Increasing defensive flexibility:

Facilitation of fear extinction by non-invasive stimulation of the brain's inhibitory pathways

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Christoph Szeska

geboren am 28.02.1993

in Grevesmühlen

Greifswald, 17.12.2021

Dekan: Prof. Dr. Gerald Kerth

1. Gutachter: Prof. Dr. Alfons O. Hamm

2. Gutachter: Prof. Dr. Erik M. Müller

3. Gutachter: Prof. Dr. André Pittig

Tag der Promotion: 28.04.2022

Say after me: It's no better to be safe than sorry.

- *a-ha, Take On Me*

LIST OF CONTENTS

Abstract

Fear is an emotional state, characterized by the activation of a defense system that is designed to ensure the organism's survival. This system enables a rapid recognition of threats and organizes defensive response patterns in order to adaptively cope with the threatening environment. Yet, to ensure its flexibility under changing environmental conditions, inhibitory pathways exist that modulate the activation of this defense system, if a previously threatening cue no longer predicts any harm – a memory-formatting process referred to as fear extinction, leading to a reduction of defensive responding. Fear extinction is presumed to at least partially underlie exposure treatment of anxiety disorders, which is why the facilitation of this learning process may promote such treatment's efficacy. Animal models suggested, that the stimulation of the vagus nerve or the superior colliculus (SC) – a midbrain structure mediating visual attentional processing – target these inhibitory extinction pathways and, thus, facilitate fear extinction. However, as it is unclear whether similar mechanisms exist in humans, this thesis manuscript examined how non-invasive stimulation of these inhibitory pathways by transcutaneous vagus nerve stimulation (tVNS) or SC-recruiting visual attentional manipulation impact on human fear extinction.

To this end, we conducted three studies using multiple-day single-cue fear conditioning and extinction paradigms. First, we elaborated on fear that is established in these paradigms by examining defensive responding that is elicited by an innocuous conditioned stimulus, which has either been paired (fear learning group) with an aversive unconditioned stimulus (US; an electric shock) or was unpaired (control group; study 1). During the following extinction training, either tVNS vs. sham stimulation was applied (study 1, study 2) or participants were instructed, to either generate saccadic eye movements (strong SC activation) vs. smooth eye pursuits (low SC activation; study 3). During subsequent sessions, extinction consolidation as well as the short- and long-term extinction recall was tested (study 2, study 3).

Conditioned fear in the fear learning group was characterized by elevated cognitive risk assessments (US-expectancy ratings), as well as increased cardiac deceleration and startle reflex potentiation compared to controls. Cardiac deceleration was positively correlated to startle potentiation, but was decoupled from cognitive risk assessments (study 1). Initial, short- and long-term extinction of these defensive responses was facilitated by tVNS on all three response levels (cognitive, physiological, behavioral; study 1, study 2). In contrast, saccades facilitated initial extinction only for physiological and behavioral elements of the defensive response pattern, while extinction consolidation and recall was impaired by any eye movement manipulation (study 3) for physiological and behavioral indicators of defensive responding.

Taken together, the data of the experimental series suggest, that on a behavioral level, conditioned fear may best be conceived as attentive immobility – a defense strategy elicited by inevitable distal threats, that is uniformly expressed across species and is accompanied by cardiac deceleration and startle reflex potentiation. In addition, it was shown that such rather automatic defensive adaptations are independent from verbally expressed threat expectancies. As expected, tVNS impacted on fear extinction on both levels, strongly in line with the suggestion, that vagal stimulation activates cortical and subcortical neural pathways involved in extinction learning, consolidation and recall. TVNS may, thus, be a promising adjuvant for exposure treatment of mental disorders. In contrast, SC-recruiting visual attentional manipulation only affected subcortically mediated defensive responding, in line with rodent findings, indicating that the SC specifically inhibits subcortical parts of the neural defense system. However, as extinction recall was impaired by any type of visual attentional manipulation, this appeared to have functioned as a form of avoidance, initially attenuating fear but preventing extinction consolidation and, thus, impairing sustained fear reduction. Both non-invasive stimulation techniques may therefore increase initial defensive flexibility in the face of no-longer threat-signaling stimuli, but only tVNS may achieve long-term effects on multiple response levels.

Zusammenfassung

Furcht ist ein affektiver Zustand, der durch die Aktivierung eines Defensivsystems gekennzeichnet ist, mit dem Ziel das Überleben des Organismus zu sichern. Dieses System ermöglicht sowohl das schnelle Erkennen von Bedrohungen, steuert aber auch defensive Reaktionsmuster um effektiv auf Bedrohungen der Umwelt zu reagieren. Um die Flexibilität dieses Systems unter wechselnden Umweltbedingungen zu gewährleisten, existieren jedoch ebenso hemmende Pfade, die die Aktivierung des Defensivsystems abschwächen, wenn ein zuvor bedrohlicher Reiz keine Gefahr mehr vorhersagt – ein Gedächtnis-bildender Prozess, der als Furchtextinktion bezeichnet wird und der zu einer Abschwächung defensiver Reaktionen führt. Es wird vermutet, dass Furchtextinktion zumindest teilweise der Expositionsbehandlung von Angststörungen zugrunde liegt, sodass eine Erleichterung dieses Lernprozesses die Effizienz dieser Therapieart steigern könnte. Tiermodelle deuten darauf hin, dass die Stimulation des Vagusnervs oder der Colliculi Superiores (SC) – eine Mittelhirnstruktur, welche die visuelle Aufmerksamkeitslenkung steuert – auf die hemmenden Extinktionspfade abzielt und somit Furchtextinktion begünstigt. Da unklar ist, ob ähnliche Mechanismen auch beim Menschen existieren, untersuchte diese Doktorarbeit, wie sich nicht-invasive transkutane Vagusnervstimulation (tVNS) oder SC-aktivierende Manipulation visueller Aufmerksamkeit auf die Furchtextinktion auswirken.

Zu diesem Zweck führten wir drei Studien durch, die mehrtägige Einfach-Reiz Furchtkonditionierungs- und -extinktionsparadigmen verwendeten. Zunächst haben wir die Furchtindikatoren, die in diesen Paradigmen gemessen werden, genauer validiert, indem wir Defensivreaktionen auf einen konditionierten Reiz auf mehreren Reaktionsebenen gemessen haben, nachdem dieser Reiz entweder gepaart (Furchtlerngruppe) mit einem aversiven unkonditionierten Reiz (US; ein elektrischer Schock) oder aber explizit ungepaart präsentiert wurde (Kontrollgruppe; Studie 1). Während des folgenden Extinktionstrainings am nächsten Tag

wurde entweder tVNS vs. eine Schein-Stimulation appliziert (Studie 1, Studie 2), oder die Probanden instruiert, Sakkaden (hohe SC Aktivierung) vs. Augenfolgebewegungen (niedrige SC Aktivierung) zu generieren (Studie 3). In nachfolgenden Sitzungen wurde die Konsolidierung der Extinktion, sowie der Kurz- und Langzeitabruf des Extinktionsgedächtnisses getestet (Studie 2, Studie 3).

In der Furchtlerngruppe löste der konditionierte Reiz eine erhöhte kognitive Bedrohungserwartung (US-Erwartungsratings), eine stärkere Herzratendezeleration sowie eine stärkere Potenzierung der Schreckreflexe, im Vergleich zur Kontrollgruppe, aus. Die Herzratendezeleration korrelierte positiv mit der Schreckreflexpotenzierung, war jedoch von der Bedrohungserwartung entkoppelt (Studie 1). Die initiale, Kurz- und Langzeitextinktion dieser Defensivreaktionen wurde durch tVNS auf allen drei Reaktionsebenen begünstigt (kognitiv, physiologisch, behavioral; Studie 1, Studie 2). Im Gegensatz dazu begünstigten Sakkaden nur die initiale Extinktion physiologischer und behavioraler Komponenten des defensiven Reaktionsmusters. Die Konsolidierung und der Abruf des Extinktionsgedächtnisses war, gemessen an physiologischen und behavioralen Defensivreaktionen, jedoch durch jede Manipulation von Augenbewegungen beeinträchtigt (Studie 3).

Zusammengenommen deuten die Daten der Versuchsreihe darauf hin, dass konditionierte Furcht funktionell am besten als ein Zustand attentiver Immobilität charakterisiert werden kann – eine Defensivstrategie, die durch unvermeidbare distale Bedrohungen ausgelöst wird und sich über verschiedene Spezies hinweg in ähnlicher Weise in einer Bradykardie und Potenzierung protektiver Reflexe manifestiert. Darüber hinaus zeigten wir, dass diese eher automatisierten Defensivreaktionen unabhängig von berichteter Bedrohungserwartung sind. Wie erwartet, begünstigte tVNS die Furchtextinktion auf beiden Ebenen, übereinstimmend mit der Annahme, dass Vagusstimulation kortikale und subkortikale Bahnen aktiviert, die bei Extinktionslernen,

-konsolidierung und -abruf beteiligt sind. TVNS könnte daher ein vielversprechender Wirkverstärker für die Expositionsbehandlung von Angststörungen darstellen. Im Gegensatz dazu beeinflusste eine SC-aktivierende visuelle Aufmerksamkeitsmanipulation nur subkortikal-gesteuerte Defensivreaktionen, übereinstimmend mit Tierbefunden, die anzeigen, dass die SC spezifisch subkortikale Teile des neuronalen Defensivsystems hemmen. Da der Abruf des Extinktionsgedächtnisses jedoch durch jede Art der visuellen Aufmerksamkeitsmanipulation beeinträchtigt war, fungierte diese scheinbar als Vermeidungsverhalten, das zunächst zwar Furcht reduziert, aber die Extinktionskonsolidierung und somit eine nachhaltige Furchtreduktion beeinträchtigt. Beide nicht-invasiven Stimulationstechniken scheinen somit im Angesicht von Reizen, die nicht länger eine Bedrohung anzeigen, die defensive Flexibilität initial zu erhöhten, aber nur tVNS kann langfristige Effekte auf mehreren Reaktionsebenen erzielen.

1 | Introduction

Fear is an emotional state, grounded in defense circuits in the brain that flexibly encode threats in the environment and organize defensive responses to efficiently evade harm (Hamm, 2020; Hamm & Flor, 2015; Lang et al., 1997). While these defensive mechanics are generally adaptive, hyperexcitability of the brain's defense system may, thus, constrain defensive flexibility and cause maladaptive, exaggerated defensive responding even towards innocuous cues (Rosen & Schulkin, 1998). Hence, such hyperexcitability was suggested to mediate the pathophysiology of anxiety, trauma- and stressor-related disorders, being largely featured by excessive fear-based symptoms (American Psychiatric Association, 2013; Rauch et al., 2006).

The preferred strategy to treat maladaptive defensive responding and restore defensive flexibility are exposure-based therapies, during which patients are repeatedly exposed to their fear-eliciting cues (Bandelow et al., 2016; Craske et al., 2014). Fear extinction learning, i.e., learning that a fear cue ("trigger") is no longer associated with the original threat, is presumed to be a central mechanism of action underlying the fear-reducing effects of this therapeutic regimen (Craske et al., 2014; D. Hermans et al., 2006). In fact, fear extinction is grounded in an inhibition of the brain's defense system (Tovote et al., 2015). Conversely, the stimulation of these inhibitory pathways might facilitate extinction and, thus, efficacy of exposure therapy.

The present thesis manuscript aimed at elaborating on how two non-invasive brain stimulation techniques, that presumably tap into neural extinction pathways, impact on human fear extinction: Transcutaneous vagus nerve stimulation and visual attentional manipulation. I will outline, how both techniques differently target these inhibitory neural circuits and discuss their effects on initial, short- and long-term extinction of defensive responding. Finally, I will highlight the clinical applicability of both techniques. To this end, the present work beforehand reviews previous research, which has examined the adaptive mechanics of the defense system regulating fear, as well as the underlying excitatory and inhibitory neural pathways.

2 | Defensive flexibility and its neural substrates

From a theoretical perspective, emotions such as fear may be conceived as action dispositions, that reflect the activation of two opponent motive systems, which evolved across species and guide behavior to sustain and protect the organism's life (Lang & Bradley, 2010; Lang & Davis, 2006). Pleasant emotions like joy reflect the activation of an appetitive motive system, driving approach responses towards potentially life sustaining stimuli (Lang & Bradley, 2010; Löw et al., 2008). Unpleasant emotions like fear, on the other hand, reflect the activation of a defense motive system, aiming to evade harm by threat and, thus, ensure survival (Lang & Bradley, 2010; Lang & Davis, 2006). To this end, the defense system orchestrates cognitive (e.g., feelings of fear and anxiety), physiological (e.g., cardiac acceleration) and behavioral defensive responses (e.g., flight) in patterns of overarching defense strategies, that may be broadly separated into defensive immobility (attentive vs. tonic immobility) and defensive action (fight or flight; Lang et al., 1998; Marks, 1987). Importantly, however, the efficacy of these strategies to cope with threats depends upon the behavioral options at hand as well as characteristics of the threat (Marks, 1987), which is why defensive flexibility is a prerequisite to efficiently evade harm. In both, animals and humans, the defense system therefore adapts the execution of defense strategies to the availability of escape options and the imminence of threat (Blanchard & Blanchard, 1989; Fanselow, 1994, 2018; Hamm, 2020; Lang et al., 1997).

2.1 Defensive flexibility: Threat imminence and escape options

Such defensive adaptation has been formalized in the *threat imminence* or *defense cascade model* for animals and humans, respectively (Fanselow, 1994; Lang et al., 1997). Upon encounters of inevitable but distal threat (e.g., a predator), animals' chief defense strategy is *attentive* immobility, during which the perceived threat cue is monitored while locomotion is inhibited (*freezing*), supported by a profound cardiac deceleration (Blanchard & Blanchard,

1989; Eilam, 2005; Fanselow, 1994; Kalin & Shelton, 1989; Kapp et al., 1979; Lang et al., 2000; Marks, 1987). At the same time, however, the animal is very responsive to abrupt changes of the environment – e.g., the startle reflex is potentiated – rendering attentive immobility a state of vigilant readiness, during which a probably fatal detection by the predator is evaded (Eilam, 2005; Fanselow, 1994; Hamm, 2020; Leaton & Borszcz, 1985; Marks, 1987; Roelofs, 2017). Attentive immobility, thus, sets the stage for rapid defensive response adaptations in case the threat imminence increases (e.g., the predator approaches) or an escape option eventually opens up. In such case, when attentive immobility would likely result in predation, the defense strategy switches to defensive action, where the animal is beyond vigilance, rather unresponsive to external stimuli and engaged in fight or flight, supported by a profound cardiac acceleration (Blanchard & Blanchard, 1989; Cannon, 1929; Eilam, 2005; Fanselow, 1994; Lang et al., 2000). Intriguingly, these defensive responses seem well preserved in humans: Inevitable distal threats have shown to evoke reduced body sway, increased attentional processing of the threat, cardiac deceleration and startle reflex potentiation alongside moderate levels of self-reported fear (Gladwin et al., 2016; Hamm, 2020; Kolassa et al., 2005; Lang et al., 1997; Löw et al., 2015; Mobbs et al., 2009). Upon increasing threat imminence and available avoidance options, yet, the defensive pattern is marked by active avoidance behavior, reduced sensory intake, cardiac acceleration and startle reflex inhibition alongside high levels of self-reported fear (Hamm, 2020; Lang et al., 1997; Löw et al., 2015; Mobbs et al., 2009).

2.2 Defensive flexibility: Fear acquisition

However, defensive flexibility not only involves the adaptation of defensive responses to threat imminence and escape options, but also the flexible activation of fear towards inherently innocuous stimuli, that have acquired the attribute of a threat signal (Hamm et al., 1993; Watson & Rayner, 1920). Such acquisition of fear towards previously innocuous stimuli has primarily

been investigated by Pavlovian fear conditioning procedures (Lonsdorf et al., 2017). Here, the organism is in a context (e.g., laboratory environment), where it receives repeated pairings of an emotionally neutral stimulus (conditioned stimulus, CS; e.g., a visual stimulus) and a harmful event (unconditioned stimulus, US; e.g., an electric shock; **Figure 1**; Lonsdorf et al., 2017; Pavlov, 1927). As a result, the CS, US, and the environmental context are encoded into an associative fear memory (see **Figure 1**), representing the predictive value of the CS concerning an upcoming US (CS = US) in due consideration of contextual aspects (Lonsdorf et al., 2017; Rescorla & Wagner, 1972). Future encounters of the CS may eventually prompt a fear memory recall, that evokes an activation of the defense system and, thus, *conditioned* defensive responding towards predicted or expected upcoming threat (Hamm & Vaitl, 1996; Lonsdorf et al., 2017; Pittig et al., 2020; Rescorla & Wagner, 1972).

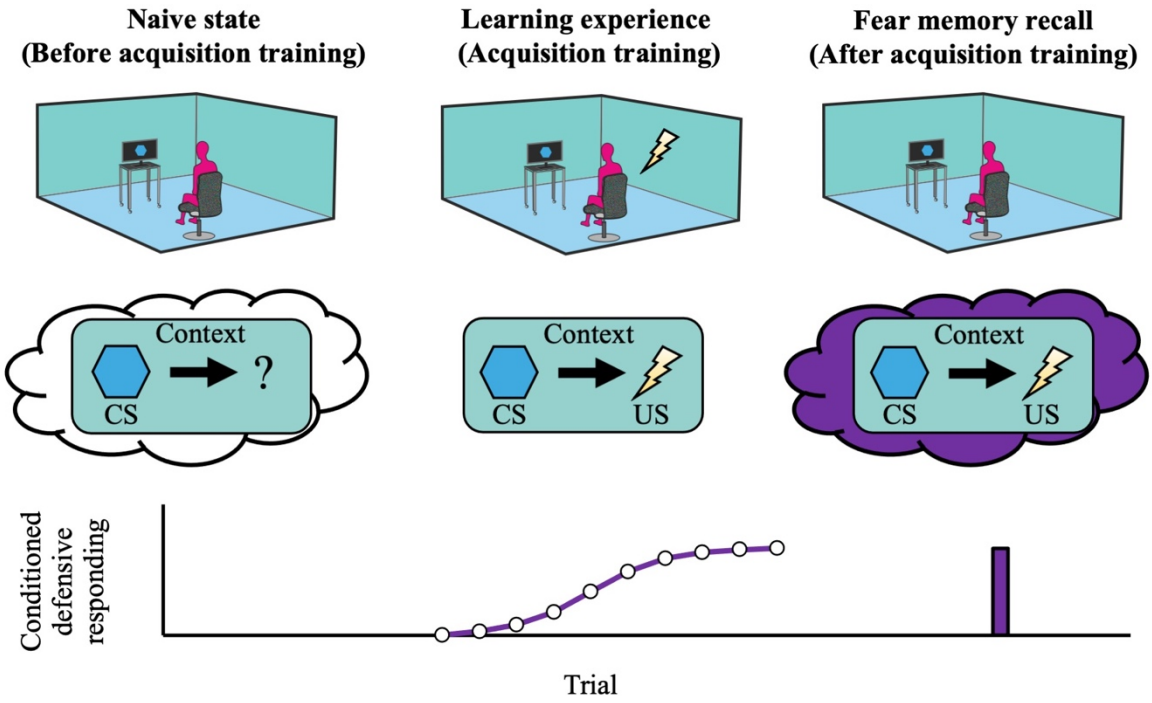


Figure 1. Stimulus presentation (upper panel), learned stimulus associations (middle panel) and conditioned responding (lower panel) before, during and after a fear acquisition training in a single experimental context. Clouds represent the respectively established memory (naive = white, fear memory = purple) and incorporate the learned stimulus associations.

2.3 Defensive flexibility: Fear extinction

Importantly, however, just as defensive flexibility embraces a facilitated fear activation towards a cue that previously predicted threat, it also comprises a reduction of defensive responses, if such previously threat-signaling cue no longer predicts any danger. Indeed, such reduction of defensive responding – referred to as fear extinction – is observed, when the organism is in a context (e.g., laboratory environment), where it is repeatedly exposed to a previously threat-signaling stimulus (CS) that is no longer followed by an aversive outcome (US; see **Figure 2**; Dunsmoor et al., 2015; D. Hermans et al., 2006; Myers & Davis, 2002).

Pavlov, who first described this phenomenon nearly a century ago, has considered an "internal inhibition" the underlying mechanism of extinction, by which the previously established conditioned responses are disrupted (Dunsmoor et al., 2015; Pavlov, 1927). Contemporary research captures this view and presumes that fear extinction – similarly to fear conditioning – is founded in a context-specific associative learning process (*fear extinction learning*), that leads to an inhibition of the defense system regulating fear (Bouton, 2014; Dunsmoor et al., 2015). Specifically, repeated prediction errors (CS ≠ US) are suggested to drive active encoding of the CS, US-omission and contextual information into a new associative extinction memory, representing the reduced predictive value of the CS concerning upcoming danger (CS = US-omission), again in due consideration of the contextual aspects (Bouton, 2014; Bouton & Woods, 2008; Dunsmoor et al., 2015). Future encounters of the previously threat-signaling CS may consequently prompt a recall of this extinction memory, which inhibits the concomitant activation of the original fear memory trace (Bouton, 2004; Craske et al., 2014; Dunsmoor et al., 2015). In this view, extinction is therefore no "forgetting" of the original fear memory. Rather, fear is still installed and extinction recall may be good or poor depending on the balance of extinction against fear memory (see **Figure 2, 3**; Dunsmoor et al., 2015; Quirk & D. Mueller, 2008).

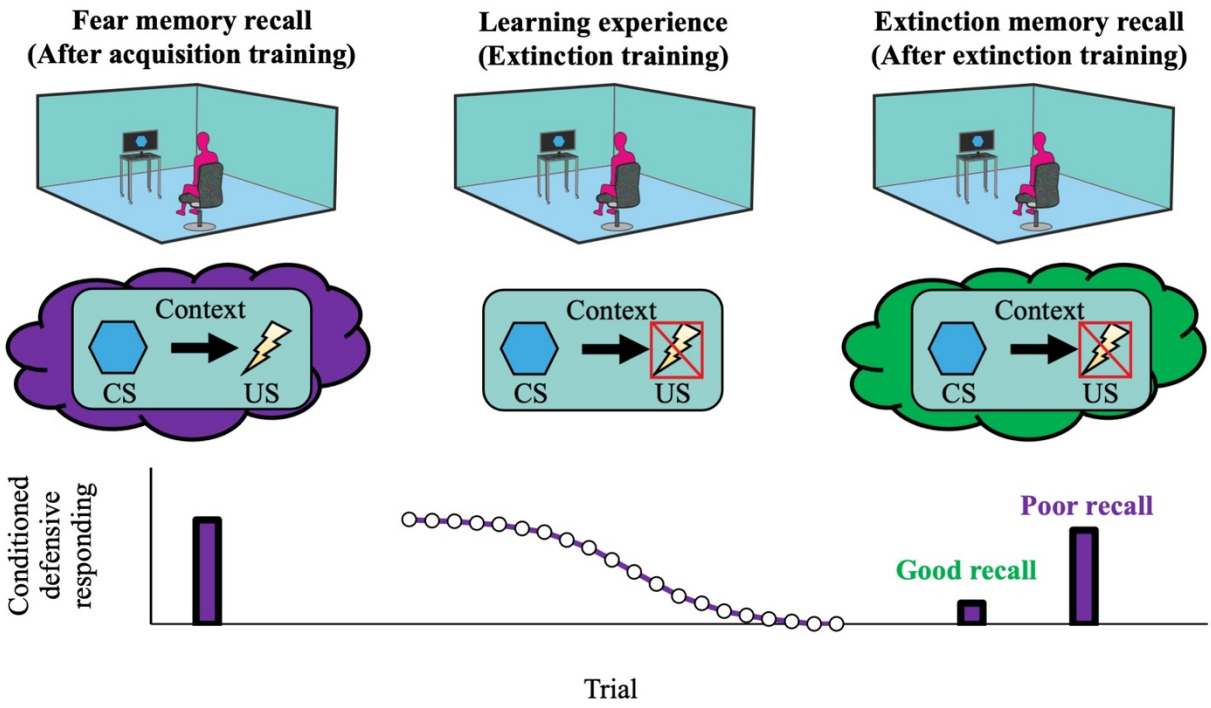


Figure 2. Stimulus presentation (upper panel), learned stimulus associations (middle panel) and conditioned responding (lower panel) before, during and after a fear extinction training in a single experimental context. Clouds represent the respectively established memory (fear memory = purple, extinction memory = green) and incorporate the learned stimulus associations.

Supporting this view, fear may rapidly return if the original fear memory is – even only partly – reactivated: Both another CS-US pairing and even a sole re-experience of the US have shown to result in *reacquisition* or *reinstatement* of fear, respectively (**Figure 3**; Haaker et al., 2014; Hollandt et al., 2020; Lonsdorf et al., 2017). On the other hand, however, fear may also return if the inhibitory impact of extinction memory is attenuated, e.g., by contextual shifts due to extinction's contextual specificity: Fear *renewal* is observed if extinction memory is recalled in a context that differs from the environment of initial extinction learning (Effting & Kindt, 2007), while *spontaneous recovery* of fear may specifically result after a temporal contextual shift, i.e., if time passed after the extinction training (**Figure 3**; Baum, 1988; Bouton, 2014;

Myers & Davis, 2002). Thus, although extinction may reduce inappropriate defensive responses, such flexible inhibition of the defense system is far more fragile and transient compared to well-conserved fear activation (Dunsmoor et al., 2015; Solomon & Wynne, 1954). Concerning defensive flexibility, fear activation is hence adaptively prioritized over extinction to avoid a potentially fatal misinterpretation of threat-signals as innocuous – anecdotally referred to as "better safe than sorry" approach – which is why only well-consolidated extinction memory may lead to long-term fear reduction (Dunsmoor et al., 2015; D. Mueller & Cahill, 2010; Quirk & D. Mueller, 2008; Van den Bergh et al., 2021).

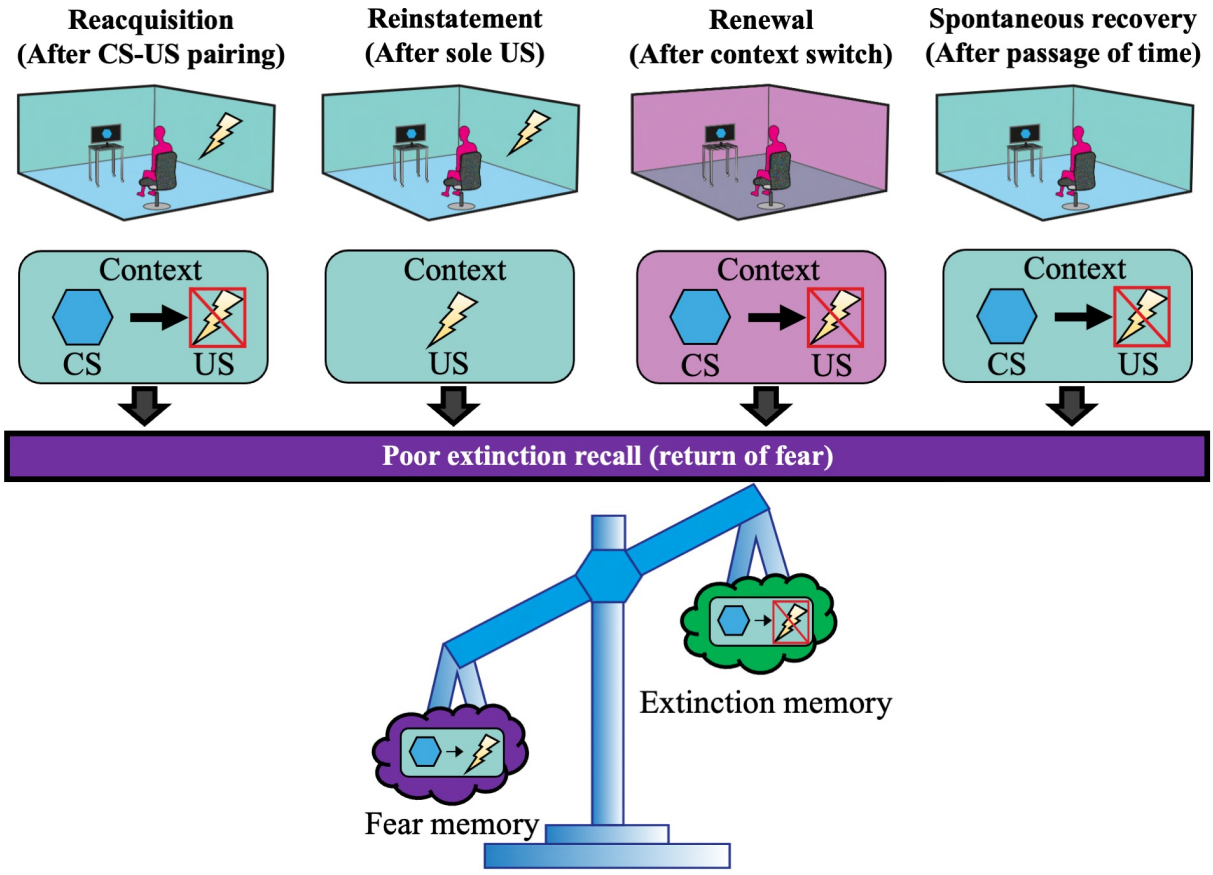


Figure 3. Stimulus presentation (upper panel) and experiences (middle panel) that may lead to poor recall of extinction memory (return of fear): Reacquisition (after another CS-US pairing), reinstatement (after re-experiencing solely the US), renewal (after contextual switches) or spontaneous recovery (after the passage of time). Lower panel: Dominance of fear (purple cloud) over extinction memory (green cloud) after the respective experiences.

2.4 The neural site of defensive flexibility: The brain's defense system

In the past decades, fear conditioning and extinction protocols have extensively been used to study the neural substrates that underlie the fear-regulating defense system as well as its inhibition, pinpointing to partially independent, yet reciprocally connected brain regions (Lang & Bradley, 2010; LeDoux & Brown, 2017; Tovote et al., 2015). The sensory interface of the defense system has been located on a subcortical level: The basolateral complex of the amygdala (BLA), comprising a lateral (LA) and basal nucleus (BA), receives threat-related sensory information by cortical and thalamic fibers (Davis & Whalen, 2001; Lang & Bradley, 2010; LeDoux & Daw, 2018; Tovote et al., 2015). The output site, yet, is the central nucleus of the amygdala (CeA), which receives threat-related information from the LA, directly or indirectly via the BA and intercalated cells (ITC), and orchestrates behavioral, autonomic and endocrine defensive responses (**Figure 4**; Davis & Whalen, 2001; Duvarci & Pare, 2014; LeDoux & Daw, 2018). In contrast to these subcortically mediated threat adaptations, feelings of fear are suggested to emerge from the activity of a general network of cognition (GNC), which embraces the posterior parietal, cingulate and frontal cortex as well as the insula (**Figure 4**; LeDoux, 1995; LeDoux & Brown, 2017).

2.4.1 Neural adaptation to threat imminence and escape options

In line with behavioral observations, the activation of these defensive circuits adapts to threat imminence and escape options to allow a flexible execution of defense strategies, if harm can thus be evaded more efficiently. At this, switching between defensive immobility and defensive action has found to be grounded in different subsets of CeA neurons (Fadok et al., 2017), which presumably project to different anatomical targets, e.g., the ventrolateral periaqueductal gray for behavioral immobility and the dorsolateral periaqueductal gray for fight/flight (**Figure 4**; LeDoux & Daw, 2018; Roelofs, 2017; Tovote et al., 2016).

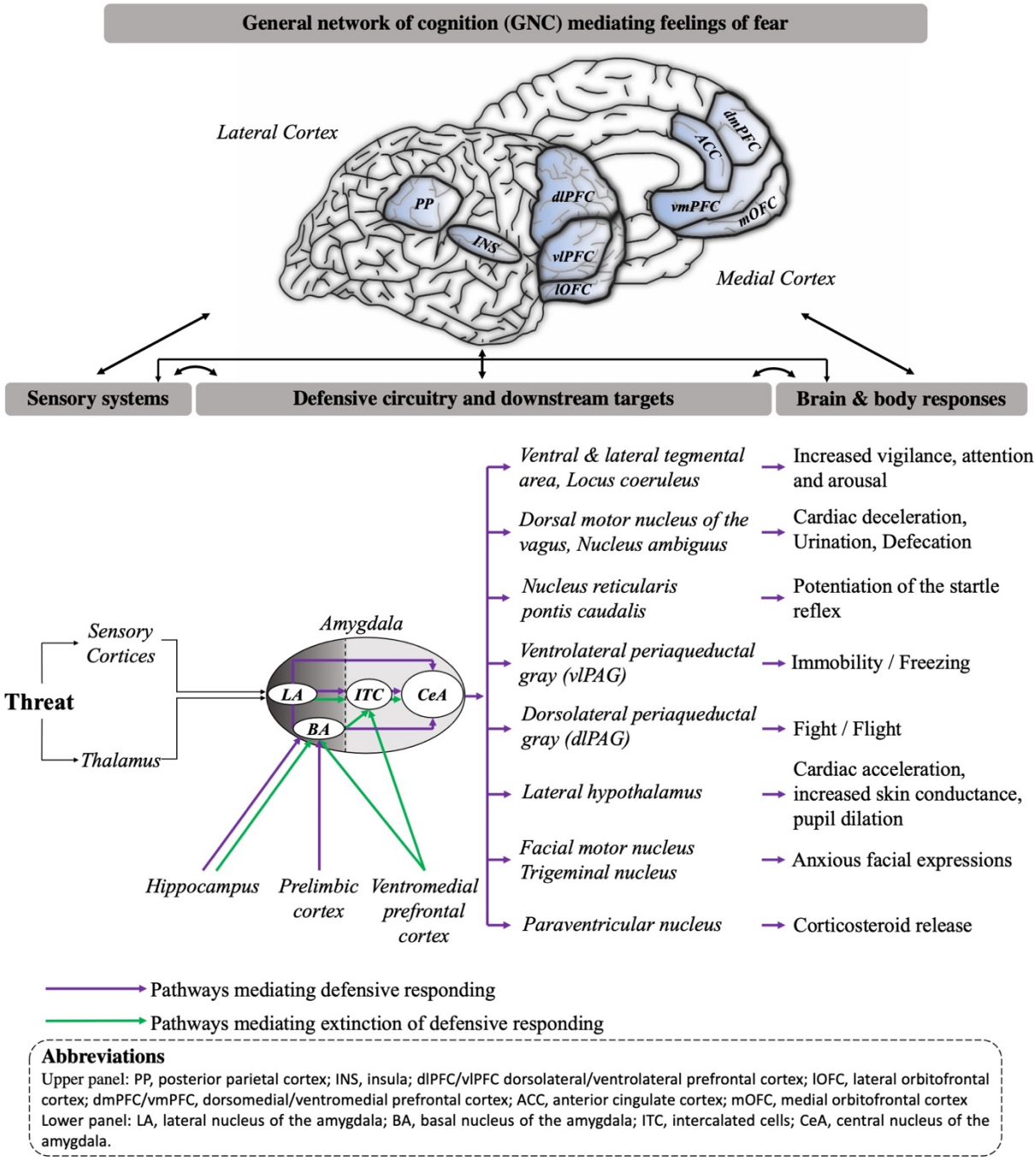


Figure 4. Upper panel: Brain areas mediating feelings of fear. Lower panel: Neural circuitry mediating behavioral and physiological defensive response patterns. Purple lines depict neural projections mediating defensive responding, while green lines depict neural projections mediating extinction of defensive responding. Note that the LA and BA combine to the basolateral complex of the amygdala, shaded dark gray in the figure. [Adapted and modified from Davis & Whalen (2001), Ledoux & Brown (2017), LeDoux & Daw (2018) and Tovote et al. (2015)].

2.4.2 Neural adaptation during fear acquisition

While the CeA therefore plays a pivotal role in adjusting defensive output, the basolateral amygdala is the critical site that adapts general fear activation to environmental conditions, e.g., towards stimuli that had acquired a threat-signaling effect via fear conditioning (Tovote et al., 2015). At this, sensory information about a conditioned threat-signal (CS) and an unconditioned threat (US) converge in the BLA (**Figure 4**; LeDoux & Daw, 2018). Here, repeated CS-US pairings increase firing and long-term excitability in a population of dedicated "*fear*" neurons, which ultimately facilitates fear activation by the CeA during CS encounters (Herry et al., 2008; Nabavi et al., 2014; Rogan et al., 1997; Sah & Westbrook, 2008; Senn et al., 2014; Tovote et al., 2015). Both the prelimbic prefrontal cortex and the hippocampus project to these "*fear*" neurons (**Figure 4**; Milad & Quirk, 2012; Senn et al., 2014). However, while the prelimbic cortex thereby mediates the recall of fear memory, the hippocampus relays additional contextual information that facilitates fear activation, e.g., whether the CS is likely to predict threat in the given environment (Herry et al., 2008; LeDoux, 1995; LeDoux & Daw, 2018; Milad & Quirk, 2012; Ramirez et al., 2013; Tovote et al., 2015).

2.4.3 Neural adaptation during fear extinction

Yet, the BLA is not only essential to regulate fear activation towards previously threat-predicting cues, but is also the critical site of fear extinction (Tovote et al., 2015). Here, the activity and long-term excitability in a population of dedicated "*extinction"* neurons is ramped up, which inhibit fear activation by the CeA during CS encounters – presumably by way of indirect projections via inhibitory intercalated cells (**Figure 4**; Amano et al., 2010; Duvarci & Pare, 2014; Herry et al., 2008; Likhtik et al., 2008; Sah & Westbrook, 2008; Senn et al., 2014; Tovote et al., 2015). Both the infralimbic region of the ventromedial prefrontal cortex (vmPFC) and the hippocampus project to these "*extinction*" neurons (**Figure 4**; Milad & Quirk, 2012;

Senn et al., 2014). However, while the vmPFC thereby mediates the recall of extinction memory (in addition to directly activating intercalated cells), the hippocampus conveys contextual information that dampens fear activation – e.g., whether the CS is likely to signal the omission of threat in the given environment (Jinzhao & Maren, 2007; LeDoux, 1995; Likhtik et al., 2008; Maren et al., 2013; Milad & Quirk, 2012; Quirk et al., 2000; Sah & Westbrook, 2008; Senn et al., 2014; Tovote et al., 2015; see also E. M. Mueller et al., 2014). At last, it is the balance between "*extinction*" and "*fear*" neuron activation in the BLA that determines the activation of the CeA and, thus, defensive output on targeted central and peripheral systems (see **Figure 4**; Herry et al., 2008; Senn et al., 2014; Tovote et al., 2015).

2.5 Increasing defensive flexibility by non-invasive brain stimulation

However, as mentioned above, the defense system not only targets these systems to regulate defensive output, but also receives information by them (Lang & Bradley, 2010). Intriguingly, such reciprocal connection exists between extinction pathways and the autonomic nervous system, which is why extinction may impact on autonomic arousal (e.g., reduced arousal towards previously threat-predicting cues; Vervliet et al., 2004), just as autonomic arousal may in turn impact on extinction (Giustino & Maren, 2018; D. Mueller & Cahill, 2010). In fact, the release of peripheral adrenaline in arousing situations evokes increased noradrenergic transmission in both the BLA and vmPFC, which improves extinction learning, consolidation and recall presumably by facilitating the underlying neural changes (Berlau & McGaugh, 2006; Duvarci & Pare, 2014; Duvarci & Paré, 2007; Giustino & Maren, 2018; McIntyre et al., 2012; D. Mueller & Cahill, 2010). The vagus nerve – a cranial nerve acting as a major autonomic communication route between the body periphery and the brain – has shown to be a critical element mediating this arousal effect (Berthoud & Neuhuber, 2000; McIntyre et al., 2012).

Eighty percent of the vagus nerve are sensory fibers, transporting information about the activity of the inner body to the brain, while only twenty percent are efferent fibers regulating this activity (Berthoud & Neuhuber, 2000; Foley & DuBois, 1937; Henry, 2002; Nemeroff et al., 2006; Silvani et al., 2016; Thayer & Sternberg, 2006). Peripheral adrenaline activates these sensory vagal fibers, which feed forward the activation via the nucleus tractus solitarius in the brainstem to the primary hub of noradrenaline in the brain – the locus coeruleus – which ultimately increases noradrenergic transmission in the BLA and vmPFC (see **Figure 5**; Jones et al., 1977; McIntyre et al., 2012; Miyashita & Williams, 2004, 2006; see McGaugh, 2018 and D. Mueller & Cahill, 2010).

Novel approaches have attempted to mimic the effects of peripheral arousal on extinction whilst circumventing the necessity of actual arousal, by invasively stimulating the vagus nerve (iVNS; **Figure 5)**: In fact, iVNS enhanced (noradrenergic) transmission in the BLA and vmPFC in animals (Follesa et al., 2007; George et al., 2000; Hassert et al., 2004; McIntyre et al., 2012; Nemeroff et al., 2006; Peña et al., 2014), and promoted extinction learning and recall: The decrease in behavioral freezing and startle potentiation in the face of a previous threat-signal (CS) was significantly faster and long-lasting in vagally stimulated rodents compared to sham-stimulated control animals (Noble et al., 2017, 2019; Peña et al., 2013, 2014).

However, the vagus nerve is not the only route, by which peripheral events may impact on the brain's extinction pathways. A further reciprocal connection has been identified between the amygdala and neural systems mediating visual attention (**Figure 5**; Gallagher & Holland, 1994; Koller et al., 2019; Öhman, 2005; Öhman et al., 2007; Vuilleumier, 2005). As a result, extinction may impact on attentional processes (e.g., disengagement from previously threat-predicting cues; Muench et al., 2016; Van Damme et al., 2006), just as attentional processes may in turn impact on extinction (see Barry et al., 2016; Panitz et al., 2019).

