selected the PD130 and P3 components to examine the processing stream of cue and safety signals because these two components are sensitive to early sensory processing and later top-down attentional processes related to stimulus categorization, respectively (Hillyard & Anllo-Vento, 1998; Kutas, McCarthy, & Donchin, 1977). Finally, the fact that the stimuli on each trial were matched in terms of their attentional properties allows more

precise investigation of whether safety signals take on positive affective qualities as suggested by Grillon and Ameli (1998).

Method

Participants

Thirty-six participants (18 female, 18 male) were recruited from the undergraduate psychology subject pool at the University of Wisconsin-Madison and the Madison, WI, community. Potential participants were screened to verify English reading and writing proficiency, and to determine any physical or psychological condition that would contraindicate study participation (e.g., uncorrected auditory or visual problems, medical condition that contraindicated electric shock administration). Participants were provided course extra credit (2 points/hour) or monetary compensation ($10/hour) for their participation in the experiment.

Instructed Cued Threat Task

The instructed cued fear task consisted of two blocks of 120 experimental trials (240 total trials) separated by a brief rest period. Each trial consisted of two stimuli (S1 and S2) with their presentation onsets separated by 2,250 ms (Figure 1). ITIs were 3 s. The S1 stimulus was one of two word categories, either an animal or body part word, presented in either red or green ink. A blue square was always used for S2. Participants were advised that an electric shock could occur 2 s after S1s written in red ink (cue+, shock administration would depend on word category/safety signal presence) and that no shocks would ever follow green S1s (cue−, regardless of word category). Cue+ and cue− trials were equiprobable (120 each).

Participants were also instructed that shocks would never be administered if the S1 was a safety signal word, regardless of the cue type (green or red ink). Safety signals (e.g., animal word for S1) were presented on 20% of trials (SS+), with the remaining 80% of S1s from the no-safety signal (SS−) category (e.g., body part words). The use of word categories as the safety signal ensured that safety signal cue detection entailed higher-order, semantic processing. Safety signal status was fully crossed with cue type such that, across 240 trials, there were 96 cue−/SS− trials, 96 cue+/SS−, 24 cue−/SS+ trials, and 24 cue+/SS+ trials.

Electric shocks were actually administered on 25% of the cue+/SS− trials (24 of 96 trials). Consistent with participant instructions, no electric shocks were administered on trials from the other three conditions (i.e., cue+/SS+ and both cue− conditions). To minimize individual differences, we measured participants' subjective shock tolerance threshold 15 min prior to the start of the instructed cued threat task following standardized procedures in our laboratory (Bradford, Shapiro, & Curtin, 2013; Curtin et al., 2001; Hefner & Curtin, 2012; Hefner, Moberg, Hachiya, & Curtin, 2013). Participants reported their response to a series of 200-ms electric shocks of increasing intensity. Shocks were administered to the distal phalanges of the index and ring fingers of the left or right hand (counterbalanced across participants). The procedure required approximately 5–10 min to complete and was stopped once participants reached the maximum level of shock that they could tolerate. In the

instructed cued threat task, we administered 200-ms shocks that were set halfway between participants' self-reported "uncomfortable" and "maximum tolerable" shock levels.

To ensure adequate fear responding, participants were instructed to attend foremost to the ink color of the S1 word and to press one of two reaction time buttons held in separate hands immediately at S2 onset (2,250 ms following S1) to indicate whether the S1 was from the cue+ or cue− category. Cue+ response was mapped to the same hand on which shock electrodes were attached to further reinforce processing of the cue+/shock contingency.

Participants were assigned to one of two pseudorandom stimulus orders. Each stimulus order was constrained such that no more than four cue+ or cue− trials were presented in a row, at least two trials separated each presentation of an SS+, and no block began with an SS+. Stimulus orders were fully crossed with shock administration hand to form four between-subjects counterbalanced task orders.

Measures

Startle response—Fifty-one startle-eliciting noise probes (50 ms, 102 dB white noise burst with instantaneous rise time) were presented at 2 s post-S1 onset. This was done to assess fear response to the S1 (12 probes in each of the four Cue Type × Safety Signal conditions; three additional probes were presented prior to the start of the task to habituate large responses that are typical in early probe presentations). Probes were never included on trials involving shock administration. Neuroscan SynAmps bioamplifiers (Compumedics Neuroscan, Charlotte, NC) sampled (2000 Hz) startle blink electromyographic response to these probes using a band-pass filter (.05–500 Hz) from miniature Ag-AgCl sensors filled with conductive gel and placed according to published guidelines (Bradford, Magruder, Korhumel, & Curtin, 2014; van Boxtel, Boelhouwer, & Bos, 1998). Offline processing was accomplished in MATLAB (Mathworks Inc., Natick, MA) using the EEGLAB toolbox (Delorme & Makeig, 2004) with the PhysBox plugin (Curtin, 2011) following published guidelines (Blumenthal et al., 2005). Processing included high-pass filtering (28 Hz, fourth-order Butterworth high-pass filter, zero phase shift), signal rectification and smoothing (30 Hz, fourth-order Butterworth low-pass filter, zero phase shift), epoching (−50–250 ms relative to probe onset), and baseline correction. We rejected trials with excessive deflections (values >± 20 μV) between −50–10 ms relative to probe onset as artifact due to unstable baseline (1.6% of trials). We rejected trials with mean amplitude less than −10 μV between 100–250 ms postprobe as artifact due to baseline overcorrection (0.1% of trials). Two participants were identified as nonresponders and were excluded from startle response analyses. Peak startle response between 20–100 ms postprobe onset was scored relative to preprobe baseline.

ERPs—Neuroscan SynAmps bioamplifiers sampled (2000 Hz) EEG activity using a band-pass filter (0.5–500 Hz) from four midline scalp sites (Fz, Fcz, Cz, Pz) referenced to linked mastoids and filled with conductive gel in Electro-Caps (Electro-Cap International, Eaton, OH). Vertical electrooculogram activity was also measured to correct for eyeblink artifact. Offline processing was accomplished in MATLAB using EEGLAB and PhysBox. Processing included low-pass filtering (30 Hz, second-order Butterworth low-pass filter, zero

phase shift), eyeblink artifact correction via regression (e.g., Semlitsch, Anderer, Schuster, & Presslich, 1986), signal epoching (−200–800 ms relative to S1 onset), and baseline correction. Trials with excessive deflections (values >± 100 μV) at any point in the epoch were rejected as artifact (4.1% of trials). Four participants were removed from ERP analyses due to excessive noise that prevented use of eyeblink artifact correction.

