Passive avoidance is linked to impaired fear extinction in humans

Brian R. Cornwell,[1,2,3] Cassie Overstreet,[1] Marissa Krimsky,[1] and Christian Grillon[1]

[1]Section on Neurobiology of Fear and Anxiety, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, USA; [2]Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, Victoria, VIC 3122, Australia

Conventional wisdom dictates we must face our fears to conquer them. This idea is embodied in exposure-based treatments for anxiety disorders, where the intent of exposure is to reverse a history of avoidant behavior that is thought to fuel a patient's irrational fears. We tested in humans the relationship between fear and avoidance by combining Pavlovian differential fear conditioning with a novel task for quantifying spontaneous passive avoidant behavior. During self-guided navigation in virtual reality following de novo fear conditioning, we observed participants keeping their distance from the feared object. At the individual level, passive avoidant behavior was highly associated with maladaptive fear expression (fear-potentiated startle) during late extinction training, indicating that extinction learning was impaired following a brief episode of avoidance. Avoidant behavior, however, was not related to initial acquired fear, raising doubt about a straightforward link between physiological fear and behavioral avoidance. We conclude that a deeper understanding of what motivates avoidance may offer a target for early intervention, before fears transition from the rational to the irrational.

Conventional wisdom dictates we must face our fears to conquer them. This idea is embodied in exposure-based treatments for anxiety disorders, where the therapeutic intent of exposure is to reverse a history of avoidant behavior that is thought to fuel a patient's irrational fears (Foa and Kozak 1986; Craske 1999). Exposure is a generally effective method of treatment (McNally 2007), and recent neuroimaging evidence provides compelling demonstration of how fear circuits are modulated in conjunction with symptom reduction following a single exposure session in the case of spider phobia (Hauner et al. 2012). However, despite being grounded in over half a century of empirical and theoretical work on animal learning (Mowrer 1947; Rescorla and Solomon 1967; Siddle and Bond 1988), the tacit assumption that avoidant behavior is motivated by fear and contributes to its pathology has never been adequately tested in humans. In fact, even in animals the link between fear and avoidance is tenuous (Mineka 1979). Here we addressed the causes and consequences of passive avoidance by combining Pavlovian fear conditioning and self-guided navigation in virtual reality (VR).

To date, few studies have investigated human avoidant behavior, which in large part reflects the difficulty in creating an experimental setting to properly observe and measure avoidance in human participants (Grillon et al. 2006; Delgado et al. 2009; Lovibond et al. 2009; Glotzbach et al. 2012). One strategy is to incorporate avoidant responding into the task structure as described to participants, such as having them learn an arbitrary response (e.g., button press) that terminates an upcoming aversive stimulus (e.g., electric shock). Lovibond et al. (2009), for instance, used this kind of active or instrumental avoidance procedure combined with Pavlovian conditioning to show that avoidant responding during extinction training preserved fear expression in a subsequent test when the opportunity to avoid was withheld. The major clinical implication of this finding is that exposure techniques in the clinic may be effective only insofar as the patient does not

engage in safety behaviors during treatment (e.g., avoiding eye contact). Generally speaking, this approach is limited in the extent to which it can establish effective precursors to avoidant behavior given that participants may be simply following task instructions rather than making avoidant responses to regulate their fear. Although this method guarantees an adequate amount of avoidant behavior to observe, the underlying motivational factors cannot be determined.

Another approach involves the use of VR to capture spontaneous (uninstructed) passive avoidant behavior, a form of avoidance which has been argued by some to be noninstrumental and thus distinct from active avoidance (for review, see Lovibond 2006). Grillon et al. (2006), for instance, had participants voluntarily return to one of three VR contexts that they had previously navigated through while receiving unsignaled shocks, signaled shocks, or no shocks, respectively. Participants were most likely to choose to re-enter the no-shock context, indicating that they deliberately avoided the other contexts in which shocks had been administered. Glotzbach et al. (2012) recently reported using a similar setup that participants who subsequently avoided a VR context associated with shock reported higher levels of subjective fear to that context than participants who did not avoid, tentatively linking experiential fear with the tendency to avoid. While this is a positive step, physiological measures of fear, which were not taken in this study, would seem preferable to establish effective predictors of avoidant behavior because of their relative immunity to experimental demand. A second limitation of these VR studies is that avoidant behavior was measured categorically (avoidant behavior was defined as all-or-nothing), which may not be optimal to elucidate its antecedents and consequences.