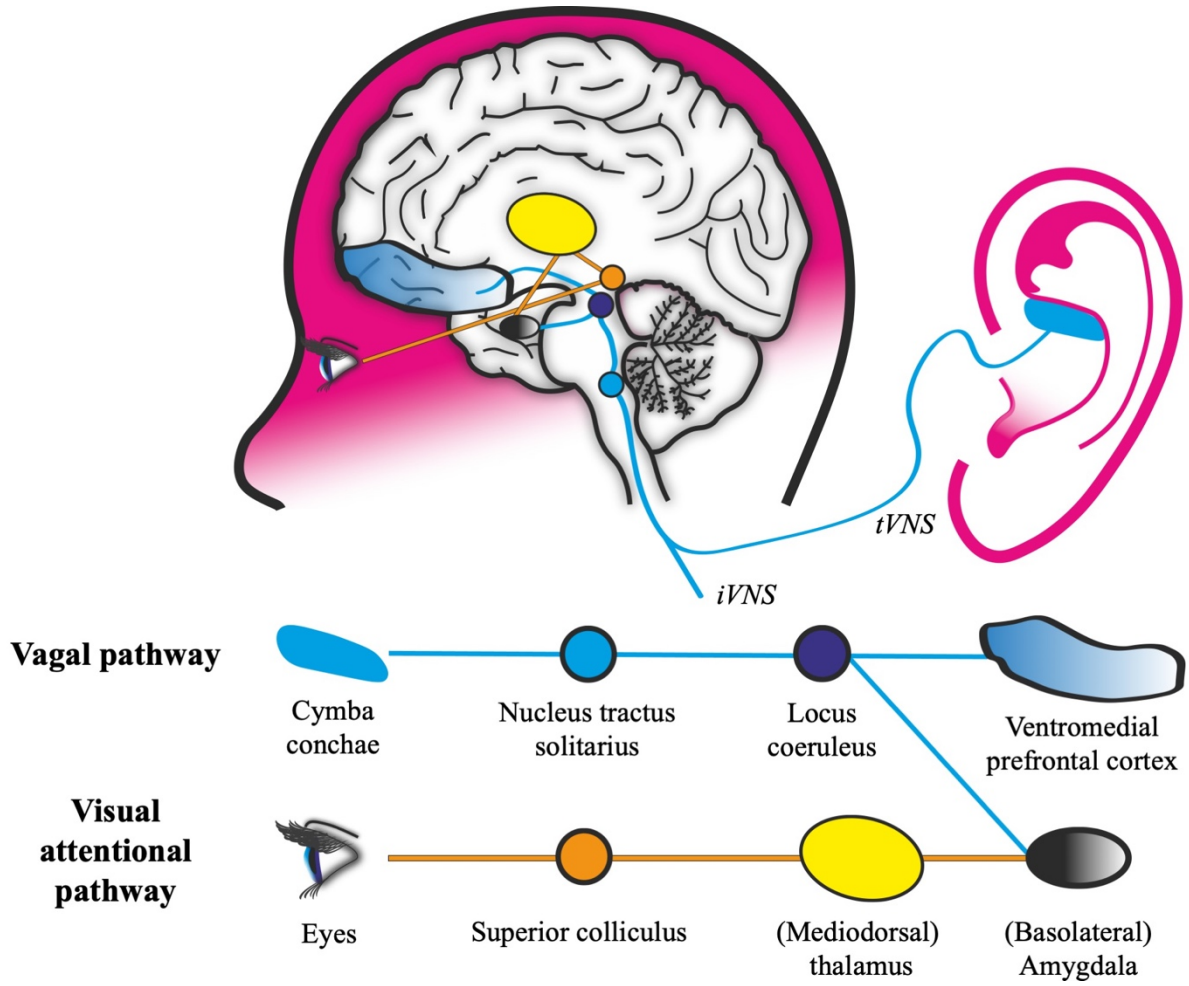


Figure 5. Neural pathways, by which vagal stimulation or visual attentional manipulation are suggested to impact on the brain's extinction network

In particular, diverting attention from previously threat-predicting cues has shown to foster the inhibition of fear responses, when they are no longer appropriate to the situation, which is likely driven by promoted disengagement from threat and, thus, attenuated fear activation (Badura-Brack et al., 2015; Barry et al., 2015; Podină et al., 2013). A critical relay underlying this effect might be the superior colliculus (SC) in the midbrain, that decodes salient stimuli, guides visual attention and has shown to indirectly project to the BLA via the mediodorsal thalamus (see **Figure 5**; Anderson & Rees, 2011; Baek et al., 2019; Herry & Garcia, 2002; Krauzlis, 2004; Krauzlis et al., 2013; Lee & Shin, 2016; Leigh & Zee, 2015; Müller et al., 2005;

Oyoshi et al., 1996; Sommer & Wurtz, 2006, 2004; White et al., 2017). Accordingly, it was suggested that SC-projections might gate threat-related sensory input from the thalamus to the BLA and, thus, inhibit fear activation when attention is diverted from previous threat-signals (Baek et al., 2019). Recent animal research tested this hypothesis and found, that alternating bilateral visual stimulation (a moving light) during the presentation of a previously threat-predicting tone in fact increased SC-activity, which evoked an inhibition of "*fear*" cell activity in the basolateral amygdala via a modulation of mediodorsal thalamic inputs (**Figure 5**; Baek et al., 2019). As a result, the decrease in behavioral freezing was significantly faster and long-lasting in SC-stimulated rodents relative to sham-stimulated control animals (Baek et al., 2019).

As both vagal stimulation and SC-recruiting visual attentional manipulation were capable to facilitate fear extinction, these techniques might be highly relevant for clinical practice (Milad et al., 2014): Extinction is considered one central mechanism underlying exposure-based treatment of anxiety, trauma- and stressor-related disorders, as during such regimen repeated fear cue exposures prompt learning, that a fear-eliciting cue is no longer associated with actual threat (Bandelow et al., 2016; Craske et al., 2014; Craske & Mystkowski, 2006; Heinig et al., 2017; Pittig et al., 2016, 2021; Richter et al., 2017; Scheveneels et al., 2016; Vervliet et al., 2013). Techniques that facilitate extinction might, hence, serve as promising treatment adjuvants to promote exposure therapy, i.e., increase responding to treatment and reduce the likelihood of relapses after the treatment (Craske et al., 2014; Heinig et al., 2017; Pittig et al., 2021). Especially concerning that extinction deficits are prevalent in anxiety, trauma- and stressor-related disorders and are viewed a risk factor for non-responding to treatment and return of fear, a facilitation of such inhibitory learning process appears even more important to cope with the hyperexcitability of the brain's defensive system, that is suggested to mediate the pathophysiology of these disorders (Craske & Mystkowski, 2006; Duits et al., 2015; Koenigs & Grafman, 2009; Lissek et al., 2005; Milad et al., 2009; Rauch et al., 2006; Rosen & Schulkin, 1998; Vervliet et al., 2013).

As potential facilitators of defensive response inhibition and exposure therapy, vagal stimulation and SC-recruiting visual attentional manipulation have accordingly started to gain increasing attention in human research as well (Cimpianu et al., 2016; George et al., 2000; Phaf et al., 2021). In human research, vagal activation is usually achieved by a non-invasive stimulation technique – transcutaneous vagus nerve stimulation (tVNS) – which refers to the electrical excitation of the auricular branch of the vagus nerve at the ear's *cymba conchae*, describing an exclusively vagally innervated skin area (**Figure 5**; Ellrich, 2011; Frangos et al., 2015; Peuker & Filler, 2002). On the other hand, SC-recruiting visual attentional manipulation may be operationalized by inducing saccadic eye movements, which directly tap into the SC (Krauzlis, 2004; Krauzlis et al., 2013; Leigh & Zee, 2015). In spite of the promising animal research, however, evidence indicating that tVNS or SC-recruiting visual attentional manipulation facilitates fear extinction in humans is scarce (see Burger et al., 2016, 2017, 2018; Genheimer et al., 2017), despite tVNS has already proven to evoke comparable changes in neural activity as its invasive counterpart (Ellrich, 2011; Frangos et al., 2015; Ventura-Bort et al., 2018, 2021)

Thus, the major aim of the current research project was to investigate the initial, short- and long-term effects of tVNS and SC-recruiting visual attentional manipulation on human fear extinction. On a longer run, the findings of the present dissertation may, thus, shed a light on the capacity of both tVNS and SC-recruiting visual attentional manipulation to promote flexible defensive response adaptation in the face of stimuli that no longer predict danger and may therefore help to evaluate both techniques as treatment adjuvants to promote exposure therapy.

3 | Facilitating fear extinction by non-invasive brain stimulation

The current research project followed up on previous human studies that endeavored a cross-species transfer of facilitated extinction by vagal stimulation, which was, however, only partly successful (Burger et al., 2016, 2017, 2018): While tVNS facilitated the extinction in cognitive indices of defensive activation, behavioral indices and the recall of extinction memory remained unaffected, which is in stark contrast to previous animal research (see Noble et al., 2017, 2019; Peña et al., 2013, 2014). Importantly, however, these human studies tested the effects of vagal stimulation by using a differential conditioning paradigm, which requires more complex discriminative learning processes between two conditioned stimuli [one paired (CS+) and the other unpaired (CS-) with the US during fear acquisition] and also targets different neural circuitry compared to single-cue conditioning and extinction designs, that were used in the animal model (Carew et al., 1983; Knight et al., 2004; Lonsdorf et al., 2017; Noble et al., 2017, 2019; Peña et al., 2013, 2014). This might be one reason for the hampered transferability of results and, thus, harmonizing methodology might be very important, if fear extinction enhancement by brain stimulation techniques is tested across species (Haaker et al., 2019).

This applies to the use of a comparable learning task. In the dissertation research program, we therefore tested the impact of non-invasive brain stimulation on fear extinction by conducting three studies that used a multiple-day single-cue fear conditioning and extinction paradigm, closely adapted to animal research (Szeska et al., 2020, 2021, under review). Here, a single CS (e.g., a visual geometric figure or acoustic 1000 Hz sine tone), was paired with an aversive US (a shock) during acquisition training in one group (the fear learning group), while the CS and US were presented explicitly unpaired in the other group (the control group). Learning effects are then tested in a between group comparison (see Lonsdorf et al., 2017; Rescorla, 1967). During the following extinction training and extinction recall phases, the CS is finally presented without the administration of the US (Lonsdorf et al., 2017).

Moreover, harmonizing methodology also applies to choosing those behavioral and physiological indicators of fear that are comparable across different mammals (Haaker et al., 2019).

3.1 Attentive immobility in the face of inevitable distal threat

In non-human animal research that examined fear extinction and extinction enhancement by brain stimulation, learned fear is indicated by a defensive response pattern that embraces expressions of attentive immobility – the chief defense strategy towards cues signaling inevitable threats, such as the CS in fear conditioning (Baek et al., 2019; Fanselow, 1994; Lang & Bradley, 2010; Noble et al., 2017, 2019; Peña et al., 2013, 2014). At this, changes in associated behavioral freezing towards a conditioned cue serve as primary read-out of learned fear (Baek et al., 2019; Noble et al., 2017, 2019; Peña et al., 2013, 2014). However, conditioned fear in animals is also often measured by two other indices, that have been associated with attentive immobility and indeed correlate to behavioral freezing: the fear potentiated startle and fear bradycardia (Gewirtz et al., 1997; Haaker et al., 2019; Plappert et al., 1993; Walker & Carrive, 2003). While the fear potentiated startle describes the facilitation of the protective startle reflex, fear bradycardia refers to a profound cardiac deceleration, both of which are reliably observed when animals process a threat-signaling CS (Brown et al., 1951; Campbell et al., 1997; Haaker et al., 2019; Landis & Hunt, 1939; Lang & Davis, 2006; Leaton & Borszcz, 1985; Plappert et al., 1993; Walker & Carrive, 2003). As both defensive responses are also modulated by the central nucleus of the amygdala, they have accordingly been viewed as respective behavioral and physiological read-outs of fear activation in animals (see **Figure 4**; Davis, 2006; Hamm, 2015; Kapp et al., 1979; Kuhn et al., 2019; Roelofs, 2017; Walker & Carrive, 2003; Weike et al., 2005).

In human conditioning research, where CSs also signal inevitable distal threats, the conditioned response pattern accordingly embraces expressions of attentive immobility, including the fear potentiated startle and fear bradycardia (Gladwin et al., 2016; Haaker et al., 2019; Hagenaars et al., 2014; Hamm, 2020; Hamm et al., 1993; Kuhn et al., 2019; Lonsdorf et al., 2017). However, while human research hence interpreted startle potentiation as a translational index of fear, this was not the case for cardiac deceleration: Instead, it has been argued for a long time that a decrease in heart rate reflects increased orienting towards the CS rather than conditioned fear, as it has been observed to emotionally neutral but significant stimuli as well (Graham, 1979; Graham & Clifton, 1966; Lonsdorf et al., 2017). However, if conditioned fear in humans is conceived as a cross-species instance of attentive immobility, which has been incent by Lang and colleagues (1997, 2000), fear and orienting are not mutually exclusive. Rather, conditioned fear may then be conceptualized as state of anxious apprehension that is critically featured by increased orienting towards a CS that signals inevitable distal threat, which is why cardiac deceleration might in fact index both emotion and attention (Bradley, 2009; Muench et al., 2016; Panitz et al., 2015).

In the first publication of the dissertation research program, we tested such hypothesis that conditioned fear in humans is conceivable as a cross-species instance of attentive immobility in order to get a more thorough understanding of the conditioned responses, whose extinction might be facilitated by non-invasive brain stimulation (Szeska, Richter, Wendt, Weymar and Hamm, 2021). In this study, we used a single cue fear conditioning paradigm and compared the coupling of cardiac deceleration and startle potentiation, evoked by a visual CS either signaling an aversive US (fear learning group) or not (control group). On the one hand, we expected a close coupling between cardiac deceleration and startle potentiation, as both responses have served as primarily subcortically mediated indices of attentive immobility in animals (see **Figure 4**; Hamm, 2020). In contrast, we expected (if at all) only a moderate correspondence between these subcortically mediated threat adaptations and more cognitive indices like the

declarative knowledge of CS-US contingencies (US-expectancy ratings), that reflects cognitive risk assessment as an antecedent of cortically mediated feelings of fear (LeDoux, 1995; LeDoux & Brown, 2017; Reisenzein, 2009).

In line with previous research (Hollandt et al., 2020), Szeska and colleagues (2021) found increased US-expectancy and startle reflex potentiation in the fear learning group relative to controls during the acquisition training. As was expected, cardiac deceleration was found in response to the motivationally significant threat-signaling CS (fear learning group) and to a safety-signaling CS (control group), but was more pronounced during the threat signaling CS in the fear learning group. Intriguingly, this stronger bradycardia was positively correlated to subcortically mediated startle potentiation, but was decoupled from cortically mediated US-expectancy ratings.

In contrast to previous views in human fear conditioning research (for a review see Lonsdorf et al., 2017), the data therefore indicate that cardiac deceleration is a valid index of attention and emotion in humans – more specifically, a primarily subcortically mediated index of attentive immobility similar to the fear potentiated startle (Bradley, 2009; Bradley et al., 2018; Lang et al., 1990). This view is in line with previous findings, showing that fear bradycardia, startle potentiation and behavioral immobility are jointly modulated by projections from the CeA to the ventrolateral periaqueductal gray and, therefore, highly interrelated during threat processing (Gladwin et al., 2016; E. J. Hermans et al., 2013; Kuhn et al., 2019; Leaton & Borszcz, 1985; LeDoux et al., 1988; Plappert et al., 1993; Walker & Carrive, 2003). Moreover, these findings are also supported by evidence from behavioral genetics, showing that the short-allelic variant of the 5-HTTLPR (serotonin transporter-linked polymorphic region) – a polymorphism linked to increased connectivity between amygdala and the periaqueductal gray – is also associated with both increased fear bradycardia and fear potentiated startle (Lonsdorf et al., 2009; Schipper et al., 2019). Importantly, however, cognitive indices of defensive

response activation seem (at least partly) independent from such closely coupled low-level indices of attentive immobility, which suggests that such cognitive indices are mediated by processes that are partially independent from those mediating low-level behavioral and physiological defensive responding (Hollandt et al., 2020; Lang et al., 1983). Future research should therefore use such multi-level approaches for a thorough investigation of fear, also allowing a disentanglement of anxious apprehension as a reflection of attentive immobility, marked by cardiac deceleration and startle potentiation, from anxious apprehension as a reflection of defensive action, that is associated with cardiac acceleration and startle inhibition (Löw et al., 2015). This is relevant for psychophysiological research, aimed at assessing mental processes by physiological and behavioral responses (Hamm, 2020). The most important implication that we draw from this study, however, is for basic science: In line with Lang's view (1997, 2000), the data indicate that conditioned fear in humans is conceivable as a cross-species instance of attentive immobility (Eilam, 2005; Marks, 1987; Walker & Carrive, 2003).

Following up on these basic paradigmatic findings we used a single-cue conditioning design to further investigate, whether extinction of behavioral and physiological indicators of attentive immobility and cognitive risk assessments could be modified by non-invasive brain stimulation.

3.2 Facilitated extinction by transcutaneous vagus nerve stimulation

In the second publication of the dissertation research program, we used the single-cue fear conditioning and extinction paradigm, that is described above, and tested the capacity of tVNS to facilitate extinction of low-level indices of attentive immobility (cardiac deceleration, startle potentiation) and cognitive risk assessments. To this end, participants first underwent an acquisition training, during which participants of the fear learning group received paired presentations of a visual CS and an aversive US, while the control group received unpaired presentations of both stimuli. Twenty-four hours later, both groups underwent an extinction

training. Here, tVNS was applied randomized and sham-controlled: Respective halves of the fear learning and control group received active stimulation of the left vagally innervated *cymba conchae* (tVNS) or a stimulation of the left non-vagally innervated earlobe (sham stimulation; Burger et al., 2016; Peuker & Filler, 2002). Extinction recall was tested after 24 hours and four weeks, which allowed the examination of short- and long-term effects of tVNS. Additionally, resistance of extinction memory against reinstatement of fear and, thus, extinction memory consolidation was tested during recall phases by non-signaled US presentations (Haaker et al., 2014). Based on previous animal findings (Noble et al., 2017, 2019; Peña et al., 2013, 2014), we expected that tVNS would facilitate initial extinction learning, but also short- and long-term extinction recall and additionally prevent a reinstatement of fear.

The data by Szeska and colleagues (2020, 2021) demonstrated specific fear extinction during the initial extinction training, indicated by stronger decreases in the measured defensive responses in the fear learning relative to the control group. As expected, tVNS promoted this process, as US-expectancy, fear bradycardia as well as the fear potentiated startle decreased more strongly and the coupling between the latter two indicators of attentive immobility was reduced (Szeska et al., 2020, 2021). For both expectancy ratings and fear bradycardia, such promoted defensive response inhibition rapidly established right at the beginning of the extinction training (Szeska et al., 2020, 2021), in line with previous reports of rapid vagally mediated anxiolytic effects that add to enhancement of learning processes (Noble et al., 2019). This tVNS-promoted attenuation of fear bradycardia did not result from elevated levels of sympathetic arousal, as indexed by skin conductance, suggesting that this effect may indeed reflect inhibition of vagal efferent output due to a central inhibition of the defensive pathways mediating fear bradycardia (see **Figure 4**; Szeska et al., 2021). Although inhibitory effects were not as rapid for startle potentiation, it was completely abolished at the end of extinction training in fear learning group subjects receiving tVNS, while startle potentiation was maintained throughout entire extinction in the sham condition, indicating that tVNS facilitated extinction

learning as well (Szeska et al., 2020). Although at the beginning of both short- and long-term recall phases extinguished fear spontaneously recovered to a similar level in both stimulation conditions of the fear learning group, as indexed by increased US-expectancies and startle potentiation, subsequent extinction of these measures was also significantly facilitated in tVNS subjects (Szeska et al., 2020). Finally, reinstatement of fear, indicated by increases in risk assessments and startle reflexes during CSs and ITIs, was attenuated on the behavioral level (startle potentiation) by tVNS (Szeska et al., 2020).

The results suggest that tVNS is capable to promote defensive flexibility in the face of stimuli that no longer signal danger, as it yields rapid anxiolytic effects, but most importantly facilitates fear extinction learning, consolidation as well as short- and long-term recall of extinction memory, persisting for at least four weeks. Concerning facilitated extinction *learning* by tVNS, the data corroborate the view that vagal projections activate the brain's extinction pathways and, upon stimulation, foster inhibition of maladaptive defensive responding (Follesa et al., 2007; Hassert et al., 2004; McGaugh, 2018; D. Mueller & Cahill, 2010; Noble et al., 2017, 2019; Peña et al., 2013, 2014). Concerning the beneficial effects of tVNS on extinction *consolidation* and *recall*, the results are also in line with previous evidence indicating that vagal signaling may be crucial for the establishment of emotional memory, extending this conception to fear extinction memory (Sellaro et al., 2018; Ventura-Bort et al., 2018, 2021). Most importantly, however, given that extinction was enhanced on multiple levels of expression (cognitive, physiological, behavioral), the results not only line up with studies indicating vagally mediated promoted cognitive flexibility (Fischer et al., 2018; Steenbergen et al., 2015), but also mark the first cross-species transfer of animal stimulation research, showing promoted extinction of low-level attentive immobility (Noble et al., 2017, 2019; Peña et al., 2013, 2014).

From a perspective of translational fear research, the results are of particular importance, as they support the notion that harmonizing cross-species methodology – in this case by using

single-cue designs – leads to stronger comparability and reproducibility of results (Haaker et al., 2019). In turn, given that previous human research partly failed to reproduce animal findings (Burger et al., 2016, 2017, 2018), the results therefore raise the question whether methodology diverging from animal research qualifies for cross-species translational fear research.

Most importantly, however, from a clinical perspective, the results help to evaluate tVNS as an adjunct for extinction-based exposure therapy of anxiety, trauma- and stressor-related disorders (Craske et al., 2014; D. Hermans et al., 2006). In fact, the data imply that tVNS may boost responding to exposure therapy by both anxiolytic effects and enhancement of learning processes, rapidly affecting cognitive and physiological fear indices, and may additionally promote the stability of cognitive and behavioral fear reduction. Thus, tVNS may serve as a treatment add-on, that targets the brain's defense system to cope with disorder-related deficits in extinction learning (Duits et al., 2015; Lissek et al., 2005) that may contribute to non-responding and relapses after treatment (Craske & Mystkowski, 2006; Duits et al., 2015; Lissek et al., 2005; Vervliet et al., 2013). However, as vagal stimulation fosters emotional memory in general (Ventura-Bort et al., 2021), tVNS might need to be used cautiously during treatment, as consolidation of aversive memories during exposure might be facilitated just as well.

3.3 Facilitated extinction by visual attentional manipulation

In contrast to vagal stimulation, visual attentional manipulation is already used as an adjuvant for the treatment of post-traumatic stress disorder to facilitate exposure effects. This therapeutic approach is referred to as eye movement desensitization and reprocessing (EMDR) and requires the patient to follow a moving object with the eyes (e.g., the therapists finger) during in-sensu exposure to trauma-related, fear-eliciting memories (Chen et al., 2014; Landin-Romero et al., 2018; Shapiro, 1989). Interestingly, this therapeutic regimen might therefore tap

into the inhibitory SC-BLA pathway suggested by animal research, by which visual attentional manipulation boosts long-term extinction of attentive immobility (Baek et al., 2019).

In the third publication of the dissertation research program (Szeska, Mohrmann & Hamm, under review), we tested whether a such mechanism may be translated to humans, by using the above described single-cue conditioning and extinction design. First, participants underwent an acquisition training, during which they received paired presentations of a tone CS and an aversive US. Twenty-four hours later, they underwent an extinction training, during which the activity of the SC was modulated by the manipulation of eye movements. It is well established, that the SC is activated more strongly during the generation of saccadic eye movements relative to smooth eye pursuits (Krauzlis, 2004; Leigh & Zee, 2015). Hence, to contrast effects of high vs. low activation of the SC on extinction, half of the participants were instructed to generate saccadic eye movements, while the other half executed smooth eye pursuits during CS presentations for control purposes, which was verified by an electrooculogram. Extinction recall was tested after 24 hours and after one week. Resistance of extinction memory against reinstatement of fear was, again, tested by non-signaled US presentations during recall phases.

The data by Szeska and colleagues (under review) showed extinction of conditioned fear, as indicated by decreasing US expectancy, cardiac deceleration and startle potentiation. However, extinction of low-level indices of attentive immobility was particularly pronounced in the saccadic eye movement condition, indicated by stronger decrease of the startle potentiation and, to a lesser extent, cardiac deceleration. Most importantly, higher saccadic accuracy and range, linked to broader and prolonged SC activation (Goossens & van Opstal, 2012; Leigh & Zee, 2015; Waitzman et al., 1988), correlated to extinction of the fear potentiated startle. As such promoted defensive response inhibition established rapidly, it may be considered evidence for rapid anxiolytic effects of eye movement desensitization, which are often reported in clinical EMDR practice (see Landin-Romero et al., 2018). However, as startle potentiation was abolished at the end of the extinction training in the saccadic, but not the

smooth eye pursuit condition, the data indicate that SC-recruiting visual attentional manipulation indeed fostered fear extinction *learning* as well. Both findings support a critical role of the SC in the inhibition of fear. However, extinguished behavioral and physiological indicators of attentive immobility strongly recovered in both eye movement conditions during the extinction recall 24 hours later and, for startle potentiation, even reached levels obtained after fear conditioning, indicating that extinction memory could not be recalled. No differences between eye movement groups were observed with respect to the long-term extinction recall or the reinstatement of fear, as well as in US-expectancy throughout the entire experiment.

The data by Szeska and colleagues (under review), suggest that SC-recruiting visual attentional manipulation facilitates initial extinction of low-level attentive immobility, in line with findings from the animal model (Baek et al., 2019). The data therefore support rodent research's notion (Baek et al., 2019), that a subcortical inhibitory SC-BLA pathway might in fact underlie fear reducing effects of EMDR in humans, yielding the implication for clinical practice, that saccadic eye movements may be more efficient to achieve therapeutic success compared to smooth eye pursuits, which are conflictingly employed just as common (Stickgold, 2002). However, as SC-recruiting visual attentional manipulation had no impact on cognitive risk assessments, results also indicate that the effects of eye movement desensitization, that are mediated by the subcortical SC-BLA pathway, are limited to low-level defensive responses and do not impact on cognitive indices of defensive activation. EMDR effects on feelings of fear (see Landin-Romero et al., 2018) might therefore not be attributed to the proposed SC-BLA pathway. Most critically and contrary to rodent findings (Baek et al., 2019), extinction recall was not facilitated, but impaired after any type of eye movement manipulation. Thus, it may be likely that attentional manipulation acted as avoidance strategy (Lovibond et al., 2009; Pittig, 2019; Vervliet & Indekeu, 2015; Wong & Pittig, 2021) interfering with the consolidation of extinction memory on a subcortical level. For clinical practice, this may imply that EMDR might indeed facilitate initial fear reduction, but at the cost of long-term treatment outcome.

4 | Summary and future directions

This work summarized research, that explored the impact of transcutaneous vagus nerve stimulation (tVNS) and visual attentional manipulation recruiting the superior colliculus (SC) on the extinction of conditioned fear in humans. We provided evidence, that conditioned fear in humans may be conceived as a cross-species instance of attentive immobility – a defensive response pattern evoked by inevitable distal threat, characterized by concordant increases in subcortically mediated cardiac deceleration and startle reflex potentiation and discordant elevations in cortically mediated threat expectancy. TVNS facilitated extinction of both cortical and subcortical defensive responses, while SC-recruiting visual attentional manipulation only affected the latter low-level indices of defensive activation. TVNS also promoted long-term extinction, but *any* visual attentional manipulation impaired extinction recall on a subcortical level, i.e., interfered with the consolidation of extinction memory. TVNS and SC-recruiting attentional manipulation, thus, initially increased defensive flexibility towards cues no longer predicting threat, but only tVNS fostered long-term fear reduction on multiple response levels.

TVNS, and to a far lesser extent SC-recruiting visual attentional manipulation, might, thus, serve as treatment adjuvants to promote extinction-based exposure therapy of mental disorders. Yet, clinical evidence for such conclusion needs to be provided, especially in the case of tVNS. Before a clinical transfer may be endeavored, however, future research should provide detailed insights about the mechanisms of action underlying such extinction enhancement in humans. Functional and structural imaging studies may uncover the involved neural pathways and chemical agents and, thus, set the stage to elaborate an optimal parametrization for both tVNS and visual attentional manipulation to achieve maximized fear extinction. Finally, to test the specificity and the clinical applicability of the observed effects, studies should examine whether facilitated extinction may also be observed for interoceptive threat (e.g., dyspnea) and as to whether the facilitation of fear extinction by a non-invasive stimulation of the brain's inhibitory pathways is only limited to attentive immobility, or extends to defensive action as well.

5 | References

Amano, T., Unal, C. T., & Paré, D. (2010). Synaptic correlates of fear extinction in the amygdala. *Nature Neuroscience*, *13*(4), 489–494. https://doi.org/10.1038/nn.2499

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. In *Arlington*. https://doi.org/10.1176/appi.books.9780890425596.744053

Anderson, E. J., & Rees, G. (2011). Neural correlates of spatial orienting in the human superior colliculus. *Journal of Neurophysiology*, *106*(5), 2273–2284. https://doi.org/10.1152/jn.00286.2011

Badura-Brack, A. S., Naim, R., Ryan, T. J., Levy, O., Abend, R., Khanna, M. M., McDermott, T. J., Pine, D. S., & Bar-Haim, Y. (2015). Effect of attention training on attention bias variability and PTSD symptoms: Randomized controlled trials in Israeli and U.S. Combat Veterans. *American Journal of Psychiatry*, *172*(12), 1233–1241. https://doi.org/10.1176/appi.ajp.2015.14121578

Baek, J., Lee, S., Cho, T., Kim, S.-W., Kim, M., Yoon, Y., Kim, K. K., Byun, J., Kim, S. J., Jeong, J., & Shin, H.-S. (2019). Neural circuits underlying a psychotherapeutic regimen for fear disorders. *Nature*, *566*(7744), 339–343. https://doi.org/10.1038/s41586-019-0931-y

Bandelow, B., Lueken, U., Wolff, J., Godemann, F., Menzler, C. W., Deckert, J., Ströhle, A., Beutel, M., Wiltink, J., Domschke, K., & Berger, M. (2016). Leitliniengerechte stationäre psychiatrisch-psychotherapeutische/psychosomatische Behandlung von Angsterkrankungen: Wieviel Personal ist erforderlich? *Nervenarzt*, *87*(3), 302–310. https://doi.org/10.1007/s00115-016-0085-1

Barry, T. J., Griffith, J. W., Vervliet, B., & Hermans, D. (2016). The Role of Stimulus Specificity and Attention in the Generalization of Extinction. *Journal of Experimental Psychopathology*, *7*(1), 143–152. https://doi.org/10.5127/jep.048615

Barry, T. J., Vervliet, B., & Hermans, D. (2015). An integrative review of attention biases and their contribution to treatment for anxiety disorders. *Frontiers in Psychology*, *6*(July), 1–15. https://doi.org/10.3389/fpsyg.2015.00968

Baum, M. (1988). Spontaneous recovery from the effects of flooding (exposure) in animals. *Behaviour Research and Therapy*, *26*(2), 185–186. https://doi.org/10.1016/0005-7967(88)90118-0

Berlau, D. J., & McGaugh, J. L. (2006). Enhancement of extinction memory consolidation: The role of the noradrenergic and GABAergic systems within the basolateral amygdala. *Neurobiology of Learning and Memory*, *86*(2), 123–132. https://doi.org/10.1016/j.nlm.2005.12.008

Berthoud, H. R., & Neuhuber, W. L. (2000). Functional and chemical anatomy of the afferent vagal system. *Autonomic Neuroscience: Basic and Clinical*, *85*(1–3), 1–17. https://doi.org/10.1016/S1566-0702(00)00215-0

Blanchard, R. J., & Blanchard, D. C. (1989). Antipredator defensive behaviors in a visible burrow system. *Journal of Comparative Psychology*, *103*(1), 70–82. https://doi.org/10.1037/0735-7036.103.1.70

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *11*(5), 485–494. https://doi.org/10.1101/lm.78804

Bouton, M. E. (2014). Why behavior change is difficult to sustain. *Preventive Medicine*, *68*, 29–36. https://doi.org/10.1016/j.ypmed.2014.06.010

Bouton, M. E., & Woods, A. M. (2008). Extinction: Behavioral Mechanisms and Their Implications. In *Learning and Memory: A Comprehensive Reference* (Third Edit, pp. 151–171). Elsevier. https://doi.org/10.1016/B978-012370509-9.00088-7

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., Zlatar, Z. Z., & Lang, P. J. (2018). Startle reflex modulation during threat of shock and "threat" of reward. *Psychophysiology*, *55*(2), 1–10. https://doi.org/10.1111/psyp.12989

Brown, J. S., Kalish, H. I., & Farber, I. E. (1951). Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *Journal of Experimental Psychology*, *41*(5), 317–328. https://doi.org/10.1037/h0060166

Burger, A. M., Van Diest, I., van der Does, W., Hysaj, M., Thayer, J. F., Brosschot, J. F., & Verkuil, B. (2018). Transcutaneous vagus nerve stimulation and extinction of prepared fear: A conceptual non-replication. *Scientific Reports*, *8*(1), 11471. https://doi.org/10.1038/s41598-018-29561-w

Burger, A. M., Verkuil, B., Fenlon, H., Thijs, L., Cools, L., Miller, H. C., Vervliet, B., & Van Diest, I. (2017). Mixed evidence for the potential of non-invasive transcutaneous vagal nerve stimulation to improve the extinction and retention of fear. *Behaviour Research and Therapy*, *97*, 64–74. https://doi.org/10.1016/j.brat.2017.07.005

Burger, A. M., Verkuil, B., Van Diest, I., Van der Does, W., Thayer, J. F., & Brosschot, J. F. (2016). The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans. *Neurobiology of Learning and Memory*, *132*, 49–56. https://doi.org/10.1016/j.nlm.2016.05.007

Campbell, B. A., Wood, G., & McBride, T. (1997). Origins of orienting and defensive responses: An evolutionary perspective. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 41–67). Erlbaum.

Cannon, W. B. (1929). *Bodily Changes in Pain, Hunger, Fear and Rage: An Account of Recent Research Into the Function of Emotional Excitement* (2nd ed.). D. Appleton and Co.

Carew, T. J., Hawkins, R. D., & Kandel, E. R. (1983). Differential Classical Conditioning of a Defensive Withdrawal Reflex in Aplysia californica. *Science*, *219*(4583), 397–400. https://doi.org/10.1126/science.6681571

Chen, Y. R., Hung, K. W., Tsai, J. C., Chu, H., Chung, M. H., Chen, S. R., Liao, Y. M., Ou, K. L., Chang, Y. C., & Chou, K. R. (2014). Efficacy of eye-movement desensitization and reprocessing for patients with posttraumatic-stress disorder: A meta-analysis of randomized controlled trials. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0103676

Cimpianu, C.-L., Strube, W., Falkai, P., Palm, U., & Hasan, A. (2016). Vagus nerve stimulation in psychiatry: a systematic review of the available evidence. *Journal of Neural Transmission*, 1–14. https://doi.org/10.1007/s00702-016-1642-2

Craske, M. G., & Mystkowski, J. L. (2006). Exposure Therapy and Extinction: Clinical Studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications.* (pp. 217–233). American Psychological Association. https://doi.org/10.1037/11474-011

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Davis, M. (2006). Neural systems involved in fear and anxiety measured with fear-potentiated startle. *American Psychologist*, *61*(8), 741–756. https://doi.org/10.1037/0003-066X.61.8.741

Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry*, *6*(1), 13–34. https://doi.org/10.1038/sj.mp.4000812

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., Van Den Hout, M. A., & Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*(4), 239–253. https://doi.org/10.1002/da.22353

Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking Extinction. *Neuron*, *88*(1), 47–63. https://doi.org/10.1016/j.neuron.2015.09.028

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

Duvarci, S., & Paré, D. (2007). Glucocorticoids enhance the excitability of principal basolateral amygdala neurons. *Journal of Neuroscience*, *27*(16), 4482–4491. https://doi.org/10.1523/JNEUROSCI.0680-07.2007

Effting, M., & Kindt, M. (2007). Contextual control of human fear associations in a renewal paradigm. *Behaviour Research and Therapy*, *45*(9), 2002–2018. https://doi.org/10.1016/j.brat.2007.02.011

Eilam, D. (2005). Die hard: A blend of freezing and fleeing as a dynamic defense - Implications for the control of defensive behavior. *Neuroscience and Biobehavioral Reviews*, *29*(8), 1181–1191. https://doi.org/10.1016/j.neubiorev.2005.03.027

Ellrich, J. (2011). Transcutaneous vagus nerve stimulation. *Eur. Neurol. Rev*, *6*(4), 254–256. https://doi.org/10.17925/ENR.2011.06.04.254

Fadok, J. P., Krabbe, S., Markovic, M., Courtin, J., Xu, C., Massi, L., Botta, P., Bylund, K., Müller, C., Kovacevic, A., Tovote, P., & Lüthi, A. (2017). A competitive inhibitory circuit for selection of active and passive fear responses. *Nature*, *542*(7639), 96–99. https://doi.org/10.1038/nature21047

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review*, *1*(4), 429–438. https://doi.org/10.3758/BF03210947

Fanselow, M. S. (2018). The role of learning in threat imminence and defensive behaviors. *Current Opinion in Behavioral Sciences*, *24*, 44–49. https://doi.org/10.1016/j.cobeha.2018.03.003

Fischer, R., Ventura-Bort, C., Hamm, A., & Weymar, M. (2018). Transcutaneous vagus nerve stimulation (tVNS) enhances conflict-triggered adjustment of cognitive control. *Cognitive, Affective and Behavioral Neuroscience*, *18*(4), 680–693. https://doi.org/10.3758/s13415-018-0596-2

Foley, J. O., & DuBois, F. S. (1937). Quantitative studies of the vagus nerve in the cat. I. The ratio of sensory to motor fibers. *The Journal of Comparative Neurology*, *67*(1), 49–67. https://doi.org/10.1002/cne.900670104

Follesa, P., Biggio, F., Gorini, G., Caria, S., Talani, G., Dazzi, L., Puligheddu, M., Marrosu, F., & Biggio, G. (2007). Vagus nerve stimulation increases norepinephrine concentration and the gene expression of BDNF and bFGF in the rat brain. *Brain Research*, *1179*(1), 28–34. https://doi.org/10.1016/j.brainres.2007.08.045

Frangos, E., Ellrich, J., & Komisaruk, B. R. (2015). Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: FMRI evidence in humans. *Brain Stimulation*, *8*(3), 624–636. https://doi.org/10.1016/j.brs.2014.11.018

Gallagher, M., & Holland, P. C. (1994). The amygdala complex: Multiple roles in associative learning and attention. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(25), 11771–11776. https://doi.org/10.1073/pnas.91.25.11771

Genheimer, H., Andreatta, M., Asan, E., & Pauli, P. (2017). Reinstatement of contextual conditioned anxiety in virtual reality and the effects of transcutaneous vagus nerve stimulation in humans. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-18183-3

George, M. S., Sackeim, H. A., Rush, A. J., Marangell, L. B., Nahas, Z., Husain, M. M., Lisanby, S., Burt, T., Goldman, J., & Ballenger, J. C. (2000). Vagus nerve stimulation: A new tool for brain research and therapy. *Biological Psychiatry*, *47*(4), 287–295. https://doi.org/10.1016/S0006-3223(99)00308-X

Gewirtz, J. C., Falls, W. A., & Davis, M. (1997). Normal conditioned inhibition and extinction of freezing and fear- potentiated startle following electrolytic lesions of medial prefrontal cortex in rats. *Behavioral Neuroscience*, *111*(4), 712–726. https://doi.org/10.1037/0735-7044.111.4.712

Giustino, T. F., & Maren, S. (2018). Noradrenergic Modulation of Fear Conditioning and Extinction. *Frontiers in Behavioral Neuroscience*, *12*(March), 1–20. https://doi.org/10.3389/fnbeh.2018.00043

Gladwin, T. E., Hashemi, M. M., van Ast, V., & Roelofs, K. (2016). Ready and waiting: Freezing as active action preparation under threat. *Neuroscience Letters*, *619*, 182–188. https://doi.org/10.1016/j.neulet.2016.03.027

Goossens, H. H. L. M., & van Opstal, A. J. (2012). Optimal control of saccades by spatial-temporal activity patterns in the monkey superior colliculus. *PLoS Computational Biology*, *8*(5). https://doi.org/10.1371/journal.pcbi.1002508

Graham, F. K. (1979). Distinguishing among orienting, defense, and startle reflexes. In H. D. Kimmel, E. H. van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in Humans* (pp. 137–167). Erlbaum.