ERPs were scored following standard guidelines (Picton et al., 2000). We focused on two ERP components, the PD130 and the P3, each of which is maximal parietally. The parietal PD130 is known to have latencies of 100–150 ms and is involved in visual processing of nonspatial features such as color (Hillyard & Anllo-Vento, 1998). The parietal P3 tends to have latencies of approximately 280–780 ms and is considered to be an index of stimulus categorization integrity and speed (Kutas et al., 1977). We identified the peak response latency and width of each component using the grand-averaged waveform, collapsed across conditions. Based on this grand-averaged waveform, we scored the PD130 as mean response between 127–152 ms and the P3 as mean response between 382–432 ms.

Response time—As indicated above, participants were instructed to make a speeded button press response on S2 onset to indicate the cue type (cue+ or cue−) of the S1. Response time was recorded in milliseconds. Trials involving incorrect (2.5% of trials) or no responses (1.6% of trials) were excluded from analyses. Trials involving electric shock administration were also excluded, as shock administration may have interfered with participants' behavioral response.

Open Science Practices

We support emerging open science guidelines (Nosek et al., 2015). Following these guidelines, we have made the data and analysis scripts associated with this report publicly available via Open Science Framework. These materials can be accessed at https://osf.io/hsjxd/

Results

General Analytic Strategy

Each dependent measure (startle response, response time, PD130, P3) is analyzed in separate general linear model with repeated measures for cue type (cue+ vs. cue−) and safety signal (SS+ vs. SS−) and between-subjects regressors for task order. If a significant Cue Type × Safety Signal interaction was observed, we tested simple effects of safety signal separately for cue+ and cue− trials. If no significant interaction was detected, we report tests of both main effects. We report both raw parameter estimates (*B*s) and partial eta-squared (η_p^2) to document effect sizes. We also report 95% confidence intervals (CIs) for the parameter estimates.

Startle Response

A significant Cue Type × Safety Signal interaction was observed, $B = -31.1$, 95% CI(B) [−43.2, −19.1], $\eta_p^2 = .48$, $t(30) = 5.26$, $p < .001$, indicating that startle potentiation (i.e., increased startle on cue+ relative to cue− trials) was significantly reduced for trials involving

safety signals (see Figure 2). Follow-up simple effects tests indicated that startle magnitude was significantly reduced on SS+ versus SS− during cue+ trials, B =−29.0, 95% CI(B) [−39.3, −18.6], η_p^2=.52, t(30) =5.70, p <.001, providing support that the safety signals decreased fear response to the cue+. In contrast, during cue− trials, there was no significant difference in startle magnitude for SS+ versus SS− trials, B =2.2, 95% CI(B) [−2.0, 6.4], η_p^2=.04, t(30) =1.06, p =.296, suggesting that safety signals did not independently inhibit the startle response when presented during an otherwise neutral (i.e., cue−) trial.

Response Time

As with startle magnitude, a significant Cue Type × Safety Signal interaction was observed, B =−41.5, 95% CI(B) [−57.1, −25.8], η_p^2=.48, t(32) =5.40, p <.001, indicating that the magnitude of response time slowing on cue+ versus cue− trials was significantly reduced for trials involving safety signals (see Figure 3). Follow-up simple effects tests indicated that response time was significantly faster on SS+ versus SS− trials during cue+ trials, B =−36.4, 95% CI(B) [−46.7, −26.0], η_p^2=.61, t(32) =7.15, p <.001, providing support that the safety signals were used successfully to reduce the behavioral interference produced by the cue+. In contrast, during cue− trials, there was no significant difference in response times for SS+ versus SS− trials, B =5.1, 95% CI(B) [−3.7, 13.9], η_p^2=.04, t(32) =1.18, p =.246, suggesting that safety signals did not independently affect behavior when presented during an otherwise neutral (i.e., cue−) trial.

ERPs

We display the grand-averaged parietal ERP waveform from which we quantified the PD130 and P3 in Figure 4. Analyses for these two components follow.

PD130—In contrast to the startle response and response time, the Cue Type × Safety Signal interaction was not significant, B =0.6, 95% CI(B) [−0.5, 1.8], η_p^2=.04, t(28) =1.11, p =.277. However, the main effect of cue type was significant, with increased PD130 on cue+ (M =5.02, SD =3.33) relative to cue− trials (M =4.03, SD =2.91), B =1.0, 95% CI(B) [0.3, 1.7], η_p^2=.21, t(28) =2.75, p =.010 (Figure 5), indicating that the threat cues modulated early visual attention. The main effect of safety signal condition was not significant, B =−0.3, 95% CI(B) [−1.0, 0.5], η_p^2=.02, t(28) =0.73, p =.474. These results suggest that, at early stages of visual processing, threat cues capture visual attention, regardless of safety signal information. The lack of a significant main effect or interaction involving safety signal indicates that the safety signal information has not yet been incorporated at this early processing stage.

P3—As with PD130, the Cue Type × Safety Signal interaction was not significant, B =0.9, 95% CI(B) [−0.8, 2.7], η_p^2=.04, t(28) =1.09, p =.285. However, the main effect of cue type was significant, with increased P3 on cue+ (M =10.52, SD =6.55) relative to cue− trials (M =8.36, SD =5.15), B =2.2, 95% CI(B) [0.9, 3.5], η_p^2=.29, t(28) =3.40, p =.002 (Figure 6). In addition, the main effect of safety signal condition was significant, with increased P3 on SS−

trials (M =9.90, SD =5.33) relative to SS+ trials (M =8.98, SD =6.18), B =0.9, 95% CI(B) [0.0, 1.8], η_p^2=.14, t(28) =2.11, p =.044. Thus, threat cues appear to capture attention both early (PD130) and later (P3) in the processing stream. In contrast, safety signal information does not appear to modulate attention until later stages of processing (P3), where the presence of a safety signal allows participants to disengage attention.

Discussion

In the present study, we examined the impact of safety cues established by verbal instruction on behavior and in downregulation of fear responses to threat of electric shock. In addition, we assessed attentional processing of cues and safety signals over time to determine when and how safety signals are used to reduce fear responses.

The startle response results indicated that fear responding was increased only on trials where a true threat was present (i.e., cue+/SS−). This suggests that, while safety signals inhibit fear (as measured by startle response) when a threat is present, they do not appear to have hedonic value when no threat exists (e.g., during cue− trials).

The response time results mirrored the startle response results, suggesting that safety signals reduced participants' performance deficit (increased time to respond to the S2). Threat of electric shock caused a performance deficit in responding to word color only when a true threat was present (i.e., cue+ and SS−). In the absence of threat (i.e., cue−), safety signals had no impact on response time, indicating that safety signals did not impact behavior in a nonthreatening context. However, it should be noted that response time measurement for the cue+/SS− trials was limited to 72 (rather than 96) trials because we excluded trials where shocks were actually administered in this condition to avoid noise due to impaired responding following actual shock administration. Seventy-two trials would be expected to be sufficient to provide reliable measurement of response time in this condition. However, it is possible that the response time slowing on these trials was not due to threat but instead to relief at not receiving an expected electric shock. Nonetheless, we believe that the comparable pattern of means across both startle response (which did not suffer from this limitation because it was measured prior to shock administration and included all trials) and response time suggests that the response time slowing resulted from threat, which was reduced by safety signals on cue+/SS+ trials.