We improved on these latter passive avoidance studies in two ways, by measuring physiological fear through startle reflex modulation and by measuring avoidant behavior quantitatively in VR. Participants were exposed to Pavlovian fear conditioning and extinction procedures. Startle responses, a cross-species defensive reflex that is potentiated during aversive states (Lang et al. 1990; Davis et al. 1993; Grillon 2002), were elicited during
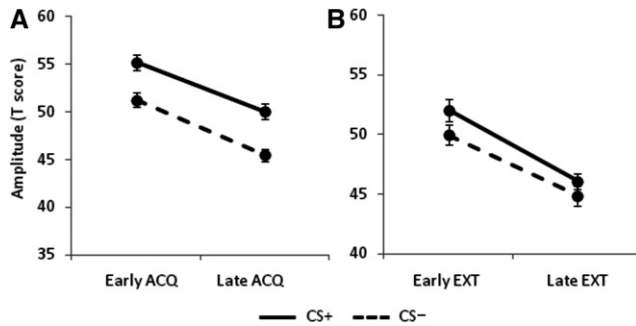
Figure 1. Participants exhibit differentially greater startle reactivity to a Pavlovian-conditioned stimulus (CS) during acquisition and extinction training. Startle reflex magnitude (expressed in standardized *T* scores) was higher in the CS+ compared to the CS− during (*A*) acquisition (ACQ) and (*B*) extinction (EXT) training. Error bars are SEM.

these procedures to establish conditioned fear and extinction learning, respectively. Between acquisition and extinction, participants performed a self-guided navigation task in a novel VR context (circular pool) in which they encountered the de novo feared object (CS+) and controlled their proximity to it. The task was to locate an escape platform in the pool and we did not explicitly draw attention to the CS+. This allowed us to address whether Pavlovian-conditioned fear manifests behaviorally as spontaneous passive avoidance of the CS+, which would be evident in participants' search behavior, and whether passive avoidance (or a lack of exposure to the feared object) is associated with subsequent fear extinction learning. Based on limited empirical evidence in humans (Lovibond et al. 2009), we predicted that a greater degree of avoidance would be related to impaired extinction learning and greater preservation of fear expression during late extinction training.

Results

Group-level analyses

To induce de novo fear, participants were exposed to a Pavlovian discriminative fear conditioning procedure. Thirty participants accurately reported the CS–US contingency at the end of the experiment (contingency aware), and were analyzed separately from those who failed to report the CS–US contingency (contingency unaware, see below). As expected, a 2 × 2 repeated-measures ANOVA revealed that startle magnitudes were higher during the CS+ compared to the CS− ($F_{CS(1,29)} = 28.89$, $P < 0.001$; Fig. 1A). There was also a significant decrease in startle magnitudes from early to late acquisition ($F_{Block(1,29)} = 40.85$, $P < 0.001$), reflecting common habituation effects on startle reactivity. The slight increase in differential startle reactivity from early to late acquisition was not significant ($F_{CS×Block(1,29)} = 0.21$, $P = 0.65$), indicating that contingency aware participants conditioned within the first few CS+ trials.

To determine whether acquired fear manifests behaviorally as passive avoidance of the CS+, participants searched two VR pools for an escape platform (Fig. 2A). Each trial series began with a platformless trial to assess spontaneous search behavior before participants had knowledge of the platform's location. In the test pool, the fear-evoking CS+ hung above the southwest (SW) quadrant. Despite no shock administration, participants showed a clear search bias in the test pool (CS+ present), spending less time searching for the platform in the SW quadrant than chance (25%, one-sample $t_{(29)} = -2.49$, $P = 0.02$, $d = -0.45$;

Fig. 2B,C). This was not the case in the control pool (CS+ absent) where no quadrant was proportionately less searched (*P*'s ≥ 0.17). A direct contrast in the proportion of time spent in the SW quadrant between the test pool (22.39 ± 5.73%) and control pool (24.39 ± 5.53%) showed a nonsignificant trend in the predicted direction ($t_{(29)} = -1.40$, $P = 0.17$, $d = -0.35$). Moreover, participants took significantly longer to first enter the SW quadrant in the test relative to the control pool (paired sample $t_{(29)} = 2.05$, $P = 0.05$, $d = 0.37$; Fig. 2D), which was not the result of delayed initiation of movement in the test (3.92 ± 1.45 sec) compared to the control pool (4.33 ± 1.58 sec, $t_{(29)} = -1.63$, $P = 0.12$). No evidence of avoidant behavior was observed for the group on the final platformless trials (time spent in SW, test pool = 30.7 ± 18.2%, control pool = 30.6 ± 13.7%) as participants' paths were increasingly determined by their knowledge of the platform's expected location (adjacent to the SW quadrant).