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin*, *65*(5), 305–320. https://doi.org/10.1037/h0023258

Haaker, J., Golkar, A., Hermans, D., & Lonsdorf, T. B. (2014). A review on human reinstatement studies: an overview and methodological challenges. *Learning & Memory*, *21*(9), 424–440. https://doi.org/10.1101/lm.036053.114

Haaker, J., Maren, S., Andreatta, M., Merz, C. J., Richter, J., Richter, S. H., Meir Drexler, S., Lange, M. D., Jüngling, K., Nees, F., Seidenbecher, T., Fullana, M. A., Wotjak, C. T., & Lonsdorf, T. B. (2019). Making translation work: Harmonizing cross-species methodology in the behavioural neuroscience of Pavlovian fear conditioning. *Neuroscience & Biobehavioral Reviews*, *107*, 329–345. https://doi.org/10.1016/j.neubiorev.2019.09.020

Hagenaars, M. A., Oitzl, M., & Roelofs, K. (2014). Updating freeze: Aligning animal and human research. *Neuroscience and Biobehavioral Reviews*, *47*, 165–176. https://doi.org/10.1016/j.neubiorev.2014.07.021

Hamm, A. O. (2015). Fear-Potentiated Startle. In *International Encyclopedia of the Social & Behavioral Sciences* (Second Edi, Vol. 8, pp. 860–867). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.55023-5

Hamm, A. O. (2020). Fear, anxiety, and their disorders from the perspective of psychophysiology. *Psychophysiology*, *57*(2), 1–14. https://doi.org/10.1111/psyp.13474

Hamm, A. O., & Flor, H. (2015). Fear learning, fear memory, and psychopathology. *International Journal of Psychophysiology*, *98*(3), 497–498. https://doi.org/10.1016/j.ijpsycho.2015.11.007

Hamm, A. O., Greenwald, M. K., Bradley, M. M., & Lang, P. J. (1993). Emotional learning, hedonic change, and the startle probe. *Journal of Abnormal Psychology*, *102*(3), 453–465. https://doi.org/10.1037/0021-843X.102.3.453

Hamm, A. O., & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, *33*(6), 698–710. https://doi.org/10.1111/j.1469-8986.1996.tb02366.x

Hassert, D. L., Miyashita, T., & Williams, C. L. (2004). The effects of peripheral vagal nerve stimulation at a memory-modulating intensity on norepinephrine output in the basolateral amygdala. *Behavioral Neuroscience*, *118*(1), 79–88. https://doi.org/10.1037/0735-7044.118.1.79

Heinig, I., Pittig, A., Richter, J., Hummel, K., Alt, I., Dickhöver, K., Gamer, J., Hollandt, M., Koelkebeck, K., Maenz, A., Tennie, S., Totzeck, C., Yang, Y., Arolt, V., Deckert, J., Domschke, K., Fydrich, T., Hamm, A., Hoyer, J., … Wittchen, H. U. (2017). Optimizing exposure-based CBT for anxiety disorders via enhanced extinction: Design and methods of a multicentre randomized clinical trial. *International Journal of Methods in Psychiatric Research*, *26*(2), 1–12. https://doi.org/10.1002/mpr.1560

Henry, T. R. (2002). Therapeutic mechanisms of vagus nerve stimulation. *Neurology*, *59*(6 Suppl 4), S3–S14. https://doi.org/10.1212/WNL.59.6_suppl_4.S3

Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in Human Fear Conditioning. *Biological Psychiatry*, *60*(4), 361–368. https://doi.org/10.1016/j.biopsych.2005.10.006

Hermans, E. J., Henckens, M. J. A. G., Roelofs, K., & Fernández, G. (2013). Fear bradycardia and activation of the human periaqueductal grey. *NeuroImage*, *66*, 278–287. https://doi.org/10.1016/j.neuroimage.2012.10.063

Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., & Lüthi, A. (2008). Switching on and off fear by distinct neuronal circuits. *Nature*, *454*(7204), 600–606. https://doi.org/10.1038/nature07166

Herry, C., & Garcia, R. (2002). Prefrontal cortex long-term potentiation, but not long-term depression, is associated with the maintenance of extinction of learned fear in mice. *Journal of Neuroscience*, *22*(2), 577–583. https://doi.org/10.1523/jneurosci.22-02-00577.2002

Hollandt, M., Wroblewski, A., Yang, Y., Ridderbusch, I. C., Kircher, T., Hamm, A. O., Straube, B., & Richter, J. (2020). Facilitating translational science in anxiety disorders by adjusting extinction training in the laboratory to exposure-based therapy procedures. *Translational Psychiatry*, *10*(1), 110. https://doi.org/10.1038/s41398-020-0786-x

Jinzhao, J., & Maren, S. (2007). Hippocampal involvement in contextual modulation of fear extinction. *Hippocampus*, *17*(9), 749–758. https://doi.org/10.1002/hipo

Jones, B. E., Halaris, A. E., McIlhany, M., & Moore, R. Y. (1977). Ascending projections of the locus coeruleus in the rat. I. axonal transport in central noradrenaline neurons. *Brain Research*, *127*(1), 1–21. https://doi.org/10.1016/0006-8993(77)90377-8

Kalin, N. H., & Shelton, S. E. (1989). Defensive behaviors in infant rhesus monkeys: Environmental cues and neurochemical regulation. *Science*, *243*(4899), 1718–1721. https://doi.org/10.1126/science.2564702

Kapp, B. S., Frysinger, R. C., Gallagher, M., & Haselton, J. R. (1979). Amygdala central nucleus lesions: Effect on heart rate conditioning in the rabbit. *Physiology and Behavior*, *23*(6), 1109–1117. https://doi.org/10.1016/0031-9384(79)90304-4

Knight, D. C., Smith, C. N., Cheng, D. T., Stein, E. A., & Helmstetter, F. J. (2004). Amygdala and hippocampal activity during acquisition and extinction of human fear conditioning. *Cognitive, Affective and Behavioral Neuroscience*, *4*(3), 317–325. https://doi.org/10.3758/CABN.4.3.317

Koenigs, M., & Grafman, J. (2009). Posttraumatic stress disorder: The role of medial prefrontal cortex and amygdala. *Neuroscientist*, *15*(5), 540–548. https://doi.org/10.1177/1073858409333072

Kolassa, I. T., Musial, F., Mohr, A., Trippe, R. H., & Miltner, W. H. R. (2005). Electrophysiological correlates of threat processing in spider phobics. *Psychophysiology*, *42*(5), 520–530. https://doi.org/10.1111/j.1469-8986.2005.00315.x

Koller, K., Rafal, R. D., Platt, A., & Mitchell, N. D. (2019). Orienting toward threat: Contributions of a subcortical pathway transmitting retinal afferents to the amygdala via the superior colliculus and pulvinar. *Neuropsychologia*, *128*(January), 78–86. https://doi.org/10.1016/j.neuropsychologia.2018.01.027

Krauzlis, R. J. (2004). Recasting the Smooth Pursuit Eye Movement System. *Journal of Neurophysiology*, *91*(2), 591–603. https://doi.org/10.1152/jn.00801.2003

Krauzlis, R. J., Lovejoy, L. P., & Zénon, A. (2013). Superior colliculus and visual spatial attention. *Annual Review of Neuroscience*, *36*, 165–182. https://doi.org/10.1146/annurev-neuro-062012-170249

Kuhn, M., Wendt, J., Sjouwerman, R., Büchel, C., Hamm, A., & Lonsdorf, T. B. (2019). The Neurofunctional Basis of Affective Startle Modulation in Humans: Evidence From Combined Facial Electromyography and Functional Magnetic Resonance Imaging. *Biological Psychiatry*, 1–11. https://doi.org/10.1016/j.biopsych.2019.07.028

Landin-Romero, R., Moreno-Alcazar, A., Pagani, M., & Amann, B. L. (2018). How does eye movement desensitization and reprocessing therapy work? A systematic review on suggested mechanisms of action. *Frontiers in Psychology*, *9:1395*. https://doi.org/10.3389/fpsyg.2018.01395

Landis, C., & Hunt, W. (1939). The startle Pattern. *Oxford: Farrar & Rinehart*.

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, *84*(3), 437–450. https://doi.org/10.1016/j.biopsycho.2009.10.007

Lang, P. J., Bradley, M. M., & Cuthbert, B. . (1990). Emotion , Attention , and the Startle Reflex. *Psychological Review*, *97*(3), 377–395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: affect, activation and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and Orienting: Sensory and Motivational Processes* (pp. 97–135). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248–1263. https://doi.org/10.1016/S0006-3223(98)00275-3

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29. https://doi.org/10.1016/S0079-6123(06)56001-7

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*(3), 137–159. https://doi.org/10.1016/S0165-0327(00)00343-8

Lang, P. J., Levin, D. N., Miller, G. a, & Kozak, M. J. (1983). Fear behavior, fear imagery, and the psychophysiology of emotion: the problem of affective response integration. *Journal of Abnormal Psychology*, *92*(3), 276–306. https://doi.org/10.1037/0021-843X.92.3.276

Leaton, R. N., & Borszcz, G. S. (1985). Potentiated Startle. Its Relation to Freezing and Shock Intensity in Rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(3), 421–428. https://doi.org/10.1037/0097-7403.11.3.421

LeDoux, J. E. (1995). Emotion: Clues from the Brain. *Annual Review of Psychology*, *46*(1), 209–235. https://doi.org/10.1146/annurev.psych.46.1.209

LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), E2016–E2025. https://doi.org/10.1073/pnas.1619316114

LeDoux, J. E., & Daw, N. D. (2018). Surviving threats: Neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nature Reviews Neuroscience*, *19*(5), 269–282. https://doi.org/10.1038/nrn.2018.22

LeDoux, J. E., Iwata, J., Cicchetti, P., & Reis, D. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *The Journal of Neuroscience*, *8*(7), 2517–2529. https://doi.org/10.1523/JNEUROSCI.08-07-02517.1988

Lee, S., & Shin, H. S. (2016). The role of mediodorsal thalamic nucleus in fear extinction. *Journal of Analytical Science and Technology*, *7*(1). https://doi.org/10.1186/s40543-016-0093-6

Leigh, R. J., & Zee, D. S. (2015). *The Neurology of Eye Movements* (5.). Oxford University Press. https://doi.org/10.1093/med/9780199969289.001.0001

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

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy*, *43*(11), 1391–1424. https://doi.org/10.1016/j.brat.2004.10.007

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lonsdorf, T. B., Weike, A. I., Nikamo, P., Schalling, M., Hamm, A. O., & Öhman, A. (2009). Genetic gating of human fear learning and extinction: Possible implications for gene-environment interaction in anxiety disorder. *Psychological Science*, *20*(2), 198–206. https://doi.org/10.1111/j.1467-9280.2009.02280.x

Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy*, *47*(8), 716–720. https://doi.org/10.1016/j.brat.2009.04.013

Löw, A., Lang, P. J., Smith, J. C., & Bradley, M. M. (2008). Both predator and prey: Emotional arousal in threat and reward. *Psychological Science*, *19*(9), 865–873. https://doi.org/10.1111/j.1467-9280.2008.02170.x

Löw, A., Weymar, M., & Hamm, A. O. (2015). When threat is near, get out of here: Dynamics of defensive behavior during freezing and active avoidance. *Psychological Science*, *26*(11), 1706–1716. https://doi.org/10.1177/0956797615597332

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews Neuroscience*, *14*(6), 417–428. https://doi.org/10.1038/nrn3492

Marks, I. M. (1987). *Fears, phobias and rituals: Panic, anxiety, and their disorders.* Oxford University Press.

McGaugh, J. L. (2018). Emotional arousal regulation of memory consolidation. *Current Opinion in Behavioral Sciences*, *19*, 55–60. https://doi.org/10.1016/j.cobeha.2017.10.003

McIntyre, C. K., McGaugh, J. L., & Williams, C. L. (2012). Interacting brain systems modulate memory consolidation. *Neuroscience and Biobehavioral Reviews*, *36*(7), 1750–1762. https://doi.org/10.1016/j.neubiorev.2011.11.001

Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., Zeidan, M. A., Handwerger, K., Orr, S. P., & Rauch, S. L. (2009). Neurobiological Basis of Failure to Recall Extinction Memory in Posttraumatic Stress Disorder. *Biological Psychiatry*, *66*(12), 1075–1082. https://doi.org/10.1016/j.biopsych.2009.06.026

Milad, M. R., & Quirk, G. J. (2012). Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. *Annual Review of Psychology*, *63*(1), 129–151. https://doi.org/10.1146/annurev.psych.121208.131631

Milad, M. R., Rosenbaum, B. L., & Simon, N. M. (2014). Neuroscience of fear extinction: Implications for assessment and treatment of fear-based and anxiety related disorders. *Behaviour Research and Therapy*, *62*, 17–23. https://doi.org/10.1016/j.brat.2014.08.006

Miyashita, T., & Williams, C. L. (2004). Peripheral arousal-related hormones modulate norepinephrine release in the hippocampus via influences on brainstem nuclei. *Behavioural Brain Research*, *153*(1), 87–95. https://doi.org/10.1016/j.bbr.2003.11.005

Miyashita, T., & Williams, C. L. (2006). Epinephrine administration increases neural impulses propagated along the vagus nerve: Role of peripheral β-adrenergic receptors. *Neurobiology of Learning and Memory*, *85*(2), 116–124. https://doi.org/10.1016/j.nlm.2005.08.013

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., Petrovic, P., Dolan, R. J., & Frith, C. D. (2009). From threat to fear: The neural organization of defensive fear systems in humans. *Journal of Neuroscience*, *29*(39), 12236–12243. https://doi.org/10.1523/JNEUROSCI.2378-09.2009

Mueller, D., & Cahill, S. P. (2010). Noradrenergic modulation of extinction learning and exposure therapy. *Behavioural Brain Research*, *208*(1), 1–11. https://doi.org/10.1016/j.bbr.2009.11.025

Mueller, E. M., Panitz, C., Hermann, C., & Pizzagalli, D. A. (2014). Prefrontal oscillations during recall of conditioned and extinguished fear in humans. *Journal of Neuroscience*, *34*(21), 7059–7066. https://doi.org/10.1523/JNEUROSCI.3427-13.2014

Muench, H. M., Westermann, S., Pizzagalli, D. A., Hofmann, S. G., & Mueller, E. M. (2016). Self-relevant threat contexts enhance early processing of fear-conditioned faces. *Biological Psychology*, *121*, 194–202. https://doi.org/10.1016/j.biopsycho.2016.07.017

Müller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 524–529. https://doi.org/10.1073/pnas.0408311101

Myers, K. M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron*, *36*(4), 567–584. https://doi.org/10.1016/S0896-6273(02)01064-4

Nabavi, S., Fox, R., Proulx, C. D., Lin, J. Y., Tsien, R. Y., & Malinow, R. (2014). Engineering a memory with LTD and LTP. *Nature*, *511*(7509), 348–352. https://doi.org/10.1038/nature13294

Nemeroff, C. B., Mayberg, H. S., Krahl, S. E., McNamara, J., Frazer, A., Henry, T. R., George, M. S., Charney, D. S., & Brannan, S. K. (2006). VNS therapy in treatment-resistant depression: clinical evidence and putative neurobiological mechanisms. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *31*(7), 1345–1355. https://doi.org/10.1038/sj.npp.1301082

Noble, L. J., Gonzalez, I. J., Meruva, V. B., Callahan, K. A., Belfort, B. D., Ramanathan, K. R., Meyers, E., Kilgard, M. P., Rennaker, R. L., & McIntyre, C. K. (2017). Effects of vagus nerve stimulation on extinction of conditioned fear and post-traumatic stress disorder symptoms in rats. *Translational Psychiatry*, *7*(8), e1217. https://doi.org/10.1038/tp.2017.191

Noble, L. J., Meruva, V. B., Hays, S. A., Rennaker, R. L., Kilgard, M. P., & McIntyre, C. K. (2019). Vagus nerve stimulation promotes generalization of conditioned fear extinction and reduces anxiety in rats. *Brain Stimulation*, *12*(1), 9–18. https://doi.org/10.1016/j.brs.2018.09.013

Öhman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology*, *30*(10), 953–958. https://doi.org/10.1016/j.psyneuen.2005.03.019

Öhman, A., Carlsson, K., Lundqvist, D., & Ingvar, M. (2007). On the unconscious subcortical origin of human fear. *Physiology and Behavior*, *92*(1–2), 180–185. https://doi.org/10.1016/j.physbeh.2007.05.057

Oyoshi, T., Nishijo, H., Asakura, T., Takamura, Y., & Ono, T. (1996). Emotional and behavioral correlates of mediodorsal thalamic neurons during associative learning in rats. *Journal of Neuroscience*, *16*(18), 5812–5829. https://doi.org/10.1523/jneurosci.16-18-05812.1996

Panitz, C., Hermann, C., & Mueller, E. M. (2015). Conditioned and extinguished fear modulate functional corticocardiac coupling in humans. *Psychophysiology*, *52*(10), 1351–1360. https://doi.org/10.1111/psyp.12498

Panitz, C., Keil, A., & Mueller, E. M. (2019). Extinction-resistant attention to long-term conditioned threat is indexed by selective visuocortical alpha suppression in humans. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-52315-1

Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex* (G. V. Anrep (ed.)). London: Oxford University Press. https://doi.org/10.2307/1134737

Peña, D. F., Childs, J. E., Willett, S., Vital, A., McIntyre, C. K., & Kroener, S. (2014). Vagus nerve stimulation enhances extinction of conditioned fear and modulates plasticity in the pathway from the ventromedial prefrontal cortex to the amygdala. *Frontiers in Behavioral Neuroscience*, *8*(September), 327. https://doi.org/10.3389/fnbeh.2014.00327

Peña, D. F., Engineer, N. D., & McIntyre, C. K. (2013). Rapid remission of conditioned fear expression with extinction training paired with vagus nerve stimulation. *Biological Psychiatry*, *73*(11), 1071–1077. https://doi.org/10.1016/j.biopsych.2012.10.021

Peuker, E. T., & Filler, T. J. (2002). The nerve supply of the human auricle. *Clinical Anatomy*, *15*(1), 35–37. https://doi.org/10.1002/ca.1089

Phaf, R. H., Hermans, M. E., Krepel, A., Lieuw-On, R. L. R., Mulder, C. B., & Weijland, S. (2021). Horizontal eye movements foster approach to negative pictures but do not change emotional valence: A dopaminergic regulation hypothesis. *New Ideas in Psychology*, *62*(September 2020), 100872. https://doi.org/10.1016/j.newideapsych.2021.100872

Pittig, A. (2019). Incentive-based extinction of safety behaviors: Positive outcomes competing with aversive outcomes trigger fear-opposite action to prevent protection from fear extinction. *Behaviour Research and Therapy*, *121*, 103463. https://doi.org/10.1016/j.brat.2019.103463

Pittig, A., Heinig, I., Goerigk, S., Thiel, F., Hummel, K., Scholl, L., Deckert, J., Pauli, P., Domschke, K., Lueken, U., Fydrich, T., Fehm, L., Plag, J., Ströhle, A., Kircher, T., Straube, B., Rief, W., Koelkebeck, K., Arolt, V., … Wittchen, H. U. (2021). Efficacy of temporally intensified exposure for anxiety disorders: A multicenter randomized clinical trial. *Depression and Anxiety*, *June*. https://doi.org/10.1002/da.23204

Pittig, A., van den Berg, L., & Vervliet, B. (2016). The key role of extinction learning in anxiety disorders. *Current Opinion in Psychiatry*, *29*(1), 39–47. https://doi.org/10.1097/YCO.0000000000000220

Pittig, A., Wong, A. H. K., Glück, V. M., & Boschet, J. M. (2020). Avoidance and its bi-directional relationship with conditioned fear: Mechanisms, moderators, and clinical implications. *Behaviour Research and Therapy*, *126*(December 2019), 103550. https://doi.org/10.1016/j.brat.2020.103550

Plappert, C. F., Pilz, P. K. D., & Schnitzler, H. U. (1993). Acoustic startle response and habituation in freezing and nonfreezing rats. *Behavioral Neuroscience*, *107*(6), 981–987. https://doi.org/10.1037/0735-7044.107.6.981

Podină, I. R., Koster, E. H. W., Philippot, P., Dethier, V., & David, D. O. (2013). Optimal attentional focus during exposure in specific phobia: A meta-analysis. *Clinical Psychology Review*, *33*(8), 1172–1183. https://doi.org/10.1016/j.cpr.2013.10.002

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, *33*(1), 56–72. https://doi.org/10.1038/sj.npp.1301555

Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *The Journal of Neuroscience*, *20*(16), 6225–6231. https://doi.org/20/16/6225 [pii]

Ramirez, S., Liu, X., Lin, P. A., Suh, J., Pignatelli, M., Redondo, R. L., Ryan, T. J., & Tonegawa, S. (2013). Creating a false memory in the hippocampus. *Science*, *341*(6144), 387–391. https://doi.org/10.1126/science.1239073

Rauch, S. L., Shin, L. M., & Wright, C. I. (2006). Neuroimaging Studies of Amygdala Function in Anxiety Disorders. *Annals of the New York Academy of Sciences*, *985*(1), 389–410. https://doi.org/10.1111/j.1749-6632.2003.tb07096.x

Reisenzein, R. (2009). Emotional experience in the computational belief-desire theory of emotion. *Emotion Review*, *1*(3), 214–222. https://doi.org/10.1177/1754073909103589

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*(1), 71–80. https://doi.org/https://doi.org/10.1037/h0024109

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton- Century-Crofts.

Richter, J., Pittig, A., Hollandt, M., & Lueken, U. (2017). Bridging the Gaps Between Basic Science and Cognitive-Behavioral Treatments for Anxiety Disorders in Routine Care. *Zeitschrift Für Psychologie*, *2017*. https://doi.org/10.1027/1866-5888/a000147

Roelofs, K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160206. https://doi.org/10.1098/rstb.2016.0206

Rogan, M. T., Stäubli, U. V, & LeDoux, J. E. (1997). Fear conditioning induces associative long-term potentiation in the amygdala. *Nature*, *390*(6660), 604–607. https://doi.org/10.1038/37601

Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review*, *105*(2), 325–350. https://doi.org/10.1037/0033-295X.105.2.325

Sah, P., & Westbrook, R. F. (2008). The circuit of fear. *Nature*, *454*(7204), 589–590. https://doi.org/10.1038/454589a

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy*, *86*, 87–94. https://doi.org/10.1016/j.brat.2016.08.015

Schipper, P., Hiemstra, M., Bosch, K., Nieuwenhuis, D., & Adinolfi, A. (2019). The association between serotonin transporter availability and the neural correlates of fear bradycardia. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(51), 25941–25947. https://doi.org/10.1073/pnas.1904843116

Sellaro, R., de Gelder, B., Finisguerra, A., & Colzato, L. S. (2018). Transcutaneous vagus nerve stimulation (tVNS) enhances recognition of emotions in faces but not bodies. *Cortex*, *99*, 213–223. https://doi.org/10.1016/j.cortex.2017.11.007

Senn, V., Wolff, S. B. E., Herry, C., Grenier, F., Ehrlich, I., Gründemann, J., Fadok, J. P., Müller, C., Letzkus, J. J., & Lüthi, A. (2014). Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron*, *81*(2), 428–437. https://doi.org/10.1016/j.neuron.2013.11.006

Shapiro, F. (1989). Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *20*(3), 211–217. https://doi.org/10.1016/0005-7916(89)90025-6

Silvani, A., Calandra-Buonaura, G., Dampney, R. A. L., & Cortelli, P. (2016). Brain–heart interactions: physiology and clinical implications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2067), 20150181. https://doi.org/10.1098/rsta.2015.0181

Solomon, R. L., & Wynne, L. C. (1954). Traumatic avoidance learning: the principles of anxiety conservation and partial irreversibility. *Psychological Review*, *61*(6), 353–385. https://doi.org/10.1037/h0054540

Sommer, M. A., & Wurtz, R. H. (2004). What the Brain Stem Tells the Frontal Cortex. I. Oculomotor Signals Sent from Superior Colliculus to Frontal Eye Field Via Mediodorsal Thalamus. *Journal of Neurophysiology*, *91*(3), 1381–1402. https://doi.org/10.1152/jn.00738.2003

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*(7117), 374–377. https://doi.org/10.1038/nature05279

Steenbergen, L., Sellaro, R., Stock, A. K., Verkuil, B., Beste, C., & Colzato, L. S. (2015). Transcutaneous vagus nerve stimulation (tVNS) enhances response selection during action cascading processes. *European Neuropsychopharmacology*, *25*(6), 773–778. https://doi.org/10.1016/j.euroneuro.2015.03.015

Stickgold, R. (2002). EMDR: A putative neurobiological mechanism of action. *Journal of Clinical Psychology*, *58*(1), 61–75. https://doi.org/10.1002/jclp.1129

Szeska, C., Mohrmann, H., & Hamm, A. O. (under review). Facilitated extinction, but impaired extinction recall by eye movement manipulation in humans - indications for action mechanisms and the applicability of eye movement desensitization.

Szeska, C., Richter, J., Wendt, J., Weymar, M., & Hamm, A. O. (2020). Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training. *Scientific Reports*, *10*(1), 1529. https://doi.org/10.1038/s41598-020-58412-w

Szeska, C., Richter, J., Wendt, J., Weymar, M., & Hamm, A. O. (2021). Attentive immobility in the face of inevitable distal threat—Startle potentiation and fear bradycardia as an index of emotion and attention. *Psychophysiology*, *58*(6), 1–17. https://doi.org/10.1111/psyp.13812

Thayer, J. F., & Sternberg, E. (2006). Beyond heart rate variability: Vagal regulation of allostatic systems. *Annals of the New York Academy of Sciences*, *1088*, 361–372. https://doi.org/10.1196/annals.1366.014

Tovote, P., Esposito, M. S., Botta, P., Chaudun, F., Fadok, J. P., Markovic, M., Wolff, S. B. E., Ramakrishnan, C., Fenno, L., Deisseroth, K., Herry, C., Arber, S., & Lüthi, A. (2016). Midbrain circuits for defensive behaviour. *Nature*, *534*(7606), 206–212. https://doi.org/10.1038/nature17996

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience*, *16*(6), 317–331. https://doi.org/10.1038/nrn3945

Van Damme, S., Crombez, G., Hermans, D., Koster, E. H. W., & Eccleston, C. (2006). The role of extinction and reinstatement in attentional bias to threat: A conditioning approach. *Behaviour Research and Therapy*, *44*(11), 1555–1563. https://doi.org/10.1016/j.brat.2005.11.008

Van den Bergh, O., Brosschot, J., Critchley, H., Thayer, J. F., & Ottaviani, C. (2021). Better Safe Than Sorry: A Common Signature of General Vulnerability for Psychopathology. *Perspectives on Psychological Science*, *16*(2), 225–246. https://doi.org/10.1177/1745691620950690

Ventura-Bort, C., Wirkner, J., Genheimer, H., Wendt, J., Hamm, A. O., & Weymar, M. (2018). Effects of Transcutaneous Vagus Nerve Stimulation (tVNS) on the P300 and Alpha-Amylase Level: A Pilot Study. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/fnhum.2018.00202

Ventura-Bort, C., Wirkner, J., Wendt, J., Hamm, A. O., & Weymar, M. (2021). Establishment of Emotional Memories Is Mediated by Vagal Nerve Activation: Evidence from Noninvasive taVNS. *The Journal of Neuroscience*, *41*(36), 7636–7648. https://doi.org/10.1523/jneurosci.2329-20.2021

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear Extinction and Relapse: State of the Art. *Annu. Rev. Clin. Psychol*, *9*(1), 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542

Vervliet, B., & Indekeu, E. (2015). Low-Cost Avoidance Behaviors are Resistant to Fear Extinction in Humans. *Frontiers in Behavioral Neuroscience*, *9*(December), 1–12. https://doi.org/10.3389/fnbeh.2015.00351

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory*, *11*(5), 555–558. https://doi.org/10.1101/lm.77404

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Waitzman, D. M., Ma, T. P., Optican, L. M., & Wurtz, R. H. (1988). Superior colliculus neurons provide the saccadic motor error signal. *Experimental Brain Research*, *72*(3). https://doi.org/10.1007/BF00250610

Walker, P., & Carrive, P. (2003). Role of ventrolateral periaqueductal gray neurons in the behavioral and cardiovascular responses to contextual conditioned fear and poststress recovery. *Neuroscience*, *116*(3), 897–912. https://doi.org/10.1016/S0306-4522(02)00744-3

Watson, J. B., & Rayner, R. (1920). Conditioned Emotional Reactions. *Journal of Experimental Psychology*, *3*(1), 1–14. https://doi.org/52, 281-302

Weike, A. I., Hamm, A. O., Schupp, H. T., Runge, U., Schroeder, H. W. S., & Kessler, C. (2005). Fear Conditioning following Unilateral Temporal Lobectomy: Dissociation of Conditioned Startle Potentiation and Autonomic Learning. *Journal of Neuroscience*, *25*(48), 11117–11124. https://doi.org/10.1523/JNEUROSCI.2032-05.2005

White, B. J., Berg, D. J., Kan, J. Y., Marino, R. A., Itti, L., & Munoz, D. P. (2017). Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nature Communications*, *8*, 2–10. https://doi.org/10.1038/ncomms14263

Wong, A. H. K., & Pittig, A. (2021). A dimensional measure of safety behavior: A non-dichotomous assessment of costly avoidance in human fear conditioning. *Psychological Research*, *0123456789*. https://doi.org/10.1007/s00426-021-01490-w

Appendix A: Publications

Manuscript 1

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A. O. (2021). Attentive immobility in the face of inevitable distal threat – fear bradycardia and startle potentiation as an index of emotion and attention. *Psychophysiology, 58*(6), 1-17. doi: 10.1111/psyp.13812

Manuscript 2

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A. O. (2020). Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training. *Scientific reports, 10*, 1529. doi: 10.1038/s41598-020-58412-w

Manuscript 3

Szeska, C., Mohrmann, H., Hamm, A. O. (under review). Facilitated extinction but impaired extinction recall by eye movement manipulation in humans – indications for action mechanisms and the applicability of eye movement desensitization

Manuscript 1

Attentive immobility in the face of inevitable distal threat – fear bradycardia and startle potentiation as an index of emotion and attention

Christoph Szeska, Jan Richter, Julia Wendt, Mathias Weymar, & Alfons O. Hamm

Psychophysiology, 58, 1-17

Published in 2021

Author contributions:

Christoph Szeska: Conceptualization, Data curation, formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing original draft, Writing review and editing

Jan Richter: Conceptualization, Methodology, Writing review and editing

Julia Wendt: Conceptualization, Methodology, Writing review and editing

Mathias Weymar: Conceptualization, Methodology, Writing review and editing

Alfons O. Hamm: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Writing review and editing

Received: 7 October 2020 | Revised: 5 February 2021 | Accepted: 2 March 2021

DOI: 10.1111/psyp.13812

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Attentive immobility in the face of inevitable distal threat—Startle potentiation and fear bradycardia as an index of emotion and attention

Christoph Szeska[1] | **Jan Richter**[1] | **Julia Wendt**[2] | **Mathias Weymar**[2,3] | **Alfons O. Hamm**[1]

[1]Department of Physiological and Clinical Psychology/Psychotherapy, University of Greifswald, Greifswald, Germany

[2]Department of Biological Psychology and Affective Science, Faculty of Human Sciences, University of Potsdam, Potsdam, Germany

[3]Faculty of Health Sciences Brandenburg, University of Potsdam, Potsdam, Germany

Correspondence
Christoph Szeska, Department of Physiological and Clinical Psychology/Psychotherapy, University of Greifswald, Franz-Mehring-Strasse 47, 17487 Greifswald, Germany.
Email: christoph.szeska@uni-greifswald.de

Funding information
This research was supported by a grant of the government of Mecklenburg-Vorpommern to Christoph Szeska. Julia Wendt and Jan Richter were supported by respective grants provided by the German Research Foundation (DFG; WE 5873/1-1; JR 2810/1-1)

Abstract

During fear conditioning, a cue (CS) signals an inevitable distal threat (US) and evokes a conditioned response that can be described as attentive immobility (freezing). The organism remains motionless and monitors the source of danger while startle responses are potentiated, indicating a state of defensive hypervigilance. Although in animals vagally mediated fear bradycardia is also reliably observed under such circumstances, results are mixed in human fear conditioning. Using a single-cue fear conditioning and extinction protocol, we tested cardiac reactivity and startle potentiation indexing low-level defensive strategies in a fear-conditioned (n = 40; paired presentations of CS and US) compared with a non-conditioned control group (n = 40; unpaired presentations of CS and US). Additionally, we assessed shock expectancy ratings on a trial-by-trial basis indexing declarative knowledge of the previous contingencies. Half of each group underwent extinction under sham or active transcutaneous vagus nerve stimulation (tVNS), serving as additional proof of concept. We found stronger cardiac deceleration during CS presentation in the fear learning relative to the control group. This learned fear bradycardia was positively correlated with conditioned startle potentiation but not with declarative knowledge of CS-US contingencies. TVNS abolished differences in heart rate changes between both groups and removed the significant correlation between late cardiac deceleration and startle potentiation in the fear learning group. Results suggest, fear-conditioned cues evoke attentive immobility in humans, characterized by cardiac deceleration and startle potentiation. Such defensive response pattern is elicited by cues predicting inevitable distal threat and resembles conditioned fear responses observed in rodents.

KEYWORDS

attentive immobility (freezing), extinction, fear bradycardia, fear conditioning, startle potentiation, transcutaneous vagus nerve stimulation

1 | INTRODUCTION

Fear can be conceived as an action disposition evoked by threat-related stimuli that activates behavioral defensive strategies to ensure the organism's survival (Hamm & Flor, 2015; Lang, 1995; Lang & Bradley, 2010). These defensive strategies can broadly be separated into defensive anticipation and immobility (*freezing*) and defensive action (e.g., active avoidance or attack; Hamm, 2020; Lang & Bradley, 2010; Marks, 1987) and are dynamically executed depending upon the imminence of the threat and the available behavioral options of the organism (e.g., chance of threat avoidance), providing flexible adaptation to the situation and, hence, increased probability of survival (Fanselow, 1994; Hamm, 2020; Lang et al., 1997, 2000; Marks, 1987; Mobbs et al., 2020).

As soon as a cue signaling a possible upcoming threat is detected, the organism is engaged in a fear-related state of attentive immobility or freezing (Eilam, 2005; Hamm, 2020; Lang et al., 1997; Marks, 1987; Roelofs, 2017). Such attentive immobility is defined by increased selective attention toward the threat-signaling cue, inhibited locomotion, a tense body posture, and potentiation of the protective startle reflex (Blanchard & Blanchard, 1969; Eilam, 2005; Fanselow, 1984; Gewirtz et al., 1997; Kalin & Shelton, 1989; Kolassa et al., 2005; Leaton & Borszcz, 1985).

Animal research shows that a wide variety of species also responds with a profound phasic deceleration of the heart rate when facing such distal threat, a phenomenon for which comparative psychophysiologists have coined the term *fear bradycardia* (Campbell et al., 1997). Supporting this view, animal fear conditioning studies showed strong and positive correlations between behavioral freezing and both prolonged heart rate deceleration (Walker & Carrive, 2003) and startle potentiation (Gewirtz et al., 1997; Leaton & Borszcz, 1985). Moreover, rodent research showed that cardiac deceleration in response to threat-signaling cues is mediated by similar underlying neural substrates, that also modulate threat-related startle potentiation and behavioral freezing, involving the central nucleus of the amygdala (CeA) and its projections to the ventrolateral periaqueductal gray (vlPAG; Applegate et al., 1983; Choi & Brown, 2003; Davis, 2006; Fendt & Fanselow, 1999; LeDoux et al., 1988; Walker & Carrive, 2003).

In human psychophysiological research, heart rate deceleration has been traditionally interpreted as an index of increased orienting toward significant stimuli that carry information (Graham, 1979). Accordingly and contrary to animal research, strong cardiac decelerations elicited by a conditioned stimulus (CS) signaling the occurrence of a threat (unconditioned stimulus, US) in early human fear conditioning experiments have been interpreted as an inhibition of the habituation of the orienting reflex (Geer, 1964; Putnam et al., 1974), rather than reflecting a fear response. In fact, although in human fear conditioning studies startle responses were found to be reliably potentiated (Grillon & Davis, 1997; Hamm et al., 1993; Lipp et al., 1994, for a review see Hamm, 2015), heart rate changes have shown to vary as a function of CS-content, US-intensity, and individual response patterns, with cardiac deceleration being observed more commonly with neutral CSs, whereas cardiac acceleration was associated with fear relevant CSs and more intense USs (Dimberg, 1987; Hamm et al., 1993; Hamm & Vaitl, 1996; Hodes et al., 1985; Lipp & Vaitl, 1990; Moratti & Keil, 2005; see Lonsdorf et al., 2017 for a review).

However, by relating stimulus significance to motivational systems that serve survival functions, Bradley (2009) presented evidence that prolonged parasympathetically dominated cardiac deceleration consistently occurs during states involving increased perceptual effort, engaged during selective attention toward motivationally significant stimuli. As this is the case during monitoring sources of inevitable danger, Bradley (2009), thus, provided a link between orienting and fear (see also Bradley et al., 2018). Supporting this view, recent research found that heart rate changes in the face of a threat in fact critically vary depending upon the behavioral options at hand, along with the actually executed defensive strategy (see Krause et al., 2018; Löw et al., 2015). As demonstrated in these two studies, there was strong cardiac deceleration as well as startle potentiation, if there was no option to actively avoid an approaching threat (moderately painful stimulus; or forced breath holding), and both measures were strongest immediately prior to the delivery of the aversive stimulus, while cardiac acceleration and startle inhibition was found when the organism was beyond vigilance and engaged in vigorous defensive action (Krause et al., 2018; Lang & Davis, 2006; Lang et al., 2000; Löw et al., 2015). These data were supported by findings of Roelofs and coworkers, showing heart rate decrease during different inevitable threat conditions including fear conditioning. In these studies heart rate deceleration was associated with reduced locomotion as measured by postural sway on a stabilometric platform, supporting the view that defensive responses acquired in human fear conditioning studies might be instances of attentive freezing (Gladwin et al., 2016; Roelofs, 2017; Roelofs et al., 2010). Accordingly, recent imaging research indicated that similar neural mechanisms that underlie behavioral freezing and cardiac deceleration in rodents also apply to humans during processing of distal inevitable threats (Wendt et al., 2017).

The current study follows up on this research and aims to provide an analysis between cardiac reactivity and startle modulation during human fear conditioning. We strived for harmonizing cross-species methodology (see Haaker et al., 2019 for a detailed discussion), by applying a multiple-day single-cue fear conditioning and extinction protocol, closely adapted to animal research (see Peña et al., 2013, 2014, but also Wong & Lovibond, 2017, 2018). Such paradigm involves between-subject comparisons of conditioned responses between a fear learning group, receiving repeated presentations of a CS paired with an aversive US during an acquisition training,

TABLE 1 Demographics and body-mass-index for the experimental groups

	Fear learning group		Control group	
	Sham	**tVNS**	**Sham**	**tVNS**
N (female/male)	20 (15/5)	20 (16/4)	20 (12/8)	20 (14/6)
Age (years)	23.75 (*SD* = 3.34)	23.30 (*SD* = 4.22)	22.05 (*SD* = 3.20)	21.90 (*SD* = 2.83)
Body-Mass-Index (kg/m^2)	21.65 (*SD* = 1.89)	22.30 (*SD* = 2.00)	21.65 (*SD* = 1.89)	22.58 (*SD* = 2.24)

and a control group, receiving explicitly unpaired presentations of both stimuli (Lonsdorf et al., 2017; Rescorla, 1967). Moreover, we expanded the extinction period to investigate the extinction of defensive responding in more detail.

We hypothesized, that human laboratory participants of the fear learning group would show stronger cardiac deceleration during the presentation of a conditioned stimulus compared with controls, as they are suggested to function at a stage of attentive immobility with easy escape blocked (e.g., by social compliance; Lang et al., 2000). More specifically, we presumed that such fear bradycardia is primarily expressed in stronger *prolonged* cardiac deceleration late during the CS presentation, which is suggested to reflect increased sensory intake or stimulus anticipation (e.g., an aversive US; Hodes et al., 1985), but not in early cardiac deceleration, which has been viewed as a transient detecting response indexing stimulus registration (Bradley, 2009; Graham, 1987; Hodes et al., 1985). Moreover, such prolonged fear bradycardia was expected to be significantly correlated to behavioral low-level correlates of attentive freezing, that is being related to increased potentiation of the startle reflex. Importantly, as primitive thalamic projections to the amygdala are particularly involved in the expression of fear during single-cue conditioning protocols, we expected that this correlation is stronger than the association between bradycardia and declarative knowledge of CS-US contingency, which has been suggested to require higher order cortical involvement (for a review see LeDoux, 1995). Additionally, defensive responding and, thus, freezing has shown to decrease during an extinction training, presumably due to reduction of CeA activity by inhibitory projections from the ventromedial prefrontal cortex (vmPFC) and the basolateral amygdala (BLA; Amano et al., 2010; Ehrlich et al., 2009; Gewirtz et al., 1997; Milad & Quirk, 2012). Thus, we hypothesized that late cardiac deceleration would extinguish in the fear learning group. Animal research has indicated that the stimulation of the vagus nerve may facilitate such extinction, possibly due to increasing noradrenergic activation of the BLA and vmPFC by way of its afferent projections to the locus coeruleus noradrenergic system (Mueller & Cahill, 2010; Peña et al., 2013, 2014). Transcutaneous vagus nerve stimulation (tVNS), involving the non-invasive stimulation of the exclusively vagally innervated left cymba conchae, is presumed to lead to an activation of afferent fibers of the left auricular vagus nerve and has shown to similarly increase activity in both the amygdala and vmPFC, correspondingly resulting in promoted fear extinction in humans when applied during extinction training (Burger et al., 2016, 2017, 2018; Frangos et al., 2015; Peuker & Filler, 2002; Szeska et al., 2020). Thus, we used transcutaneous vagus nerve stimulation as an additional proof of concept in our paradigm and expected that tVNS would promote the extinction of prolonged cardiac deceleration during extinction training by facilitating inhibition of human neural freezing circuitry. Importantly, attentive immobility as well as fear bradycardia are suggested to be primarily parasympathetically dominated defensive responses (Campbell et al., 1997; Roelofs, 2017). Thus, we hypothesized that a potential tVNS-induced attenuation of cardiac deceleration would be driven by an inhibition of parasympathetic control of the heart, rather than by an increase in sympathetic nervous activity, as indexed by the skin conductance level (SCL).

2 | METHOD

2.1 | Participants

The study included 80 participants, primarily students of the University of Greifswald (*M* = 22.75, *range* = 18 to 34 years; 57 women; see Table 1 for further information). During a phone interview, all participants reported to be in the desired age range (18–35 years), to have a body-mass-index in normal range (18.5 kg/m^2 to 27 kg/m^2), and to be free from any previous or current medical or mental condition, which would have been associated with an affection of any of the outcome variables or would have contraindicated the use of tVNS (i.e., cochlear implants or pregnancy, checked by a pregnancy test). The sample and data set is the same as has been reported by Szeska et al., (2020). Each participant gave her/his informed consent and received either monetary reward (34 €) or partial course credits. The study was approved by the ethical committee of the German Society for Psychology ("Deutsche Gesellschaft für Psychologie; DGPs").

2.2 | Stimulus materials

Figure 1a gives an overview of the used stimulus materials. All visual stimuli were presented on a 24-inch computer monitor

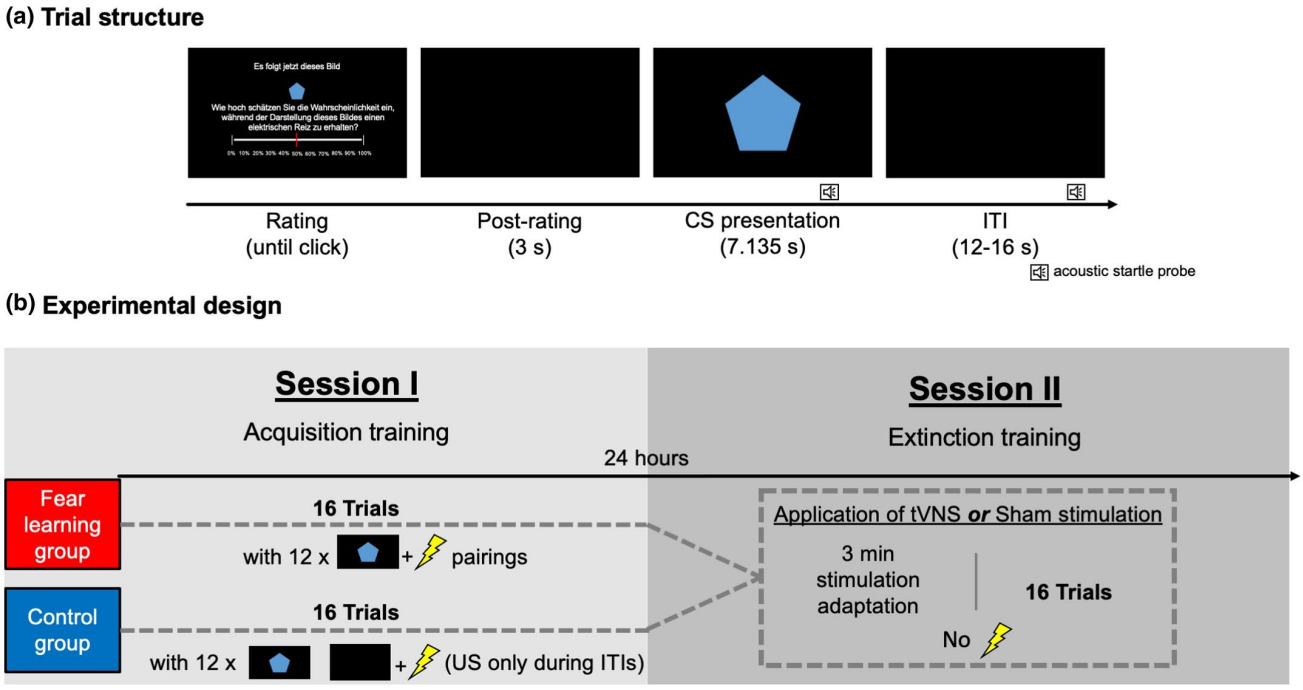


FIGURE 1 (a) Trial structure during the single-cue fear conditioning and extinction paradigm. Each trial began with a CS-US contingency rating, where the CS (blue pentagon) was previewed in smaller size, and participants were instructed to rate the probability, that this cue would be followed by the US during the upcoming CS presentation in full size (English translation of the German instruction: "Next, this picture will follow. How likely do you think is it to receive an electrical shock during the upcoming presentation of this picture?"). Three seconds after completing the rating, the cue was presented in full size on the screen, ensuring that physiological fear responses were not affected by any parallel cognitive evaluation task. (b) Schematic presentation of the analyzed experimental sessions. The acquisition and extinction training consisted of 16 trials each. The fear learning group received paired presentations of the CS and US in 12 of the 16 trials (75% CS-US contingency) during acquisition (i.e., four CS presentations without US), whereas individuals of the control group received 16 presentations of the CS and 12 shocks during the inter-trial interval (ITI) so that CS and US were explicitly unpaired (0% CS-US contingency). Extinction started with a 3 min adaption period to the stimulation device (for both sham stimulation and tVNS, respectively). During extinction, 16 CSs were presented in both groups without any US. Half of the fear learning and control group underwent the extinction training under the influence of tVNS, whereas the other half received a sham stimulation of the earlobe. Throughout each experimental session, acoustic startle probes were presented during the presentations of the CS and during the ITIs

(1,024 × 768 pixel resolution) 1.45 m in front of the participant. The CS was a blue pentagon on a black background, which was displayed for 7.135 s, whereas a black screen, presented for 12, 14, or 16 s ($M = 14$ s), served as inter-trial interval (ITI).