The ERP results indicate that, very early in processing (PD130), threatening stimuli attract more attention than nonthreating stimuli, regardless of safety signal presence. This indicates that the two sources of information—salient threat cue (word color) and a more complex safety signal (semantic category)—have not yet been synthesized by PD130. This early in processing, safety signals do not impact attention. At approximately 400 ms, however, P3 results indicate that, while potential threats (cue+) still garner attention, the presence of safety signals allowed participants to disengage attention at this stage of processing. In other words, threat and safety information were disambiguated by this time point, and this guided participants' attention. Taken together, these ERP results suggest that threat and safety information were processed elementally (i.e., independently) as opposed to configurally (both cue and safety signal information combined to form a single entity; Jovanovic et al.,

2005; Williams et al., 1994). Had threat and safety information been processed configurally, one might approach the task as having only two types of stimuli, one that threatened shock (cue+/SS−) versus the other three. Near the end of the trial, immediately before shock administration, this was the pattern observed for affective response as indicated by the Cue Type × Safety Signal interactions for startle potentiation and behavioral interface (response time slowing). For these indices at this late point, participants manifested strong negative affect to the cue+/SS− trials but little response to the other three conditions. However, analysis of ERPs earlier during the trial suggested that participants first processed threat information as indicated by the PD130 and then later (P3 at 400 ms) displayed attentional modulation independently based on both threat and safety signals.

Together, these data suggest that safety signals established by instruction can be used to effectively downregulate fear responses to threatening stimuli, as evidenced by startle response results. This finding is further bolstered by behavioral data, as safety signals improved performance during trials on which a true threat existed, but not in the absence of threat. Because startle response and response time were unaffected by safety signal presence during cue− trials, safety signals do not appear to hold hedonic value in the absence of threat.

In addition, ERP data also suggest that, early in the processing stream (i.e., by 400 ms), participants used safety signals to adaptively disengage their attention in an otherwise ambiguously threatening environment. Together, our results suggest that, at least in an instructed fear paradigm, the fear-inhibition properties of safety signals may result from attentional disengagement rather than from positive hedonic qualities attributed to the safety signals per se. However the later affective consequences of this early attentional disengagement by safety signals (i.e., the inhibition of fear documented with startle response and response time) manifest only on cue+ trials because the cue− trials do not elicit any strong affect to start.

Thus, safety signal presence appears to be important only in the context of threatening cues; safety signals are not affectively charged, except when they provide information about an affective (threatening) stimulus that is present simultaneously. This stands in contrast to the hypothesis that safety signals confer positive rewarding properties (Dinsmoor, 2001; Grillon & Ameli, 1998); safety signals did not appear to alter participants' affective response independent of their modulatory effects during threat. Our findings corroborate animal work suggesting that safety signals do not acquire affective qualities in a neutral condition (Josselyn et al., 2005). However, it does not rule out the possibility that safety signals could acquire qualities independent of the specific context in which they are acquired and could be generalized to a different threatening context/stimulus, thereby attenuating an unconditioned response (Pollak et al., 2008). Moreover, a recent review suggests that safety signals do not provide such immunization; while they reduce the impact of current stressors, they do not moderate the impact of future stressors (Christianson & Greenwood, 2014).

Implications for Psychopathology and Treatment

The degree to which individuals utilize top-down modulation of emotion has potential significance for the development, maintenance, and treatment of psychopathological

disorders characterized by emotional dysregulation. Our findings highlight the potential utility of treatments involving safety signals as clinical tools for patients with otherwise poor emotion regulation (e.g., anxiety disorders, depression, PTSD).

Although clinical lore has suggested that extinction in exposure therapy may be deterred if the client avoids experiencing the full fear response during exposure (e.g., by using a safety signal such as a pill bottle or rabbit's foot; Foa & Kozak, 1985), more recent commentary (McKay, 2010) focuses on the inherent contradictions in current conceptualization of safety signals and safety-seeking behaviors as unilaterally undermining such therapies, particularly given the natural inclination of humans and animals to seek safety (e.g., Woody & Rachman, 1994). In Rachman's (1984) safety signal perspective of agoraphobia, patients are actively encouraged to seek out signs of safety and security or to travel through danger toward a safety person, while avoidance behaviors are actively discouraged. Indeed, judicious use of safety signals (such as a safe person or therapist assistance) in agoraphobics can lead to better clinical gains such as decreased depression and increased mobility (Sartory et al., 1989). Similarly, among patients with panic disorder with agoraphobia exposed to CO_2 inhalation, the mere presence of a safe person decreased both subjective experiences of anxiety as well as related physiological arousal (Carter et al., 1995). Among healthy individuals, spousal hand-holding (but not stranger handholding) effectively reduced subjective unpleasantness and arousal, as well as neural activation to threat of electric shock (Coan et al., 2006). Additionally, learned safety in rodents decreases not only fear responding, but also depression-like behavior, sharing neurobiological hallmarks of pharmacological antidepressants (i.e., increased expression of brain-derived neurotrophic factor [BDNF]; Pollak et al., 2008). Taken together, this work suggests clinical safety signals could take the form of a safe person (e.g., Carter et al., 1995; Coan et al., 2006; Sartory et al., 1989), or anything else that provides information about alternative behavior, coping, and potentially positive outcomes (Lohr, Olatunji, & Sawchuk, 2007). The results of the present study suggest that, in these otherwise threatening situations, individuals process threat and safety information in parallel, synthesizing them over time to effectively downregulate negative emotions.

Furthermore, existing empirically validated therapies already benefit from the use of what are essentially safety signals. It has been suggested that the "active ingredient" of safety signals is to instill a sense of controllability to existing stressors (Quirk & Beer, 2006), and recent research indicates that individuals high in trait anxiety have an impoverished ability to learn about action-outcome associations in unpredictable and/or uncontrollable environments, impairing judgments about outcome likelihood and contributing to poor decision making (Browning, Behrens, Jocham, O'Reilly, & Bishop, 2015). Thus, treatments could also be developed to improve prediction of external stressors and cues that predict them, as well as to enhance control over internal cues that have essentially become danger signals (Lohr et al., 2007). Predictability and perceived control could be increased through a combination of effortful downregulation of psychophysiological cues (danger signals), and recordkeeping to identify precipitating events for panic attacks and to monitor panic symptoms. Given the well-documented role of unpredictability in contributing to anxiety and the psychophysiological response to unpredictable threat (Grupe & Nitschke, 2013; Herry et al., 2007; Sarinopoulos et al., 2010), and the potential utility of safety signals to

adaptively shift attention to safety to aid emotion regulation, it appears that broadening the field's current conceptualization of safety signals would be fruitful.