Following the virtual navigation task, participants were exposed to a fear extinction procedure in the original conditioning context. A 2 × 2 ANOVA revealed that startle magnitudes were higher during the CS+ relative to the CS− ($F_{CS(1,28)} = 5.85$, $P = 0.02$; Fig. 1B). There was also a significant decrease in startle magnitudes from early to late extinction ($F_{Block(1,28)} = 31.47$, $P < 0.001$). However, the slight decrease in differential startle reactivity between the CS+ and CS− from early to late extinction was not significant ($F_{CS×Block(1,28)} < 1$), indicating that participants, overall, continued to express fear to the CS+ throughout extinction training despite the absence of shocks.
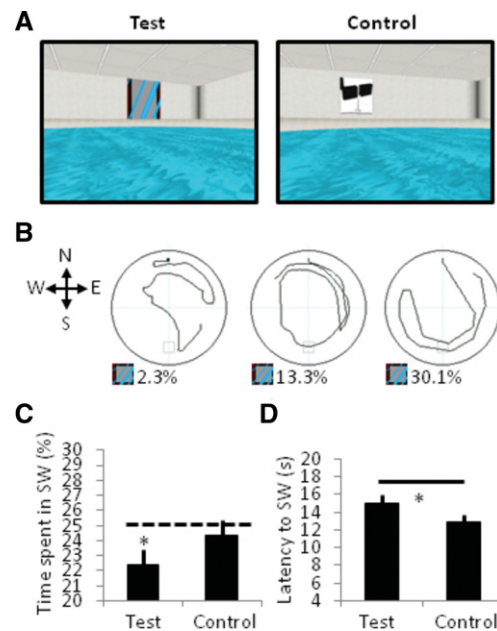

Figure 2. Participants exhibit passive avoidance of the CS+ in VR following de novo fear conditioning. (*A*) Participants searched for an escape platform in two VR pool contexts. Each pool snapshot shows a view of the southwest (SW) quadrant from a northeast position, with the CS+ present (test pool) or absent (control pool). Participants were instructed that shocks were possible during the navigation task, but none was actually administered. (*B*) Aerial views of search paths taken by three participants depicting variable amounts of search bias in the test pool (2.3%, 13.3%, or 30.1% time spent in the SW quadrant). The approximate location of the CS+ on the surrounding walls is indicated. (*C*) In the test pool, participants spent less time searching in the SW quadrant and took longer to first enter the SW quadrant of the test pool relative to the control pool. (*) *P* < 0.05. Error bars are SEM.
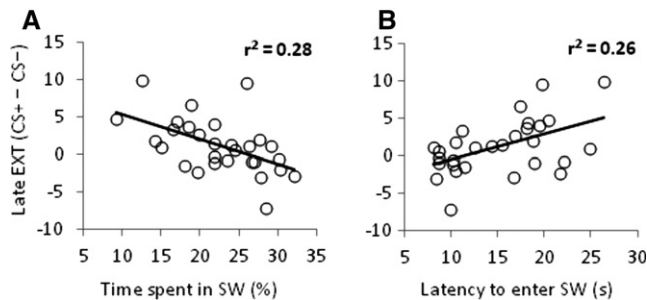
Figure 3. Avoidant behavior and fear expression during late extinction are robustly correlated. Scatterplots with least squares lines (*A*) depict the negative association between time spent in the SW quadrant of the test pool and differential startle reactivity to the CS+ during late extinction, and (*B*) depict the positive association between latency to enter the SW quadrant of the test pool and differential startle reactivity to the CS+ during late extinction.

Correlation analyses

To test whether physiological fear was related to avoidant behavior at the individual level, we calculated Pearson correlations between avoidance variables (time spent in the SW quadrant and latency to first enter the SW quadrant on the first platformless trials) and fear-potentiated startle (CS+ minus CS−) during early and late acquisition and extinction. To control false positive rates, a Bonferroni correction was applied to these correlation results ($\alpha = 0.0063$). No correlations emerged for fear-potentiated startle during acquisition or early extinction (time spent in the SW quadrant, r's ≥ -0.18, P's ≥ 0.35; latency to enter the SW quadrant, r's ≤ 0.26, $P \geq 0.17$), suggesting that the level of acquired fear to the CS+ was not related to the degree of avoidant behavior. In contrast, fear-potentiated startle during late extinction was highly associated with time spent in the SW quadrant ($r_{(27)} = -0.53$, $P = 0.003$, Fig. 3A) and latency to enter the SW quadrant ($r_{(27)} = 0.51$, $P = 0.005$, Fig. 3B). For descriptive purposes, we also split the sample into avoiders and nonavoiders to illustrate the distinct relationship between avoidance (i.e., time spent in the SW quadrant) and fear-potentiated startle during late extinction (Fig. 4). Importantly, no relationships were observed between the same behavioral variables in the control pool and fear-potentiated startle during late extinction (time spent in the SW quadrant, $r_{(27)} = -0.03$, $P = 0.87$; latency to enter the SW quadrant, $r_{(27)} = 0.27$, $P = 0.16$), confirming the relevance of the CS+ in driving the relationship between biased search behavior and continued fear expression during late extinction.