An unpleasant, individually adjusted electrical shock with a duration of 625 ms, consisting of 125 single pulses, each with a duration 2 ms and a 3 ms break between pulses was applied by an S-48K stimulator (Grass instruments, West Warwick, RI, USA) and was used as US. Importantly, there was no significant difference between the tVNS and sham condition in adjusted US intensity ($M_{tVNS} = 3.41$ mA, $SD = 1.53$; $M_{sham} = 3.44$ mA, $SD = 1.47$; Stimulation and Stimulation × Group, all Fs < 1.09, all ps > .30).

A binaurally presented 95 dB(A) burst of white noise with a duration of 50 ms and an instant rise/fall time (<1 ms), presented by AKG K66 headphones, was used as acoustic startle probe to elicit the startle eyeblink response.

The device for transcutaneous vagus nerve stimulation (CMO2, Cerbomed, Erlangen, Germany) was applied during session 2 (extinction training) at the left auricle with two titanium electrodes positioned in either of two locations: In the tVNS condition, the electrodes were positioned at the cymba conchae, which is exclusively innervated by the auricular branch of the vagus nerve (ABVN), whereas in the sham condition, electrodes were placed in the center of the earlobe, which is free of vagal innervation because of being innervated by the great auricular nerve (GAN; Peuker & Filler, 2002). Electrical stimulation was delivered during the stimulation adaptation period (3 min), as well as throughout the following extinction training (session 2; approximately 10 min) with a pulse width of 200–300 μs at a rate of 25 Hz, applying a 30 s ON and 30 s OFF procedure. Ensuring the activation of either the ABVN or GAN, participants were required to individually adjust the stimulation intensity at the beginning of session 2 to be clearly perceivable, but below the pain

threshold (see also 2.3.2). Importantly, the mean stimulation intensity did not differ between the tVNS and sham condition ($M_{tVNS} = 2.28$ mA, $SD = 1.13$; $M_{sham} = 2.53$ mA, $SD = 1.11$; stimulation and stimulation × group, all Fs < 1.31, all ps > .25).

2.3 | Experimental design and procedure

Figure 1 provides a linear depiction of the trial structure (a), the experimental design and procedure (b). We used a 2 × 2 between-subject design to test our hypotheses, with Group (fear learning vs. control group) and Stimulation (tVNS vs. sham stimulation) as between-subject factors. Consequently, eligible participants were allocated to one of four conditions: a fear learning group receiving tVNS ($n = 20$), a fear learning group receiving sham stimulation ($n = 20$), a control group receiving tVNS ($n = 20$), and a control group receiving sham stimulation ($n = 20$). The allocation to either of the four conditions was randomized and single-blind sham controlled.

Participants were seated in a dimly lit, sound-attenuated room during each experimental session. Sensors for physiological recording as well as the electrodes to deliver the electrical shock at the non-dominant hand's wrist were attached prior to any experimental manipulation. Each session began with a startle habituation phase, during which six acoustic startle probes were presented (inter-stimulus intervals of 7, 9, 10, 6, and 8 s; $M = 8$ s; duration: 84 s), ensuring the adaptation of startle magnitudes to a stable baseline.

2.3.1 | Acquisition training (session 1)

Prior to acquisition training, participants underwent a shock workup, during which the experimenter individually adjusted the US intensity following a standardized protocol to a level, which the participant perceived as clearly unpleasant but not painful. The workup consisted of a number of sample shocks, starting at an intensity of 2.0 mA. After each shock administration, participants were asked to rate the shock intensity on a continuous 5-point visual analog scale, ranging from "1 (not painful/annoying)" to "5 (very painful/annoying)." After each rating, the shock intensity was increased to finally achieve an intensity that was rated as "4 (unpleasant/quite annoying)." As soon as the shock was rated as "4 (unpleasant/quite annoying)," the shock workup was terminated and the respective shock intensity was used for the experiment (see also Klumpers et al., 2010).

After the shock workup and just before the acquisition training, all participants were instructed, that the CS, the US and acoustic startle probes may be presented at any time, with no explicit information given with regard to the CS-US contingencies.

During the acquisition training, all participants received 16 presentations of the CS. In the fear learning group, the CS was paired with the aversive US in 12 of the 16 trials (6.5 s after CS onset; 75% CS-US contingency) to induce a reliable and robust conditioned fear response while increasing its resistance to extinction. By contrast, the control group received explicitly unpaired presentations of CS and US, which was delivered 12 times only during the inter-trial intervals (3, 4, 5, 6, 7, 8, 11, or 12 s after ITI onset, $M = 6.98$ s; 0% CS-US contingency). Startle probes were delivered during the CS in 12 out of 16 trials for both groups either at 4.5, 5 or 5.5 s (four trials for each probe time) after the CS onset arranged in eight experimental orders, participants were randomly assigned to. Moreover, 12 startle probes were presented during the inter-trial intervals (ITIs). By using a minimum interval of 1 s between startle probe and US onset, we ensured that the presentation of both stimuli was not confounded in any of the experimental groups.

2.3.2 | Extinction training (session 2)

The second experimental session took place 24 ± 4 hr after the acquisition training. After the electrodes for physiological assessment and US-application were refitted, the tVNS/sham stimulation device was positioned at the participants' left ear in the desired location and a tVNS/sham stimulation workup began, where participants were instructed to set the stimulation intensity to be clearly perceivable, but without being painful. Following the same protocol as Ventura-Bort and colleagues (2018), adjustment started at an intensity of 0.1 mA and after each up- or down-adjustment of 0.1 mA participants were asked to rate their subjective sensation of the stimulation intensity on a visual 11-point scale, ranging from "nothing (0)," "light tingling (3)," "strong tingling (6)" to "painful (10)". The workup lasted until a "strong tingling" sensation of 8 was reported by the participant, after which a full 30 s ON and 30 s OFF stimulation protocol was run in order to provide an experience of the stimulation, as it would be during the extinction training (see also Ventura-Bort et al., 2018). Only if the participants still rated the sensation as 8 after the protocol, the adjusted stimulation intensity would be used for the extinction training—otherwise the workup went on until that point was reached.

Subsequently, participants were informed that the upcoming second experimental session would begin with a 3-min period to adapt to the stimulation (either tVNS or sham). No other stimuli (startle probes or CSs) were presented during this period. Participants were further instructed that after the adaptation period any of the stimuli might be presented, that have also been presented during session 1. Again, no explicit information was given with regard to the CS-US contingencies. During extinction

training, the CS was presented 16 times without any US. Presentation of acoustic startle probes was similar to order of the acquisition training.

2.4 | Assessments and data reduction

2.4.1 | Electrocardiogram (ECG; Heart rate)

The ECG was measured using an Einthoven Lead II setup with two electrolyte filled (Marquette Hellige, Freiburg, Germany) standard Ag/AgCl electrodes (8 mm diameter). Using a Coulbourn system, the raw signal was filtered with an 8–13 Hz band-pass filter and amplified by the factor 2000. ECG data were digitally sampled at 400 Hz and artifact corrected using ANSLAB (v. 2.4; Autonomic Nervous System Laboratory, University of Basel, Switzerland), and subsequently converted to heart rate in beats per minute for every half-second of the sampling period (Graham, 1978). Finally, allowing to quantify baseline-independent cardiac responding during the CS, heart rate during the CS was subtracted from base period heart rate (mean of the first two half-seconds after CS onset) for every half-second after the CS onset for the full duration of CS presentation (14 data points for the 7.135 s CS duration). These half-second bins were averaged across all trials for each experimental session and additionally for each half of the extinction training to analyze the time course of extinction learning.

As conditioned cardiac responses have shown to follow a triphasic course, we additionally identified average peaks of early cardiac deceleration (*D1*, slowest half-second between 1 and 2 s after CS onset), acceleration (*A1*; fastest half-second between 2 and 5 s after CS onset), and late cardiac deceleration (*D2*; slowest half-second between 5 and 7 s after CS onset) for each experimental session adapted from the rules of Gatchel and Lang (1973). Average peaks of the heart rate responses are expressed in beats per minute change scores, deviated from the base period (Δ bpm).

2.4.2 | Electromyography (EMG; Startle eyeblink response)

We measured the eyeblink component of the startle response, elicited by the acoustic startle probe, by recording the electromyographic activity of the orbicularis oculi muscle underneath the left eye by using two electrolyte filled (Marquette Hellige, Freiburg, Germany) Ag/AgCl miniature surface electrodes (3 mm diameter, Sensormedic, Yorba Linda, CA, USA), which were attached on the skin over the muscle. The EMG signal was amplified by a Coulbourn S75-01 amplifier and filtered with a 30 Hz high-pass and a Kemo LEM-VBf8-03 400 Hz low-pass filter (smoothing the rectified signal with a time constant of 10 ms). Moreover, a notch filter (50 Hz) was used. The signal was digitally sampled at a rate of 1,000 Hz between 100 ms before and 400 ms after the startle probe onset. Startle eyeblink responses were scored semi-automatically with a computer program, identifying blink onset and peak amplitude (Globisch et al., 1993). Each detected startle eyeblink response was additionally visually inspected for artifacts (Blumenthal et al., 2005) and manually corrected if necessary. Only blinks were scored as valid startle responses, which started 20–120 ms after the startle probe onset and peaked within 150 ms, with a minimum amplitude of 1.954 μV. If no blink was detected, the trials were scored as zero responses. Based on previously published guidelines, we set trials as missing if clear movement artifacts, excessive baseline activity, or artifacts due to tVNS/sham stimulation were found (Blumenthal et al., 2005). For acquisition training (session 1), 0.5% were scored as zero responses ($M = 0.15$), and 2.2% of all probed trials were set as missing ($M = 0.66$). For the extinction training (session 2), 0.3% of all probed trials were scored as zero responses ($M = 0.09$), and 31.2% were set as missing ($M = 9.38$; higher rate of missings due to tVNS/sham stimulation-induced noise). After scoring, raw blink magnitudes were z-transformed and finally T-standardized ($50 + (z \times 10)$) individually for each participant to control for individual differences in overall startle magnitude. Finally, we computed the mean startle potentiation (difference of T-transformed CS startle and T-transformed ITI startle magnitude) for each experimental session.

2.4.3 | Shock expectancy ratings

Before each CS presentation, participants were required to rate their expectancy to receive an US during the upcoming CS on a continuous 11-point visual analog scale (ranging from "0%" to "100%") by shifting a red cursor and pushing the left mouse button (see Figure 1a). During this rating, the CS was presented in smaller size above the line rating. This procedure is very much comparable to clinical practice during exposure-based treatments, during which patients are asked to rate the likelihood, that their central concern might become true (e.g., fainting), before the exposure exercise begins (see Hollandt et al., 2020). There was no time restriction for completing the rating. After the rating was completed, a three second post-rating period (black screen) followed. After the post-rating period, the CS was displayed in full size on the screen. Thus, we ensured that physiological responses evoked by the CS were not affected by a parallel cognitive evaluation task. Equivalently to the startle responses, we computed the mean shock-expectancy rating separately for each experimental session as an index of declarative knowledge of CS-US contingency.

2.4.4 | Skin conductance level

The SCL was measured from the hypothenar eminence of the palmar surface of the participant's non-dominant hand to provide an index of sympathetic nervous activity. Two Ag/AgCl electrodes (8 mm diameter) were filled with a 0.05 M sodium chloride electrolyte medium before attachment. The signal was amplified by a Coulbourn S71-22 skin conductance coupler, which provided a constant current of 0.5 V across the two electrodes, sampled at a rate of 10 Hz and processed with a resolution of 0.01 μS. SCL for every half-second after the CS onset for the full CS duration (14 points for analysis during the 7.135 s stimulus presentation) was subtracted from the base period SCL (mean of the first two half-seconds after CS onset) and averaged across all trials for the extinction training, thus using the same scoring procedure as for heart rate.

2.4.5 | Statistical analyses and Figure creation

We analyzed the course of conditioned cardiac responses during each session (acquisition training and extinction training) using linear mixed models with only fixed effects included (see Bagiella et al., 2000; Duricki et al., 2016). On the one hand, such linear mixed regression models advantageously also include participants with missing values, whereas on the other hand it is possible to model the error covariance structure in a way that best fits the data, providing higher statistical power of analysis (Bagiella et al., 2000; Duricki et al., 2016). We created all linear mixed regression models using restricted maximum likelihood estimation, to include all available data (Duricki et al., 2016), and modeled the error covariance structure of the repeated measurements by specifying a first-order autoregressive covariance structure with heterogeneous variances (ARH1). This type of covariance structure was chosen because it provided the best fit to the sample data according to Akaike's information criterion while being parsimonious in parameter estimation, thus following recommendations of Duricki and colleagues (2016). *Group* (fear learning vs. control group) and *Stimulation* (tVNS vs. sham stimulation) served as between-subject factors, and *Time* (14 half-second bins during the CS presentation) served as within-subject factor. Moreover, to specifically examine the impact of the extinction training on cardiac responding, we compared cardiac waveforms during the first half with cardiac waveforms during the second half of the extinction training, by additionally including the within-subject factor *Half* into analyses.

If significant differences between average, heart rate curves were found across groups during acquisition or extinction, univariate analyses of variances were carried out, analyzing between-subject differences between average peaks of cardiac decelerative (*D1* and *D2*) and accelerative response components (*A1*) of phasic heart rate changes during each session with *Group* and *Stimulation* as between-subject factors. In a second step, we compared cardiac peak component scores between the first and second half of extinction to analyze the impact of the extinction in more detail. We used the linear mixed regression models as described above, with *Group* and *Stimulation* as between-subject factors and *Half (first vs. second)* as within-subject factor. To examine the potential influence of startle probes on late cardiac deceleration (*D2*) during acquisition and extinction (probed vs. non-probed trials), we did additional analyses including the within-subject factor *Probe (probed vs. non-probed)*.

Moreover, we computed Spearman rank correlations between cardiac response components and mean startle potentiation as well as mean shock expectancy ratings for each experimental session to evaluate the association between heart rate changes, startle potentiation, and CS-US shock expectancy ratings. In a second step, we tested whether these correlations were significantly different following the procedure recommended by Meng and colleagues (Meng et al., 1992). We also calculated correlations during extinction for each of the four experimental groups to assess the influence of the stimulation conditions and the learning experience on the correlational pattern (see Eid et al., 2011 for the procedure used to compare independent correlations).

Changes in SCL during the extinction training were analyzed as an index of sympathetic nervous activity by using linear mixed regression models as described above, with *Group* (fear learning vs. control group) and *Stimulation* (tVNS vs. sham stimulation) as between-subject factors and *Time* (14 half-second bins) serving as within-subject factor.

Partial eta-squared was computed following recommendations by Lakens (2013). Bonferroni correction was applied, when relevant. All statistical analyses were conducted using IBM SPSS Statistics 25. Microsoft Excel and Microsoft PowerPoint were used for Figure creation.

3 | RESULTS

3.1 | Acquisition training

The overall heart rate response to the CS during acquisition training in the fear learning (red line) and the control group (blue line) is depicted in panel (a) of Figure 2. All participants showed a significant heart rate deceleration to the CS (Time, $F(13,320.25) = 19.58$, $p < .001$, $\eta^2_p = .44$; Figure 2a). However, a significant time by group interaction indicated stronger cardiac deceleration in the fear learning relative to the control group prior to the delivery of the US (Time × Group, $F(13,320.25) = 3.68$, $p < .001$, $\eta^2_p = .13$; significant group differences 6–7 s after CS onset). Correspondingly, although

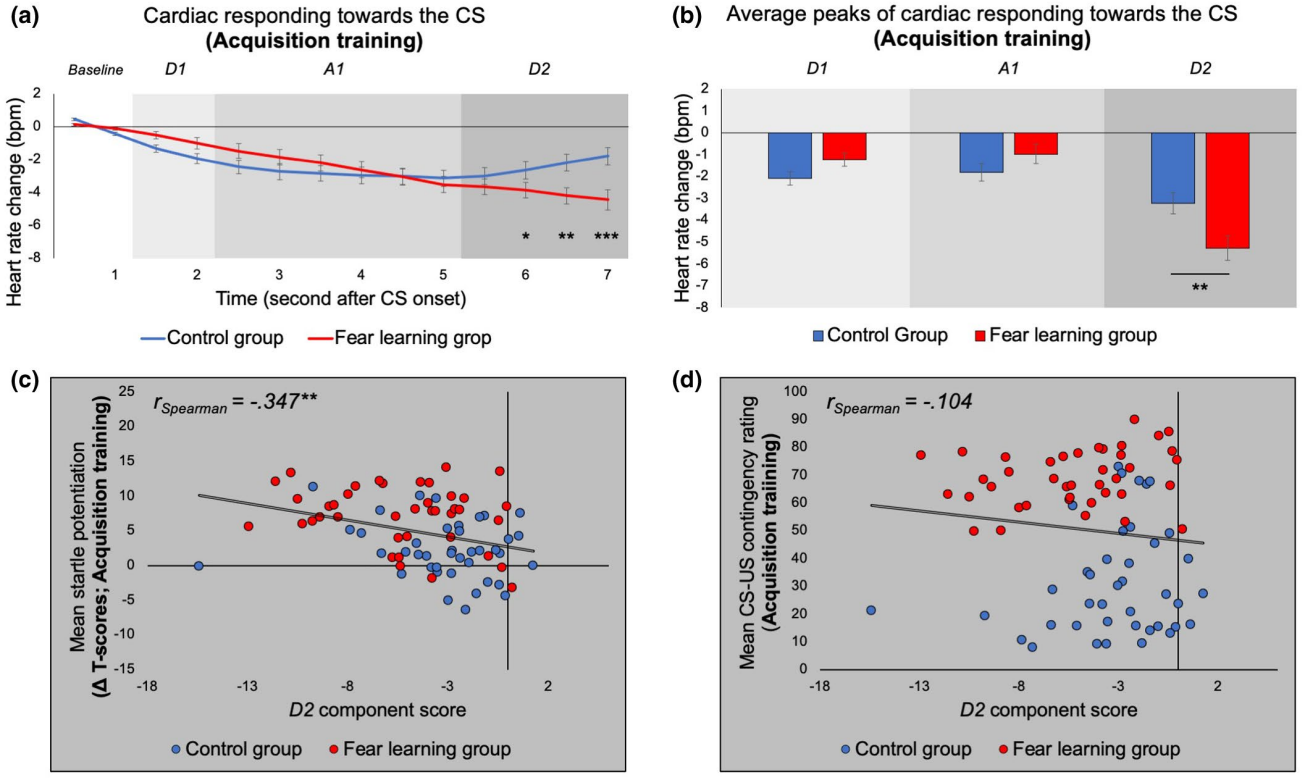


FIGURE 2 (a) Heart rate change after CS onset during the acquisition training for the fear learning (red line) and control group (blue line), averaged across all 16 acquisition trials, depicted in half-second bins. (b) Average peaks of cardiac response components during the acquisition training for the fear learning (red bars) and control group (blue bars). *D1* indicates the mean slowest half-second between 1 and 2 s after CS onset. *A1* indicates the mean fastest half-second between 2 and 5 s after CS onset. *D2* indicates the mean slowest half-second between 5 and 7 s after CS onset. (c) Scatter plot of mean startle potentiation (standardized [T-scores] startle magnitudes elicited during the CS minus standardized [T-scores] startle magnitudes elicited during the ITI, averaged across all probed trials) as a function of *D2* component score variation during the acquisition training for the fear learning (red dots) and control group (blue dots). (d) Scatter plot of mean CS-US contingency rating (averaged across all trials) as a function of *D2* component score variation during the acquisition training for the fear learning (red dots) and control group (blue dots). For all graphs: Time windows for the analyses of the average peaks of cardiac response components are depicted in different gray scales ranging from light (*D1*) to medium (*A1*) and dark gray (*D2*). Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations with $^{*}p \leq .05$, $^{**}p \leq .01$, and $^{***}p < .001$

no significant differences were found between the fear learning and control group in early decelerative (*D1*: Group, $F(1,76) = 3.94$, $p = .051$; Figure 2b) or accelerative cardiac peak components (*A1*: Group, $F(1,76) = 1.95$, $p = .167$; Figure 2b), participants of the fear learning group showed increased *late* cardiac decelerative peaks (*D2*) compared with participants of the control group (Group, $F(1,76) = 7.66$, $p = .007$, $\eta^2_p = .09$; Figure 2b). Although cardiac deceleration was overall smaller in probed relative to the non-probed trials (Probe, $F(1,76) = 5.79$, $p = .018$, $\eta^2_p = .07$; Figure S1a), such effect was significantly smaller in the fear learning group (Probe × Group, $F(1,76) = 5.42$, $p = .023$, $\eta^2_p = .07$; Figure S1b). As predicted, overall correlation between late heart rate deceleration (*D2*) and fear potentiated startle was significant ($r_{Spearman}$ (80) = −.347, $p = .002$). By contrast, no significant overall correlation between cardiac deceleration and CS-US expectancy ratings was observed ($r_{Spearman}$ (80) = −.104, $p = .358$; Figure 2c,d). Further testing revealed that the correlation between late cardiac deceleration (*D2*) and startle potentiation was significantly stronger than the correlation between the *D2* component and CS-US contingency ratings ($z = -1.80$, p(one-tailed) = .036). By contrast, neither early decelerative (*D1*) nor accelerative cardiac peak responses (*A1*) were significantly associated with startle potentiation (*D1*: $r_{Spearman}$ (80) = .010, $p = .933$; *A1*: $r_{Spearman}$ (80) = −.091, $p = .420$) or CS-US expectancy ratings (*D1*: $r_{Spearman}$ (80) = .196, $p = .086$; *A1*: $r_{Spearman}$ (80) = .085, $p = .454$). As expected, no effects of stimulation were found during the acquisition training, as no stimulation was yet applied (all Fs < .66, all ps > .427).

3.2 | Extinction training

During the extinction training 24 hr later, all participants continued to show a significant heart rate deceleration in response

to the CS, indicating that the CS has acquired relevance to the participants (Time, $F(13,319.59) = 19.25$, $p < .001$, $\eta^2_p = .44$; Figure 3a,b). Yet, the fear learning group displayed significantly stronger heart rate deceleration relative to controls during the presentation of the conditioned stimulus, particularly during the late phase of the CS (Time × Group, $F(13,319.59) = 3.27$, $p < .001$, $\eta^2_p = .12$; Figure 3a,b). As expected, tVNS significantly attenuated prolonged cardiac deceleration but only in subjects of the fear learning group, whereas cardiac reactivity was unaffected by stimulation in controls (Group × Stimulation, $F(1,99.43) = 5.81$, $p = .018$, $\eta^2_p = .65$; Time × Group × Stimulation, $F(13,319.59) = 2.99$, $p < .001$, $\eta^2_p = .11$; Figure 3a,b).

In the sham condition, fear learning group participants showed significantly stronger overall cardiac deceleration relative to controls (Group, $F(1,48.87) = 16.33$, $p < .001$, $\eta^2_p = .25$; Figure 3a), again with strongest deceleration during the late phase of CS-processing (Time × Group, $F(13,97.45) = 4.32$, $p < .001$, $\eta^2_p = .37$; Figure 3a; significant group differences 2.5–7 s after CS onset). Accordingly, although both groups of the sham condition did not differ in early deceleration (*D1*; Group, $F(1,38) = 1.79$, $p = .189$; Figure 3c), the fear learning group displayed lower accelerative and stronger late decelerative peak responding relative to controls (*A1*: Group, $F(1,38) = 21.24$, $p < .001$, $\eta^2_p = .36$; *D2*: Group, $F(1,38) = 14.51$, $p < .001$, $\eta^2_p = .28$; Figure 3c). Importantly, these between-group differences in the sham condition were significantly stronger at the beginning of the experimental session and declined throughout the extinction training (Half × Group, $F(1,212.68) = 5.47$, $p = .020$, $\eta^2_p = .03$; Figure 4a,b), which resulted from extinguished heart rate deceleration in fear learning group participants (Half, $F(1,112.34) = 3.49$, $p = .064$, $\eta^2_p = .03$), rather than from increased deceleration in controls (Half, $F(1,95.06) = 1.92$, $p = .169$). Accordingly, although early decelerative peak responding remained stable in all sham-stimulated subjects (*D1*: Half × Group, $F(1,38) = .76$, $p = .388$; Figure 4e), fear learning group participants showed lower cardiac accelerative and

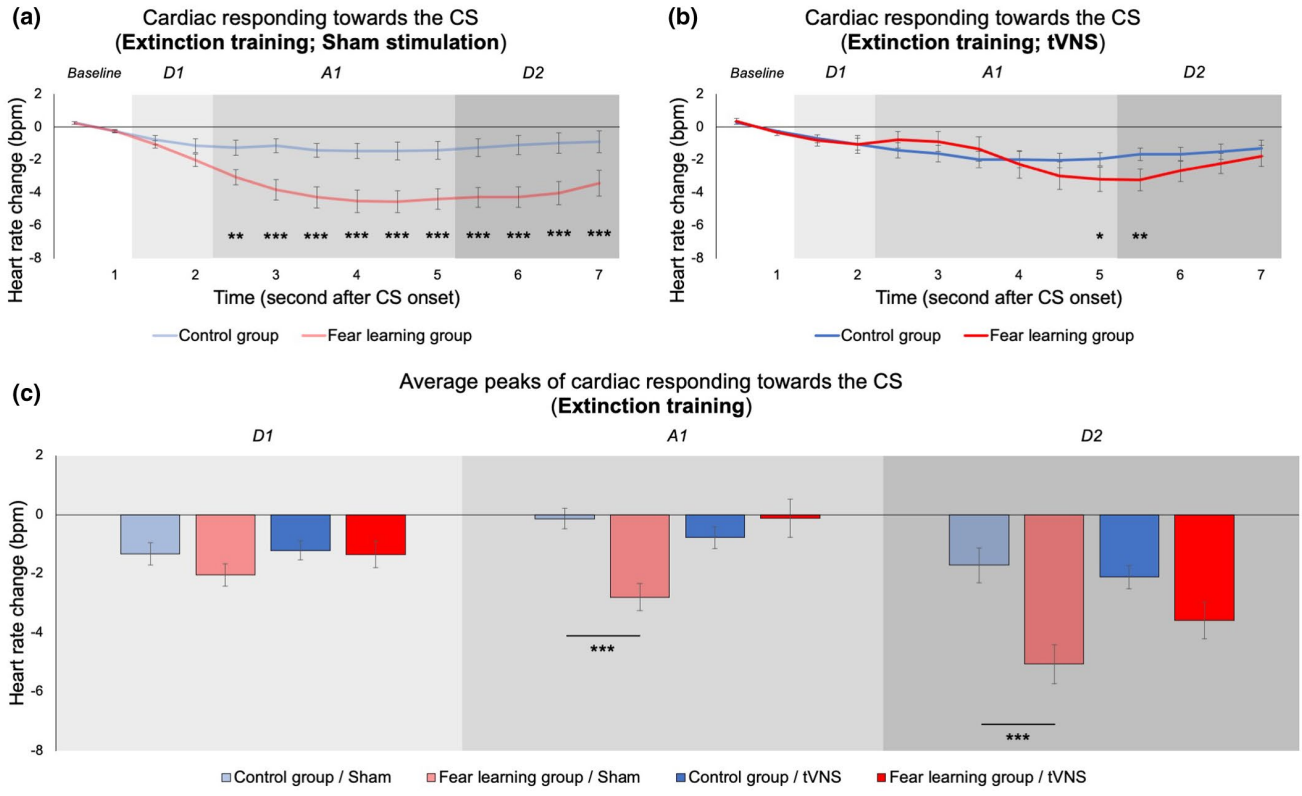


FIGURE 3 (a) Heart rate change after CS onset during the extinction training in the sham condition, averaged across all 16 trials, depicted in half-second bins for the fear learning (light red line) and control group (light blue line). (b) Heart rate change after CS onset during the extinction training in the tVNS condition, averaged across all 16 trials, depicted in half-second bins for the fear learning (red line) and control group (blue line). (c) Average peaks of cardiac response components during the extinction training for both the fear learning and control group in the sham condition (light red and light blue bars) and tVNS condition (red and blue bars). *D1* indicates the mean slowest half-second between 1 and 2 s after CS onset. *A1* indicates the mean fastest half-second between 2 and 5 s after CS onset. *D2* indicates the mean slowest half-second between 5 and 7 s after CS onset. For all graphs: Time windows for the analyses of the average peaks of cardiac response components are depicted in different gray scales ranging from light (*D1*) to medium (*A1*) and dark gray (*D2*). Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations with $*p \leq .05$, $**p \leq .01$, and $***p < .001$

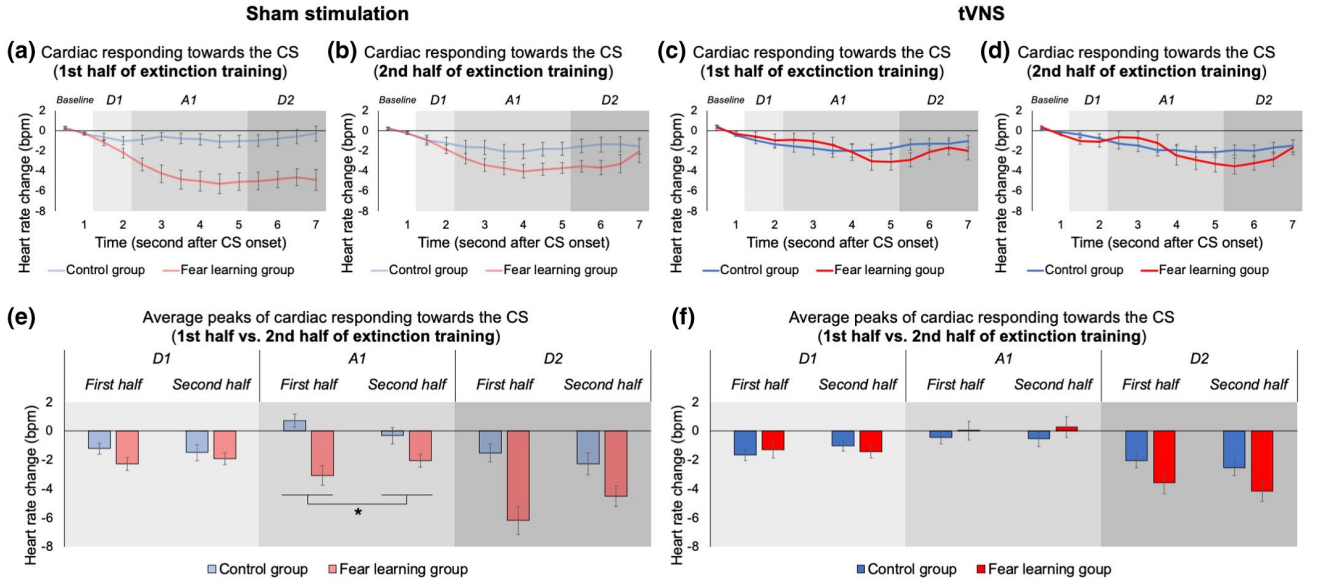


FIGURE 4 (a) and (b) Heart rate change after CS onset during the first half (a) and second half (b) of the extinction training in the sham condition, averaged across all 16 trials, depicted in half-second bins for the fear learning (light red line) and control group (light blue line). (c) and (d) Heart rate change after CS onset during the first half (a) and second half (b) of the extinction training in the tVNS condition, averaged across all 16 trials, depicted in half-second bins for the fear learning (red line) and control group (blue line). (e) and (f) Average peaks of cardiac response components during the first half (e) and second half (f) of the extinction training for both the fear learning and control group in the sham condition (light red and light blue bars, respectively) and tVNS condition (red and blue bars, respectively). *D1* indicates the mean slowest half-second between 1 and 2 s after CS onset. *A1* indicates the mean fastest half-second between 2 and 5 s after CS onset. *D2* indicates the mean slowest half-second between 5 and 7 s after CS onset. For all graphs: Time windows for the analyses of the average peaks of cardiac response components are depicted in different gray scales ranging from light (*D1*) to medium (*A1*) and dark gray (*D2*). Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations with $*p \leq .05$, $**p \leq .01$, and $***p < .001$

stronger late decelerative peaks during the first half of extinction, which both appeared to decline throughout the session (*A1*: Half × Group, $F(1,38) = 4.78$, $p = .035$, $\eta^2_p = .11$; *D2*: Half × Group, $F(1,38) = 4.08$, $p = .051$, $\eta^2_p = .09$; Figure 4e).

In contrast to the sham condition, the application of tVNS abolished differences in overall cardiac responding between the fear learning and control group during the extinction training (Group, $F(1,50.87) = .36$, $p = .554$; Figure 3b), which was evident right from the beginning of the experimental session (heart rate curves: Half × Group, $F(1,187.37) = .03$, $p = .854$; Figure 4b,c; peak components: Half × Group, all $Fs < 1.930$, all $ps > .172$; Figure 4f). Nevertheless, the fear learning group still displayed stronger cardiac deceleration during the late phase of the CS processing (Time × Group, $F(13,179.93) = 2.15$, $p = .014$, $\eta^2_p = .14$; significant group differences 5–5.5 s after CS onset; Figure 3b). Accordingly, we observed no significant group differences in early decelerative peaks (*D1*; $F(1,38) = .06$, $p = .807$) or cardiac acceleration (*A1*; $F(1,38) = .78$, $p = .382$), whereas we found a trend for stronger late decelerative peaks in the fear learning group relative to controls in the tVNS condition (*D2*: $F(1,38) = 3.89$, $p = .056$, $\eta^2_p = .09$; Figure 3c).

Further analyses revealed that the abolished group differences in the tVNS condition did not result from attenuated early or late decelerative peak responding (*D1*: Stimulation × Group, $F(1,76) = .57$, $p = .454$; *D2*: Stimulation × Group, $F(1,76) = 2.69$, $p = .105$; Figure 3c), but from elevated cardiac acceleration in vagally stimulated fear learning group subjects compared with the sham condition (*A1*: Stimulation × Group, $F(1,76) = 12.47$, $p = .001$, $\eta^2_p = .14$; Figure 3c). Thus, although both active and sham-stimulated subjects of the fear learning group displayed significant cardiac deceleration immediately prior to the US (*D2* component or late cardiac deceleration), tVNS resulted in a significantly delay of such increased cardiac deceleration due to transient cardiac acceleration 2 s after CS onset. As no differences between stimulation conditions were found in SCL change during the CS presentation (all $Fs < 1.66$, all $ps > .070$; Figure 5a,b), our data indicate that such elevated cardiac acceleration was not accompanied by an increase in sympathetic nervous activity.

As during acquisition, late cardiac deceleration (*D2*) was overall reduced during probed compared with no probed trials (Probe, $F(1,76) = 5.40$, $p = .023$, $\eta^2_p = .07$; Figure S2a) – an effect that did not differ between fear learning and control participants (Probe × Group, $F(1,76) = .048$, $p = .827$; Figure S2b).

As expected, stronger prolonged cardiac deceleration continued to be significantly correlated with increased startle potentiation (*D2*: $r_{Spearman}(80) = -.376$, $p < .001$;

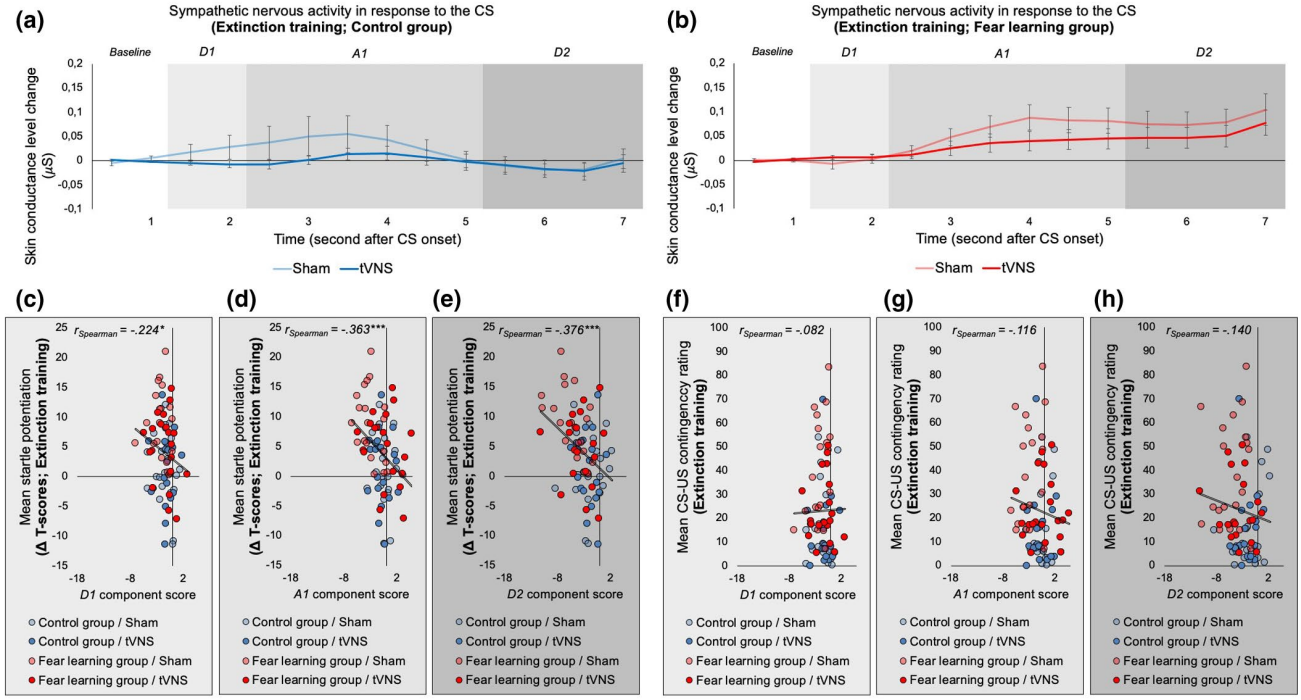


FIGURE 5 (a) and (b) Skin conductance level change after CS onset during the extinction training, averaged across all 16 trials, depicted in half-second bins for the tVNS (blue) and sham condition (light blue) of the control group (a) and fear learning group (b). (c), (d), and (f) Scatter plot of mean startle potentiation (standardized [T-scores] startle magnitudes elicited during the CS minus standardized [T-scores] startle magnitudes elicited during the ITI, averaged across all probed trials) as a function of *D1* component score variation (c), *A1* component score variation (d) and *D2* component score variation (e) during the extinction training for both the fear learning and control group in the sham (light red and light blue dots, respectively) and tVNS condition (red and blue dots, respectively). (f), (g), and (h) Scatter plot of mean CS-US contingency rating (averaged across all trials) as a function of *D1* component score variation (f), *A1* component score variation (g), and *D2* component score variation (h) during the extinction training for both the fear learning and control group in the sham (light red and light blue dots, respectively) and tVNS condition (red and blue dots, respectively). For all graphs: Time windows for the analyses of the average peaks of cardiac response components are depicted in different gray scales ranging from light (*D1*) to medium (*A1*) and dark gray (*D2*). Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations with $*p \leq .05$, $**p \leq .01$, and $***p < .001$

Figure 5e, but not with CS-US expectancy ratings (*D2*: $r_{Spearman}$ (80) = −.116, p = .305; Figure 5h) in the overall analysis, and the *D2* was significantly stronger related to startle potentiation than to CS-US contingency ratings during the extinction training (z = −1.99, p(one-tailed) = .023; Figure 5e,h). Analyses for the separate groups showed that the association between *D2* and startle potentiation was not significant and not modulated by tVNS in the control group (control group/tVNS: $r_{Spearman}$ = −.245, p = .298; control group/sham: $r_{Spearman}$ = .165, p = .486; z = −1.215, p(one-tailed) = .112), whereas it was significant in fear learning group participants in the sham condition (fear learning group/sham: $r_{Spearman}$ = −.627, p = .003; Figure 6a). As expected, this significant correlation was abolished by tVNS (fear learning group/tVNS: $r_{Spearman}$ = −.146, p = .539; Figure 6b; between group comparison of both correlations was significant z = −1.718, p(one-tailed) = .043).

In contrast to the acquisition training, however, we also found that earlier onset of cardiac deceleration (i.e., lower early decelerative and accelerative component scores between 1 and 5 s after stimulus onset) was significantly correlated with increased startle potentiation during the CS (*D1*: $r_{Spearman}$ (80) = −.224, p = .046; A1: $r_{Spearman}$ (80) = −.363, p < .001; Figure 5c,d), whereas not being related to CS-US expectancy ratings (*D1*: $r_{Spearman}$ (80) = −.082, p = .471;*A1*: $r_{Spearman}$ (80) = −.140, p = .216; Figure 5f,g). However, comparing both correlations for significant differences, there were no significant differences for D1 (z = −1.06, p(one-tailed) = .144; Figure 5c,f), whereas the correlation between *A1* and startle potentiation was significantly stronger than the correlation between *A1* and shock expectancy ratings (z = −1.72, p(one-tailed) = .043; Figure 5d,g).