Moreover, cognitive therapy techniques involving coping skills training to reduce the negativity of one's emotions to establish emotion regulatory control may function similarly to safety signals (Maier, 2015). Indeed, it has been argued that cognitive techniques recruit the PFC to inhibit maladaptive emotional responses in a top-down fashion (DeRubeis, Siegle, & Hollon, 2008). Furthermore, grounding techniques encourage patients with PTSD to attend to neutral tangible objects in the present environment, and to acknowledge their present safety to effectively reduce anxiety and fear (Najavits, 2001). This is not dissimilar from the focusing of attention to the present moment in mindfulness practices, which has also been shown to aid emotional regulatory processes in dysregulated patients (Goldin & Gross, 2010; Vøllestad, Sivertsen, & Nielsen, 2011). Cognitive therapy, grounding, and mindfulness may be effective because they allow patients to effortfully downregulate negative affective responses by detracting attention from otherwise subjectively threatening stimuli (e.g., intrusive thoughts or worries that are future oriented, involving catastrophic outcomes with objectively low probability of occurring), instead focusing on neutral stimuli; over time, this effortful downregulation becomes a more automatic part of the individuals' behavioral repertoire, or a skill.

Limitations and Future Directions

Future research in this area should examine the effects of safety signals paired with aversive stimuli established via true conditioning (e.g., learning), the time course of affective response (via assessing startle response at earlier time points), and clinically relevant individual differences. Our results indicate that, in healthy individuals, safety signals established via instruction can effectively be used to downregulate fear responses; however, it may be the case that individuals with heightened or pathological anxiety are not able to use such cues as effectively. In fact, a recent study in adolescents at risk for developing emotional disorders indicated that startle responses during safe conditions predicted the development of anxiety disorders over the next 4 years (Craske et al., 2012). Tasks such as the one used in the present experiment have the potential to inform such individual differences in fear potentiation versus inhibition of fear responding within clinical populations. For example, PTSD patients may exhibit normal fear potentiation, but impaired ability to impair fear responding (Grillon & Morgan, 1999). Recent research has suggested that this may be related to impaired ability to learn safety signals (Jovanovic et al., 2012), and it is possible that other anxiety disorders share this dysregulation. As generalized anxiety has been conceptualized as an "unsuccessful search for safety" (Woody & Rachman, 1994) that is characterized by hypervigilance to threat, future research should assess the ability of these patients to utilize safety signals to down-regulate fear and anxiety. Furthermore, paradigms similar to the one used in the present experiment could be used to assess the impact of pharmacological treatments on processes of fear potentiation and fear inhibition, respectively. Such work has the potential to shed light on prevention and treatment of PTSD and other anxiety disorders (Christianson et al., 2011).

As aforementioned, our results indicate that safety signals do not possess affective qualities outside of their fear-diminishing effect during threatening conditions. However, our conclusions regarding the affective properties of safety signals rely primarily on the startle response-dependent measure. Future research could use a broader array of alternative indices of affective response, and conclusions regarding hedonic properties of safety signals should be drawn from this study with some discretion.

Finally, is also possible that how safety signals established first via instruction operated in this study is actually quite different than how safety signals function in true conditioned inhibition, learned safety, and/or inhibitory fear-learning paradigms. Although available neuroscientific research in humans suggests that verbally instructed fear paradigms induce both amygdala activation and physiological responses to threat comparably to acquired fear paradigms in animals (e.g., Phelps et al., 2001), it is possible that processing of instructed safety signals may operate differently. It should also be explored whether safety signals established via instruction or through therapeutic techniques such as cognitive behavioral therapy, grounding, mindfulness, and other techniques can engender the same antidepressant effects as do those established by experiential learning processes in humans and animals (e.g., Pollak et al., 2008). If so, this could be a very powerful clinical tool for numerous psychiatric disorders characterized by emotional dysregulation.

Acknowledgments

Writing of this manuscript was partially supported by the Office of Academic Affiliations, Advanced Fellowship Program in Mental Illness Research and Treatment, Department of Veterans Affairs. Blaise Worden provided technical assistance on this study.

References

Blumenthal TD, Cuthbert BN, Filion DL, Hackley S, Lipp OV, van Boxtel A. Committee report: Guidelines for human startle eyeblink electromyographic studies. Psychophysiology. 2005; 42(1):1–15. DOI: 10.1111/j.1469-8986.2005.00271.x [PubMed: 15720576]

Bradford DE, Magruder KP, Korhumel RA, Curtin JJ. Using the threat probability task to assess anxiety and fear during uncertain and certain threat. Journal of Visualized Experiments: JoVE. 2014; 91:51905.doi: 10.3791/51905

Bradford DE, Shapiro BL, Curtin JJ. How bad could it be? Alcohol dampens stress responses to threat of uncertain intensity. Psychological Science. 2013; 24(12):2541–2549. DOI: 10.1177/0956797613499923 [PubMed: 24145332]

Brennan FX, Beck KD, Servatius RJ. Leverpress escape/avoidance conditioning in rats: Safety signal length and avoidance performance. Integrative Physiological and Behavioral Science. 2003; 38(1): 36–44. [PubMed: 12814195]

Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ. Anxious individuals have difficulty learning the causal statistics of aversive environments. Nature Neuroscience. 2015; 18(4):590–596. DOI: 10.1038/nn.3961 [PubMed: 25730669]

Butler EA, Egloff B, Wilhelm FH, Smith NC, Erickson EA, Gross JJ. The social consequences of expressive suppression. Emotion. 2003; 3(1):48–67. [PubMed: 12899316]

Campeau S, Falls WA, Cullinan WE, Helmreich DL, Davis M, Watson SJ. Elicitation and reduction of fear: Behavioural and neuroendocrine indices and brain induction of the immediate-early gene c-fos. Neuroscience. 1997; 78(4):1087–1104. [PubMed: 9174076]

Carter MM, Hollon SD, Carson R, Shelton RC. Effects of a safe person on induced distress following a biological challenge in panic disorder with agoraphobia. Journal of Abnormal Psychology. 1995; 104(1):156–163. [PubMed: 7897039]

Christianson JP, Greenwood BN. Stress-protective neural circuits: Not all roads lead through the prefrontal cortex. Stress. 2014; 17(1):1–12. DOI: 10.3109/10253890.2013.794450 [PubMed: 23574145]