Regression analyses

To further explore this distinctive relationship between avoidant behavior and fear expression during extinction, we regressed fear-potentiated startle during early and late extinction separately on the same two predictor variables, fear-potentiated startle during late acquisition and time spent in the SW quadrant. The model fit was significant for late extinction ($F_{(2,26)} = 5.36$, adjusted $r^2 = 0.24$, $P = 0.01$), but not early extinction ($F_{(2,26)} = 2.12$, adjusted $r^2 = 0.07$, $P = 0.14$). Degree of avoidance predicted fear-potentiated startle during late extinction ($\beta = -0.51$, $t_{(26)} = -3.07$, $P = 0.005$), but fear-potentiated startle during late acquisition did not ($\beta = 0.10$, $P = 0.55$).

Contingency unaware subject analyses

Seven participants failed to report the CS–US relationship at the end of the experiment. These participants showed no evidence

by visual inspection of greater startle magnitudes during the CS+ compared to the CS− during early or late acquisition (Fig. 5A), which supports evidence that awareness of the CS–US contingency is necessary for fear learning (Purkis and Lipp 2001; Lovibond and Shanks 2002; Cornwell et al. 2007; Mitchell et al. 2009; but also see Wiens and Ohman 2002; Weike et al. 2007). Given this outcome, we expected that these contingency unaware participants would not show passive avoidance of the CS+. As predicted, unlike the 30 contingency aware participants, contingency unaware participants did not show a delay in first entering the SW quadrant in the test pool relative to the control pool (Fig. 5B). For time spent in the SW quadrant on the initial platformless trials, contingency unaware participants did significantly deviate from the expected chance value (25%) in the test pool (13.9 ± 11.5%, $t_{(6)} = -2.52$, $P = 0.04$). However, this was also true in the control pool (16.3 ± 8.8%, $t_{(6)} = -2.52$, $P = 0.04$), suggesting participants showed a similar search bias in the presence and absence of the CS+.

Discussion

By combining Pavlovian conditioning and a virtual reality (VR) task, our data provide critical insight into the role of passive avoidance in linking the development of fear to its abnormal persistence. During self-guided navigation in VR following de novo fear conditioning, we observed participants keeping their distance from the feared object. Importantly, this was not the case for those who were unable to report the CS–US contingency at the end of the experiment, suggesting that passive avoidant behavior was the direct result of associative learning. Moreover, while fear expression during acquisition, indexed by fear-potentiated startle, was not correlated with the degree of passive avoidance, the latter was highly predictive of fear expression during late extinction. Based on this outcome, we conclude that passive avoidant behavior may negatively impact subsequent extinction learning (similar to active avoidant behavior [Lovibond et al. 2009]), and thus lead to a preservation of fear despite the absence of threat. Given that exaggerated fear responding is a hallmark characteristic of pathological anxiety, our findings suggest that avoidance may contribute to the genesis of irrational fears.
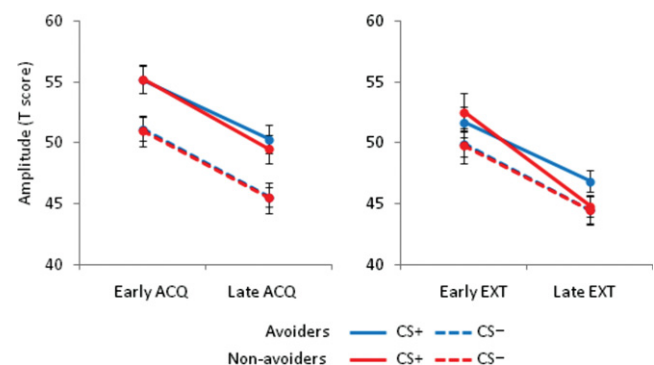

Figure 4. Passive avoidance is linked to continued fear expression. The sample was split, for illustrative purposes, into avoiders ($N = 18$) and non-avoiders ($N = 11$) based on time spent in the SW quadrant of the test pool (<25% and ≥25%, respectively). Both groups show similar levels of differential startle reactivity during acquisition and early extinction, but only the avoiders continued to show elevated startle to the CS+ during late extinction. Note that the blue dotted line is partially occluded by the red one because startle responding during the CS− was nearly identical for avoiders and nonavoiders.
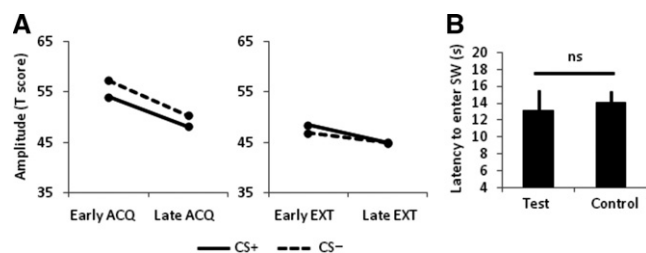
Figure 5. Contingency unaware participants did not show fear-potentiated startle and did not show biased search indicative of avoidance of the CS+. (*A*) Fear-potentiated startle to the CS+ is absent during both acquisition and extinction phases. (*B*) There was no difference in the latency to first enter the SW quadrant between the test (CS+ present) and control (CS+ absent) pools.