4 | DISCUSSION

Prolonged cardiac deceleration has commonly been interpreted as an index of increased orienting toward motivationally significant stimuli (Bradley, 2009), but not necessarily also indexing defensive responding and fear (see Lonsdorf

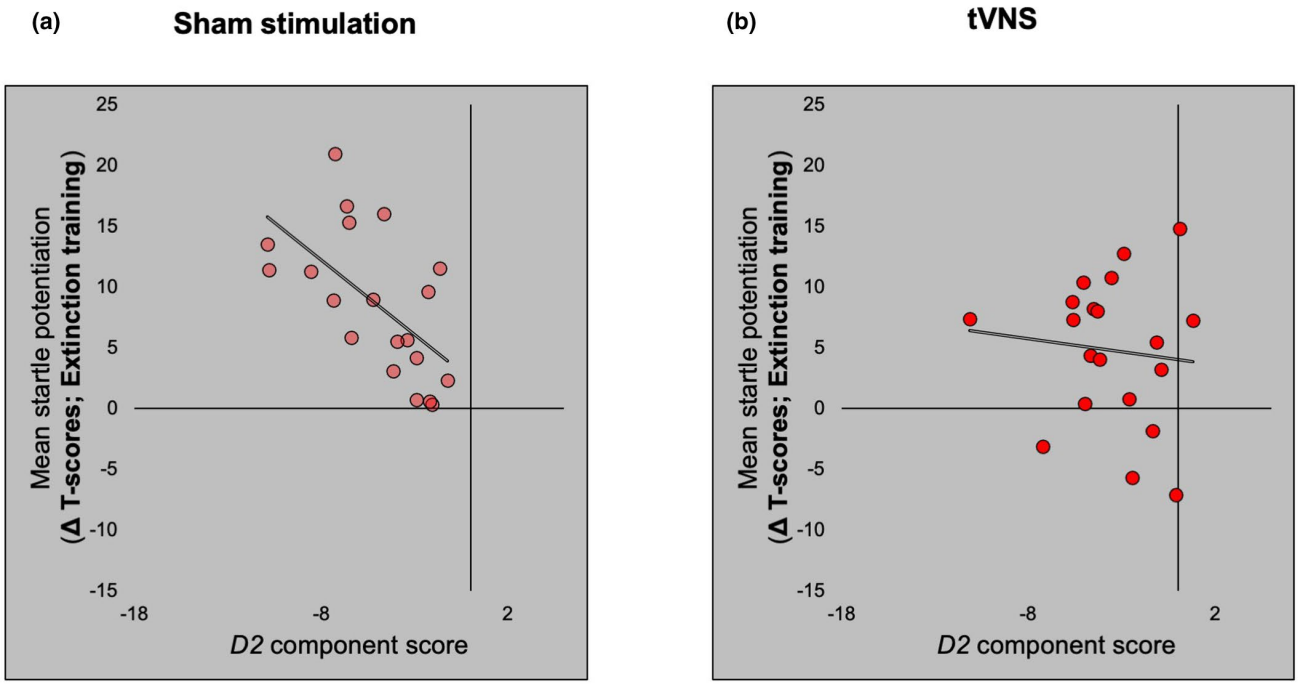


FIGURE 6 (a) and (b) Scatter plot of mean startle potentiation (standardized [T-scores] startle magnitudes elicited during the CS minus standardized [T-scores] startle magnitudes elicited during the ITI, averaged across all probed trials) as a function of *D2* component score variation during extinction training for the fear learning group receiving sham stimulation (a) and tVNS (b). Asterisks indicate statistical significance of correlations with $*p \leq .05$, $**p \leq .01$, and $***p < .001$

et al., 2017 for a review). However, previous research has consistently found cardiac deceleration during attentive immobility (freezing), a defense strategy toward distal threats when easy escape is blocked (Hamm, 2020; Krause et al., 2018; Löw et al., 2015; Marks, 1987; Roelofs, 2017). Following up on this research, we investigated cardiac reactivity as an index of defensive responding by applying a single-cue multiple-day human fear conditioning and extinction protocol. Transcutaneous vagus nerve stimulation (tVNS)—a non-invasive brain stimulation technique which has proven to facilitate the reduction of defensive responding during extinction compared with a sham stimulation of the earlobe (Burger et al., 2016, 2017, 2018; Szeska et al., 2020)—was used during extinction training as an additional proof of concept. We found stronger cardiac deceleration during the late phase of a conditioned stimulus predicting an approaching threat. Such conditioned "fear bradycardia" (Campbell et al., 1997) was significantly associated with potentiation of the startle response—a low level protective brain stem reflex (Davis, 2006), suggesting that human fear conditioning evokes a defensive response pattern that can best be characterized as attentive immobility, a defensive strategy observed in animals when withdrawal from danger is not possible or helpful (Marks, 1987). TVNS promoted the extinction of this conditioned defensive response pattern including a decoupling of the autonomic and protective reflex indices of the conditioned response.

In fact, the presentation of the conditioned stimulus evoked prolonged heart rate deceleration during both acquisition and extinction training in all participants, regardless of whether they underwent a fear conditioning protocol (fear learning group) or not (control group). Previous studies consistently found such prolonged cardiac deceleration in contexts involving perceptual processing and, thus interpreted heart rate deceleration as an index of increased orienting toward informative or motivationally significant stimuli (Bradley, 2009; Graham, 1979; Graham & Clifton, 1966; Lacey & Lacey, 1970). Accordingly, our data indicate that both groups showed increased orienting toward relevant stimuli, either threat- (fear learning group) or safety-signaling (control group).

As expected, both groups did not differ in the early decelerative component of CS-evoked cardiac changes (*D1*; initial 2 s after stimulus onset), supporting previous notions that this component of conditioned cardiac responding might be considered as a reflexive index of stimulus registration occurring in response to any low- or moderate-intensity stimulus regardless of repetition or motivational significance (Bradley, 2009; Graham, 1987). Correspondingly, this early decelerative responding did not correlate with startle potentiation or shock expectancy ratings during conditioning, did not extinguish and was not affected by tVNS.

Importantly, however, both groups in fact differed in the strength of *prolonged* cardiac deceleration. The fear learning

group displayed stronger cardiac deceleration compared with controls, which was more evident later during CS presentation (*D2* component) both during acquisition and extinction. The strength, and during extinction also earlier onset, of this heart rate deceleration was positively correlated with the potentiation of the startle response. In animal studies, heart rate deceleration correlated with behavioral freezing (Walker & Carrive, 2003), and in two other studies behavioral freezing was associated with startle potentiation in rats (Leaton & Borszcz, 1985; Plappert et al., 1993). In humans, mean cardiac deceleration during anticipation of shock was associated with mean decrease in body sway (Gladwin et al., 2016). The data of the current experiment, thus, complement the picture in showing a significant relationship between heart deceleration and startle potentiation and suggest that in humans—like in other animals – the same pattern of attentive immobility is evoked during the anticipation of a mild but aversive, inevitable unconditioned stimulus. Supporting this view, our data show that cardiac deceleration was strongest immediately prior to the delivery of the inevitable threat, thus mirroring findings from Löw et al. (2015) and Krause et al. (2018), which suggested that attentive immobility becomes stronger with increasing imminence of an inevitable threat. Moreover, during conditioning such late cardiac deceleration has shown to be particularly robust despite the effects of the presented startle probes, which previously have found to cause transient heart rate acceleration (Chen et al., 2014; Cook et al., 1992).

Furthermore, as the association between heart rate deceleration and startle potentiation was stronger compared with the correlation between cardiac deceleration and CS-US expectancy ratings, the correlational pattern of the data implies that cardiac deceleration and startle potentiation are both indices of the same defensive response strategy that is activated independently of the explicit declarative knowledge of the exact contingencies. In fact, this is in line with previous animal research, showing that cardiac deceleration in the face of an inevitable threat is driven by subcortical projections from the central nucleus of the amygdala (CeA) to the midbrain ventrolateral periaqueductal gray (vlPAG), which are also involved in the mediation of behavioral freezing and startle potentiation during threat processing (Applegate et al., 1983; Choi & Brown, 2003; Davis, 2006; Fendt & Fanselow, 1999; LeDoux, 1995; LeDoux et al., 1988; Walker & Carrive, 2003). This view is further supported by genetic research, showing increased threat-induced connectivity between the amygdala and PAG along with stronger heart rate deceleration during the presentation of a threat-predicting CS in carriers of the short allelic variant of the 5-HTTLPR (serotonin transporter-linked polymorphic region; Schipper et al., 2019). Individuals carrying this genetic variant also showed increased potentiation of the startle response during fear conditioning relative to l-allele carriers (Lonsdorf et al., 2009).

Animal research further showed that the CeA is inhibited by the basolateral amygdala and ventromedial prefrontal cortex during extinction learning, resulting in successive attenuation of attentive immobility – a process, which may be facilitated by stimulation of vagal afferents (Amano et al., 2010; Dunsmoor et al., 2015; Ehrlich et al., 2009; Peña et al., 2013, 2014). In sham-stimulated subjects, consequently, cardiac deceleration declined throughout extinction, although it did not fully extinguish, indicating a well consolidated fear memory which can possibly be attributed to partial CS reinforcement during initial fear conditioning (Hilton, 1969; Lonsdorf et al., 2017) and following a 24 hr memory consolidation period until extinction training began (Norrholm et al., 2008). TVNS, on the other hand, resulted in a faster attenuation of cardiac deceleration early during extinction, reflecting both its memory enhancing and anxiolytic effects (Noble et al., 2019) and also abolished the significant correlation between cardiac deceleration and startle potentiation. Importantly, the attenuation of cardiac deceleration was not accompanied by elevated levels of sympathetic nervous activity, as indicated by the SCL changes during extinction. These data suggest that the extinction of cardiac deceleration probably results from inhibition of parasympathetic heart rate control, driven by inhibition of the CeA-PAG neural freezing circuit—a process which may further be facilitated by tVNS. Previous research proposed that the defensive response pattern during attentive immobility (freezing) is primarily parasympathetically dominated (Roelofs, 2017) and, thus, our data again foster the view that cardiac deceleration functions as an autonomic expression of such defensive responding.

Importantly, cardiac responses to threat-signaling stimuli might therefore critically distinguish between different modes of fear-related defensive responding. As previous research indicated and is supported by our study, distal and inevitable threats elicit a defensive response pattern of parasympathetically dominated attentive immobility defined by fear bradycardia, during which orienting to and monitoring the source of danger is the best strategy to ensure survival, as it allows optimal preparation for defensive action in case threatening confrontation becomes increasingly imminent (see Roelofs, 2017). However, if the threat-signaling cue has become sufficiently imminent and defensive action is required to ensure survival, the defensive response pattern switches and the organism is under sympathetic control, defined by cardiac acceleration to support flight or fight (Cannon, 1929; Eilam, 2005; Lang et al., 1997, 2000; Lang & Davis, 2006; Roelofs, 2017). Moreover, using cues that are more complex or motivationally significant (i.e., facial expressions or pictures of snakes and spiders) might also activate cardiac acceleration (Hamm et al., 1993; Hodes et al., 1985) whereas

simple geometric figures as used in the current experiment evoke a constant cardiac deceleration in anticipation of the shock. More importantly, although both attentive immobility (freezing) and active withdrawal are defensive states accompanied by the feeling of fear (see Hamm, 2020 and Mobbs et al., 2009), the switch from parasympathetic fear bradycardia to sympathetic cardiac acceleration might therefore mark a transition of defensive strategies. We, thus, want to encourage future research on human fear to include threat-related heart rate changes into analyses, which possibly yield valuable information about the actual defensive state of the organism.

ACKNOWLEDGEMENT

We thank Dr. Heino Mohrmann for technical support. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTION

Christoph Szeska: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Visualization; Writing-original draft; Writing-review & editing. **Jan Richter:** Conceptualization; Methodology; Writing-review & editing. **Julia Wendt:** Conceptualization; Methodology; Writing-review & editing. **Mathias Weymar:** Conceptualization; Methodology; Writing-review & editing. **Alfons O. Hamm:** Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing-review & editing.

ORCID

Christoph Szeska https://orcid.org/0000-0002-2819-8536
Alfons O. Hamm https://orcid.org/0000-0002-3352-6231

REFERENCES

Amano, T., Unal, C. T., & Paré, D. (2010). Synaptic correlates of fear extinction in the amygdala. *Nature Neuroscience*, *13*(4), 489–494. https://doi.org/10.1038/nn.2499

Applegate, C. D., Kapp, B. S., Underwood, M. D., & McNall, C. L. (1983). Autonomic and somatomotor effects of amygdala central N. stimulation in awake rabbits. *Physiology and Behavior*, *31*(3), 353–360. https://doi.org/10.1016/0031-9384(83)90201-9

Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, *37*(1), 13–20. https://doi.org/10.1017/S0048577200980648

Blanchard, R. J., & Blanchard, D. C. (1969). Crouching as an index of fear. *Journal of Comparative and Physiological Psychology*, *67*(3), 370–375. https://doi.org/10.1037/h0026779

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*(1), 1–15. https://doi.org/10.1111/j.1469-8986.2005.00271.x

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., Zlatar, Z. Z., & Lang, P. J. (2018). Startle reflex modulation during threat of shock and "threat" of reward. *Psychophysiology*, *55*(2), 1–10. https://doi.org/10.1111/psyp.12989

Burger, A. M., Van Diest, I., van der Does, W., Hysaj, M., Thayer, J. F., Brosschot, J. F., & Verkuil, B. (2018). Transcutaneous vagus nerve stimulation and extinction of prepared fear: A conceptual non-replication. *Scientific Reports*, *8*(1), 11471. https://doi.org/10.1038/s41598-018-29561-w

Burger, A. M., Verkuil, B., Fenlon, H., Thijs, L., Cools, L., Miller, H. C., Vervliet, B., & Van Diest, I. (2017). Mixed evidence for the potential of non-invasive transcutaneous vagal nerve stimulation to improve the extinction and retention of fear. *Behaviour Research and Therapy*, *97*, 64–74. https://doi.org/10.1016/j.brat.2017.07.005

Burger, A. M., Verkuil, B., Van Diest, I., Van der Does, W., Thayer, J. F., & Brosschot, J. F. (2016). The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans. *Neurobiology of Learning and Memory*, *132*, 49–56. https://doi.org/10.1016/j.nlm.2016.05.007

Campbell, B. A., Wood, G., & McBride, T. (1997). Origins of orienting and defensive responses: An evolutionary perspective. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 41–67). Erlbaum.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear and rage: An account of recent research into the function of emotional excitement* (2nd ed.). D. Appleton and Co.

Chen, K. H., Aksan, N., Anderson, S. W., Grafft, A., & Chapleau, M. W. (2014). Habituation of parasympathetic-mediated heart rate responses to recurring acoustic startle. *Frontiers in Psychology*, *5*(November), 1–10. https://doi.org/10.3389/fpsyg.2014.01288

Choi, J. S., & Brown, T. H. (2003). Central amygdala lesions block ultrasonic vocalization and freezing as conditional but not unconditional responses. *Journal of Neuroscience*, *23*(25), 8713–8721. https://doi.org/10.1523/jneurosci.23-25-08713.2003

Cook, E. W., Davis, T. L., Hawk, L. W., Spence, E. L., & Gautier, C. H. (1992). Fearfulness and startle potentiation during aversive visual stimuli. *Psychophysiology*, *29*(6), 633–645. https://doi.org/10.1111/j.1469-8986.1992.tb02038.x

Davis, M. (2006). Neural systems involved in fear and anxiety measured with fear-potentiated startle. *American Psychologist*, *61*(8), 741–756. https://doi.org/10.1037/0003-066X.61.8.741

Dimberg, U. (1987). Facial reactions, autonomic activity and experienced emotion: A three component model of emotional conditioning. *Biological Psychology*, *24*(2), 105–122. https://doi.org/10.1016/0301-0511(87)90018-4

Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking extinction. *Neuron*, *88*(1), 47–63. https://doi.org/10.1016/j.neuron.2015.09.028

Duricki, D. A., Soleman, S., & Moon, L. D. F. (2016). Analysis of longitudinal data from animals with missing values using SPSS. *Nature Protocols*, *11*(6), 1112–1129. https://doi.org/10.1038/nprot.2016.048

Ehrlich, I., Humeau, Y., Grenier, F., Ciocchi, S., Herry, C., & Lüthi, A. (2009). Amygdala inhibitory circuits and the control of fear memory. *Neuron*, *62*(6), 757–771. https://doi.org/10.1016/j.neuron.2009.05.026

Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statistik und Forschungsmethoden* (2nd ed.). Beltz. https://doi.org/10.17877/DE290R-12739

Eilam, D. (2005). Die hard: A blend of freezing and fleeing as a dynamic defense—Implications for the control of defensive behavior. *Neuroscience and Biobehavioral Reviews*, *29*(8), 1181–1191. https://doi.org/10.1016/j.neubiorev.2005.03.027

Fanselow, M. S. (1984). What is conditioned fear? *Trends in Neurosciences*, *7*(12), 460–462. https://doi.org/10.1016/S0166-2236(84)80253-2

Fanselow, M. S. (1994). Neural orginization of the defensive behavior system responsible for fear. *Psychonomic Bulletin and Review*, *1*(4), 429–438.

Fendt, M., & Fanselow, M. S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neuroscience and Biobehavioral Reviews*, *23*(5), 743–760. https://doi.org/10.1016/S0149-7634(99)00016-0

Frangos, E., Ellrich, J., & Komisaruk, B. R. (2015). Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: FMRI evidence in humans. *Brain Stimulation*, *8*(3), 624–636. https://doi.org/10.1016/j.brs.2014.11.018

Gatchel, R. J., & Lang, P. J. (1973). Accuracy of psychophysical judgments and physiological response amplitude. *Journal of Experimental Psychology*, *98*(1), 175–183. https://doi.org/10.1037/h0034312

Geer, J. H. (1964). Measurement of the conditioned cardiac response. *Journal of Comparative and Physiological Psychology*, *57*(3), 426–433. https://doi.org/10.1037/h0041736

Gewirtz, J. C., Falls, W. A., & Davis, M. (1997). Normal conditioned inhibition and extinction of freezing and fear- potentiated startle following electrolytic lesions of medial prefrontal cortex in rats. *Behavioral Neuroscience*, *111*(4), 712–726. https://doi.org/10.1037/0735-7044.111.4.712

Gladwin, T. E., Hashemi, M. M., van Ast, V., & Roelofs, K. (2016). Ready and waiting: Freezing as active action preparation under threat. *Neuroscience Letters*, *619*, 182–188. https://doi.org/10.1016/j.neulet.2016.03.027

Globisch, J., Hamm, A. O., Schneider, R., & Vaitl, D. (1993). A computer program for scoring reflex eyeblink and electrodermal responses written in PASCAL. *Psychophysiology*, *30*(1), 30.

Graham, F. K. (1978). Constraints on measuring heart rate and period sequentially through real and cardiac time. *Psychophysiology*, *15*(5), 492–495. https://doi.org/10.1111/j.1469-8986.1978.tb01422.x

Graham, F. K. (1979). Distinguishing among orienting, defense, and startle reflexes. In H. D. Kimmel, E. H. van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in Humans* (pp. 137–167). Erlbaum.

Graham, F. K. (1987). Sokolov registered, model evicted. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychiphysiology* (Vol. *2*, pp. 211–232). JAI Press.

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin*, *65*(5), 305–320. https://doi.org/10.1037/h0023258

Grillon, C., & Davis, M. (1997). Fear-potentiated startle conditioning in humans: Explicit and contextual cue conditioning following paired versus unpaired training. *Psychophysiology*, *34*(4), 451–458. https://doi.org/10.1111/j.1469-8986.1997.tb02389.x

Haaker, J., Maren, S., Andreatta, M., Merz, C. J., Richter, J., Richter, S. H., Meir Drexler, S., Lange, M. D., Jüngling, K., Nees, F., Seidenbecher, T., Fullana, M. A., Wotjak, C. T., & Lonsdorf, T. B. (2019). Making translation work: Harmonizing cross-species methodology in the behavioural neuroscience of Pavlovian fear conditioning. *Neuroscience & Biobehavioral Reviews*, *107*, 329–345. https://doi.org/10.1016/j.neubiorev.2019.09.020

Hamm, A. O. (2015). Fear-potentiated startle. In J. D. Wright (Ed.). *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. *8*). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.55023-5

Hamm, A. O. (2020). Fear, anxiety, and their disorders from the perspective of psychophysiology. *Psychophysiology*, *57*(2), 1–14. https://doi.org/10.1111/psyp.13474

Hamm, A. O., & Flor, H. (2015). Fear learning, fear memory, and psychopathology. *International Journal of Psychophysiology*, *98*(3), 497–498. https://doi.org/10.1016/j.ijpsycho.2015.11.007

Hamm, A. O., Greenwald, M. K., Bradley, M. M., & Lang, P. J. (1993). Emotional learning, hedonic change, and the startle probe. *Journal of Abnormal Psychology*, *102*(3), 453–465. https://doi.org/10.1037/0021-843X.102.3.453

Hamm, A. O., & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, *33*(6), 698–710. https://doi.org/10.1111/j.1469-8986.1996.tb02366.x

Hilton, A. (1969). Partial reinforcement of a conditioned emotional response in rats. *Journal of Comparative and Physiological Psychology*, *69*(2), 253–260. https://doi.org/10.1037/h0028235

Hodes, R. L., Cook, E. W., & Lang, P. J. (1985). Individual differences in autonomic response: Conditioned association or conditioned fear? *Psychophysiology*, *22*(5), 545–560. https://doi.org/10.1111/j.1469-8986.1985.tb01649.x

Hollandt, M., Wroblewski, A., Yang, Y., Ridderbusch, I. C., Kircher, T., Hamm, A. O., Straube, B., & Richter, J. (2020). Facilitating translational science in anxiety disorders by adjusting extinction training in the laboratory to exposure-based therapy procedures. *Translational Psychiatry*, *10*(1), 1–10. https://doi.org/10.1038/s41398-020-0786-x

Kalin, N. H., & Shelton, S. E. (1989). Defensive behaviors in infant rhesus monkeys: Environmental cues and neurochemical regulation. *Science*, *243*(4899), 1718–1721. https://doi.org/10.1126/science.2564702

Klumpers, F., Raemaekers, M. A. H. L., Ruigrok, A. N. V., Hermans, E. J., Kenemans, J. L., & Baas, J. M. P. (2010). Prefrontal mechanisms of fear reduction after threat offset. *Biological Psychiatry*, *68*(11), 1031–1038. https://doi.org/10.1016/j.biopsych.2010.09.006

Kolassa, I. T., Musial, F., Mohr, A., Trippe, R. H., & Miltner, W. H. R. (2005). Electrophysiological correlates of threat processing in spider phobics. *Psychophysiology*, *42*(5), 520–530. https://doi.org/10.1111/j.1469-8986.2005.00315.x

Krause, E., Benke, C., Koenig, J., Thayer, J. F., Hamm, A. O., & Pané-Farré, C. A. (2018). Dynamics of defensive response mobilization to approaching external versus interoceptive threat. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(6), 525–538. https://doi.org/10.1016/j.bpsc.2017.12.002

Lacey, J. I., & Lacey, B. C. (1970). Some autonomic-central nervous system interrelationships. In P. Black (Ed.), *Physiological correlates of emotion* (1st ed., pp. 205–227). Academic Press.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(November), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lang, P. J. (1995). The emotion probe. *American Psychologist Association*, *50*(5), 372–385. https://doi.org/10.1037/0003-066X.50.5.372

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, *84*(3), 437–450. https://doi.org/10.1016/j.biopsycho.2009.10.007

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97–135). Lawrence Erlbaum Associates Inc.

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29. https://doi.org/10.1016/S0079-6123(06)56001-7

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*(3), 137–159. https://doi.org/10.1016/S0165-0327(00)00343-8

Leaton, R. N., & Borszcz, G. S. (1985). Potentiated startle. Its relation to freezing and shock intensity in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(3), 421–428. https://doi.org/10.1037/0097-7403.11.3.421

LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, *46*(1), 209–235. https://doi.org/10.1146/annurev.psych.46.1.209

LeDoux, J. E., Iwata, J., Cicchetti, P., & Reis, D. J. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *8*(7), 2517–2529. https://doi.org/10.1523/JNEUROSCI.08-07-02517.1988

Lipp, O. V., Sheridan, J., & David, A. T. S. (1994). Human blink startle during aversive and nonaversive Pavlovian conditioning. *Journal of Experimental Psychology*, *20*(4), 380–389.

Lipp, O. V., & Vaitl, D. (1990). Reaction time task as unconditional stimulus—Comparing aversive and nonaversive unconditional stimuli. *The Pavlovian Journal of Biological Science*, *25*(2), 77–83. https://doi.org/10.1007/BF02964606

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lonsdorf, T. B., Weike, A. I., Nikamo, P., Schalling, M., Hamm, A. O., & Öhman, A. (2009). Genetic gating of human fear learning and extinction: Possible implications for gene-environment interaction in anxiety disorder. *Psychological Science*, *20*(2), 198–206. https://doi.org/10.1111/j.1467-9280.2009.02280.x

Löw, A., Weymar, M., & Hamm, A. O. (2015). When threat is near, get out of here: Dynamics of defensive behavior during freezing and active avoidance. *Psychological Science*, *26*(11), 1706–1716. https://doi.org/10.1177/0956797615597332

Marks, I. M. (1987). *Fears, phobias, and rituals. Panic, anxiety, and their disorders*. Oxford University Press.

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172–175. https://doi.org/10.1037/0033-2909.111.1.172

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, *63*(1), 129–151. https://doi.org/10.1146/annurev.psych.121208.131631

Mobbs, D., Headley, D. B., Ding, W., & Dayan, P. (2020). Space, time, and fear: Survival computations along defensive circuits. *Trends in Cognitive Sciences*, *24*(3), 228–241. https://doi.org/10.1016/j.tics.2019.12.016

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., Petrovic, P., Dolan, R. J., & Frith, C. D. (2009). From threat to fear: The neural organization of defensive fear systems in humans. *Journal of Neuroscience*, *29*(39), 12236–12243. https://doi.org/10.1523/JNEUROSCI.2378-09.2009

Moratti, S., & Keil, A. (2005). Cortical activation during Pavlovian fear conditioning depends on heart rate response patterns: An MEG study. *Cognitive Brain Research*, *25*(2), 459–471. https://doi.org/10.1016/j.cogbrainres.2005.07.006

Mueller, D., & Cahill, S. P. (2010). Noradrenergic modulation of extinction learning and exposure therapy. *Behavioural Brain Research*, *208*, 1–11. https://doi.org/10.1016/j.bbr.2009.11.025

Noble, L. J., Meruva, V. B., Hays, S. A., Rennaker, R. L., Kilgard, M. P., & McIntyre, C. K. (2019). Vagus nerve stimulation promotes generalization of conditioned fear extinction and reduces anxiety in rats. *Brain Stimulation*, *12*(1), 9–18. https://doi.org/10.1016/j.brs.2018.09.013

Norrholm, S. D., Vervliet, B., Jovanovic, T., Boshoven, W., Myers, K. M., Davis, M., Rothbaum, B., & Duncan, E. J. (2008). Timing of extinction relative to acquisition: A parametric analysis of fear extinction in humans. *Behavioral Neuroscience*, *122*(5), 1016–1030. https://doi.org/10.1037/a0012604

Peña, D. F., Childs, J. E., Willett, S., Vital, A., McIntyre, C. K., & Kroener, S. (2014). Vagus nerve stimulation enhances extinction of conditioned fear and modulates plasticity in the pathway from the ventromedial prefrontal cortex to the amygdala. *Frontiers in Behavioral Neuroscience*, *8*(September), 327. https://doi.org/10.3389/fnbeh.2014.00327

Peña, D. F., Engineer, N. D., & McIntyre, C. K. (2013). Rapid remission of conditioned fear expression with extinction training paired with vagus nerve stimulation. *Biological Psychiatry*, *73*(11), 1071–1077. https://doi.org/10.1016/j.biopsych.2012.10.021

Peuker, E. T., & Filler, T. J. (2002). The nerve supply of the human auricle. *Clinical Anatomy*, *15*(1), 35–37. https://doi.org/10.1002/ca.1089

Plappert, C. F., Pilz, P. K. D., & Schnitzler, H. U. (1993). Acoustic startle response and habituation in freezing and nonfreezing rats. *Behavioral Neuroscience*, *107*(6), 981–987. https://doi.org/10.1037/0735-7044.107.6.981

Putnam, L. E., Ross, L. E., & Graham, F. K. (1974). Cardiac orienting during "good" and "poor" differential eyelid conditioning. *Journal of Experimental Psychology*, *102*(4), 563–573. https://doi.org/10.1037/0097-7403.29.3.C2

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*(1), 71–80. https://doi.org/10.1037/h0024109

Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160206. https://doi.org/10.1098/rstb.2016.0206

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: Social threat induces bodily freeze in humans. *Psychological Science*, *21*(11), 1575–1581. https://doi.org/10.1177/0956797610384746

Schipper, P., Hiemstra, M., Bosch, K., Nieuwenhuis, D., Adinolfi, A., Glotzbach, S., Borghans, B., Lopresto, D., Fernández, G., Klumpers,

F., Hermans, E. J., Roelofs, K., Henckens, M. J. A. G., & Homberg, J. R. (2019). The association between serotonin transporter availability and the neural correlates of fear bradycardia. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(51), 25941–25947. https://doi.org/10.1073/pnas.1904843116

Szeska, C., Richter, J., Wendt, J., Weymar, M., & Hamm, A. O. (2020). Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training. *Scientific Reports*, *10*(1), 1–16. https://doi.org/10.1038/s41598-020-58412-w

Ventura-Bort, C., Genheimer, H., Wirkner, J., Wendt, J., Hamm, A. O., & Weymar, M. (2018). Effects of transcutaneous vagus nerve stimulation (tVNS) on the P300 and alpha-amylase level: A Pilot Study. *Frontiers in Human Neuroscience*, *12*, 1–12. https://doi.org/10.3389/fnhum.2018.00202

Walker, P., & Carrive, P. (2003). Role of ventrolateral periaqueductal gray neurons in the behavioral and cardiovascular responses to contextual conditioned fear and poststress recovery. *Neuroscience*, *116*(3), 897–912. https://doi.org/10.1016/S0306-4522(02)00744-3

Wendt, J., Löw, A., Weymar, M., Lotze, M., & Hamm, A. O. (2017). Active avoidance and attentive freezing in the face of approaching threat. *NeuroImage*, *158*, 196–204. https://doi.org/10.1016/j.neuroimage.2017.06.054

Wong, A. H. K., & Lovibond, P. F. (2017). Rule-based generalisation in single-cue and differential fear conditioning in humans. *Biological Psychology*, *129*, 111–120. https://doi.org/10.1016/j.biopsycho.2017.08.056

Wong, A. H. K., & Lovibond, P. F. (2018). Excessive generalisation of conditioned fear in trait anxious individuals under ambiguity. *Behaviour Research and Therapy*, *107*(May), 53–63. https://doi.org/10.1016/j.brat.2018.05.012

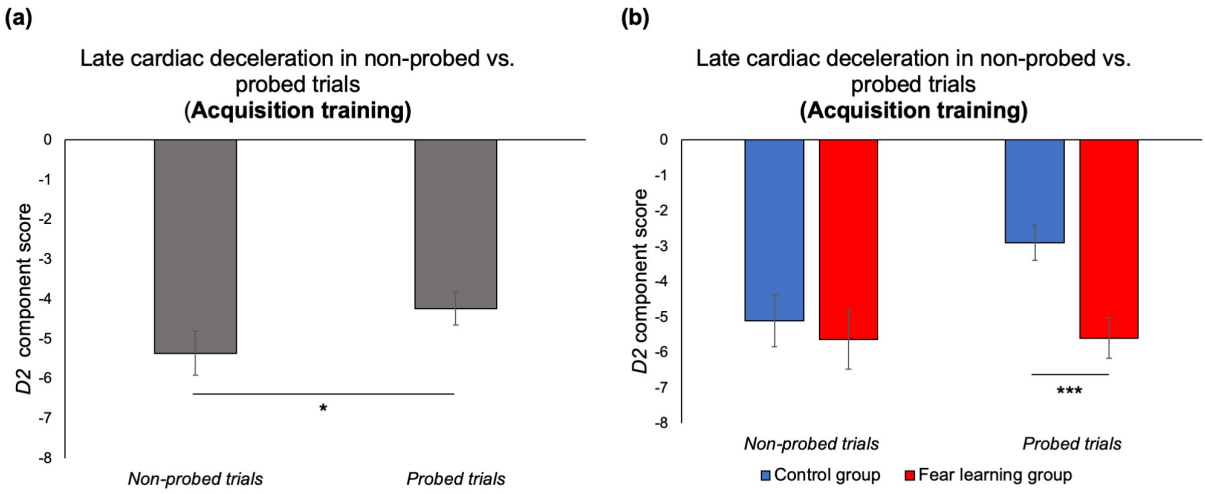
SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Figure S1. (a) Mean *D2* component score (slowest half-second between 5 and 7 s after CS onset) during the acquisition training, depicted for trials during which no startle probe was administered (non-probed trials) and for trials during which startle probes were delivered (probed trials). (b) Mean *D2* component score for non-probed and probed trials for both the control group (blue bars) and the fear learning group (red bars) during the acquisition training. Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations, with * for $p \leq .05$, ** for $p \leq .01$ and *** for $p < .001$

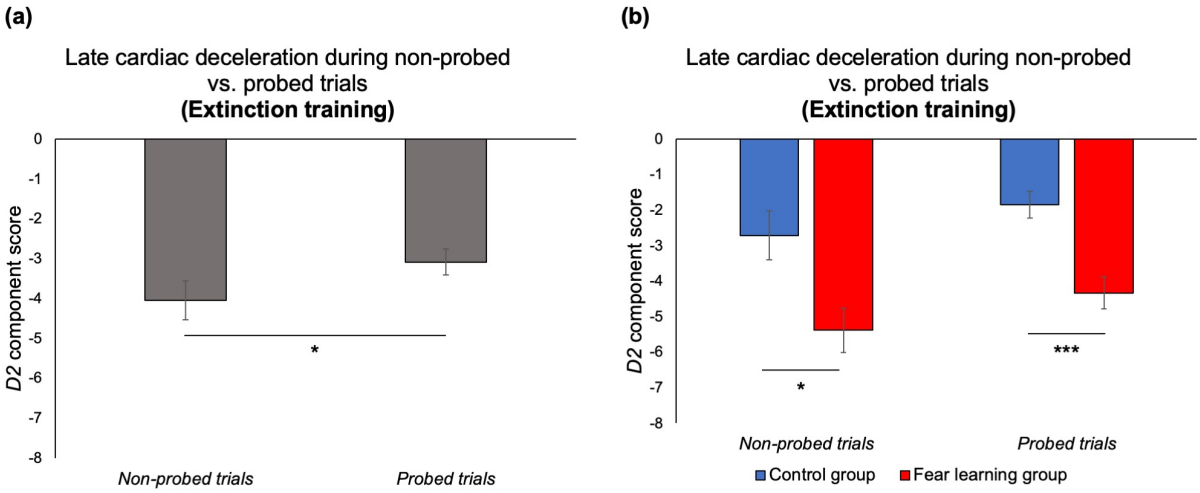
Figure S2. (a) Mean *D2* component score (slowest half-second between 5 and 7 s after CS onset) during the extinction training, depicted for trials during which no startle probe was administered (non-probed trials) and for trials during which startle probes were delivered (probed trials). (b) Mean *D2* component score for non-probed and probed trials for both the control group (blue bars) and the fear learning group (red bars) during the extinction training. Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations, with * for $p \leq .05$, ** for $p \leq .01$ and *** for $p < .001$

How to cite this article: Szeska C, Richter J, Wendt J, Weymar M, Hamm AO. Attentive immobility in the face of inevitable distal threat—Startle potentiation and fear bradycardia as an index of emotion and attention. *Psychophysiology*. 2021;00:e13812. https://doi.org/10.1111/psyp.13812



Supplementary Figure 1. (a) Mean *D2* component score (slowest half-second between 5 and 7 s after CS onset) during the acquisition training, depicted for trials during which no startle probe was administered (non-probed trials) and for trials during which startle probes were delivered (probed trials). (b) Mean *D2* component score for non-probed and probed trials for both the control group (blue bars) and the fear learning group (red bars) during the acquisition training.

Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations, with * for $p \leq .05$, ** for $p \leq .01$ and *** for $p < .001$.



Supplementary Figure 2. (a) Mean *D2* component score (slowest half-second between 5 and 7 s after CS onset) during the extinction training, depicted for trials during which no startle probe was administered (non-probed trials) and for trials during which startle probes were delivered (probed trials). (b) Mean *D2* component score for non-probed and probed trials for both the control group (blue bars) and the fear learning group (red bars) during the extinction training.

Error bars, when depicted, represent standard error of the mean. Asterisks indicate statistical significance of correlations, with * for $p \leq .05$, ** for $p \leq .01$ and *** for $p < .001$.

Manuscript 2

Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training

Christoph Szeska, Jan Richter, Julia Wendt, Mathias Weymar, & Alfons O. Hamm

Scientific Reports, 1529

Published in 2020

Author contributions:

Christoph Szeska: Conceptualization, Data curation, formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing original draft, Writing review and editing

Jan Richter: Conceptualization, Methodology, Writing review and editing

Julia Wendt: Conceptualization, Methodology, Writing review and editing

Mathias Weymar: Conceptualization, Methodology, Writing review and editing

Alfons O. Hamm: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Writing review and editing

www.nature.com/scientificreports

SCIENTIFIC REPORTS
nature research

OPEN

Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training

Christoph Szeska[1*], Jan Richter[1], Julia Wendt[2], Mathias Weymar[2] & Alfons O. Hamm[1]

Inhibiting fear-related thoughts and defensive behaviors when they are no longer appropriate to the situation is a prerequisite for flexible and adaptive responding to changing environments. Such inhibition of defensive systems is mediated by ventromedial prefrontal cortex (vmPFC), limbic basolateral amygdala (BLA), and brain stem locus-coeruleus noradrenergic system (LC-NAs). Non-invasive, transcutaneous vagus nerve stimulation (tVNS) has shown to activate this circuit. Using a multiple-day single-cue fear conditioning and extinction paradigm, we investigated long-term effects of tVNS on inhibition of low-level amygdala modulated fear potentiated startle and cognitive risk assessments. We found that administration of tVNS during extinction training facilitated inhibition of fear potentiated startle responses and cognitive risk assessments, resulting in facilitated formation, consolidation and long-term recall of extinction memory, and prevention of the return of fear. These findings might indicate new ways to increase the efficacy of exposure-based treatments of anxiety disorders.

Learning to inhibit defensive responses to cues which are no longer associated with aversive events is crucial for flexible and appropriate responding towards changing environments[1]. During such extinction learning the previously acquired association is not erased from memory, rather the organism has to learn to actively inhibit the previously acquired response – in this case – a fear response[2,3]. There has been extensive research to delineate the neural systems underlying this inhibitory learning process. These efforts have pointed out the crucial role of two interconnected brain regions inhibiting the activity of the central nucleus of the amygdala, the key structure orchestrating fear responses:[4–6] the basolateral complex of the amygdala (BLA) and the ventromedial prefrontal cortex (vmPFC)[4–11]. While activity of the BLA is involved in the acquisition and consolidation of extinction memory[5,6,12], increased activity of the vmPFC is found during recall of extinction memory, pinpointing its importance for the long-term inhibition of defensive responses[13,14].

Invasive peripheral stimulation of the vagus nerve leads to specific noradrenergic activation of both the BLA and vmPFC, by way of activating its afferent projections via the nucleus tractus solitarius to the locus coeruleus-noradrenergic system (LC-NAs)[15–17]. Consequently, invasive stimulation of vagal afferents resulted in enhanced inhibition of defensive freezing during fear extinction learning and promoted extinction memory consolidation and recall in rodents[18–20], presumably by increasing the activity in the BLA-vmPFC pathway[18,20]. Transcutaneous vagus nerve stimulation (tVNS) serves as a non-invasive equivalent, activating the same brain regions as its invasive counterpart by stimulation of the cymba conchae of the human auricle – a skin area exclusively innervated by the vagus nerve[21,22]. However, transfer from the promising animal findings to humans was only partly successful. While tVNS enhanced extinction of cognitive risk assessments in humans (US expectancy ratings)[23,24], no effect of tVNS was found for behavioral or physiological measures of fear[23–25]. One reason for these mixed findings might be, that previous human research used the same fixed and rather low tVNS

[1]University of Greifswald, Department of Physiological and Clinical Psychology/Psychotherapy, Franz-Mehring-Strasse 47, 17487, Greifswald, Germany. [2]University of Potsdam, Department of Biological Psychology and Affective Science, Karl-Liebknecht Strasse 24/25, 14476, Potsdam, Germany. *email: christoph.szeska@uni-greifswald.de

intensities for each participant[23–25], not taking individual differences in sensitivity into account that might have resulted in reduced stimulation of the vagus nerve in some subjects. More important, most of these human studies used differential-cue conditioning tasks, while aversive single-cue conditioning was used in animal research. Differential-cue conditioning, however, requires more complex discriminative learning during both acquisition and extinction probably involving other neural circuitry[26–28].

Thus, in the current study we investigated the impact of individually adjusted tVNS on extinction learning in humans using such a single-cue conditioning and extinction paradigm, closely adapted to previous animal research[18,19]. This protocol requires between-subject comparisons of conditioned responses between a fear learning group, in which the CS is repeatedly paired with the unconditioned stimulus (US) and a control group, in which the CS is presented explicitly unpaired with the US[27,29]. Consequently, participants of the fear learning group are expected to acquire conditioned fear responses to the CS, which should extinguish during extinction training. Participants of the control group are expected not to show any conditioned responses to the CS[30], thus being unaffected by an extinction training. Consequently, this single-cue conditioning design allows to distinguish conditioned fear, which should only develop in the fear learning group, from elevated general defensive alertness due to the mere presentation of the US, which may also be observed in the control group. Hence, the design further allows to disentangle potential fear reducing effects of tVNS due to facilitated extinction of conditioned fear, which should only be observed in the fear learning group. In contrast a reduction of general defensive alertness by tVNS, should be also observed in the control group.

We presumed tVNS to improve the extinction of cognitive and behavioral indices of fear only in the fear learning group, but not in the control group. Moreover, based on animal findings, we hypothesized that tVNS would also facilitate the recall of fear extinction memory in the fear learning group, but not in the control group[19]. Finally, we also wanted to investigate whether tVNS would also reduce the return of fear after reinstatement. Indeed, even a fully extinguished fear response can be reinstated after re-experiencing the threatening event[2,31,32], likely contributing to relapses after successful exposure therapy[33,34]. If tVNS would reduce reinstatement of fear, long-term efficacy of exposure-based treatments might be augmented using this neuroscience-based intervention.

Results

TVNS was applied in a 2 × 2 between-group design with fear learning (group: fear learning vs. control) and type of stimulation (stimulation: stimulation of the cymba conchae [tVNS] vs. sham stimulation of the earlobe) as two between subject factors (for a linear depiction of the experimental phases please see Fig. 1).

During acquisition training (**session 1**) subjects of the fear learning group ($n = 40$) received sixteen paired presentations of a single visual cue (conditioned stimulus, CS; geometric figure on a black background) with an aversive electrical shock (unconditioned stimulus, US) during 12 of the 16 trials (75% reinforcement rate). In the control group ($n = 40$) the CS was also presented sixteen times but explicitly unpaired with the US, which was presented 12 times during the inter-trial intervals (ITI; black screen).