Christianson JP, Jennings JH, Ragole T, Flyer JGN, Benison AM, Barth DS, … Maier SF. Safety signals mitigate the consequences of uncontrollable stress via a circuit involving the sensory insular cortex and bed nucleus of the stria terminalis. Biological Psychiatry. 2011; 70(5):458–464. DOI: 10.1016/j.biopsych.2011.04.004 [PubMed: 21684526]

Coan JA, Schaefer HS, Davidson RJ. Lending a hand: Social regulation of the neural response to threat. Psychological Science. 2006; 17(12):1032–1039. DOI: 10.1111/j.1467-9280.2006.01832.x [PubMed: 17201784]

Craske MG, Wolitzky-Taylor KB, Mineka S, Zinbarg R, Waters AM, Vrshek-Schallhorn S, … Ornitz E. Elevated responding to safe conditions as a specific risk factor for anxiety versus depressive disorders: Evidence from a longitudinal investigation. Journal of Abnormal Psychology. 2012; 121(2):315–324. DOI: 10.1037/a0025738 [PubMed: 21988452]

Curtin, JJ. PhysBox: The Psychophysiology toolbox. An open source toolbox for psychophysiological data reduction within EEGLab. 2011. Retrieved from http://dionysus.psych.wisc.edu/PhysBox.htm

Curtin JJ, Lang AR, Patrick CJ, Stritzke WGK. Alcohol and fear-potentiated startle: The role of competing cognitive demands in the stress-reducing effects of intoxication. Journal of Abnormal Psychology. 1998; 107(4):547–565. [PubMed: 9830242]

Curtin JJ, Patrick CJ, Lang AR, Cacioppo JT, Birbaumer N. Alcohol affects emotion through cognition. Psychological Science. 2001; 12(6):527–531. DOI: 10.1111/1467-9280.00397 [PubMed: 11760143]

Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods. 2004; 134(1):9–21. DOI: 10.1016/j.jneumeth.2003.10.009 [PubMed: 15102499]

DeRubeis RJ, Siegle GJ, Hollon SD. Cognitive therapy versus medication for depression: Treatment outcomes and neural mechanisms. Nature Reviews Neuroscience. 2008; 9(10):788–796. DOI: 10.1038/nrn2345 [PubMed: 18784657]

Dinsmoor JA. Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. Journal of the Experimental Analysis of Behavior. 2001; 75(3):311–333. DOI: 10.1901/jeab.2001.75-311 [PubMed: 11453621]

Dinsmoor JA, Sears GW. Control of avoidance by a response-produced stimulus. Learning and Motivation. 1973; 4(3):284–293. DOI: 10.1016/0023-9690(73)90018-0

Falls WA, Bakken KT, Heldt SA. Lesions of the perirhinal cortex interfere with conditioned excitation but not with conditioned inhibition of fear. Behavioral Neuroscience. 1997; 111(3):476–486. [PubMed: 9189262]

Falls, WA., Davis, M. Behavioral and physiological analysis of fear inhibition: Extinction and conditioned inhibition. In: Friedman, MJ., Charney, DS., editors. Neurobiological and clinical consequences of stress: From normal adaptation to post-traumatic stress disorder. Philadelphia, PA: Lippincott Williams & Wilkins Publishers; 1995. p. 177-202.

Fernando ABP, Urcelay GP, Mar AC, Dickinson A, Robbins TW. Comparison of the conditioned reinforcing properties of a safety signal and appetitive stimulus: Effects of D-amphetamine and anxiolytics. Psychopharmacology. 2013; 227(2):195–208. DOI: 10.1007/s00213-012-2952-1 [PubMed: 23299096]

Foa, EB., Kozak, MJ. Treatment of anxiety disorders: Implications for psychopathology. In: Tuma, AH., Maser, JD., editors. Anxiety and the anxiety disorders. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1985. p. 421-452.

Friedman RS, Förster J. Implicit affective cues and attentional tuning: An integrative review. Psychological Bulletin. 2010; 136(5):875–893. DOI: 10.1037/a0020495 [PubMed: 20804240]

Genud-Gabai R, Klavir O, Paz R. Safety signals in the primate amygdala. Journal of Neuroscience. 2013; 33(46):17986–17994. DOI: 10.1523/JNEUROSCI.1539-13.2013 [PubMed: 24227710]

Gewirtz JC, Falls WA, Davis M. Normal conditioned inhibition and extinction of freezing and fear-potentiated startle following electrolytic lesions of medical prefrontal cortex in rats. Behavioral Neuroscience. 1997; 111(4):712–726. [PubMed: 9267649]

Goldin PR, Gross JJ. Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. Emotion. 2010; 10(1):83–91. DOI: 10.1037/a0018441 [PubMed: 20141305]

Gray JR. Emotional modulation of cognitive control: Approach-withdrawal states double-dissociate spatial from verbal two-back task performance. Journal of Experimental Psychology General. 2001; 130(3):436–452. [PubMed: 11561919]

Grillon C, Ameli R. Effects of threat and safety signals on startle during anticipation of aversive shocks, sounds, or airblasts. Journal of Psychophysiology. 1998; 12(4):329–337.

Grillon C, Ameli R. Conditioned inhibition of fear-potentiated startle and skin conductance in humans. Psychophysiology. 2001; 38(5):807–815. [PubMed: 11577904]

Grillon C, Morgan CA. Fear-potentiated startle conditioning to explicit and contextual cues in Gulf War veterans with posttraumatic stress disorder. Journal of Abnormal Psychology. 1999; 108(1): 134–142. [PubMed: 10066999]

Gross JJ. Emotion regulation: Affective, cognitive, and social consequences. Psychophysiology. 2002; 39(3):281–291. [PubMed: 12212647]

Grupe DW, Nitschke JB. Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. Nature Reviews Neuroscience. 2013; 14(7):488–501. DOI: 10.1038/nrn3524 [PubMed: 23783199]

Hefner KR, Curtin JJ. Alcohol stress response dampening: Selective reduction of anxiety in the face of uncertain threat. Journal of Psychopharmacology. 2012; 26(2):232–244. DOI: 10.1177/0269881111416691 [PubMed: 21937686]

Hefner KR, Moberg CA, Hachiya LY, Curtin JJ. Alcohol stress response dampening during imminent versus distal, uncertain threat. Journal of Abnormal Psychology. 2013; 122(3):756–769. DOI: 10.1037/a0033407 [PubMed: 24016014]

Herry C, Bach DR, Esposito F, Salle FD, Perrig WJ, Scheffler K, … Seifritz E. Processing of temporal unpredictability in human and animal amygdala. Journal of Neuroscience. 2007; 27(22):5958–5966. [PubMed: 17537966]

Hillyard SA, Anllo-Vento L. Event-related brain potentials in the study of visual selective attention. Proceedings in the National Academy of Sciences of the United States of America. 1998; 95:781–787.