In the VR navigation task, we relied on two measures to quantify spontaneous passive avoidant behavior, time spent in the SW quadrant and latency to first enter the SW quadrant. Both measures operationalized avoidant behavior in terms of (virtual) spatial proximity to the CS+, which hung on the wall above the SW quadrant in the test pool. In addition, we included a control condition in which participants performed the same task in a different pool in the absence of the CS+ to allow for interpreting specific search behaviors in the test pool as passive avoidance. At the group level, effect sizes were moderate, and we only observed avoidance on the first platformless trial of each series. Several factors can account for these results, including the lack of shock administration during the VR task and the change in context. On this latter point, it should be emphasized that any avoidant behavior observed during VR navigation was predicated on participants recognizing the CS+ in the new context, where it was embedded in the background and may not have been as perceptually salient, and generalizing their fear accordingly. To our knowledge, this is the first demonstration that conditioned fear can transfer contexts as measured by passive avoidance. The fact that search biases did not persist to the final platformless trial of each series could be attributable to a change in shock expectancy, given the lack of reinforcement during the navigation task. However, search behavior was increasingly determined by greater knowledge of the platform's location, which complicates interpretation of this search behavior in terms of passive avoidance beyond the first platformless trials. This suggests that, in future investigations, the virtual navigation task could be scaled back to one or two platformless trials to measure passive avoidance.

The main goal of quantitatively measuring avoidant behavior was to study individual differences. Here is where we found robust correlations between the degree of avoidance on the first platformless trials and fear expression during late extinction. In categorical terms, participants who tended to preferentially search for the platform outside the SW quadrant containing the CS+ and who increasingly delayed entry into the SW quadrant continued to show fear to the CS+ throughout subsequent extinction training (Fig. 4). Those who did not show these search biases exhibited a clear reduction of fear-potentiated startle by late extinction. While we cannot establish a causal link, these data offer compelling evidence that passive avoidance of a feared object impairs extinction learning, consistent with the effects of active avoidance (Lovibond et al. 2009). In contrast, we found no evidence that individual differences in initial acquired fear and avoidant behavior were related. Although this might seem counterintuitive, it is not too surprising in light of evidence that physiological and behavioral correlates of fear often diverge (Riccio and Silvestri 1973; Mineka 1979), which, altogether, sug-

gests that fear is best construed as a set of loosely coupled response systems (Lang 1994).

That the degree of passive avoidance was linked to fear expression during late extinction and not also to early extinction requires some discussion. For those who showed little or no passive avoidant behavior, the virtual navigation task should have served as initial extinction training given that they navigated in close spatial proximity and exposed themselves to the CS+ without receiving the US. However, it is important to consider evidence that extinction learning is context dependent (Bouton 2004; Alvarez et al. 2007). Extinction learning in a novel context may not transfer to the original context in which fear was acquired, manifesting as renewal of fear when retested in the original context. This could explain in our data why the degree of passive avoidance, or conversely the amount of potential extinction learning in the novel VR context, did not correlate with fear expression upon returning to the original acquisition context for extinction training. Fear renewal in participants who showed little or no passive avoidance may be obscuring the link between passive avoidance and fear expression during early extinction training. Nevertheless, the mechanism linking passive avoidance to subsequent impaired extinction learning will require further work to be clarified. This may provide for a more refined theoretical framework for understanding how passive and active avoidance may distinctly impact extinction learning (Lovibond 2006).

Although initial acquired fear and avoidant behavior were not related at the individual level, it is plausible to conclude that passive avoidance stemmed from successful differential fear conditioning. Contingency unaware participants neither showed fear-potentiated startle to the CS+ nor a clear pattern of CS+ avoidance (Fig. 5), indicating that mere exposure to the CS+ before the virtual navigation task is not likely to have driven passive avoidance in contingency aware subjects. The contingency unaware participants did not delay their first entries into the SW quadrant containing the CS+, but did spend a significant proportion of time in other quadrants in both the test and control pools. In fact, one participant never entered the SW quadrant in either pool. Rather than reflecting avoidance of the CS+, biased search in these contingency unaware participants is likely due to a general lack of exploratory behavior irrespective of the presence of the CS+, given that starting positions were always outside the SW quadrant. This problem of appearance would be rectified by instituting random starting positions, but the nature of the task would be altered. Because our aim was to capture *passive* avoidance without providing shock reinforcement during navigation, we could not use locations close to the CS+ as starting positions, which would require an instrumental response (i.e., escape) to potential shocks. Future modifications of the task may, nonetheless, provide additional insights and more definitive conclusions regarding the antecedent conditions of passive avoidance. Importantly, given the small number of contingency unaware participants, these data should be interpreted cautiously in terms of whether contingency awareness is necessary to acquire and express conditioned fear physiologically or behaviorally.