Extinction training started 24 hours later (**session 2**), where the CS was now presented without the US in all groups. Half of the subjects of each group underwent extinction training receiving individually adjusted transcutaneous stimulation of the cymba conchae (**tVNS**; $n = 20$ of the fear learning and n = 20 of the control group), while the other half of both groups was given individually adjusted sham stimulation of the earlobe, not stimulating any vagal afferents[22] (see Fig. 1). Allocation to the stimulation condition (tVNS vs. sham stimulation) was randomized and its administration single-blind sham controlled (see also **Methods**).

Twenty-four hours after extinction training (**session 3**), subjects underwent a subsequent extinction test, during which the CS was presented 16 times without the US, followed by a reinstatement procedure (US was presented alone on three successive trials) and a reinstatement test session (16 CS presentations without US) to investigate the effects of tVNS on short-term extinction memory recall and the effect of tVNS on reinstatement of fear.

Long-term effects of tVNS were tested approximately 28 days after the third assessment (**session 4**), using the same procedure as during session 3.

We assessed shock expectancy ratings as an index of cognitive risk assessments and the potentiation of the startle-eyeblink reflex elicited by a white noise probe stimulus (95 dB, 50 ms) – a low level brainstem reflex reflecting fear in humans and other animals as it is indexing amygdala-dependent automatic freezing[35–38]. Correspondingly, potentiation of startle responses during CS trials compared to ITIs reflects the level of fear elicited by the CS.

Between-group single-cue fear conditioning results in stable conditioned fear responses in a fear learning group, but not in a control group (session 1; acquisition training). Acquisition training established a reliable fear response in all response systems in the fear learning group but not in the control group. This was indicated by increasing shock expectancy ratings from the start to the end of acquisition training and increased startle potentiation during CS presentations relative to the ITI at the end of the acquisition training in the fear learning, but not in the control group, which in contrast showed decreasing shock expectancy ratings and a lack of potentiation of the startle response during the CS (trials × group, $F_{1,76} = 71.13$, $P < 0.001$, $\eta^2_p = 0.48$; see Fig. 2a, b; potentiation × group, $F_{1,228} = 20.08$, $P < 0.001$, $\eta^2_p = 0.08$; see Fig. 2c, d). As expected, no differences were found between stimulation groups (tVNS vs. sham stimulation) in shock expectancy ratings (stimulation, stimulation × group, trials × stimulation, trials × stimulation × group, all $Fs < 1$, all $Ps > 0.33$) or fear potentiated startle responses (potentiation × stimulation, potentiation × trials × stimulation, potentiation × group × stimulation, potentiation × trials × group × stimulation, all $Fs < 1.1$, all $Ps > 0.31$).

Transcutaneous vagus nerve stimulation facilitates initial extinction learning (session 2; extinction training). Throughout the extinction training 24 hours later, shock expectancy ratings were still elevated

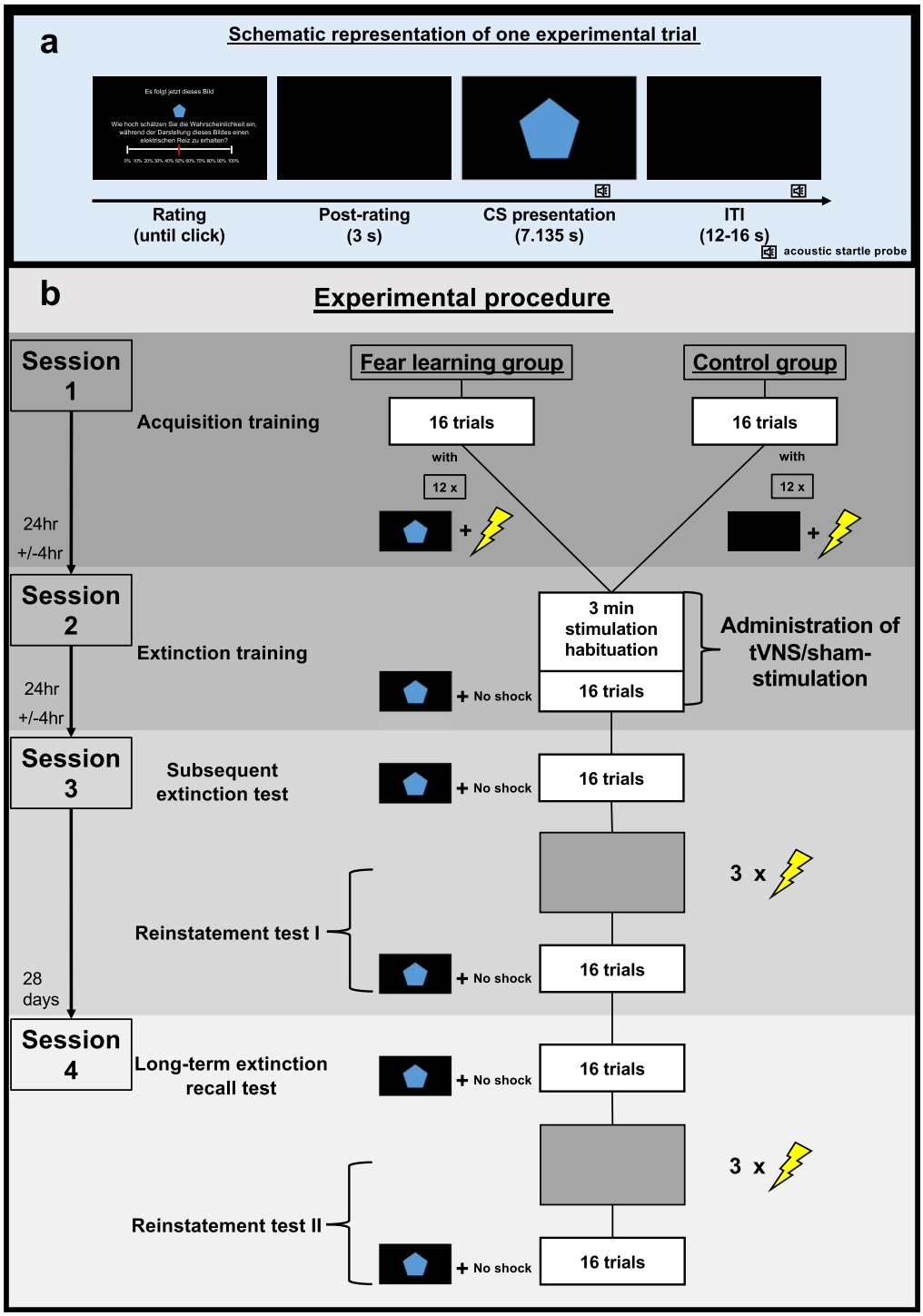


Figure 1. Schematic representation of an experimental trial and the experimental procedure of the single-cue fear conditioning and extinction paradigm. Upper Panel (**a**): Schematic representation of an experimental trial. Each trial started with a shock expectancy rating, where the CS was presented in smaller size and participants were instructed to rate the probability that this cue would be followed by the shock during the upcoming trial (English translation of the German instruction: "Next, this picture will follow. How likely do you think is it, to receive an electrical shock during the upcoming presentation of this picture?"). This procedure is very much comparable to the clinical procedure during exposure therapy, during which the patient is first asked to rate the probability that the central concern might become true (e.g., fainting) before the exposure exercise begins. Three seconds after completing the rating, the cue is presented in full size on the screen, making sure that the physiological fear response is not affected by any simultaneous cognitive evaluation task. Acoustic startle probes were presented both during the CS and ITIs. Lower Panel (**b**): Schematic representation of the experimental procedure. During each experimental phase sixteen trials were presented. During acquisition training participants of the fear learning group received twelve paired presentations of the conditioned stimulus and

an electrical shock, while in the control group CS and shocks were presented explicitly unpaired (shocks only during inter-trial intervals). During extinction training, the subsequent extinction test and long-term extinction recall test no electrical shocks were presented. Both reinstatement tests started with three non-signaled electrical shocks, followed by sixteen unreinforced CS presentations. Acoustic startle probes were presented during the CS and ITIs throughout each experimental phase.

in the fear learning relative to the control group (group, $F_{1,76} = 16.90$, $P < 0.001$, $\eta^2_p = 0.18$; Fig. 3a, b). Moreover, blink magnitudes elicited during the CS were also significantly potentiated relative to the ITI in the fear learning, but not in the control group (potentiation × group, $F_{1,1058.90} = 31.77$, $P < 0.001$, $\eta^2_p = 0.03$; see Fig. 3c, d). These results indicate a robustly consolidated fear memory.

Group differences diminished during course of extinction for expectancy ratings (group × trials, $F_{15,76} = 7.86$, $P < 0.001$, $\eta^2_p = 0.09$) but remained stable for startle potentiation (potentiation × group × trials, $F < 1$), indicating that cognitive measures of fear adapt faster to changing stimulus outcome relationships compared to subcortically low-level indicators of defensive freezing.

As expected, tVNS modulated extinction learning in the fear learning but not in the control group. Shock expectancy ratings were significantly reduced in the fear learning group receiving tVNS compared to the sham condition (stimulation, $F_{1,38} = 7.13$, $P = 0.011$ $\eta^2_p = 0.16$; see Fig. 3a), while tVNS had no overall effect in the control group (stimulation, $F < 1$; see Fig. 3b). This differential effect of vagus nerve stimulation tended to be stronger during the first half of extinction (group × stimulation, $F_{1,76} = 3.73$, $P = 0.057$, $\eta^2_p = 0.05$; see Fig. 3a), being strongest at the beginning of extinction training (significant interaction group × stimulation during trials 1–2; $F_{1,76} = 5.99$, $P = 0.017$, $\eta^2_p = 0.07$; Fig. 3a), indicating rapid anxiolytic effects of tVNS right from the start of extinction training.

At the end of extinction training, differences in shock expectancy between the fear learning and control group were no longer evident in tVNS participants, while we found a trend for higher shock expectancy ratings in sham subjects of the fear learning group compared to sham-controls (trials 15–16; group, $F_{1,38} = 3.08$, $P = 0.087$, $\eta^2_p = 0.08$).

Besides robust fear potentiated startle throughout extinction training (Fig. 3c, d), participants of the fear learning group receiving tVNS showed significantly reduced potentiation of the startle reflex *at the end* of extinction training in comparison to subjects receiving sham stimulation (see Fig. 4a, b). While blink magnitudes were no longer potentiated during CS trials relative to ITI for the last two extinction trials in the fear learning group receiving tVNS (trials 15–16; potentiation, $F < 1$; see Fig. 4a), potentiation persisted for fear learning group receiving sham stimulation (trials 15–16; potentiation, $F_{1,14.38} = 11.88$, $P = 0.004$, $\eta^2_p = 0.45$, see Fig. 4a). As for cognitive risk assessments, differences in fear potentiated startle responses between the fear learning and control group were no longer evident in tVNS participants at the end of extinction training (last two probed trials; potentiation × group, $F_{1,35.10} = 1.51$, $P = 0.227$; Fig. 4a, b), while sham subjects of the fear learning group still showed higher fear potentiated startle responses compared to sham-controls (last two probed trials; potentiation × group, $F_{1,33.79} = 5.91$, $P = 0.020$, $\eta^2_p = 0.15$; Fig. 4a, b). Thus, while declarative fear might have been reduced due to the anxiolytic effects of tVNS right from the start of extinction training, tVNS boosted the extinction of behavioral fear responses only late during extinction training.

TVNS-paired extinction facilitates the subsequent extinction of conditioned fear responses after 24 hours (session 3; subsequent extinction test). Initial short-term recall of extinction memory was unaffected by tVNS, both for cognitive risk assessments (trials 1–2; stimulation, stimulation × group, all $Fs < 1$, all $Ps > 0.49$; see Fig. 5a, b) and potentiation of the startle responses (probes 1–2; potentiation × stimulation, potentiation × group × stimulation, all $Fs < 1.20$, all $Ps > 0.27$; see Fig. 5c, d). However, *during* the subsequent extinction test 24 hours later, we found that previous tVNS-paired extinction further facilitated extinction learning in the fear learning group, indicated by a stronger decrease in shock expectancy ratings in the tVNS relative to the sham condition (trials × stimulation, $F_{15,570} = 2.34$, $P = 0.003$, $\eta^2_p = 0.06$; Fig. 5a). No such effect was observed in the control group (trials × stimulation, $F < 1$.; Fig. 5b). This effect was even more pronounced for potentiation of the startle response magnitudes, showing a significant group by stimulation by trials interaction (potentiation × trials × group × stimulation, $F_{11,1654.52} = 2.02$, $P = 0.023$, $\eta^2_p = 0.01$; Fig. 5c, d). Thus, subsequent extinction of fear potentiated startle responses was facilitated by previous tVNS-paired extinction in the fear learning group relative to sham and this effect was not present in the control group (see also Supplementary Fig. 2). In sum, these findings indicate that previous tVNS-paired extinction facilitated subsequent extinction in both cognitive risk assessments and defensive brain reflex measures.

TVNS-paired extinction results in the prevention of behavioral reinstatement of fear (session 3; reinstatement test I). As depicted in Fig. 6, re-experience of the aversive event (three repetitive US presentations without CS) increased anxious apprehension and defensive response mobilization, as indicated by a significant increase in shock expectancy (Fig. 6a, b) and blink magnitudes (Fig. 6c, d) from the last two subsequent extinction test trials to the first two reinstatement test I trials (reinstatement, $Fs = 48.72; 79.53$; all $Ps < 0.001$, all $\eta^2_p > 0.26$ for shock expectancy ratings and blink magnitudes, respectively). Previous tVNS-paired extinction training did not affect the reinstatement effect for shock expectancy ratings (trials × stimulation, trials × stimulation × group, all $Fs < 1.21$, all $Ps > 0.27$). However, tVNS-paired extinction training resulted in significantly reduced startle response sensitization relative to the sham condition in the fear learning group but not in the control group (reinstatement × group × stimulation, $F_{1,224.56} = 4.29$, $P = 0.039$, $\eta^2_p = 0.02$; Fig. 6c, d). Moreover, while potentiation of the blink magnitudes during CS presentations relative to ITIs was reinstated in the sham

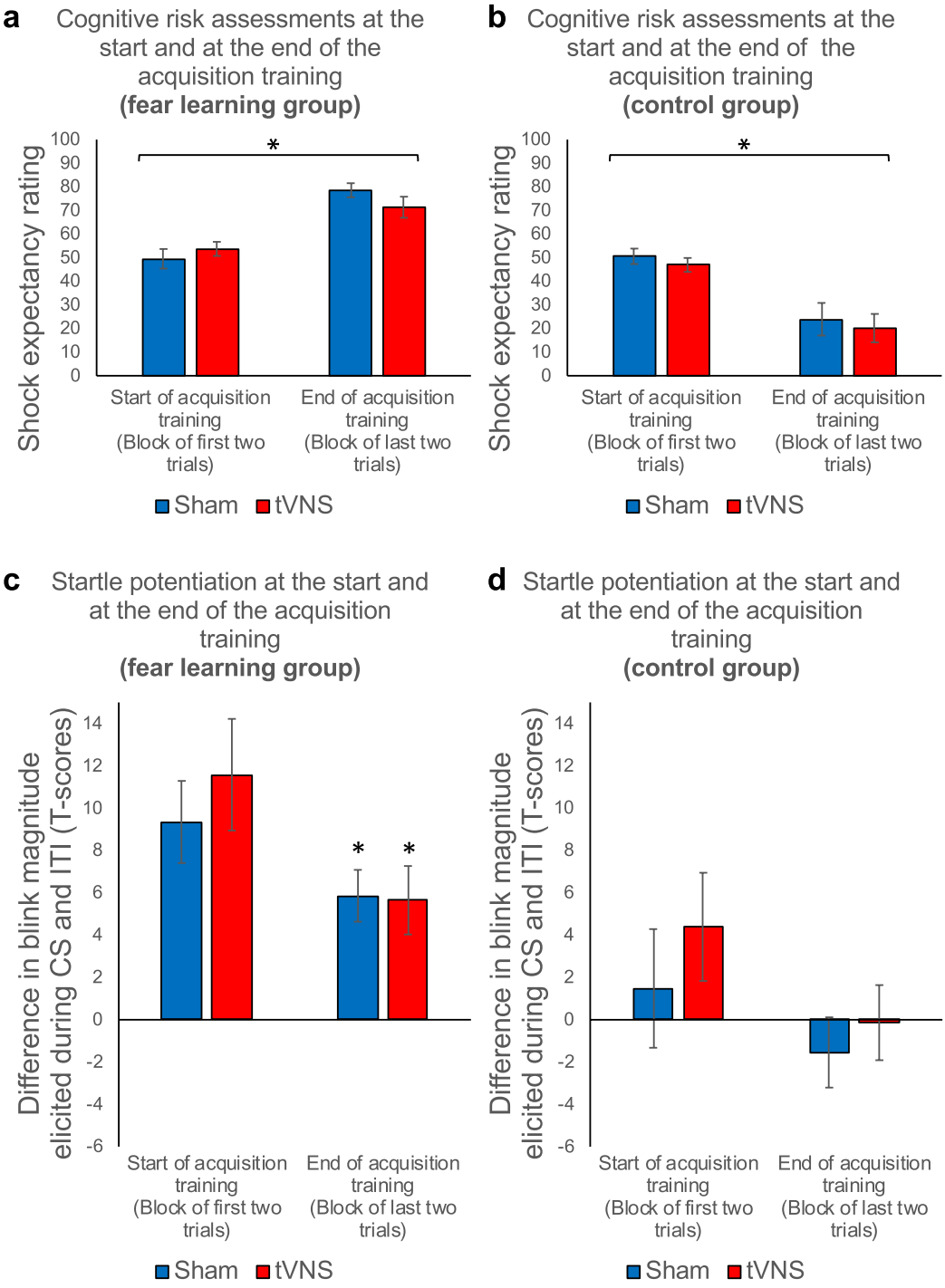


Figure 2. Paired presentations of CS and US result in stable conditioned fear responses (fear learning group) in comparison to explicitly unpaired presentations of CS and US (control group). Upper Panel (**a** and **b**): Mean shock expectancy ratings at the start of acquisition training (first two trials averaged) and at the end of the acquisition training (last two trials averaged) for the sham stimulation (*blue*) and tVNS condition (*red*) in the fear learning group (**a**: left) and in the control group (**b**: right). Error bars represent the standard error of the means (SEM). Participants of the fear learning group, receiving paired presentations of the CS and the electrical shock, showed an increase in rated shock expectancy from the start to the end of acquisition training. Participants of the control group, receiving explicitly unpaired presentations of the CS and the electrical shock, showed a decrease in rated shock expectancy from the start to the end of acquisition training. Lower Panel (**c** and **d**): Mean startle potentiation scores (standardized (T-scores) blink magnitudes elicited during the CS minus standardized (T-scores) blink magnitudes elicited during the inter-trial intervals) averaged across the magnitudes at the start of acquisition (first two probes in either condition (CS and ITI)) and at the end of acquisition training (last two probes in either condition (CS and ITI)). Potentiation scores are presented for sham stimulation (*blue*) and the tVNS condition (*red*) in the fear learning group (**c**: left) and in the control

group (**d**: right). Again, error bars represent SEM. Participants of the fear learning group, receiving paired presentations of the CS and the electrical shock, showed stable fear potentiated startle responses at the end of the acquisition training. Participants of the control group, receiving explicitly unpaired presentations of the CS and the electrical shock, did not show fear potentiated startle responses at the end of the acquisition training.

condition (potentiation, $F_{1,18} = 5.75$, $P = 0.028$, $\eta^2_p = 0.24$; Fig. 6c), extinction of fear potentiated startle survived reinstatement in the tVNS group (potentiation, $F_{1,18.58} = 1.11$, $P = 0.31$; Fig. 4d).

After reinstatement, during the following reinstatement test I phase, shock expectancy ratings were still elevated in the fear learning group relative to the control group (group, $F_{1,76} = 5.22$, $P = 0.025$, $\eta^2_p = 0.06$; Supplementary Fig. 1). Moreover, blink magnitudes were significantly potentiated in the fear learning but not in the control group (potentiation × group, $F_{1,1644.79} = 9.59$, $P = 0.002$, $\eta^2_p = 0.006$; Supplementary Fig. 2), indicating that the fear memory trace was still evident.

TVNS-paired extinction facilitates long-term recall of extinction memory (session 4; long-term extinction recall test and reinstatement test II). Even after 28 days, the fear memory trace was still evident, indicated by elevated shock expectancy ratings and startle potentiation in the fear learning group relative to the control group (group and potentiation × group, $Fs = 11.02$; 11.05, all $Ps < 0.05$, all $\eta^2_p > 0.007$ for shock expectancy ratings and blink magnitudes, respectively; see Fig. 7a–d). Initial long-term recall of extinction memory was unaffected by tVNS, both for cognitive risk assessments (trials 1–2; stimulation, stimulation × group, all $Fs < 1.79$, all $Ps > 0.18$; see Fig. 7a) and potentiation of the startle responses (probes 1–2; potentiation × stimulation, potentiation × group × stimulation, all $Fs < 2.45$, all $Ps > 0.12$; see Fig. 7c). Importantly, participants who received tVNS, but not sham stimulation on day 2, showed better recall of fear extinction memory *during* the long-term extinction recall test, 28 days after session 3. This was indicated by stronger reduction of shock expectancy in the fear learning group receiving tVNS compared to the sham condition (trials × stimulation, $F_{15,570} = 1.97$, $P = 0.016$, $\eta^2_p = 0.05$; Fig. 7a). Moreover, extinction of the potentiation of the startle response magnitudes in the fear learning group was again facilitated by previous (4 weeks ago) tVNS relative to the sham condition (potentiation × group × stimulation, $F_{1,1668.68} = 5.09$, $P = 0.024$, $\eta^2_p = 0.003$; Fig. 7c, d).

After 28 days, re-experience of the aversive event also led to a significant defensive sensitization, indicated by a significant increase in shock expectancy and blink magnitudes from the last two long-term extinction recall test trials to the first two reinstatement test II trials in the fear learning and the control group (reinstatement, $Fs = 25.33$; 13.27; all $Ps < 0.001$, all $\eta^2_p > 0.05$ for shock expectancy ratings and blink magnitudes, respectively; Supplementary Figs. 1 and 2). Importantly, differences between the fear learning group and the control group were no longer significant in shock expectancy ratings (group, $F_{1,76} = 3.64$, $P = 0.060$, $\eta^2_p = 0.01$; Supplementary Fig. 1) and startle potentiation (potentiation × group, $F < 1$; Supplementary Fig. 2), indicating fully extinguished fear in the fear learning group.

Discussion

Using a multiple-day single-cue conditioning and extinction paradigm, adapted to the experimental protocols used in animal studies[18,19], we found that an extinction training paired with transcutaneous vagus nerve stimulation resulted in both rapid anxiolytic effects as well as a facilitation of fear extinction learning in comparison to a sham stimulation. During the extinction training, anxiolytic effects were observed for cognitive indicators of anxious apprehension (shock expectancy ratings), while tVNS facilitated the extinction of subcortically mediated indicators of defensive freezing (fear potentiated startle response). Moreover, tVNS boosted the subsequent extinction, as well as long-term recall of extinction memory and also prevented reinstatement of fear. Thus, our findings provide evidence, that tVNS not only facilitates the inhibition of fear itself, but is also capable of facilitating further subsequent extinction of cognitive and behavioral indices of conditioned fear and long-term recall of extinction memory in humans.

It has been repeatedly proposed that transcutaneous vagus nerve stimulation results in both anxiolytic effects and enhanced cognitive flexibility in humans and rodents[39–42]. In fact, our results support the hypothesis, that tVNS promotes both an overall reduction of cognitive risk assessments and a flexible cognitive adaptation to changing aspects of the environment. Accordingly, active tVNS significantly reduced cognitive risk assessments relative to sham stimulation, when a cue was no longer predicting an aversive event. As this effect was strongest early during extinction learning and was even evident right from the start of extinction training, we may assume rapid anxiolytic effects, as proposed by previous rodent research[41]. However, even more remarkable, although the initial spontaneous recovery of fear after a passage of 24 hours has not been prevented by a tVNS-paired extinction training, it in fact promoted flexible cognitive adaptation and further facilitated the extinction of cognitive risk assessments 24 hours later during a subsequent extinction test.

Extending these findings, that are also in line with previous human data from Burger and coworkers[23,24], we also observed long-term effects of tVNS on recall of extinction memory. Again, although a spontaneous recovery of fear after 28 days was not prevented by tVNS-paired extinction, these new finding suggest that tVNS may also facilitate the reduction of long-term risk assessments by consolidating the extinction memory that a cue is no longer associated with an aversive event. Since elevated expectations of aversive events in a certain context seem to be important for motivating avoidance behavior[43,44] transcutaneous vagus nerve stimulation might be an important adjunct to cognitive-behavioral exposure-based therapy, aimed to reduce persistent avoidance behavior in patients with anxiety disorders.

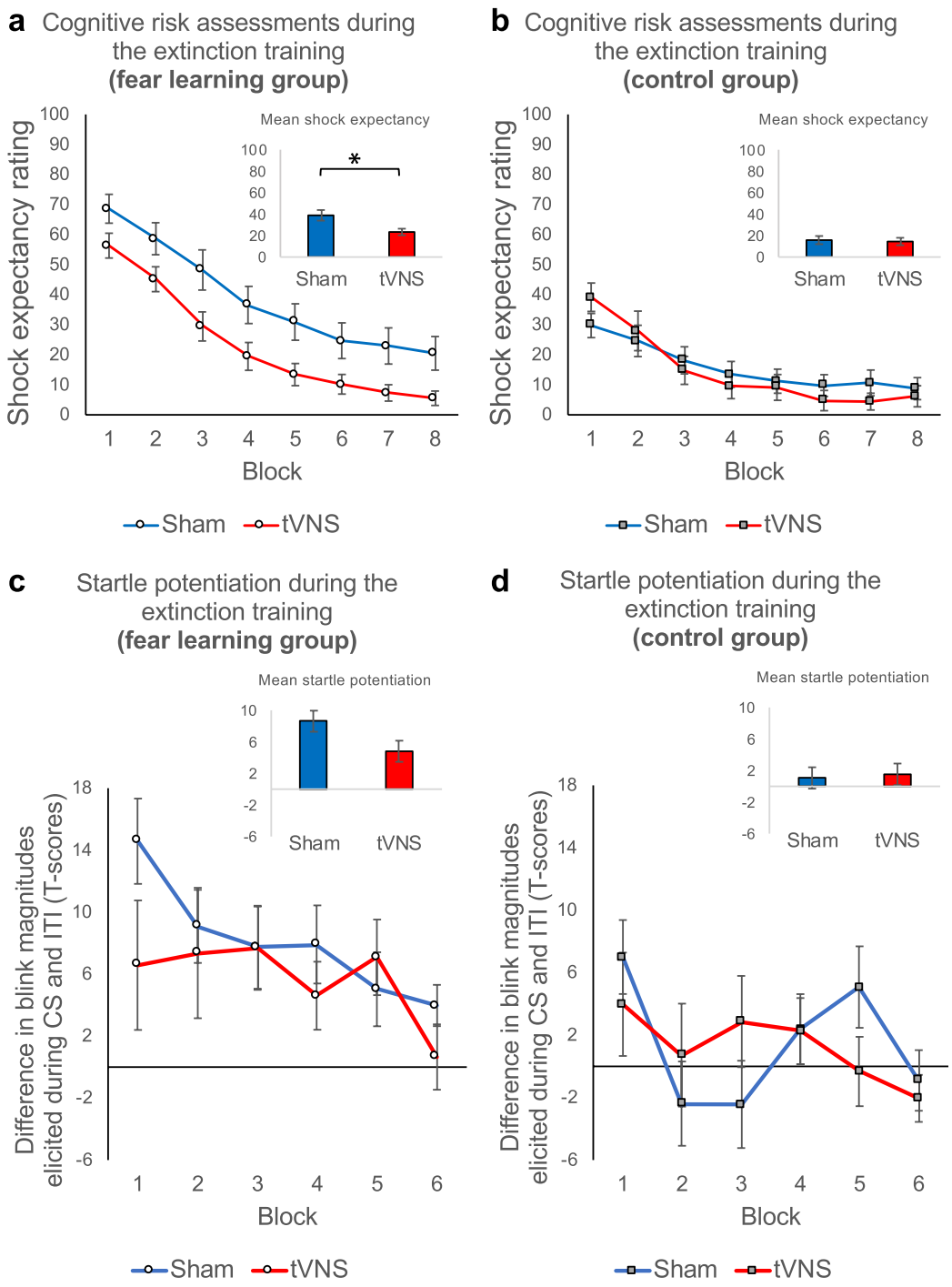


Figure 3. Transcutaneous vagus nerve stimulation facilitates fear extinction learning. Upper Panel (**a** and **b**): Mean shock expectancy ratings averaged across blocks of two extinction trials for the sham (*blue*) and tVNS condition (*red*) in the fear learning group (**a**: left) and in the control group (**b**: right). Error bars represent the standard error of the means (SEM). Participants of the fear learning group receiving tVNS showed facilitated reduction of fear in comparison to participants receiving sham stimulation (see **a**), indicated by a stronger decrease in shock expectancy ratings. Extinction processes in control group participants were not modulated by tVNS (see **b**). Lower Panel (**c** and **d**): Mean startle potentiation (standardized (T-scores) blink magnitudes elicited during the CS minus standardized (T-scores) blink magnitudes elicited during the inter-trial intervals) averaged across two probe stimuli presented during extinction training for the sham stimulation (*blue*) and tVNS condition (*red*) in the fear learning group (**c**: left) and the control group (**b**: right). Again, error bars represent SEM. We did not find effects of tVNS on the extinction of fear responses throughout the *whole* extinction training in neither group. However, we found that tVNS boosted extinction of fear responses in the fear learning group, but not the control group, late at the end of the extinction training (see Fig. 4).

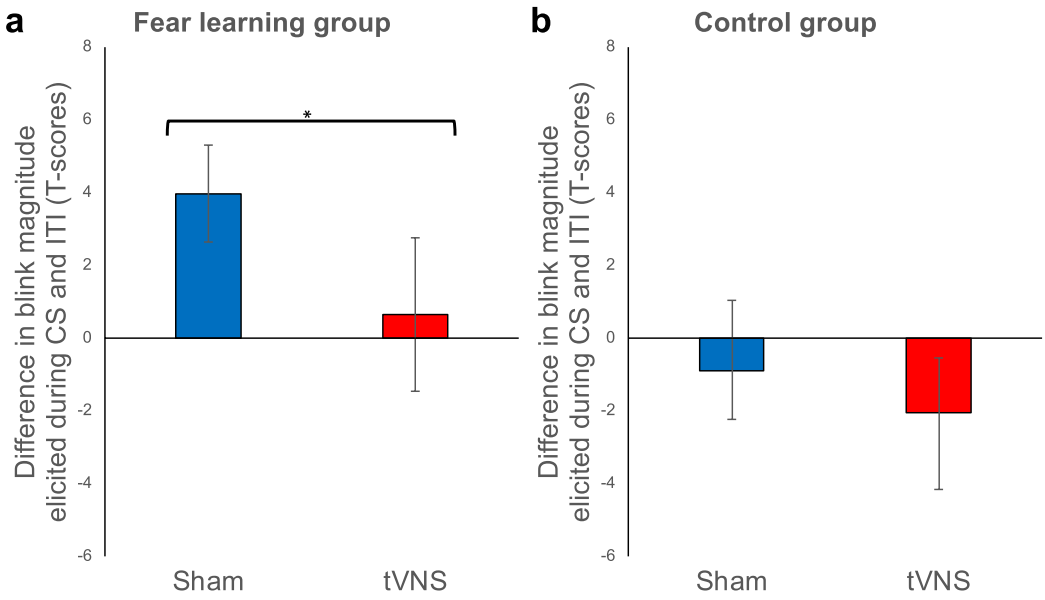


Figure 4. Transcutaneous vagus nerve stimulation boosts extinction of fear potentiated startle responses *late* during extinction training. Mean startle potentiation scores (standardized (T-scores) blink magnitudes elicited during the CS minus standardized (T-scores) blink magnitudes elicited during the inter-trial intervals) averaged across the magnitudes elicited by the last two probes in either condition (CS and ITI) during extinction training. Potentiation scores are presented for the sham stimulation (*blue*) and the tVNS condition (*red*) in the fear learning group (**a**: left) and the control group (**b**: right). Error bars represent SEM. At the end of the extinction training blink magnitudes were no longer potentiated during CS trials relative to ITI in the tVNS condition of the fear learning group (**a**: red bar). However, potentiation persisted for fear learning group receiving sham stimulation (**a**: blue bar). TVNS had no impact on fear potentiated startle responses in the control group (**b**: right) at the end of the extinction training.

In contrast to previous human research[23–25], we found that tVNS also promoted extinction of fear potentiated startle responses, thus in line with animal findings showing that stimulation of vagal afferents promoted extinction of defensive freezing[18–20]. One reason for this first successful translation to human research might be, that we used individually adjusted intensity of vagal stimulation, presumably coping better with participants' individual differences in skin resistance. More importantly, we used a simple learning task highly comparable to the one used in rodent research[18–20]. In fact, using a single-cue conditioning and extinction paradigm, we found that tVNS facilitated inhibition of fear potentiated startle during extinction training and, more remarkably, during a subsequent extinction test after 24 hours. Most strikingly, tVNS improved recall of extinction memory when tested after 28 days. While the fear learning group still showed elevated shock expectancy ratings and potentiated startle during the conditioned cue in the sham condition, no such effect was observed for the fear learning group receiving tVNS during initial extinction training. Finally, behavioral reinstatement of fear was not only inhibited, but even prevented by tVNS, indicating enhanced consolidation of extinction memory due to vagal stimulation.

Importantly the extinction of conditioned startle reflex potentiation took longer compared to the reduction of shock expectancy ratings, suggesting that the extinction process of subcortically mediated defensive behaviors might take more time than learning on a cognitive level, that the environment has changed. This is particularly important for a better understanding of the mechanisms, that might be responsible for the behavioral return of fear. The current data suggest, that it takes more trials to extinguish defensive action dispositions than cognitive risk assessments. Correspondingly, behavioral fear reducing effects of tVNS took longer to evolve during extinction training, indicating that behavioral fear reducing effects of tVNS boost extinction learning late during extinction training. These data therefore suggest, that while tVNS might have an inherent anxiolytic effect on cognitive fear responses, this effect may not be evident for low level behavioral indices of fear responses. It is therefore even more remarkable, that the current findings indicate that tVNS facilitates long-term extinction of defensive behavior in humans.

We want to further emphasize, that tVNS did not affect the startle responses of the control group during the ITIs during extinction training and subsequent test sessions. Since the control group received US presentations that were explicitly unpaired with the CS, such procedure might also have resulted in increased contextual anxiety in this group, as the US occurred unpredictably during the ITI[45]. However, as tVNS resulted in stronger effects in the fear learning group, we may assume that the extinction enhancing effects of tVNS might be specific to the extinction of cued fear, and may not affect the extinction of contextual anxiety, thus supporting previous work by Genheimer and coworkers[46].

Elaborating the mechanisms of action, animal data suggest that the stimulation of vagal afferents results in specific noradrenergic activation of the basolateral amygdala (BLA) and the ventromedial prefrontal cortex

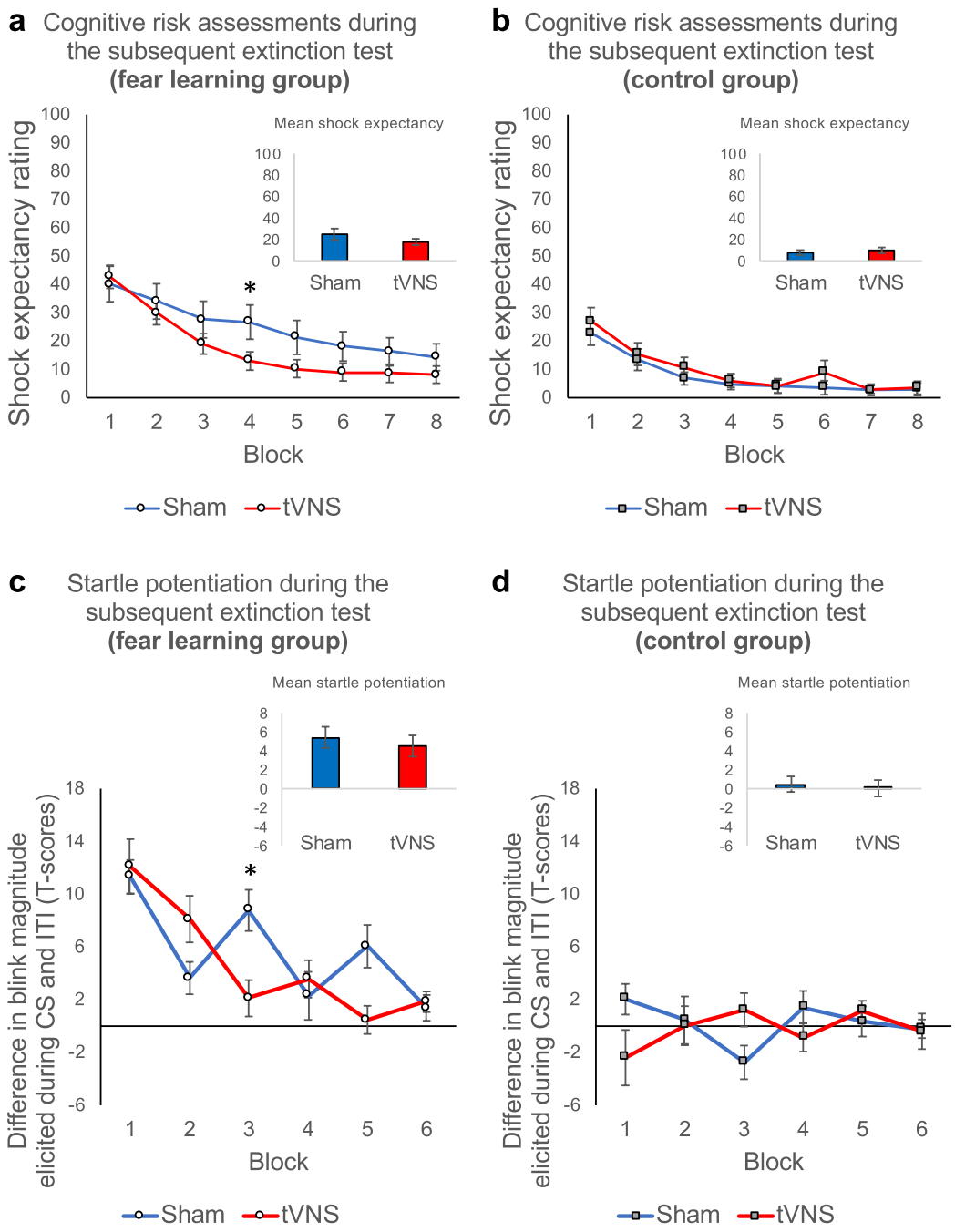


Figure 5. TVNS-paired extinction facilitates subsequent extinction of fear 24 hours later. Upper Panel (**a** and **b**): Mean shock expectancy ratings averaged across blocks of two subsequent extinction test trials for the sham stimulation (*blue*) and tVNS condition (*red*) in the fear learning group (**a**: left) and in the control group (**b**: right). Error bars represent SEM. Fear learning group participants who underwent extinction training under tVNS showed facilitated subsequent extinction 24 hours after the extinction training, indicated by a stronger decrease in shock expectancy ratings in the tVNS relative to the sham condition (see **a**). No such effect was observed in the control group (see **b**). Lower Panel (**c** and **d**): Mean startle potentiation (standardized (T-scores) blink magnitudes elicited during the CS minus standardized (T-scores) blink magnitudes elicited during the inter-trial intervals) averaged across two probe stimuli presented during the subsequent extinction test. Potentiation scores are presented for the sham stimulation (*blue*) and the tVNS condition (*red*) in the fear learning group (**c**: left) and the control group (d: right). Again, error bars represent SEM. Subsequent extinction of fear potentiated startle was accelerated by tVNS relative to sham (see **c**). No such effect was observed in the control group (see **d**).

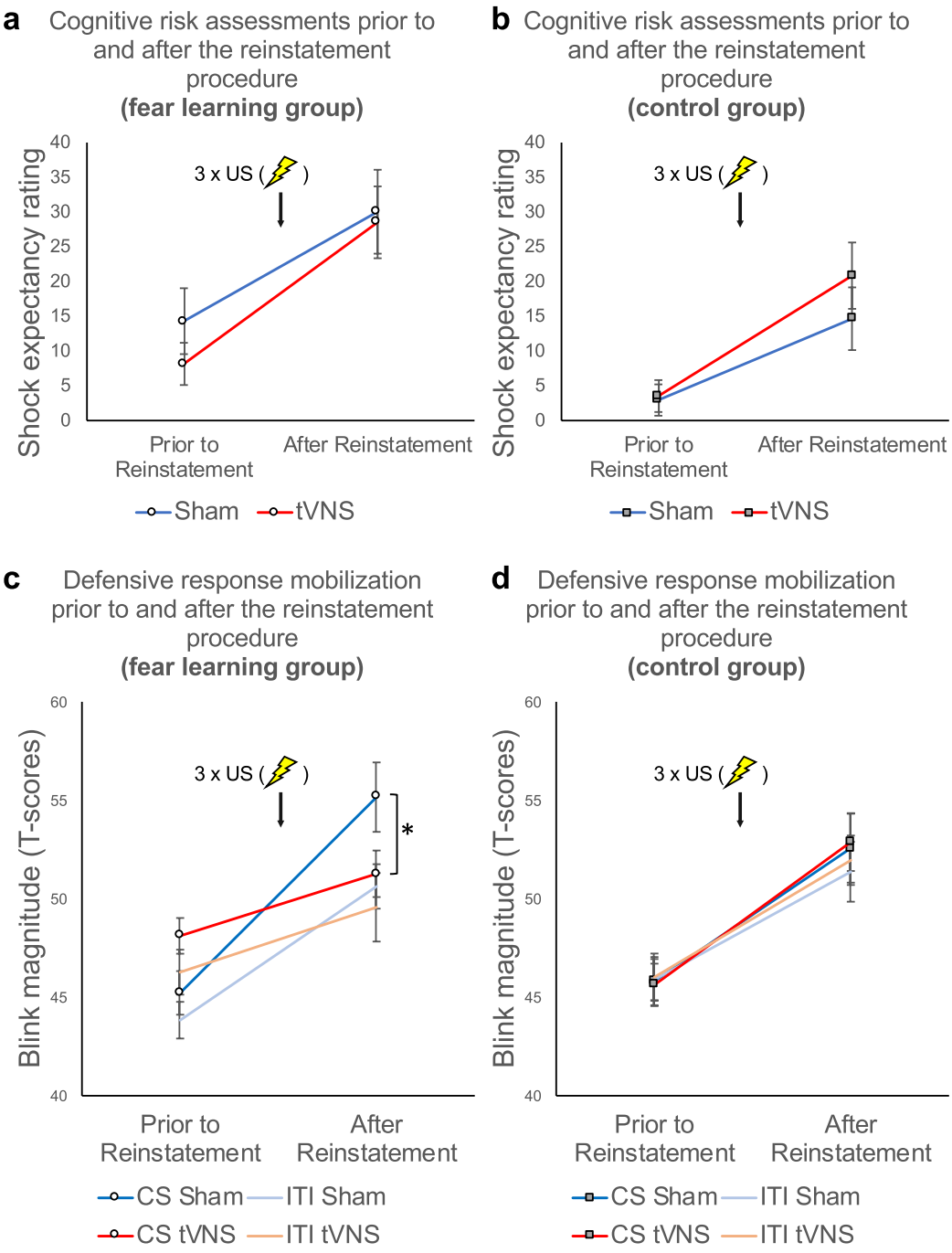


Figure 6. TVNS-paired extinction training results in the prevention of behavioral reinstatement of fear. Upper Panel (**a** and **b**): Mean shock expectancy ratings averaged across the last two subsequent extinction test trials prior to reinstatement and the first two trials following reinstatement for the sham stimulation (*blue*) and tVNS condition (*red*) in the fear learning group (a: left) and in the control group (b: right) Error bars represent SEM. TVNS-paired extinction did not affect the reinstatement of fear responses 24 hours later in neither group. Lower Panel (c and d): Blink magnitudes (standardized T-scores) elicited by the last two probes presented during the CS or the ITI prior to reinstatement for the sham stimulation (*blue shaded*) and the tVNS condition (*red shaded*) in the fear learning group (c: left) and the control group (d: right). Again, error bars represent SEM. TVNS-paired extinction resulted in a prevention of behavioral reinstatement of fear, as fear potentiated startle responses were only re-instated in the sham condition of the fear learning group, but not in the tVNS condition (see c). Reinstatement of fear was not modulated by tVNS in the control group (see d).