Johnstone T, van Reekum CM, Urry HL, Kalin NH, Davidson RJ. Failure to regulate: Counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. Journal of Neuroscience. 2007; 27(33):8877–8884. DOI: 10.1523/JNEUROSCI.2063-07.2007 [PubMed: 17699669]

Josselyn SA, Falls WA, Gewirtz JC, Pistell P, Davis M. The nucleus accumbens is not critically involved in mediating the effects of a safety signal on behavior. Neuropsychopharmacology. 2005; 30(1):17–26. [PubMed: 15257308]

Jovanovic T, Kazama A, Bachevalier J, Davis M. Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology. 2012; 62(2):695–704. DOI: 10.1016/j.neuropharm.2011.02.023 [PubMed: 21377482]

Jovanovic T, Keyes M, Fiallos A, Myers KM, Davis M, Duncan EJ. Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. Biological Psychiatry. 2005; 57:1559–1564. [PubMed: 15953493]

Kaye JT, Bradford D, Curtin JJ. Psychometric properties of startle and corrugator response in NPU, affective picture viewing, and resting state tasks. Psychophysiology. (in press).

Kong E, Monje FJ, Hirsch J, Pollak DD. Learning not to fear: Neural correlates of learned safety. Neuropsychopharmacology. 2014; 39(3):515–527. DOI: 10.1038/npp.2013.191 [PubMed: 23963118]

Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. Science. 1977; 197:792–795. [PubMed: 887923]

Larson CL, Schaefer HS, Siegle GJ, Jackson CAB, Anderle MJ, Davidson RJ. Fear is fast in phobic individuals: Amygdala activation in response to fear-relevant stimuli. Biological Psychiatry. 2006; 60(4):410–417. DOI: 10.1016/j.biopsych.2006.03.079 [PubMed: 16919528]

Lissek S, Rabin SJ, McDowell DJ, Dvir S, Bradford DE, Geraci M, … Grillon C. Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. Behaviour Research and Therapy. 2009; 47(2):111–118. DOI: 10.1016/j.brat.2008.10.017 [PubMed: 19027893]

Lohr JM, Olatunji BO, Sawchuk CN. A functional analysis of danger and safety signals in anxiety disorders. Clinical Psychology Review. 2007; 27(1):114–126. DOI: 10.1016/j.cpr.2006.07.005 [PubMed: 16997437]

Maier SF. Behavioral control blunts reactions to contemporaneous and future adverse events: Medial prefrontal cortex plasticity and a corticostriatal network. Neurobiology of Stress. 2015; 1:12–22. DOI: 10.1016/j.ynstr.2014.09.003 [PubMed: 25506602]

McKay D. Safe, but exposed: Inherent conflicts in safety signal conceptualization. Clinical Psychology: Science and Practice. 2010; 17(3):234–237. DOI: 10.1111/j.1468-2850.2010.01214.x

Najavits, LM. Seeking safety: A treatment manual for PTSD and substance abuse. New York, NY: Guilford Press; 2001.

Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, … Yarkoni T. Scientific standards. Promoting an open research culture. Science. 2015; 348(6242):1422–1425. DOI: 10.1126/science.aab2374 [PubMed: 26113702]

Ochsner KN, Bunge SA, Gross JJ, Gabrieli JDE. Rethinking feelings: An FMRI study of the cognitive regulation of emotion. Journal of Cognitive Neuroscience. 2002; 14(8):1215–1229. DOI: 10.1162/089892902760807212 [PubMed: 12495527]

Ochsner KN, Gross JJ. The cognitive control of emotion. Trends in Cognitive Sciences. 2005; 9(5): 242–249. DOI: 10.1016/j.tics.2005.03.010 [PubMed: 15866151]

Ochsner KN, Ray RD, Cooper JC, Robertson ER, Chopra S, Gabrieli JDE, Gross JJ. For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. NeuroImage. 2004; 23(2):483–499. DOI: 10.1016/j.neuroimage.2004.06.030 [PubMed: 15488398]

Pessoa L. To what extent are emotional visual stimuli processed without attention and awareness? Current Opinion in Neurobiology. 2005; 15(2):188–196. [PubMed: 15831401]

Pessoa L, Kastner S, Ungerleider LG. Attentional control of the processing of neural and emotional stimuli. Brain Research. Cognitive Brain Research. 2002; 15(1):31–45. [PubMed: 12433381]

Phelps EA, O'Connor KJ, Gatenby JC, Gore JC, Grillon C, Davis M. Activation of the left amygdala to a cognitive representation of fear. Nature Neuroscience. 2001; 4(4):437–441. [PubMed: 11276236]

Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson RJ, … Taylor MJ. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. Psychophysiology. 2000; 37(2):127–152. [PubMed: 10731765]

Pollak DD, Monje FJ, Zuckerman L, Denny CA, Drew MR, Kandel ER. An animal model of a behavioral intervention for depression. Neuron. 2008; 60(1):149–161. DOI: 10.1016/j.neuron.2008.07.041 [PubMed: 18940595]

Quirk GJ, Beer JS. Prefrontal involvement in the regulation of emotion: Convergence of rat and human studies. Current Opinion in Neurobiology. 2006; 16(6):723–727. doi:16/j.conb.2006.07.004. [PubMed: 17084617]

Rachman S. Agoraphobia: A safety-signal perspective. Behaviour Research and Therapy. 1984; 22(1): 59–70. [PubMed: 6696714]

Rogan MT, Leon KS, Perez DL, Kandel ER. Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron. 2005; 46:309–320. [PubMed: 15848808]

Sadeh N, Verona E. Visual complexity attenuates emotional processing in psychopathy: Implications for fear-potentiated startle deficits. Cognitive, Affective & Behavioral Neuroscience. 2012; 12(2): 346–360. DOI: 10.3758/s13415-011-0079-1

Sarinopoulos I, Grupe DW, Mackiewicz KL, Herrington JD, Lor M, Steege EE, Nitschke JB. Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. Cerebral Cortex. 2010; 20(4):929–940. DOI: 10.1093/cercor/bhp155 [PubMed: 19679543]

Sartory G, Master D, Rachman S. Safety-signal therapy in agoraphobics: A preliminary test. Behaviour Research and Therapy. 1989; 27(2):205–209. [PubMed: 2930448]

Semlitsch HV, Anderer P, Schuster P, Presslich O. A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. Psychophysiology. 1986; 23(6):695–703. [PubMed: 3823345]

van Boxtel A, Boelhouwer AJW, Bos AR. Optimal EMG signal bandwidth and interelectrode distance for the recording of acoustic, electrocutaneous and photic blink reflexes. Psychophysiology. 1998; 35(6):690–697. [PubMed: 9844430]