Finally, our regression analyses in contingency aware participants revealed that fear expression during late extinction was associated with degree of avoidance even after controlling for fear-potentiated startle during late acquisition. This indicates that the maladaptive persistence of fear appears to be uniquely related to the passive avoidant behavior exhibited during navigation, irrespective of the fear that was initially acquired. Notably, this finding supports the claim that avoidant behavior is not just an expression of fear, but may also inadvertently maintain fear, even following only a brief period of avoidance. The results further support the use of cognitive–behavioral techniques aimed at having patients confront the source of their fears (Kashdan

et al. 2012), and indicate that early intervention may be critical to reducing fear in addition to its behavioral symptoms. Moreover, the current paradigm may provide a valuable tool for studying pharmacological effects of various compounds on fear and avoidance. For instance, recent interest in augmenting fear extinction learning during exposure therapy with D-cycloserine may benefit from further de novo fear conditioning studies that include an avoidance component to better simulate the pathological condition (Grillon 2009).

To summarize, by allowing participants to freely navigate around a virtual environment that was distinct from the fear acquisition/extinction context, we observed that it was not the most fearful participants who avoided the threat, indicating a more complex set of dispositional factors (such as general risk aversion, for example) that motivates avoidant behavior (Mineka 1979). Yet, what began as an ostensibly adaptive response to a perceived threat seems to become a barrier to learning that the threat is no longer real. Consequently, fear persists in those who avoided, much like it does for the social phobic who routinely skips social gatherings or the agoraphobic who never wanders far from home. Given that avoidance does not appear to be a simple function of the initial fear that was acquired during conditioning, a deeper understanding of what motivates avoidance may offer a target for early intervention, before fears turn from the rational to the irrational.

Materials and Methods

Participants

A total of 40 participants volunteered and were paid for participation. Thirty were included in the main analyses ($N = 30$, 19 women, age [mean $\pm$ SD] = 27 $\pm$ 5 yr). Three were excluded because of equipment failure. Seven were separately analyzed because they failed to show awareness of the CS–US contingency (see below). Participants were medically and psychiatrically healthy based on a physical exam and clinical interview (First et al. 1995). All procedures were approved by the Combined Neuroscience Institutional Review Board of the National Institutes of Health. We obtained informed consent from all participants after procedures were fully explained.

Psychophysiological recording

A commercial system was used for physiological data collection and analyses (Contact Precision Instruments). Two 2-mm tin-cup electrodes were attached beneath the left eye to record orbicularis oculi muscle activity. Electromyographic (EMG) data were sampled at 1000 Hz with a bandwidth of 30–500 Hz. EMG data were rectified offline and smoothed using a 20-msec kernel.

Startle reflexes were elicited by bursts of white noise (40 msec, 104 dB(A), near instantaneous rise/fall) delivered binaurally through headphones. Shocks were administered through two (Ag/AgCl) electrodes attached to the wrist by a constant current stimulator (100 msec duration, 3–5 mA). Shock intensity was set using a work-up procedure to identify a subjective level of moderate discomfort. Attachment of the electrodes and shock work-up followed a recording in which participants were presented nine startle probes alone to habituate startle responses (data not presented).

Fear acquisition

Participants were instructed that they would be presented images on the computer screen and that shocks may occur, and that if they paid attention they might learn to predict the shocks. Two visual objects (blue-striped square, brown-striped square) were presented nine times each in pseudo-random order with no more than two sequential presentations of the same object and an intertrial interval (ITI) of 10–15 sec. One object (CS+) always co-terminated with a shock (100% reinforcement, counterbalanced across subjects). The other object (CS−) never co-terminat-

ed with a shock. Each CS was presented for 10 sec. For the first 5 sec of each presentation, the CS increased in size to the approximate size in which it would appear in the VR task. This feature was included to simulate increasing proximity to the participant in an effort for them to associate "being close" to the CS+ with shock. Startle probes were delivered 6–8 sec following CS onset on every trial. In addition, two startle probes were presented alone before the first CS trial and three startle probes were delivered during ITI (data not presented). The overall schedule of startle probe delivery was maintained at 18–22 sec.