(vmPFC)[15–17]. Human research indicates that tVNS also activates the BLA, the vmPFC and also the locus coeruleus[21,47], the central hub for releasing noradrenaline located in the brainstem[48,49]. These circuits are also involved in extinction of conditioned fear responses[1,9,12,50]. Thus, we may assume that tVNS facilitates fear extinction by

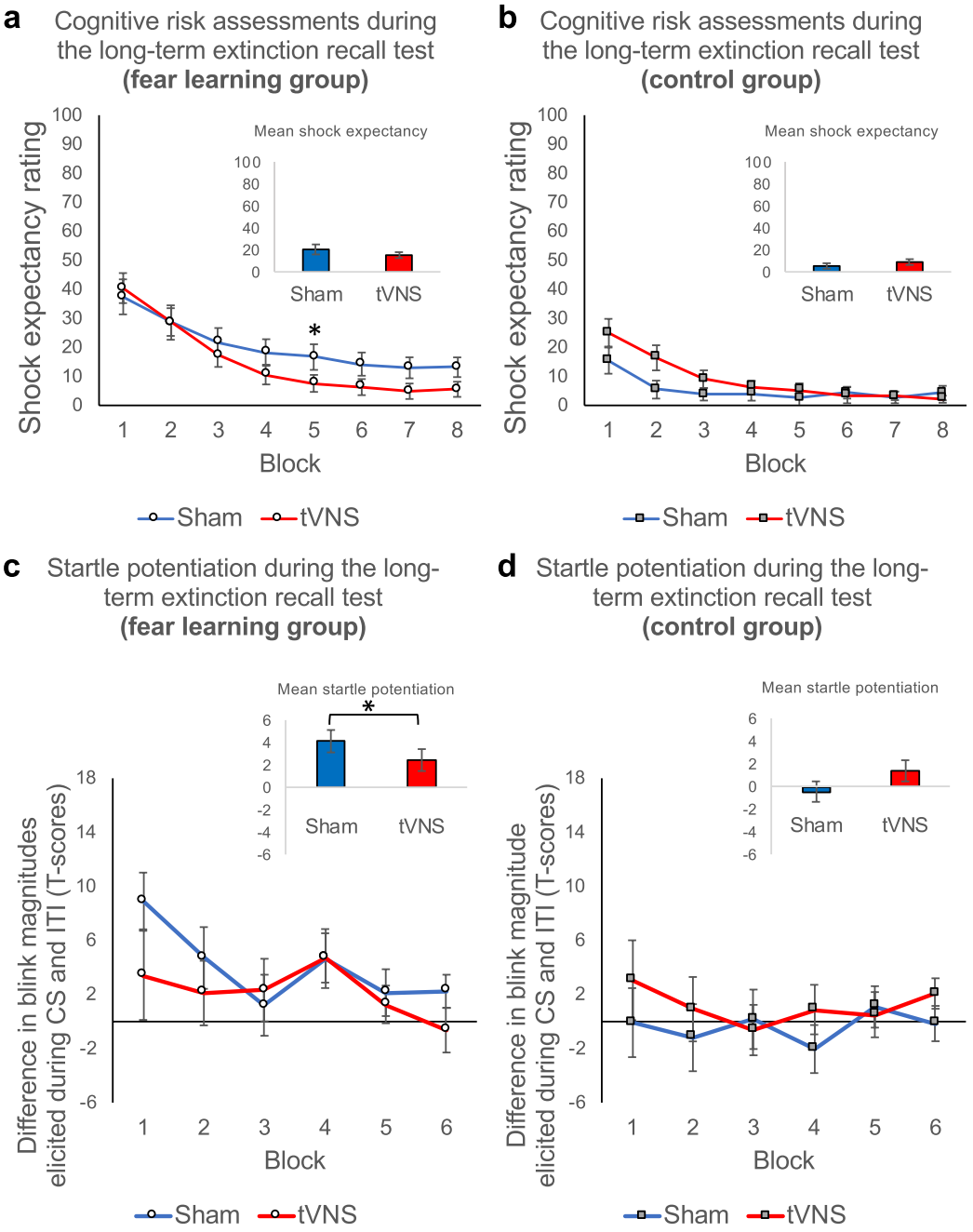


Figure 7. TVNS-paired extinction facilitates long-term recall of fear extinction memory (28 days after the subsequent extinction test). Upper Panel (**a** and **b**): Mean shock expectancy ratings averaged across blocks of two long-term extinction recall test trials for the sham stimulation (*blue*) and tVNS condition (*red*) in the fear learning group (**a**: left) and in the control group (**b**: right). Error bars represent SEM. Fear learning group participants who underwent extinction training under tVNS showed accelerated long-term extinction recall extinction 28 days after the subsequent extinction test, indicated by a stronger decrease in shock expectancy ratings in the tVNS relative to the sham condition (see **a**). No such effect was observed in the control group (see **b**).
Lower Panel (**c** and **d**): Mean startle potentiation (standardized (T-scores) blink magnitudes elicited during the CS minus standardized (T-scores) blink magnitudes elicited during the inter-trial intervals) averaged across two probe stimuli presented during the long-term extinction recall test. Potentiation scores are presented for the sham stimulation (*blue*) and the tVNS condition (*red*) in the fear learning group (**c**: left) and the control group (**d**: right). Again, error bars represent SEM. Fear learning group participants who underwent extinction training under tVNS showed accelerated long-term extinction recall extinction 28 days after the subsequent extinction test, indicated by a stronger decrease in fear potentiated startle responses in the tVNS relative to the sham condition (see **a**). No such effect was observed in the control group (see **b**).

augmentation of noradrenergic signaling in the BLA and the vmPFC, presumably resulting in subsequent inhibition of the central nucleus of the amygdala by their projections to GABAergic intercalated cells[4–6]. In fact, animal research has shown that synaptic plasticity in the vmPFC-BLA pathway – probably critical for long-term fear inhibition - is indeed modulated by vagus nerve stimulation[18–20]. As our data also show improved (long-term) extinction of conditioned fear after tVNS, the same mechanism of action might work in humans.

The current data suggest that tVNS may be a promising adjuvant for exposure-based treatments, the preferred strategy in treating anxiety disorders. Patients with anxiety disorders do show impairments in extinction learning[51,52], probably contributing to the high percentage of relapses after exposure-based treatments and therapy non-responders[33,34,53,54]. As tVNS resulted in anxiolytic effects as well as a facilitation of extinction learning and long-term extinction recall, transcutaneous vagus nerve stimulation may be helpful to reduce the proportion of therapy non-responders but might also be helpful to reduce relapses after successful therapy.

Methods

Subject details and ethical approval. Eighty subjects (23 men, 57 women), mainly chosen from the student population of the University of Greifswald, ranging in age from 18 to 34 ($M = 22.75$, SD = 3.46), participated in the study.

Each participant gave her/his informed consent. All participants either received partial course credits or monetary reward (34 €). This study was approved by the ethical committee of the German Society for Psychology ("Deutsche Gesellschaft für Psychologie; DGPs"). The experiment was performed in accordance to relevant guidelines and regulations.

Eligibility Criteria. Each participant underwent a screening interview by phone to check for in- and exclusion criteria. All participants were assessed to be 18–35 years old with a body-mass-index in normal range (18.5 kg/m^2 to 27 kg/m^2) and to be free from any current or previous medical condition or mental disorder (including nicotine addiction) that would have affected any of the outcome measures or would have contraindicated the application of tVNS (i.e., pregnancy checked by a pregnancy test prior to the study or electrical implants).

Randomization. Eligible participants were then randomly assigned to one of four groups: a fear learning group ($n = 40$) receiving either tVNS ($n = 20$) or sham stimulation to the earlobe ($n = 20$) and a control group ($n = 40$) receiving either tVNS ($n = 20$) or sham stimulation to the earlobe ($n = 20$). The sample size was determined based on previously published articles in the field of fear conditioning and extinction[55].

Single-cue fear conditioning paradigm. It was recently noted, that fear conditioning and extinction protocols often substantially differ between rodent and human research, which in consequence hampered the scientific transfer[56]. We therefore used a multiple-day single-cue fear conditioning and extinction paradigm as it is commonly used in rodent research (for further details, see **Experimental procedure**, see also Fig. 1).

Stimulus Materials

Conditioned stimulus (CS). A blue pentagon on a black background, displayed for 7.135 s on a 24-inch computer monitor with a pixel resolution of 1024 × 768, served as conditioned stimulus (see Fig. 1).

Inter-trial interval (ITI). A black screen, displayed for either 12, 14 or 16 s ($M = 14$ s) on a 24-inch computer monitor with a pixel resolution of 1024 × 768, served as inter-trial interval (see Fig. 1).

Unconditioned stimulus (US; electrical stimulation). An individually adjusted electrical shock served as unconditioned stimulus and was delivered by an S-48K stimulator (Grass instruments) and characterized by a 625 ms period of stimulation, consisting of 125 single pulses with a duration of 2 ms each and a 3 ms break between the stimuli. The average US intensity for the tVNS condition was 3.41 mA (SD = 1.53) and 3.44 mA (SD = 1.47) for the sham condition. No significant difference could be observed in US intensities between the tVNS and sham condition (stimulation, group × stimulation, all $Fs < 1.09$, $P > 0.30$).

Acoustic startle probe. A binaurally presented burst of white noise (95 dB, duration of 50 ms, rise/fall time < 1 ms) served as the acoustic startle probe stimulus.

Transcutaneous vagus nerve stimulation (tVNS). The tVNS stimulation device (CMO2, Cerbomed, Erlangen, Germany) was located in the left auricle where two titan electrodes were placed in one of two positions: In the group receiving active vagus nerve stimulation (tVNS) the electrodes were placed in the cymba conchae, an area exclusively innervated by the auricular branch of the vagus nerve (ABVN)[22], hence stimulating only vagal afferents. In the group receiving a sham stimulation the electrodes were positioned in the center of the ear lobe, an area free of vagal innervation since it is innervated by the great auricular nerve (GAN). The stimulation was delivered during the whole extinction training (session 2), applying a 30 s ON and 30 s OFF procedure with a pulse width of 200–300 μs at a rate of 25 Hz. To ensure the activation of the ABVN or GAN, the stimulus intensity of the stimulation was set to be clearly perceived but with no associated discomfort. All participants therefore underwent a stimulation workup prior to the first extinction training on session 2, where they adjusted the stimulation intensity on their own to be perceptible but below the pain threshold. The average stimulation intensity for both groups was 2.28 mA (SD = 1.13) for the tVNS and 2.53 mA (SD = 1.11) for the sham condition. No significant difference could be observed in stimulation intensities between the tVNS and sham condition (stimulation, group × stimulation, all $Fs < 1.31$, all $Ps > 0.25$). Stimulation was administered for approximately 10 min.

Experimental Procedure

Experimental setting. During each session the participants sat in a dimly lit, sound-attenuated room 1.45 m in front of a computer monitor. First sensors were attached for recording physiological signals as well as the electrodes to deliver the electrical shock (at the non-dominant hand's wrist). A shock workup procedure followed, in which the intensity of the US was individually set at a level the participant defined to be clearly unpleasant, but not painful[57]. Then participants were instructed that during the subsequent learning sessions (acquisition training, extinction training, subsequent extinction test with following reinstatement test I or long-term extinction recall test with following reinstatement test II), the CS, the US and acoustic startle probes may be presented at any time without any explicit information with regard to the contingencies. Each learning session began with a 74 s startle habituation phase, where 6 acoustic startle probes (95 dB, duration 50 ms) were presented repetitively to adapt startle magnitudes to a stable baseline.

Acquisition training (session 1). During the acquisition training the CS was presented 16 times, with a respective duration of 7.135 s and an inter-trial interval of 12, 14 or 16 s ($M = 14$ s, see stimulus materials). To establish a robust and reliable fear response and to increase resistance to extinction, the electrical shock (US) was delivered during the CS in 12 out of 16 trials in the fear learning group (reinforcement rate: 75%). Electrical shocks lasted for 625 ms and were delivered 6.5 s after CS-onset. In the control group 12 shocks were explicitly unpaired with the CS presentation and delivered during the inter-trial intervals (ITI) to ensure that sensitization effects did not differ between groups (US onset of 3, 4, 5, 6, 7, 8, 11 or 12 s after ITI onset ($M = 6.98$ s)). Acoustic startle probes were presented during the CS in 12 out of 16 trials in both the fear learning and control group either 4.5, 5 or 5.5 s after CS onset ($M = 5$ s). Moreover, 12 probes were presented during the inter-trial intervals in 12 out of 16 trials in both the fear learning and control group either 6, 7 or 8 s after ITI onset ($M = 7$ s).

Extinction training (session 2). The second day of investigation took place 24 +/− 4 hours after the acquisition training. After refitting the electrodes for physiological recording and US-delivery, tVNS/sham stimulation was applied. The participants were instructed to adjust the transcutaneous vagus nerve/sham stimulation intensity to be perceptible without being painful starting with an intensity of 0.1 mA. Participants were required to rate their subjective sensation of the stimulation intensity after each adjustment of 0.1 mA on a visual 11-point scale, ranging from nothing (0), light tingling (3), strong tingling (6) to painful (10). The tVNS workup lasted until the participants reported a "tingling" sensation of 8, after which they underwent a 30 s ON and 30 s OFF stimulation protocol to experience the stimulation as it would be during the extinction training[47]. If the participant still reported a sensation of 8 after the protocol, the adjusted intensity was used during extinction training. Then participants were instructed that the session would begin with a 3-minute period to adjust to the tVNS and sham stimulation before CSs, USs and acoustic startle probes might be presented again as during the previous day. Again, no explicit references to the contingencies were made, i.e., participants were not informed that the US was no longer going to be delivered during the following extinction training. The CS was presented 16 times without any US. Acoustic startle probes were presented similar to session 1 (see Fig. 1).

Subsequent extinction test and reinstatement test I (session 3). The subsequent extinction test took place 24 +/- 4 hours after the extinction training. After electrodes for physiological recordings and stimulation were attached, participants were again instructed as prior to the extinction training. During the subsequent extinction test the CS was presented 16 times without any presentation of the US. Then the reinstatement procedure (reinstatement test I) followed during which the US was presented repetitively three times without any CS. However, the background color of the monitor changed to grey in order to exclude conditioning to the background color of the monitor during the ITI. Afterwards the CS was presented again for 16 trials without any presentation of the US (see Fig. 1).

Tests for long-term extinction recall and reinstatement test II (session 4). Approximately 28 days ($M = 28.26$, SD = 3.46) after the third session, with a minimum interval of 21 days, the same procedure as on session 3 was repeated (see Fig. 1) to assess long-term recall of extinction memory.

Assessments

Shock expectancy ratings. Prior to each CS presentation, participants were asked to rate their subjective expectancy of US occurrence during the upcoming CS presentation on a continuous 11-point rating scale (ranging from "0%" to "100%") by shifting a red cursor and pushing the left mouse button (see Fig. 1). During this rating procedure, the CS was presented in smaller size and participants were instructed to rate the probability, that this cue would be followed by the shock during the upcoming trial. This procedure is very much comparable to the clinical procedure during exposure therapy, during which the patient is first asked to rate the probability that the central concern might become true (e.g., fainting) before the exposure exercise begins. Three seconds after completing the rating the conditioned stimulus is presented in full size on the screen. Thus, we made sure that the physiological fear response is not affected by a parallel cognitive evaluation task.

Startle eyeblink response. To measure the eyeblink component of the startle response as an amygdala-dependent indicator of defensive freezing, the electromyographic activity was recorded using two electrolyte filled (Marquette Hellige, Freiburg, Germany) Ag/AgCl miniature surface electrodes (3 mm diameter, Sensormedic, Yorba Linda, CA) attached over the orbicularis oculi muscle underneath the left eye. The EMG signal was amplified by a Coulbourn S75–01 amplifier and filtered with a 30 Hz high-pass and a Kemo LEM-VBF8–03 400 Hz low-pass filter (smoothing the rectified signal with a time constant of 10 ms). Digital sampling was carried out at a rate of 1000 Hz between 100 ms before and 400 ms after the onset of the startle probe. Startle eyeblink responses

were scored semi-automatically using a computer program[58], identifying blink onset and peak amplitude. Valid startle response was scored, if blinks started 20–120 ms after the onset of the acoustic startle probe and peaked within 150 ms with a minimum amplitude of 1.954 μV. If no blink was detected in the defined time window, the trials were scored as zero responses. Trials with clear movement artifacts or excessive baseline shifts were set as missings[59]. Raw blink magnitudes were standardized individually for each participant using a z-score transformation. Subsequently, these standardized responses were converted to T-scores ($50 + (z \times 10)$) for each participant individually. Thus, we made sure that every participant contributed equally to the groups' mean. As the stimulation device for tVNS and sham stimulation produced some noise during extinction training on session 2, only trials were scored, if no stimulation artifacts could be obtained. For acquisition training (session 1), 2.2% of all probed trials were set as missing ($M = 0.66$) and 0.5% were scored as zero responses ($M = 0.15$). For extinction training (session 2), on average 31.2% of all probed trials were set as missing ($M = 9.38$; higher rate of missings due to tVNS induced noise) and 0.3% were scored as zero responses ($M = 0.09$). For the subsequent extinction test (session 3), 4.1% of all probed trials were set as missing ($M = 1.25$) and 1.4% were scored as zero responses ($M = 0.43$). For short-term reinstatement test I (session 3), 4.5% of all probed trials were set as missing ($M = 1.35$) and 3.2% were scored as zero responses ($M = 0.96$). For long-term extinction recall test (session 4), 3.5% of all probed trials were set as missing ($M = 1.06$) and 1.4% were scored as zero responses ($M = 0.43$). For reinstatement test II (session 4), 4.4% of all probed trials were set as missing ($M = 1.33$) and 3.5% were scored as zero responses ($M = 1.05$). So, except for extinction training (during which stimulation induced more noise) proportion of missing data were below 5% during all phases of the experiment.

Electrocardiogram and skin conductance level. We also recorded an electrocardiogram (ECG) and the skin conductance level (SCL) for all experimental sessions. However, the focus of this manuscript firstly lies on the translation of the extinction enhancing effects of vagal stimulation on behavioral indicators of fear, which were found in rodents[18–20]. Second, we wanted to replicate previous results in humans, showing extinction enhancing effects of tVNS on cognitive indicators of fear[23–25]. Thus, the ECG and SCL data are out of this manuscript's scope.

Statistical analyses and graphical representation. Shock expectancy ratings and startle response magnitudes were analyzed using linear mixed regression models with only fixed effects included and an underlying compound symmetry covariance matrix, offering a flexible and powerful analysis of the repeated measures data with missing values. This type of linear mixed regression was chosen, because it is identical to a repeated-measures ANOVA model, the common strategy of analysis for physiological data with multiple time points, with the advantage of also including participants with missing values[60,61], thus, providing stronger comparability to previous fear conditioning studies. Group (fear learning vs. control) and Type of Stimulation (tVNS vs. sham) served as between-subject factors with Trials as within-subject factor. For analyses of startle response magnitudes, the factor Potentiation (CS vs. inter-trial intervals) was included as an additional within-subject factor. Partial Eta-squared was computed following recommendations by Lakens[62]. We conducted the statistical analyses using IBM SPSS Statistics 25. Microsoft Excel was used for creating the figures.

Data and software availability

The data, that support the findings of this study, are available from the corresponding author (Christoph Szeska (christoph.szeska@uni-greifswald.de)) upon reasonable request.

Received: 8 October 2019; Accepted: 15 January 2020;
Published online: 30 January 2020

References

1. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).
2. Bouton, M. E. Why behavior change is difficult to sustain. *Prev. Med. (Baltim).* **68**, 29–36 (2014).
3. Dunsmoor, J. E., Niv, Y., Daw, N. & Phelps, E. A. Rethinking Extinction. *Neuron* **88**, 47–63 (2015).
4. Duvarci, S. & Pare, D. Amygdala microcircuits controlling learned fear. *Neuron* **82**, 966–980 (2014).
5. Amano, T., Unal, C. T. & Paré, D. Synaptic correlates of fear extinction in the amygdala. *Nat. Neurosci.* **13**, 489–494 (2010).
6. Ehrlich, I. *et al.* Amygdala Inhibitory Circuits and the Control of Fear Memory. *Neuron* **62**, 757–771 (2009).
7. Milad, M. R. & Quirk, G. J. Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. *Annu. Rev. Psychol* **63**, 129–151 (2012).
8. Marek, R., Xu, L., Sullivan, R. K. P. & Sah, P. Excitatory connections between the prelimbic and infralimbic medial prefrontal cortex show a role for the prelimbic cortex in fear extinction. *Nat. Neurosci.* **21**, 654–658 (2018).
9. Mueller, D. & Cahill, S. P. Noradrenergic modulation of extinction learning and exposure therapy. *Behav. Brain Res.* **208**, 1–11 (2010).
10. Phelps, E. A., Delgado, M. R., Nearing, K. I. & Ledoux, J. E. Extinction Learning in Humans: Role of the Amygdala and vmPFC. *Neuron* **43**, 897–905 (2004).
11. Burgos-Robles, A., Vidal-Gonzalez, I., Santini, E. & Quirk, G. J. Consolidation of Fear Extinction Requires NMDA Receptor-Dependent Bursting in the Ventromedial Prefrontal Cortex. *Neuron* **53**, 871–880 (2007).
12. Berlau, D. J. & McGaugh, J. L. Enhancement of extinction memory consolidation: The role of the noradrenergic and GABAergic systems within the basolateral amygdala. *Neurobiol. Learn. Mem.* **86**, 123–132 (2006).
13. Hamm, A. O. & Flor, H. Fear learning, fear memory, and psychopathology. *Int. J. Psychophysiol.* **98**, 497–498 (2015).
14. Milad, M. R. *et al.* Recall of Fear Extinction in Humans Activates the Ventromedial Prefrontal Cortex and Hippocampus in Concert. *Biol. Psychiatry* **62**, 446–454 (2007).
15. Hassert, D. L., Miyashita, T. & Williams, C. L. The effects of peripheral vagal nerve stimulation at a memory-modulating intensity on norepinephrine output in the basolateral amygdala. *Behav. Neurosci.* **118**, 79–88 (2004).
16. Follesa, P. *et al.* Vagus nerve stimulation increases norepinephrine concentration and the gene expression of BDNF and bFGF in the rat brain. *Brain Res.* **1179**, 28–34 (2007).

17. Roosevelt, R. W., Smith, D. C., Clough, R. W., Jensen, R. A. & Browning, R. A. Increased extracellular concentrations of norepinephrine in cortex and hippocampus following vagus nerve stimulation in the rat. *Brain Res.* **1119**, 124–132 (2006).
18. Peña, D. F. *et al.* Vagus nerve stimulation enhances extinction of conditioned fear and modulates plasticity in the pathway from the ventromedial prefrontal cortex to the amygdala. *Front. Behav. Neurosci.* **8**, 327 (2014).
19. Peña, D. F., Engineer, N. D. & McIntyre, C. K. Rapid remission of conditioned fear expression with extinction training paired with vagus nerve stimulation. *Biol. Psychiatry* **73**, 1071–1077 (2013).
20. Alvarez-Dieppa, A. A. C., Griffin, K., Cavalier, S. & Mcintyre, C. K. Vagus nerve stimulation enhances extinction of conditioned fear in rats and modulates Arc protein, CaMKII, and GluN2B-containing NMDA receptors in the basolateral amygdala. *Neural Plast.* **2016**, 1–19 (2016).
21. Frangos, E., Ellrich, J. & Komisaruk, B. R. Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: FMRI evidence in humans. *Brain Stimul.* **8**, 624–636 (2015).
22. Peuker, E. T. & Filler, T. J. The nerve supply of the human auricle. *Clin. Anat.* **15**, 35–37 (2002).
23. Burger, A. M. *et al.* Mixed evidence for the potential of non-invasive transcutaneous vagal nerve stimulation to improve the extinction and retention of fear. *Behav. Res. Ther.* **97**, 64–74 (2017).
24. Burger, A. M. *et al.* The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans. *Neurobiol. Learn. Mem.* **132**, 49–56 (2016).
25. Burger, A. M. *et al.* Transcutaneous vagus nerve stimulation and extinction of prepared fear: A conceptual non-replication. *Sci. Rep.* **8**, 11471 (2018).
26. Carew, T. J., Hawkins, R. D. & Kandel, E. R. Differential Classical Conditioning of a Defensive Withdrawal Reflex in Aplysia californica. *Science (80-.).* **219**, 397–400 (1983).
27. Lonsdorf, T. B. *et al.* Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neurosci. Biobehav. Rev.* **77**, 247–285 (2017).
28. Knight, D. C., Smith, C. N., Cheng, D. T., Stein, E. A. & Helmstetter, F. J. Amygdala and hippocampal activity during acquisition and extinction of human fear conditioning. *Cogn. Affect. Behav. Neurosci.* **4**, 317–325 (2004).
29. Rescorla, R. A. Pavlovian conditioning and its proper control procedures. *Psychol. Rev.* **74**, 71–80 (1967).
30. Grillon, C. & Davis, M. Fear-potentiated startle conditioning in humans: Explicit and contextual cue conditioning following paired versus unpaired training. *Psychophysiology* **34**, 451–458 (1997).
31. Haaker, J., Golkar, A., Hermans, D. & Lonsdorf, T. B. A review on human reinstatement studies: An overview and methodological challenges. *Learning and Memory* **21**, 424–440 (2014).
32. Hermans, D., Craske, M. G., Mineka, S. & Lovibond, P. F. Extinction in Human Fear Conditioning. *Biol. Psychiatry* **60**, 361–368 (2006).
33. Craske, M. G. & Mystkowski, J. L. Exposure Therapy and Extinction: Clinical Studies. in *Fear and Learning: From basic processes to clinical implications* (eds. Craske, M. G., Hermans, D. & Vansteenwegen, D.) 217–233 (Amercian Psychological Association, 2006).
34. Ginsburg, G. S. *et al.* Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA Psychiatry* **71**, 310–318 (2014).
35. Davis, M. & Whalen, P. J. The amygdala: Vigilance and emotion. *Molecular Psychiatry* **6**, 13–34 (2001).
36. Davis, M. Neural systems involved in fear and anxiety measured with fear-potentiated startle. *Am. Psychol.* **61**, 741–756 (2006).
37. Hamm, A. O. *Fear-Potentiated Startle. International Encyclopedia of the Social & Behavioral Sciences: Second Edition* **8**, (Elsevier, 2015).
38. Kuhn, M. *et al.* The Neurofunctional Basis of Affective Startle Modulation in Humans: Evidence From Combined Facial Electromyography and Functional Magnetic Resonance Imaging. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2019.07.028 (2019).
39. Steenbergen, L. *et al.* Transcutaneous vagus nerve stimulation (tVNS) enhances response selection during action cascading processes. *Eur. Neuropsychopharmacol.* **25**, 773–778 (2015).
40. Fischer, R., Ventura-Bort, C., Hamm, A. & Weymar, M. Transcutaneous vagus nerve stimulation (tVNS) enhances conflict-triggered adjustment of cognitive control. *Cogn. Affect. Behav. Neurosci.* **18**, 680–693 (2018).
41. Noble, L. J. *et al.* Vagus nerve stimulation promotes generalization of conditioned fear extinction and reduces anxiety in rats. *Brain Stimul.* **12**, 9–18 (2019).
42. George, M. S. *et al.* A pilot study of vagus nerve stimulation (VNS) for treatment-resistant anxiety disorders. *Brain Stimul.* **1**, 112–121 (2008).
43. Hamm, A. O. *et al.* Panic disorder with agoraphobia from a behavioral neuroscience perspective: Applying the research principles formulated by the Research Domain Criteria (RDoC) initiative. *Psychophysiology* **53**, 312–322 (2016).
44. Benke, C., Krause, E., Hamm, A. O. & Pané-Farré, C. A. Predictors of behavioral avoidance during respiratory symptom provocation. *Behav. Res. Ther.* **112**, 63–67 (2019).
45. Schmitz, A. & Grillon, C. Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nat. Protoc.* **7**, 527–32 (2012).
46. Genheimer, H., Andreatta, M., Asan, E. & Pauli, P. Reinstatement of contextual conditioned anxiety in virtual reality and the effects of transcutaneous vagus nerve stimulation in humans. *Sci. Rep.* **7**, (2017).
47. Ventura-Bort, C. *et al.* Effects of Transcutaneous Vagus Nerve Stimulation (tVNS) on the P300 and Alpha-Amylase Level: A Pilot Study. *Front. Hum. Neurosci.* **12**, (2018).
48. Jones, B. E., Halaris, A. E., McIlhany, M. & Moore, R. Y. Ascending projections of the locus coeruleus in the rat. I. axonal transport in central noradrenaline neurons. *Brain Res.* **127**, 1–21 (1977).
49. Bouret, S. & Sara, S. J. Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci.* **28**, 574–582 (2005).
50. Milad, M. R. R. & Quirk, G. J. J. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**, 70–74 (2002).
51. Duits, P. *et al.* Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depress. Anxiety* **32**, 239–253 (2015).
52. Lissek, S. *et al.* Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behav. Res. Ther.* **43**, 1391–1424 (2005).
53. Stewart, R. E. & Chambless, D. L. Cognitive-Behavioral Therapy for Adult Anxiety Disorders in Clinical Practice: A Meta-Analysis of Effectiveness Studies. *J. Consult. Clin. Psychol.* **77**, 595–606 (2009).
54. Craske, M. G. *Anxiety disorders: Psychological approaches to theory and treatment*. (Westview Press, 1999).
55. Kindt, M., Soeter, M. & Vervliet, B. Beyond extinction: erasing human fear responses and preventing the return of fear. *Nat. Neurosci.* **12**, 256–258 (2009).
56. Haaker, J. *et al.* Making translation work: Harmonizing cross-species methodology in the behavioural neuroscience of Pavlovian fear conditioning. *Neurosci. Biobehav. Rev.* **107**, 329–345 (2019).
57. Hamm, A. O. & Vaitl, D. Affective learning: Awareness and aversion. *Psychophysiology* **33**, 698–710 (1996).
58. Globisch, J., Hamm, A. O., Schneider, R. & Vaitl, D. A computer program for scoring reflex eyeblink and electrodermal responses written in PASCAL. *Psychophysiology* **30**, 1, 30 (1993).
59. Blumenthal, T. D. *et al.* Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology* **42**, 1–15 (2005).
60. Bagiella, E., Sloan, R. P. & Heitjan, D. F. Mixed-effects models in psychophysiology. *Psychophysiology* **37**, 13–20 (2000).

61. O'Connell, M. Generalized Linear Mixed Models: A Pseudo-Likelihood Approach. *J. Stat. Comput. Simul.* **48**, 233–243 (1993).
62. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* **4**, 1–12 (2013).

Acknowledgements
We thank Dr. Heino Mohrmann for technical support. This research was supported by a grant of the government of Mecklenburg-Vorpommern to Christoph Szeska. Julia Wendt was supported by a grant provided by the German Research Foundation (DFG; WE 5873/1–1). We acknowledge support for the Article Processing Charge from the DFG (German Research Foundation, 393148499) and the Open Access Publication Fund of the University of Greifswald.

Author contributions
All authors designed the study. C.S. collected and analyzed the data. C.S. wrote the first draft of the manuscript. All authors read, commented on and extensively edited the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-58412-w.

Correspondence and requests for materials should be addressed to C.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplemental information

Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training

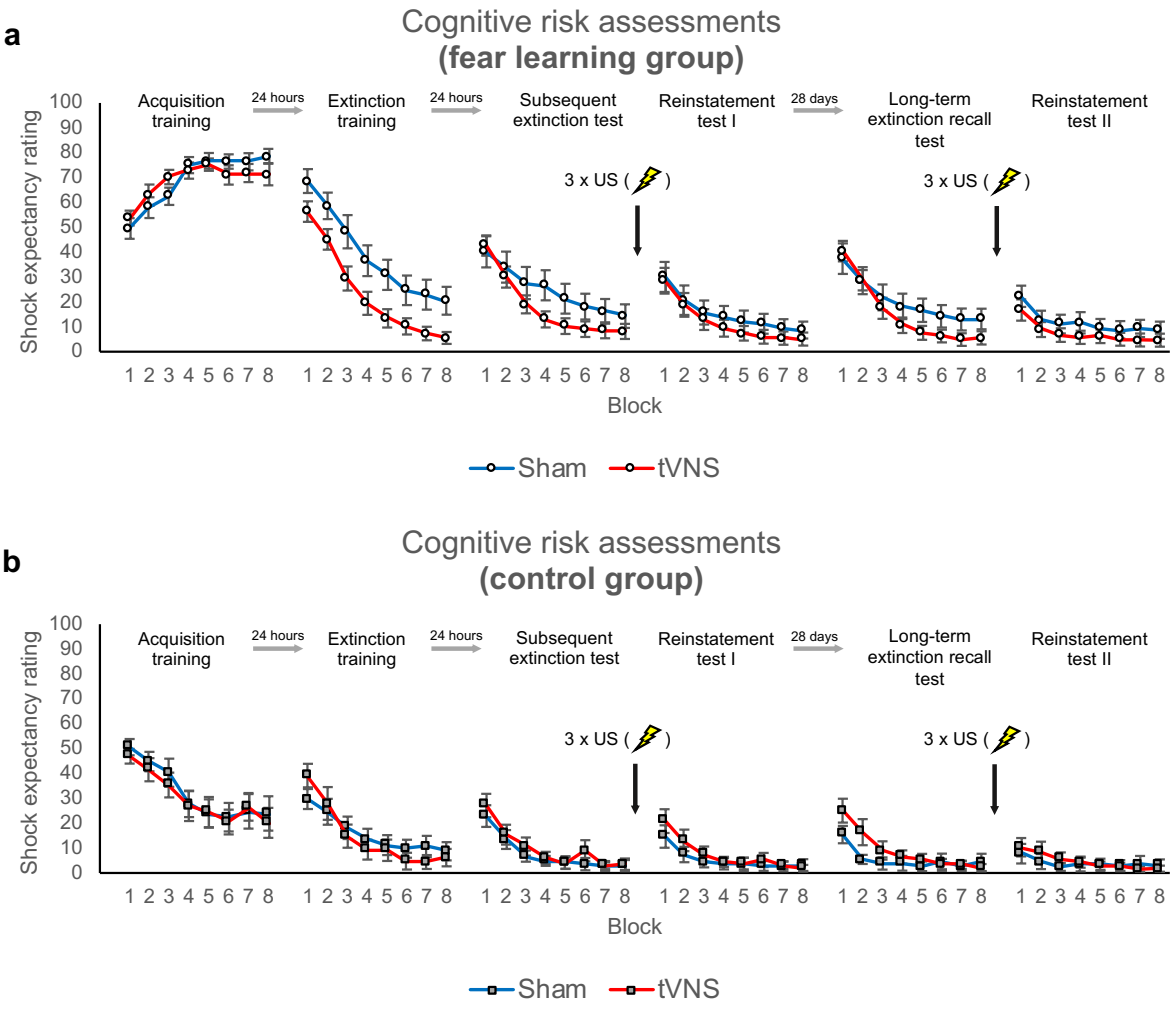
Christoph Szeska*[1], Jan Richter[1], Julia Wendt[2], Mathias Weymar[2], & Alfons O. Hamm[1]

*Corresponding author

[1] University of Greifswald, Department of Physiological and Clinical Psychology / Psychotherapy, Franz-Mehring-Strasse 47, 17487 Greifswald, Germany

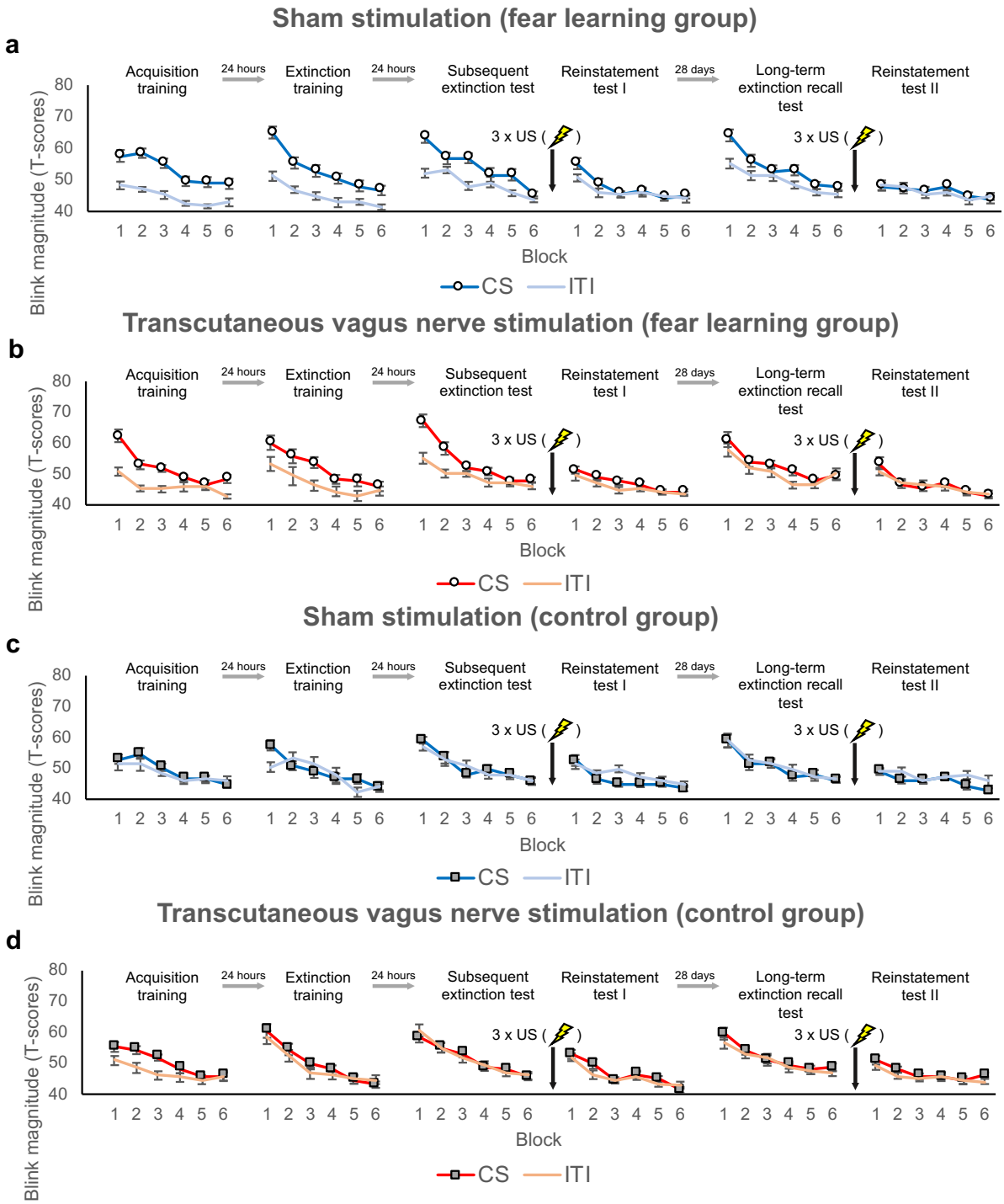
[2] University of Potsdam, Department of Biological Psychology and Affective Science, Karl-Liebknecht Strasse 24/25, 14476 Potsdam, Germany

Corresponding author:
Christoph Szeska
University of Greifswald
Franz-Mehring-Strasse 47
17487 Greifswald
Phone: +49 3834 420 3797
Fax: +49 3834 420 3763
Email: christoph.szeska@uni-greifswald.de



Supplementary Figure 1. Cognitive risk assessments during the multiple-day single-cue fear conditioning and extinction paradigm.

Mean shock expectancy ratings during CS trials averaged across blocks of 2 trials, during acquisition training (session 1), extinction training 24 hours later (session 2), the subsequent extinction test 24 hours after extinction training (session 3), reinstatement test I immediately following the subsequent extinction test (session 3), the long-term extinction-recall test 28 days after the subsequent extinction test (session 4) and reinstatement test II immediately following the long-term extinction recall test for the sham stimulation (*blue*) and tVNS condition (*red*) of the fear learning group (upper panel: a) and the control group (lower panel: b). Error bars represent SEM.


a
Sham stimulation (fear learning group)
b
Transcutaneous vagus nerve stimulation (fear learning group)
c
Sham stimulation (control group)
d
Transcutaneous vagus nerve stimulation (control group)
Acquisition training
24 hours
Extinction training
24 hours
Subsequent extinction test
Reinstatement test I
28 days
Long-term extinction recall test
Reinstatement test II
3 x US ( )
Blink magnitude (T-scores)
80
70
60
50
40
1 2 3 4 5 6
Block
CS
ITI

Supplementary Figure 2. Startle responses during the multiple-day single-cue fear conditioning and extinction paradigm.

Mean blink magnitudes (T-scores) elicited during the CS and during inter-trial intervals averaged in blocks across two startle probes during acquisition training (session 1), extinction training 24 hours later (session 2), the subsequent extinction test 24 hours after extinction training (session 3), reinstatement test I immediately following the subsequent extinction test (session 3), the long-term extinction-recall test 28 days after the subsequent extinction test (session 4) and reinstatement test II immediately following the long-term extinction recall test for the sham stimulation (*blue shaded lines;* panels: a and c) and the tVNS condition (*red shaded lines;* panels: b and d). Error bars represent SEM.