Vøllestad J, Sivertsen B, Nielsen GH. Mindfulness-based stress reduction for patients with anxiety disorders: Evaluation in a randomized controlled trial. Behaviour Research and Therapy. 2011; 49(4):281–288. DOI: 10.1016/j.brat.2011.01.007 [PubMed: 21320700]

Walasek G, Wesierska M, Zieliński K. Conditioning of fear and conditioning of safety in rats. Acta Neurobiologiae Experimentalis. 1995; 55(2):121–132. [PubMed: 7660862]

Williams DA, Sagness KE, McPhee JE. Configural and elemental strategies in predictive learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20(3):694–709. DOI: 10.1037/0278-7393.20.3.694

Woody S, Rachman S. Generalized anxiety disorder (GAD) as an unsuccessful search for safety. Clinical Psychology Review. 1994; 14(8):743–753. DOI: 10.1016/0272-7358(94)90040-X

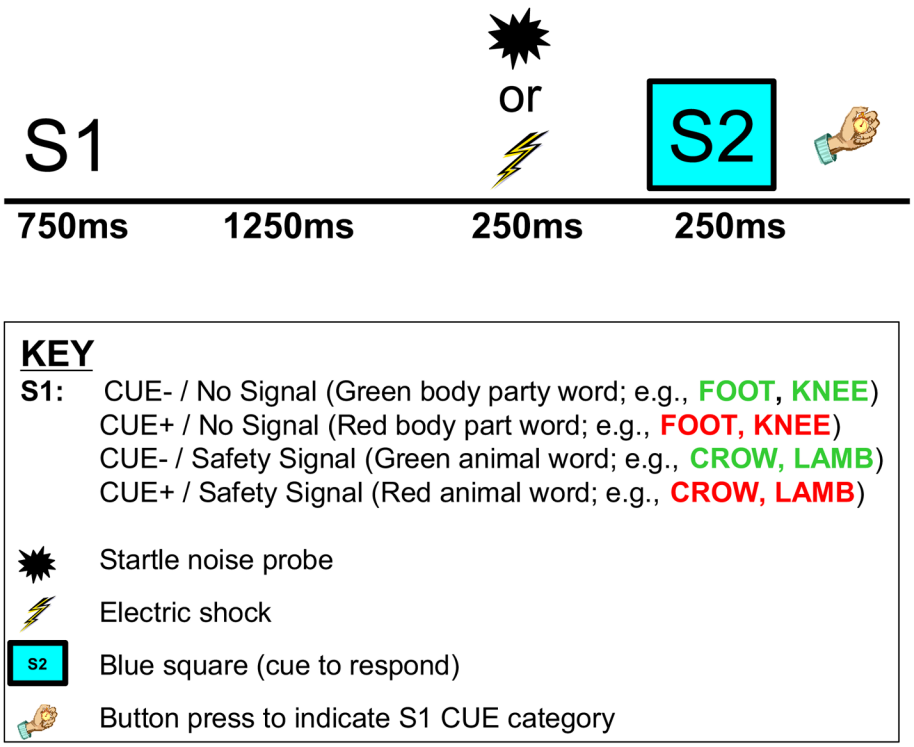

Figure 1.

Each trial consisted of two stimuli (S1 and S2) with their presentation onsets separated by 2,250 ms. The S1 stimulus was one of two word categories, either an animal or body part word, presented in either red or green ink. A blue square was always used for S2. Participants were advised that an electric shock could occur 2 s after S1s written in red ink (cue+, regardless of word category) and that no shocks would ever follow green S1s (cue−, regardless of word category).

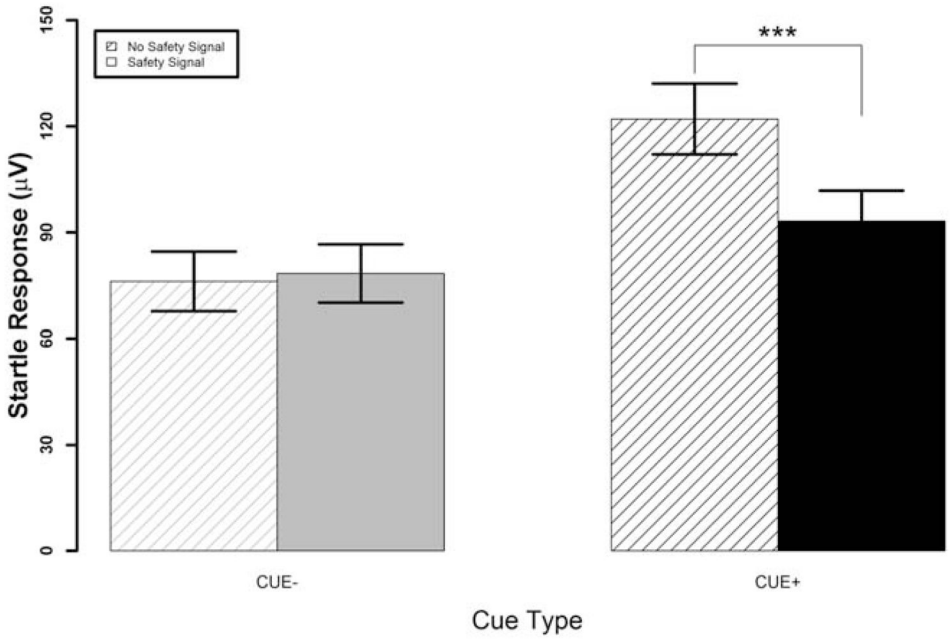

Figure 2.

Startle response magnitude by cue type and safety signal presence is displayed. Startle response magnitude was significantly potentiated during cue+ trials relative to cue− trials (p <.001). Follow-up tests indicated the main effect of cue type was moderated by safety signal condition, p <.001, such that fear-potentiated startle to threat was reduced for trials involving safety signals. Error bars represent the standard errors for the startle potentiation point estimates from the general linear model. ***p <.001.

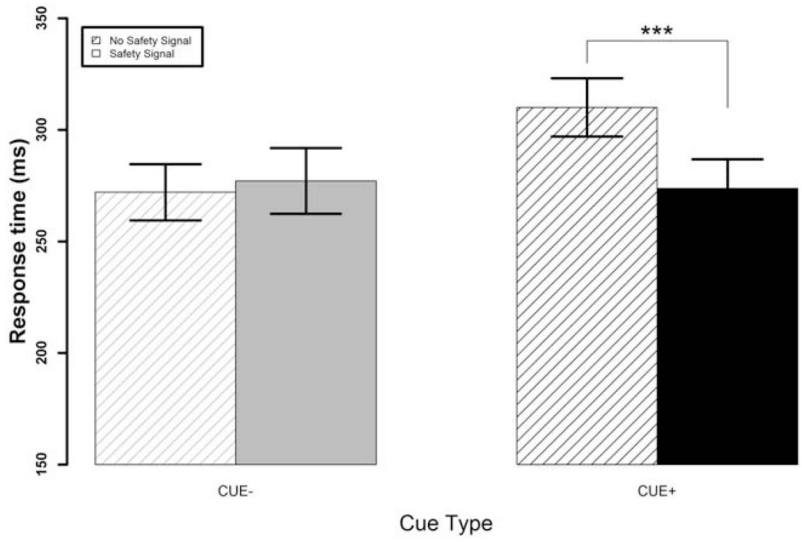

Figure 3.