Virtual navigation (avoidance) task

The virtual navigation task was implemented with commercial VR software (NeuroInvestigations, Inc.). Participants navigated with button presses around two virtual circular pools (test and control) to locate an escape platform. This task is a simulated version of the Morris water maze, a well-established translational paradigm to study hippocampal-dependent spatial navigation in rodents (Morris et al. 1982) and humans (Cornwell et al. 2008, 2012). Participants were told that shocks may occur and electrodes remained attached (no shocks were administered). Two series of trials in the test (CS+ present) and control (CS+ absent) pools were completed. In the test pool, the CS+ hung above the southwest (SW) quadrant, and three novel objects were displayed on other walls to aid navigation. Four novel objects were displayed in the control pool. Participants performed a series in one pool before switching to the other pool. This was then repeated after informing participants that the platform had moved location in each pool. The starting pool was counterbalanced across participants.

Each series contained seven trials, beginning and ending with a platformless trial. Trial compositions (starting location, platform location) were identical between test and control pool series. Platform location was fixed for each series, at the pool's edge due west or due south. Starting position at the pool's edge varied across trials, but participants never started in the SW quadrant. For the main analyses, we focused on the first platformless trial because it allowed for measurement of spontaneous search behavior before participants gained knowledge of the platform's location. The middle trials included four with a visible platform (trials 2, 3, 5, 6) and one in which the platform was hidden initially, but became visible if not found within 20 sec (trial 4). These trials served to reinforce to participants that the platform exists and can be reached even if it is initially hidden, not to measure avoidant behavior. The final platformless trial was used to assess whether the CS+ continued to bias search paths despite participants having gained knowledge of the platform's location.

Fear extinction

Participants were instructed that they would again be presented images on the screen and that shocks may occur. Procedures were identical to acquisition with the exception that no shocks were administered.

Awareness check

At the end of extinction, participants were administered a short recognition questionnaire to determine awareness of the CS–US contingency (Dawson and Reardon 1973). Given substantial controversy over whether fear conditioning is dependent on conscious awareness (for extended discussion, see Lovibond and Shanks 2002; Wiens and Ohman 2002), we aimed to ensure that our main analyses concerning the relationship between avoidance and extinction learning included only participants with an adequate level of contingency awareness which we could assume would show robust differential fear during acquisition. Participants were instructed to recall the "first part" of the experiment when answering the following two questions: (1) Shocks usually followed (a) light blue-striped image, (b) reddish-brown striped image, (c) it wasn't systematic, (d) I couldn't tell; and (2) The degree of certainty I have in my above answer is (a) completely uncertain, (b) fairly uncertain, (c) fairly certain, (d) completely

certain. Those who answered the first question correctly were deemed contingency aware regardless of their degree of certainty and included in the main analyses (two of 30 contingency aware subjects lacked confidence in their answer). Finally we also confirmed at the end that each participant recognized the CS+ during the virtual navigation task.

Data analysis

For startle quantification, peak amplitude was defined between 20 and 100 msec following probe onset and normalized by a 50-msec baseline average. Trials with excessive baseline EMG noise were excluded (2% trials discarded). One subject's extinction data was excluded because of excessive EMG noise. Startle data were standardized (T scores) to control for interindividual differences in baseline startle reactivity. We excluded the first CS+ and CS− trial for acquisition data analyses. Thus, means for early acquisition, late acquisition, and late extinction are based on four trials, while the mean for early extinction is based on five trials.

For avoidant behavior, we computed three variables from the first and final platformless trials. Time spent in the SW quadrant is expressed as a percentage of the total trial time and was automatically calculated by the software. Latency to enter the SW quadrant was determined by the first time point in which the participant's position was in the SW quadrant. We also computed latency to first movement, not as a measure of avoidance, but to determine whether latency to first enter the SW quadrant was driven by how quickly movement was initiated. Data were averaged between two series in each pool.

Acknowledgments

References

Alvarez RP, Johnson L, Grillon C. 2007. Contextual-specificity of short-delay extinction in humans: Renewal of fear-potentiated startle in a virtual environment. *Learn Mem* **14:** 247–253.

Bouton ME. 2004. Context and behavioral processes in extinction. *Learn Mem* **11:** 485–494.

Cornwell BR, Echiverri AM, Grillon C. 2007. Sensitivity to masked conditioned stimuli predicts conditioned response magnitude under masked conditions. *Psychophysiology* **44:** 403–406.

Cornwell BR, Johnson LL, Holroyd T, Carver FW, Grillon C. 2008. Human hippocampal and parahippocampal θ during goal-directed spatial navigation predicts performance on a virtual Morris water maze. *J Neurosci* **28:** 5983–5990.

Cornwell BR, Arkin N, Overstreet C, Carver FW, Grillon C. 2012. Distinct contributions of human hippocampal θ to spatial cognition and anxiety. *Hippocampus* **22:** 1848–1859.

Craske MG. 1999. *Anxiety disorders: Psychological approaches to theory and treatment*. Westview Press, Boulder, CO.

Davis M, Falls WA, Campeau S, Kim M. 1993. Fear-potentiated startle: A neural and behavioral analysis. *Behav Brain Res* **58:** 175–198.

Dawson ME, Reardon DP. 1973. Construct validity of recall and recognition post-conditioning measures of awareness. *J Exp Psychol* **98:** 308–315.

Delgado MR, Jou RL, LeDoux JE, Phelps EA. 2009. Avoiding negative outcomes: Tracking the mechanisms of avoidance learning in humans during fear conditionings. *Front Behav Neurosci* **3:** 33.

First MB, Spitzer RI, Williams JBW, Gibbon M. 1995. *Structured clinical interview for DSM-IV (SCID)*. American Psychiatric Association, Washington, DC.

Foa EB, Kozak MJ. 1986. Emotional processing of fear: Exposure to corrective information. *Psychol Bull* **99:** 20–35.

Glotzbach E, Ewald H, Andreatta M, Pauli P, Muhlberger A. 2012. Contextual fear conditioning predicts subsequent avoidance behaviour in a VR environment. *Cogn Emot* **26:** 1256–1272.

Grillon C. 2002. Startle reactivity and anxiety disorders: Aversive conditioning, context, and neurobiology. *Biol Psychiatry* **52:** 958–975.

Grillon C. 2009. DCS facilitation of fear extinction and exposure-based therapy may rely on lower-level, automatic mechanisms. *Biol Psychiatry* **66:** 636–641.

Grillon C, Baas JMP, Cornwell B, Johnson L. 2006. Context conditioning and behavioral avoidance in a virtual reality environment: Effect of predictability. *Biol Psychiatry* **60:** 752–759.

Hauner KK, Mineka S, Voss JL, Paller KA. 2012. Exposure therapy triggers lasting reorganization of neural fear processing. *Proc Natl Acad Sci* **109:** 9203–9208.

Kashdan TB, Adams L, Read J, Hawk L Jr. 2012. Can a one-hour session of exposure treatment modulate startle response and reduce spider fears? *Psychiatry Res* **196:** 79–82.

Lang PJ. 1994. The varieties of emotional experience: A meditation on James–Lange theory. *Psychol Rev* **101:** 211–221.

Lang PJ, Bradley MM, Cuthbert BN. 1990. Emotion, attention, and the startle reflex. *Psychol Rev* **97:** 1–19.

Lovibond PF. 2006. Fear and avoidance: An integrated expectancy model. In *Fear and learning: From basic processes to clinical implications* (ed. MG Craske, et al.), pp. 117–132. American Psychological Association, Washington, DC.

Lovibond PF, Shanks DR. 2002. The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *J Exp Psychol: Anim Behav Process* **28:** 3–26.

Lovibond PF, Mitchell CJ, Minard E, Brady A, Menzies RG. 2009. Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behav Res Ther* **47:** 716–720.

McNally RJ. 2007. Mechanisms of exposure therapy: How neuroscience can improve psychological treatments for anxiety disorders. *Clin Psychol Rev* **27:** 750–759.

Mineka S. 1979. The role of fear in theories of avoidance learning, flooding, and extinction. *Psychol Bull* **86:** 985–1010.

Mitchell CJ, De Houwer J, Lovibond PF. 2009. The propositional nature of human associative learning. *Behav Brain Sci* **32:** 183–246.

Morris RGM, Garrud P, Rawlins JN, O'Keefe J. 1982. Place navigation impaired in rats with hippocampal lesions. *Nature* **297:** 681–683.

Mowrer OH. 1947. On the dual nature of learning: A reinterpretation of "conditioning" and "problem solving". *Harvard Educ Rev* **17:** 102–148.

Purkis HM, Lipp OV. 2001. Does affective learning exist in the absence of contingency awareness? *Learn Motiv* **32:** 84–99.

Rescorla RA, Solomon RL. 1967. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychol Rev* **75:** 151–181.

Riccio DC, Silvestri R. 1973. Extinction of avoidance behavior and the problem of residual fear. *Behav Res Ther* **11:** 1–9.

Siddle DAT, Bond NW. 1988. Avoidance learning, Pavlovian conditioning, and the development of phobias. *Biol Psychol* **27:** 167–183.

Weike AI, Schupp HT, Hamm AO. 2007. Fear acquisition requires awareness in trace but not delay conditioning. *Psychophysiology* **44:** 170–180.

Wiens S, Ohman A. 2002. Unawareness is more than a chance event: Comment on Lovibond and Shanks (2002). *J Exp Psychol: Anim Behav Process* **28:** 27–31.