Response time by cue type and safety signal presence is displayed. Response time was increased during cue+ trials relative to cue-trials (p =.002). Follow-up tests revealed this main effect of cue type was strongly moderated by safety signal condition, p <.001, indicating that the magnitude of response time slowing on cue+ versus cue− trials was reduced for trials involving safety signals. Error bars represent the standard errors for the response time point estimates from the general linear model. ***p <.001.

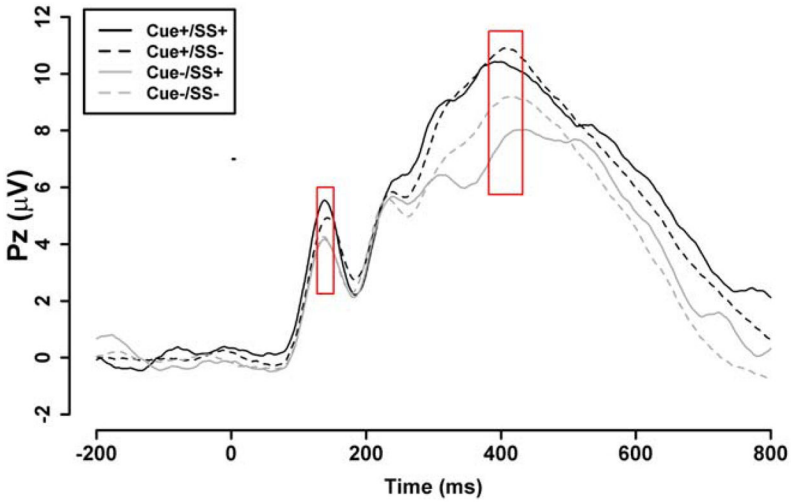

Figure 4.

Midline parietal ERP waveform is displayed from 200 ms pre- to 800 ms postcue onset. PD130 (mean response between 127–152 ms) and P3 (mean responses between 382–432 ms) are indicated by boxes surrounding the relevant window.

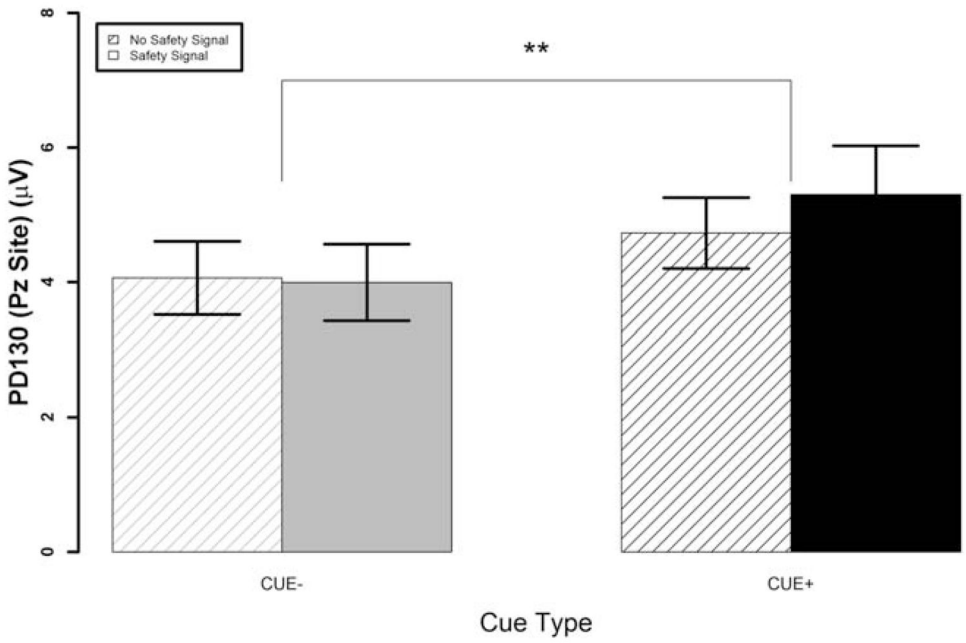

Figure 5.

ERP response during PD130 (mean response 127–152 ms) by cue type and safety signal presence is presented. The main effect of cue type was significant, with increased PD130 on cue+ relative to cue− trials, $p = .010$, indicating greater attention toward threat cues early in the processing stream, regardless of safety signal presence. Error bars represent the standard errors for the PD130 point estimates from the general linear model. $^{**}p < .01$.

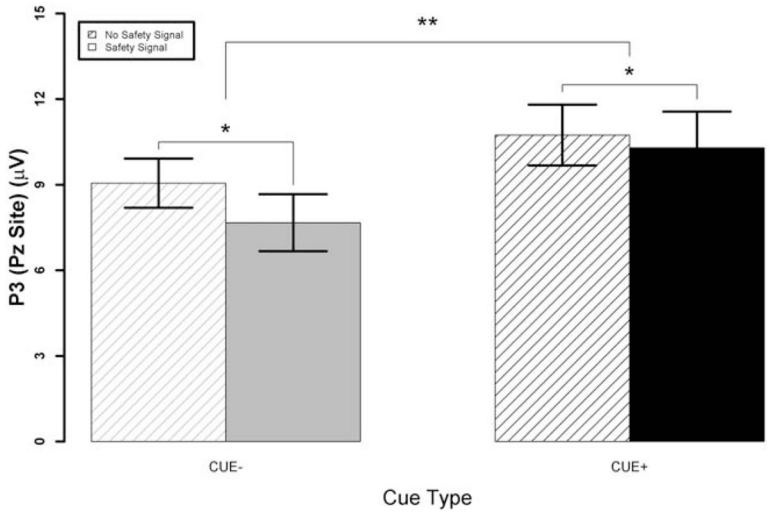

Figure 6.

ERP response during P3 (mean response 382–432 ms) by cue type and safety signal presence is presented. The main effect of cue type was significant, with increased P3 on cue + relative to cue− trials, $p = .002$. The main effect of safety signal condition was also significant, with increased P3 on SS− relative to SS+ trials, $p = .044$, indicating that participants are able to disengage attention in the presence of safety signals by P3. Error bars represent the standard errors for the P3 point estimates from the general linear model. *$p < .05$; **$p < .01$.