Manuscript 3

Facilitated extinction but impaired extinction recall by eye movement manipulation in humans – indications for action mechanisms and clinical applicability of eye movement desensitization

Christoph Szeska, Heino Mohrmann, & Alfons O. Hamm

Under review

Author contributions:

Christoph Szeska: Conceptualization, Data curation, formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing original draft, Writing review and editing

Heino Mohrmann: Conceptualization, Methodology, Writing review and editing

Alfons O. Hamm: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Writing review and editing

Facilitated extinction but impaired extinction recall by eye movement manipulation in humans – indications for action mechanisms and the applicability of eye movement desensitization

Christoph Szeska[1*], Heino Mohrmann[1] & Alfons O. Hamm[1]

[1] University of Greifswald, Department of Physiological and Clinical Psychology / Psychotherapy, Franz-Mehring-Strasse 47, 17487 Greifswald, Germany

*Corresponding author

Christoph Szeska

University of Greifswald

Department of Physiological and Clinical Psychology / Psychotherapy

Franz-Mehring-Strasse 47

17487 Greifswald

Phone: +49 3834 420 3740

Fax: +49 3834 420 3790

E-Mail: christoph.szeska@uni-greifswald.de

Abstract

Eye movement desensitization and reprocessing (EMDR) therapy utilizes the manipulation of eye movements to reduce affective distress during fear-exposure. Animal research recently suggested the potential neural mechanism underlying these effects, by which increased activity of the superior colliculus (SC), mediating visual attention, increases the inhibition of the basolateral amygdala (BLA), mediating defensive plasticity.

We tested such mechanism in forty healthy humans using a multiple-day single-cue fear conditioning and extinction paradigm. The activity of the SC during extinction was experimentally manipulated by eye movements, as half of the participants executed saccadic eye movements (n = 20; major SC involvement), while the other half executed smooth eye pursuits (n = 20; minor SC involvement). Low-level amygdala-mediated fear potentiated startle responses, fear bradycardia, as well as cognitive risk assessments were analyzed.

Saccadic eye movements facilitated the extinction of low-level fear potentiated startle, and higher saccadic accuracy and range positively correlated with stronger startle response inhibition. Cognitive risk assessments were not affected by different eye movements. However, during extinction recall, startle potentiation and fear bradycardia resurged and partly reached levels obtained after fear acquisition. In contrast, cognitive risk assessments were not elevated during extinction recall.

Within limitations, results support an inhibitory SC-BLA pathway in humans by which eye movements may reduce low-level defensive responding, but not cognitive risk assessments. On the other hand, manipulating eye movements during extinction learning seems to impair extinction recall. Thus, increasing SC activity may enhance initial efficacy of exposure treatment, but additional strategies seem necessary for sustained fear attenuation.

Keywords: extinction, startle potentiation, fear bradycardia, EMDR, saccades, pursuits

1. Introduction

Since its first application in 1989, eye movement desensitization and reprocessing (EMDR) has been proven an efficient psychotherapeutic regimen for post-traumatic stress disorder (Bradley et al., 2005; Landin-Romero et al., 2018; Shapiro, 1989). During such treatment, patients typically follow a moving object with their eyes (e.g., the therapist's finger), while recalling trauma-related and fear-eliciting memories, ultimately resulting in a reduction of affective distress (Bradley et al., 2005; Landin-Romero et al., 2018). However, despite its efficacy, the critical role of eye movements during trauma exposure remained elusive and has been subject of contentious debates (Landin-Romero et al., 2018), until recent rodent research explored the neural mechanisms that might be involved in the modulatory effects of eye movement desensitization (Baek et al., 2019). Baek and colleagues (2019) found stronger reduction of defensive freezing in mice during extinction learning, when the presentation of a fear-eliciting stimulus was additionally paired with alternating bilateral sensory stimulation (ABSS; a moving light). They hypothesized, that the superior colliculus (SC), a midbrain structure critically involved in visual attentional processing (Leigh and Zee, 2015; Müller et al., 2005), might be responsible for these effects. In fact, using optogenetic methods, they found that increased extinction of freezing during sensory stimulation was associated with increased SC firing, which is transmitted to the mediodorsal thalamus (MD) and eventually leads to suppressed activity of the basolateral amygdala - the main hub orchestrating plasticity of defensive responding (Amano et al., 2011; Baek et al., 2019; Herry et al., 2008). Following up on these findings, the current study investigated whether ABSS would also promote extinction of behavioral freezing in humans and whether this effect would also be driven by increased SC activation.

Since we could not stimulate the SC using optogenetic methods in humans, we intended to manipulate SC activation experimentally as an independent variable by manipulating eye movements. Previous studies revealed, that the SC plays a major role in

generating saccadic eye movements, with two different cell populations being crucial: Build-up cells, which start to discharge when a visual stimulus becomes the target for a saccade, and burst-cells, whose activity is related to the generation and termination of a saccadic movement (see Leigh and Zee, 2015 for a review). Albeit a role of the SC in the generation of smooth eye pursuits has been discussed as well, such involvement appears limited to its rostral part, corresponding to the representation of the central visual field and containing fixation neurons (Krauzlis, 2004; Leigh and Zee, 2015). Thus, we may assume, that the SC is stronger involved in controlling saccadic eye movements compared to smooth eye pursuits (Krauzlis, 2004; Leigh and Zee, 2015) and in fact, electrical stimulation of the SC does interfere with saccadic eye movements (Gandhi and Keller, 1999) but has little effect on smooth eye pursuits (Krauzlis et al., 2012).

Following up on Baek and colleagues (2019), we therefore hypothesized that the execution of saccadic eye movements during exposure towards a conditioned fear stimulus (CS), previously paired with an aversive event (unconditioned stimulus; US), would result in reduced freezing in humans relative to smooth eye pursuits. In addition, stronger inhibitory effects were expected to be associated with higher saccadic accuracy (few stops during a target jump) and range, as such eye movements have been linked to broader and more enduring activity in the SC (Goossens and van Opstal, 2012; Leigh and Zee, 2015; Waitzman et al., 1988). However, since eye movement manipulation should primarily suppress autonomic and behavioral indicators of freezing, we did not expect similar inhibitory effects on shock expectancy ratings, an indicator of cognitive risk assessment (LeDoux, 1995; LeDoux and Pine, 2016; Weike et al., 2005).

To test our hypotheses, we manipulated eye movements (saccadic eye movements vs. smooth eye pursuits) during the exposure towards a feared cue using a multiple-day single-cue fear conditioning and extinction paradigm, closely adapted to rodent research and particularly suited for examination of freezing in humans (Haaker et al., 2019; LeDoux,

1995). Defensive freezing was assessed by measuring the potentiation of the startle response - a primary protective brain stem reflex modulated by a circuit dependent on amygdala signaling (see Davis et al., 2010 and Hamm, 2015 for reviews) - and cardiac deceleration (*fear bradycardia*). Rodent and human research identified both responses as primarily subcortically mediated amygdala-driven indicators of defensive responding in the face of inevitable threat (Davis, 2006; Hermans et al., 2013; Kapp et al., 1979; Kuhn et al., 2019; see Hamm, 2020 and Roelofs, 2017 for reviews). In fact, both responses have shown to be not only correlated to each other (Szeska et al., 2021) but also to defensive freezing, the primary behavioral index of fear in animal research (Leaton and Borszcz, 1985; Walker and Carrive, 2003). On the other hand, cognitive risk assessments (i.e., US-expectancy ratings) were measured as a predominantly cognitive and, thus, cortically-mediated index of conscious anxious apprehension (LeDoux, 1995; LeDoux and Pine, 2016; Weike et al., 2005).

2. Materials and Methods

2.1 Participants.

Fifty subjects, selected from a student population of the University of Greifswald, participated in the study (*mean age* = 21.86*, range* = 18 to 31 years, 32 women). During a telephone screening all participants reported to be 18-35 years old, have a body-mass-index in normal range (18.5 kg/m^2 -27 kg/m^2) and to be free from any previous or current medical or mental condition (including nicotine addiction), which would have affected any of the outcome measurements or would have caused problems with electrical stimulation. All participants agreed on not using any medication around the time of experimental participation (i.e., one week). Each participant stated non-impaired ability of hearing, eye movement and vision (in general or by visual aid), gave her/his informed consent and either received partial course

credits or a monetary reward (30 €). Eligible participants were randomly allocated to one of two stimulation conditions, differing only in the applied protocol of alternating bilateral sensory stimulation during extinction (session 2): smooth eye pursuit (n = 26) or saccadic eye movement (n = 24). Allocation to either condition was single-blinded. Five participants were excluded from final analyses due to technical problems during physiological recordings (2 women), three participants were excluded due to errors in the experimental procedure (3 women) and two participants prematurely terminated their participation due to personal reasons, resulting in a final sample size of 40 subjects (*mean age* = 21.87*, range* = 18-31 years, 25 women). Twenty participants were allocated to the smooth eye pursuit condition (*mean age* = 22.50*, range* = 18-29 years, 14 women) twenty subjects were allocated to the saccadic eye movement condition (*mean age* = 21.25*, range* = 18-31 years, 11 women). Both stimulation conditions did not differ significantly with regard to relevant sample characteristics (see **Supplementary Table 1**) (Buhr and Dugas, 2002; Hoyer and Gloster, 2013; Spielberger, 2010). The study was approved by the local ethical committee (ethical committee of the Greifswald Medical School) and we complied with all relevant ethical regulations.

2.2 Stimulus materials.

A tone (75 dB(A); 1000 Hz sine wave) with a duration of 7.135 s served as conditioned stimulus (CS) and was binaurally presented by AKG K66 headphones while participants watched a black screen (24-inch, resolution of 1024 x 768 pixels). The black screen was also continuously presented during the 12, 14 or 16 s (M = 14 s) inter-trial intervals (ITI). An electrical shock with an individually adjusted amperage and a duration of 625 ms (train of 125 x 2 ms electrical pulses with a 3 ms break between pulses) served as aversive unconditioned stimulus (US) and was delivered by an S-48K stimulator (Grass Instruments, West Warwick, RI, USA) to the participant's non-dominant hand's wrist. The individually adjusted intensity

of the US did not differ between both stimulation conditions ($F_{1,38}$ = .595, P = .445). A burst of white noise (95 dB(A)) with an instant rise/fall time (< 1 ms) served as acoustic probe stimulus eliciting the startle eyeblink response and was binaurally presented for 50 ms by AKG K66 headphones.

2.2.1 Alternating bilateral sensory stimulation (ABSS).

Alternating bilateral sensory stimulation was delivered by the EMDR Kit Wireless Light Tube (SE Factory, Groningen, NE), employing a horizontal bar displaying a white light (5000 Lux at zero distance to the light), placed right above the black screen monitor at a distance of 115 cm in front of the participant. While the light remained in middle position during the ITIs, light's movement was started by the experimenter via a Bluetooth-powered mobile application (SE Factory, 2020) during the post-rating period and carried on during the subsequent presentation of the CS. In the smooth eye pursuit condition, the light swept from the right to the left end of the bar and vice versa at a rate of 0.96 Hz (In-App speed of 10/20; In-App mode "Sweep") while it was displayed alternatingly at a rate of 0.96 Hz at the right and left end of the bar in the saccadic eye movement condition (In-App speed of 10/20; In-App mode "Blink"; see **Supplementary Movie 1**). After the presentation of the auditory CS ceased, light's movement was stopped by the experimenter.

2.3 Experimental design and procedure.

Figure 1 gives an overview of the applied single-cue fear conditioning and extinction design (**A**) and the structure of each experimental trial (**B**). We tested our hypotheses applying a single-cue fear conditioning and extinction protocol using a between-subject design with Stimulation (smooth eye pursuit vs. saccadic eye movement) serving as between-group factor, closely following the procedure used in previous animal research (Baek et al., 2019). The subjects were seated in a sound-attenuated and dimly lit room adjacent to the experimenter's

room. Sensors for recording the physiological signals as well as the electrodes for electrical stimulation (US) were attached at the beginning of each experimental session. For all experimental sessions, no explicit information was given with regard to the stimulus contingencies.

2.3.1 Acquisition (session 1).

After attaching the stimulation electrode and the sensors, the first session began with a shock-workup using a standardized protocol (Szeska et al., 2021) until the participants rated the unconditioned stimulus as unpleasant but not painful. After this shock-workup, six probe stimuli were presented to ensure the adaptation of the participant's startle magnitudes to a stable baseline. Then, the single-cue fear conditioning protocol started, during which the conditioned stimulus was presented sixteen times, with twelve CS-presentations being paired with the electrical shock 6.5 s after CS onset (CS-US contingency: 75%). Acoustic startle probes were presented during twelve of the sixteen CSs either at 4.5, 5 or 5.5 s after CS onset Twelve probes were presented during the ITIs.

2.3.2 Extinction (session 2).

The extinction session was set to take place after 24 hours after the first session (to control for circadian effects, the second session was aimed to take place during the same time of the previous day; *mean difference* = 20 min, *range* = 0-3 h). After refitting the sensors, the stimulation electrode and presentation of six startle probes, extinction training began, during which the tone CS was presented again for sixteen trials without administerin the US. Acoustic startle probes were presented during CSs and ITIs, similarly to session 1. In addition, participants were now instructed to follow any movements of the displayed light (see Stimulus materials) with their eyes without moving their head, which would result in

smooth eye pursuits during CSs in the smooth-eye-pursuit condition, and in saccadic eye movements in the saccadic-eye-movement condition.

2.3.3 Short-term extinction recall and return of fear tests (session 3).

The third session was set to take place 24 hours after session 2 (*mean difference* = 32.5 min, *range* = 0-2 h). After preparation (see sessions 1 and 2) the CS was presented again in sixteen trials without any administration of electrical shocks to test recall of extinction memory. Then, a reinstatement procedure followed (reinstatement I), during which three electric shocks administered repeatedly after 3.5, 18.5 and 26.5 s while the color of the screen was changed to gray to avoid counter-conditioning to the ITI. The CS was not presented during this period. A short-term return of fear test followed, during which the CS was presented sixteen times without any US. Startle probes during both the short-term extinction recall and the short-term return of fear test were administered following the structure of sessions 1 and 2.

2.3.4 Long-term extinction recall and return of fear tests (session 4).

The fourth session was set to take place *seven* days after the third appoint (*mean difference* = 7 days; *range* = 0), also at the same time of day as session 3 (*mean difference* = 25 min; *range* = 0-2 h). The experimental procedure during session 4 was kept similar to session 3.

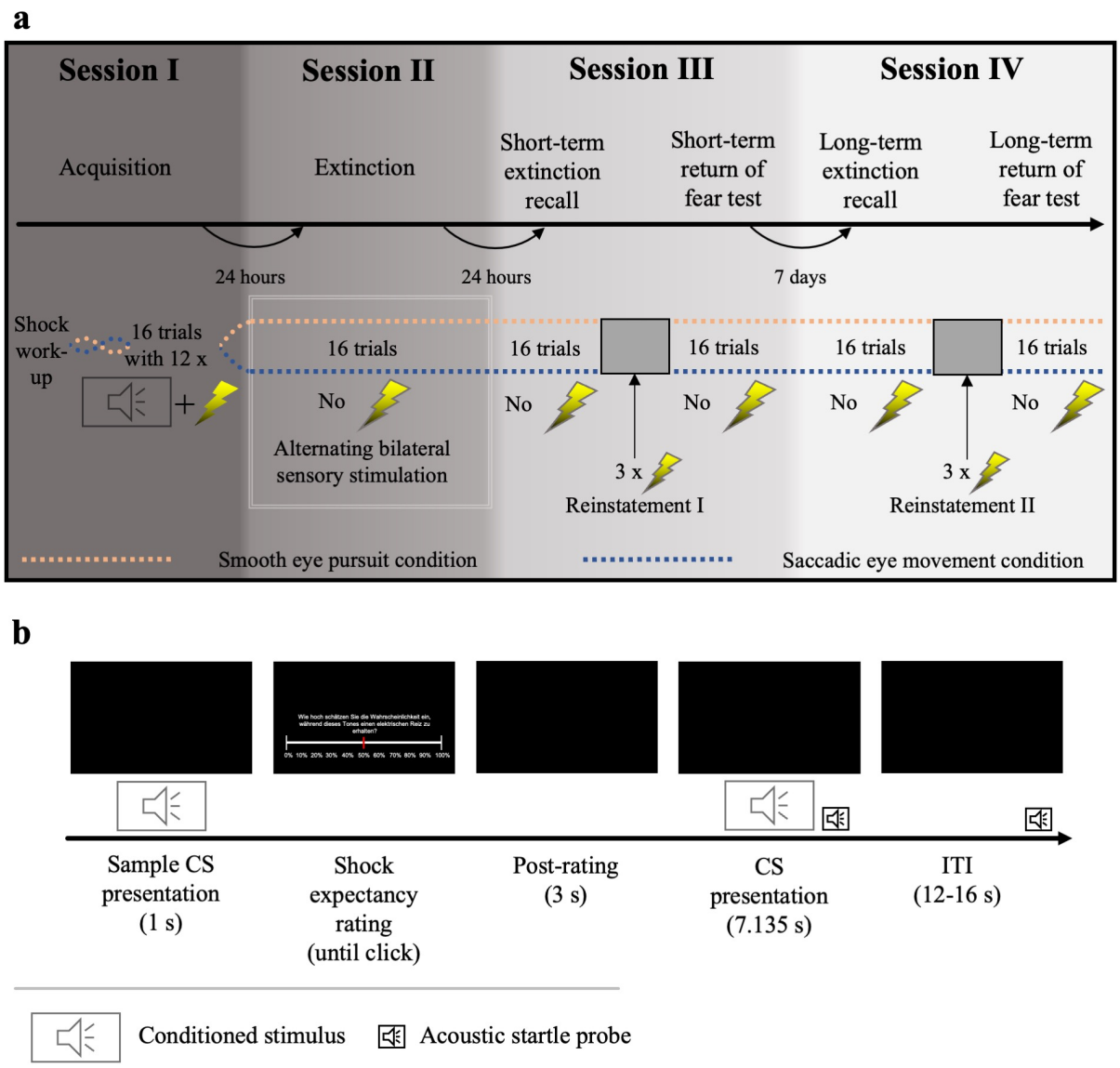


Figure 1. Overview of the experimental design and trial structure.

(A) Experimental sessions: The experiment started with a shock work-up, during which an electrical stimulus was individually adjusted to a level which was perceived as unpleasant, but not painful by the participant. Next an acquisition training started, during which 16 tones were presented, of which 12 were paired with the electrical shock (session 1). During the extinction training (session 2; 24 hours after acquisition), short-term extinction recall and return of fear tests (session 3; 24 hours after extinction) and long-term extinction recall and return of fear test (session 4; 7 days after session 3) each participant received 16 presentations of the tone, which was never paired with an electrical shock. During extinction, participants received alternating bilateral sensory stimulation according to the condition they were allocated to

(saccadic eye movement vs. smooth eye pursuit). During two reinstatement procedures (session 3 and session 4), three non-signaled electrical shocks were applied, while the background color of the monitor changed from black to gray.

(B) Schematic representation of a single experimental trial. Each trial started with a 1-s sample of the auditory conditioned stimulus during a black screen, after which participants were asked to rate their subjective expectancy to receive an electrical shock during the upcoming presentation of this stimulus on a 11-point visual analogue scale ranging from "0%" to "100%" (English translation of the German instruction: "How likely do you think is it, to receive an electrical shock during the presentation of this tone?"). Three seconds after the rating was completed (post-rating period; black screen) the conditioned stimulus was presented for 7.135 s during a black screen, ensuring that physiological responses during its presentation were unbiased by any parallel cognitive evaluation task. After the presentation ofthe conditioned stimulus ceased, an inter-trial interval followed (black screen for 12-16 s). Acoustic startle probes were presented both during the presentation of the conditioned stimulus and during inter-trial intervals (for more information with regard to the rationale of this trial structure see also Holland et al., 2020).

2.4 Dependent measures

2.4.1 Startle eyeblink magnitudes

The eyeblink component of the startle reflex was measured by recording the electromyographic activity of the orbicularis oculi muscle beneath the left eye, using two electrolyte-filled (Marquette Hellige, Freiburg, Germany) Ag/AgCl miniature surface electrodes (3 mm diameter, Sensormedic, Yorba Linda, CA, USA), attached on the skin over the muscle. A Coulbourn V75-04 amplifier filtered and amplified the electromyographic signal with a Coulbourn V75-48 30 Hz high-pass 400 Hz low-pass filter, smoothing and

rectifying the signal with a time constant of 10 ms. The signal was sampled at a rate of 2000 Hz using Acqknowledge (Biopac Systems Inc., Goleta, CA, USA). For subsequent analyses, the signal was sampled down at 1000 Hz. A digital 60 Hz notch filter was applied offline before the scoring procedure started. The time window for scoring startle blink response started 100 ms before and 400 ms after the onset of the probe stimulus. Using a computer program, identifying blink onset and peak amplitude, startle responses were scored semi-automatically (Globisch et al., 1993) only, if starting 20-120 ms and peaking within 150 ms after the probe onset with a minimum amplitude of 1.954 μV. Additionally, each detected amplitude was visually inspected an corrected if necessary. Magnitudes were scored as zero, if no blink was detected in the defined time-window, or were set as missing, if clear movement artifacts or excessive baseline shifts were found (Blumenthal et al., 2005) (see **Supplementary Table 2** for an overview of zero responses and missing values). Raw blink magnitudes were z-transformed and subsequently T-standardized (50+(z x 10)) individually for each participant to ensure that each participant equally contributed to the mean potentiation scores irrespective of their biological baseline differences.

2.4.2 Shock expectancy ratings.

Prior to each presentation of the conditioned stimulus, participants were informed of an upcoming CS presentation, aided by a 1-s sample of the CS (German translation of the instruction: “Next, this tone is going to be presented.”). Participants were then asked to rate their expectancy to receive an electrical shock during the upcoming CS-presentation on an 11-point visual analogue scale, ranging from “0%” to “100 %” (German translation of the instruction: “How likely do you think it is to receive an electrical shock during the presentation of this tone?”; see also **Figure 1B**). This procedure is very much comparable to clinical exposure-based therapy, during which participants are asked to rate the likelihood,

that their central concern (e.g., fainting) might become true, before the exposure begins (Hollandt et al., 2020).

2.4.3 Heart rate.

An electrocardiogram (ECG), using an Einthoven Lead II setup with two electrolyte-filled (Marquette Hellige, Freiburg, Germany) standard Ag/AgCl surface electrodes (8 mm diameter) was recorded. A Coulbourn V75-04 system amplified the signal by the factor 2000 and filtered (8-13 Hz band-pass) the signal, which was subsequently sampled at a rate of 2000 Hz using AcqKnowledge (Biopac Systems Inc., Goleta, CA, USA). For subsequent analyses, the ECG data were digitally sampled down at 400 Hz and corrected for artifacts using ANSLAB software (v. 2.4; Autonomic Nervous System Laboratory, University of Basel, Switzerland). The data were converted into heart rate in beats per minute for every half-second bin of the sampling period (Graham, 1978). To quantify baseline-independent cardiac changes during the presentation of the conditioned stimulus, heart rate during CS presentation was subtracted from base period heart rate (mean of the first two half-seconds after CS onset) for every half-second bin after CS onset (14 data points for the 7.135 s CS duration). Finally, these half-second difference scores were averaged across all trials separately for each experimental stage, allowing to analyze the average time course of the conditioned heart rate response.

2.4.4 Eye movements.

An electrooculogram (EOG) was recorded to measure horizontal eye movements for each participant during extinction (session 2). Two electrolyte-filled (Marquette Hellige, Freiburg, Germany) standard Ag/AgCl surface electrodes (8 mm diameter) were attached at the temples of the participants. A 12-channel Isolated Bioelectric AC/DC Amplifier System (San Diego Instruments, San Diego, CA, USA) amplified the measured EOG signal by the factor 1000

and filtered the signal using a 35 Hz low-pass filter. The signal was sampled at a rate of 2000 Hz using AcqKnowledge (Biopac Systems Inc., Goleta, CA, USA). To allow a thorough investigation of the eye movements during ABSS, a hardware reset was executed at the beginning of each post-rating period (see **Figure 1B**) just prior to the start of ABSS, resetting the EOG to a baseline of zero to control for baseline-drifts in the EOG signal. Second, to cope with different onset latencies of eye movements across trials and participants due to the manual start of the ABSS by the experimenter, the signal was sampled down offline at 10 Hz and the individual onset of eye movements was scored for each trial and each participant. This procedure allowed us to run a custom MATLAB-based script, computing a grand average of the EOG signal for each participant across all post-rating periods and CS presentations, which was corrected for individual differences in scored latencies across trials and corrected for linear trends, e.g., due to remaining drifts in the EOG signal (see **Supplementary Figure 5**). Based on the latency of this averaged EOG signal (equals to the mean of individual latencies across trials within the participant), the average time window of CS presentation after onset of eye movements was calculated for each participant. The averaged and de-trended raw EOG signal during this time window was extracted from the data. Next, we performed a z-transformation of the extracted EOG data for each participant so that changes from baseline were available on a comparable scale. Z-transformed EOG data were then used to create a distribution of gaze durations (in ms) across eye positions for all participants (z-scores; see **Figure 4A**). Next, we assessed the interquartile range of the individual distribution of gaze durations across eye positions for each participant, with higher interquartile ranges indicating higher saccadic accuracy (i.e., longer gaze durations at extreme eye positions at the expense of intermediate ones). Finally, we assessed the range of the distribution of gaze durations across eye positions for each participant, providing an index of the individual amplitude of eye movements.

2.5 Data Analyses.

To analyze the change of the outcome measures during and across each experimental stage, we used linear mixed regression models with only fixed effects included, which were created using restricted maximum likelihood estimation to include all available data, while the covariance structure of the repeated measurements was modelled to best fit the sample data for highest statistical power (see **Supplementary Dataset 1**) (Duricki et al., 2016). *Stimulation* (smooth eye pursuit vs. saccadic eye movement) served as between-group factor.

To evaluate the change of startle responses and shock expectancy ratings *during* each experimental stage, we included the repeated-measures factor *Trial*. For analyses of startle responses, we additionally included the within-subject factor *Potentiation* (CS vs. ITI). As heart rate responses during single trials are very noisy, we abstained from trial-based analyses of heart rate responses and included *Time* (in half second bins) as within-subject factor to examine the course of average heart rate change during CS presentation for each experimental session.

To evaluate initial recall of fear/extinction memory *across* experimental sessions, we compared the last trial of each experimental session with the first trial of the subsequent session by including the within-subject factor *Recall*. To compare average physiological and verbal measures across two experimental sessions, we included the within-subject factor *Phase*.

To test for differences in the distribution of gaze durations across eye positions between stimulation conditions, we used the two-sample Kolmogorov-Smirnov test. The interquartile range of the individual distribution of gaze durations as an index of saccadic accuracy was compared by univariate analysis of variances (ANOVA). To evaluate the differential impact of saccadic accuracy in interaction with saccadic range on the assessments in both stimulation conditions, we performed univariate analysis of variances with covariates (ANCOVA) using *Saccadic accuracy* (interquartile range of the distribution of eye

movements, see Assessments) and *Range of eye movements* as covariates. If significant interactions between the predictors were found with regard to the dependent variables, Pearson correlation between the product of *Saccadic accuracy x Range of eye movements* and the dependent variable was calculated examining the association between larger and more accurate saccades and the dependent measures.

We used SPSS Statistics Ver. 27 for all statistical analyses and Microsoft Excel 2016, Microsoft PowerPoint 2016 and SPSS Statistics Ver. 27 were used for figure creation. Bonferroni correction was applied when relevant.

3. Results

3.1 Saccadic eye movements facilitate extinction of conditioned freezing

First, all participants learned to associate a tone (CS) with an unpleasant electrical shock (US) during a single-cue fear conditioning procedure (acquisition; session 1), indicated by increased shock expectancy ratings (Trial, $F_{15,355.433} = 4.371$, $P < .001$; left panel of **Figures 2A** and **3A**), a reliable potentiation of the brain stem startle reflex (Potentiation, $F_{1,73.763} = 353.255$, $P < .001$; left panel of **Figures 2B, 3B and C**) and a profound fear bradycardia evoked by the CS (Time, $F_{13,74.454} = 7.815$, $P < .001$; left panel of **Figure 2C**), indexing robust conditioned behavioral freezing elicited by the auditory CS. Importantly, the two experimental groups (saccadic eye movement vs. smooth eye pursuit) did not differ in their acquired defensive responses and cognitive risk assessments after acquisition training (all Fs ≤ 1.277, all $Ps \geq .248$; **Figures 2 and 3**).

After a twenty-four-hours consolidation period, participants underwent an extinction training (session 2), during which the conditioned auditory stimulus was presented again without the aversive US and ABSS was applied during tone presentation according to the

assigned stimulation condition (see Materials and Methods and **Supplementary Movie 1**). Indicating compliance to the instruction, we found significant differences between both groups in the distribution of gaze durations across eye positions (Kolmogorov-Smirnov test, $P < .001$; **Figure 4A**), with longer gaze durations at extreme eye positions towards the left and right field of view at the expense of intermediate positions in the saccadic eye movement group compared to the smooth eye pursuit group. Likewise, the individual interquartile range of the distribution of gaze durations across eye positions during the CS presentation (i.e., the accuracy of saccadic eye movement) was increased in the saccadic eye movement condition compared to smooth eye pursuit group (Stimulation, $F_{1,38} = 10.242$, $P = .003$; **Figure 4B**).

At the beginning of extinction, we found comparable shock expectancy ratings relative to the last trial of acquisition in both groups (Recall, $F_{1,38} = 1.251$, $P = .270$; middle panel of **Figure 2A**), indicating successful recall of previously acquired fear memory. Throughout the extinction training, shock expectancies significantly declined (Trial, $F_{15,38} = 36.631$, $P < .001$; middle panel of **Figure 2A**), and were overall reduced compared to acquisition (Phase, $F_{1,76} = 147.651$, $P < .001$; middle panel of **Figure 3A**). As expected, reduction of shock expectancies was not affected by the type of eye movements during session 2 (all $Fs \leq 1.34$, all $Ps \geq .227$; middle panel of **Figures 2A, 3A**).

Startle blink magnitudes also declined throughout extinction (Trial, $F_{11,109.703} = 11.608$, $P < .001$; middle panel of **Figure 2B**, **Supplementary Figure 1**). Yet, there was still a strong potentiation of the startle reflex elicited during the CS compared to blinks evoked during the inter-trial interval (Potentiation, $F_{1,77.101} = 89.626$, $P < .001$; middle panel of **Figure 3C**). Remarkably, extinction of fear potentiated startle was significantly stronger in the saccadic eye movement condition compared to the smooth eye pursuit group (Potentiation x Stimulation, $F_{1,77.101} = 4.043$, $P = .048$; middle panel of **Figures 2B, 3B**). Moreover, while startle responses were still potentiated during late extinction (last third of the trials) in the smooth eye pursuit group (Potentiation, $F_{1,19} = 6.192$, $P = .022$; middle panel of **Figure 2B**),

this was not the case in the saccadic eye movement condition (Potentiation, $F_{1,19} = 2.020$, $P = .171$; middle panel of **Figure 2B)** . Correspondingly, startle potentiation during extinction was overall reduced compared to acquisition (Phase x Potentiation, $F_{1,38} = 48.850$, $P < .001$; middle panel of **Figure 3B**) and this overall inhibition of defensive responding was particularly fostered by saccadic eye movements (Phase x Potentiation x Stimulation, $F_{1,38} = 6.717$, $P = .013$; middle panel of **Figure 3B**).

As during acquisition, the CS also elicited a strong cardiac deceleration during extinction (Time, $F_{13,73.936} = 3.688$, $P < .001$; middle panel of **Figure 2C**). This fear bradycardia was less pronounced during CS processing (3 – 5 s after CS-onset) in the saccadic eye movement compared to the smooth eye pursuit condition (Time x Stimulation, $F_{13,73.936} = 2.109$, $P = .023$; middle panel of **Figure 2C**). However, comparing acquisition and extinction, cardiac deceleration was significantly attenuated during extinction for both eye movement interventions (Phase, $F_{1,38} = 50.985$, $P < .001$; Phase x Stimulation $F_{1,38} = .007$, $P = .935$; middle panel of **Figure 2C**).

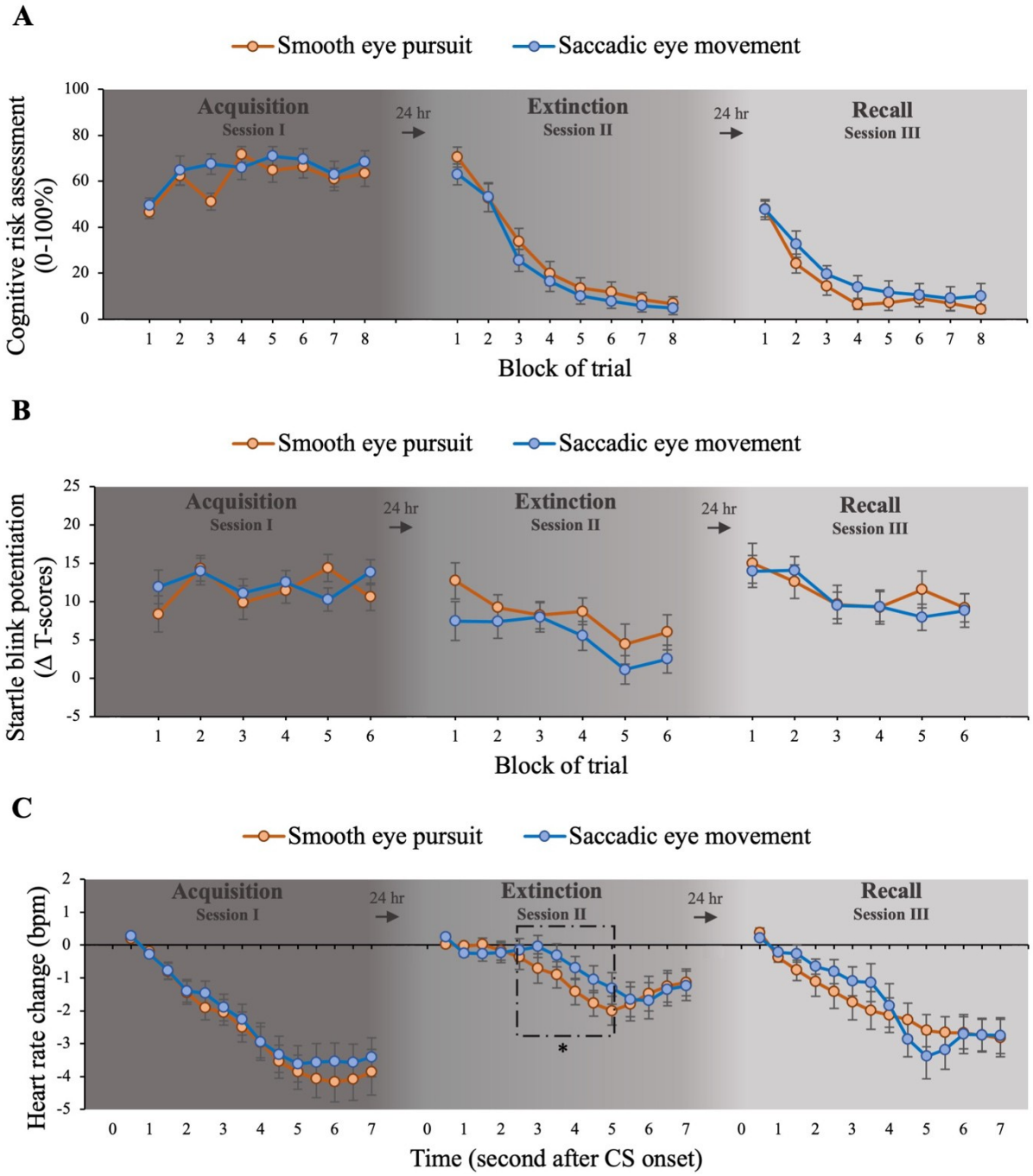


Figure 2. (A) Cognitive risk assessment during the CS presentation throughout acquisition (session 1), extinction (24 hours later; session 2) and short-term extinction recall test (session 3) averaged across two trials.

(B) Standardized (T-scores) potentiation of the startle response (CS minus ITI) throughout acquisition (session 1), extinction (session 2) and short-term extinction recall test (session 3),averaged across two probes.

(C) Heart rate change after CS onset during acquisition (session 1), extinction (session 2) andshort-term extinction recall test (session 3), averaged across all trials, depicted in half-secondbins. The time window, during which we found strongest differences between both stimulation conditions during extinction, is framed by a dashed line.

For all graphs: Orange lines and bars represent data of the smooth eye pursuit condition, blue lines and bars represent data of the saccadic eye movement condition. Error bars represent thestandard error of the mean. Asterisks indicate statistical significance of between-group analyses, with * for $P < .05$, ** for $P < .01$, *** for $P < .001$.

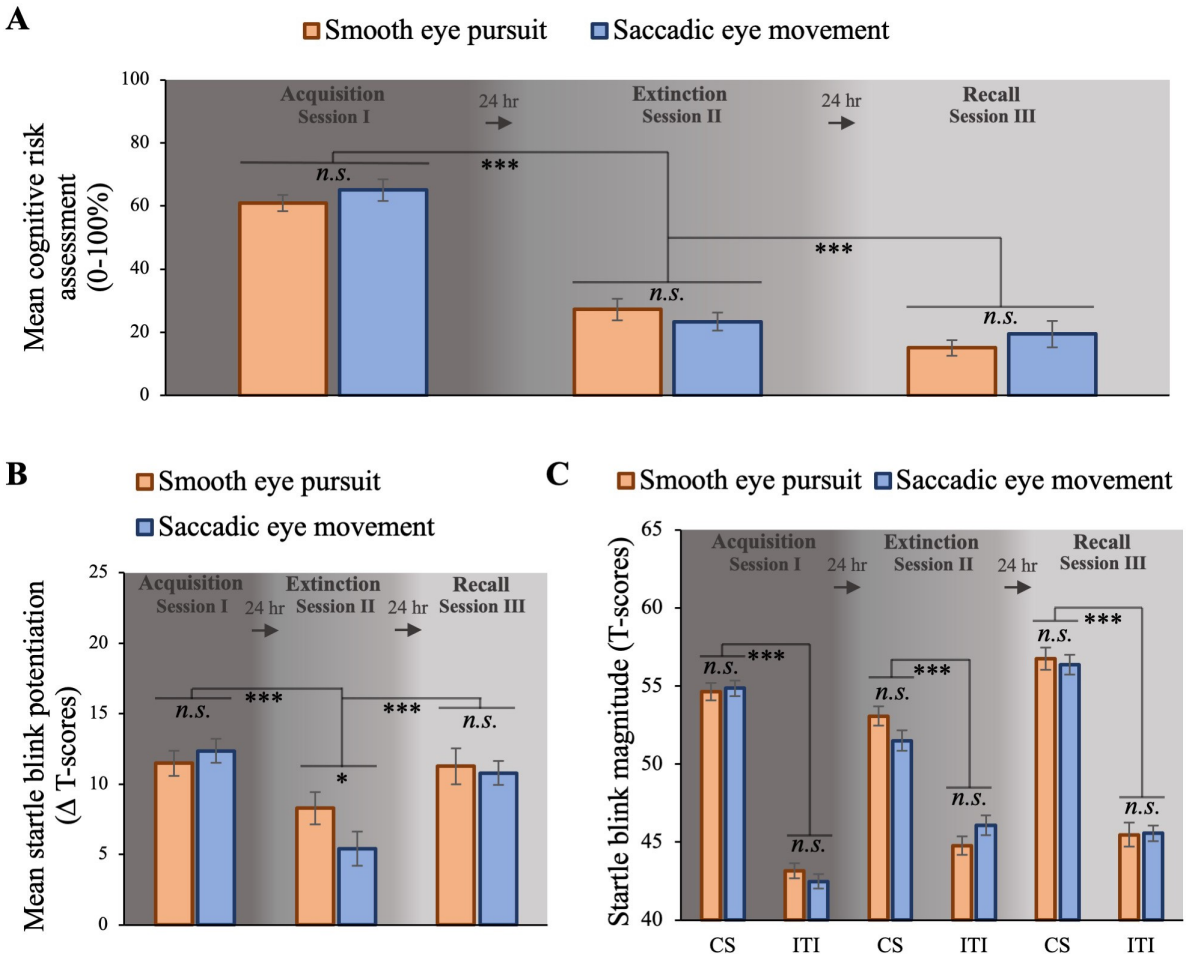


Figure 3. (A) Mean cognitive risk assessment during the CS presentation during acquisition (session 1), extinction (24 hours later; session 2) and short-term extinction recall test (session 3) averaged across all trials.

(B) Mean standardized (T-scores) potentiation of the startle response (CS minus ITI) during acquisition (session 1), extinction (session 2) and short-term extinction recall test (session 3), averaged across all probed trials.

(C) Mean standardized (T-scores) magnitudes of the startle blink response during the presentation of the CS and ITI during acquisition (session 1), extinction (session 2) and short-term extinction recall test (session 3), averaged across all probed trials.

For all graphs: Orange lines and bars represent data of the smooth eye pursuit condition, blue lines and bars represent data of the saccadic eye movement condition. Error bars represent the standard error of the mean. Asterisks indicate statistical significance of between-group analyses, with *n.s.* for $P > .05$, * for $P < .05$, ** for $P < .01$, *** for $P < .001$.

3.2 Higher saccadic accuracy and range are associated with stronger extinction of fear potentiated startle

As expected and in line with the above described results of between-group analyses, saccadic accuracy and range of eye movements did not differentially impact on cognitive risk assessments or heart rate in the stimulation conditions (shock expectancy: Stimulation x Saccadic accuracy x Range, $F_{1,32} = 2.683$, $P = .111$; heart rate: Stimulation x Saccadic accuracy x Range, $F_{1,32} = .236$, $P = .630$).

In contrast, higher saccadic accuracy and range was significantly associated with stronger reduction of the fear potentiated startle in the saccadic eye movement, but not in the smooth eye pursuit condition (Stimulation x Saccadic accuracy x Range, $F_{1,32} = 4.05$, $P = .034$; **Figure 4C, D**). Correlational analyses in the saccadic eye movement condition supported this finding, showing that the arithmetic product of saccadic accuracy and range corresponded to stronger extinction of the fear potentiated startle reflex ($r = -.346$, P(one-tailed) = .064; **Figure 4C**), driven by specifically reduced startle responses elicited during CSs ($r = -.410$, P(one-tailed) = .036), but not by modulation of blink magnitudes elicited

during the ITI (r = .241, P(one-tailed) = .153). This negative correlation between saccadic accuracy and extinction of startle potentiation was most pronounced in participants who executed saccades with highest range (upper tertile of ranges in the saccadic eye movement condition, n = 7; correlation with startle potentiation: r = -.729, P(one-tailed) = .032; **Supplementary Figure 2A**; correlation with CS startles: r = -.783, P(one-tailed) = .019; **Supplementary Figure 2B**; correlation with ITI startles: r = .603, P(one-tailed) = .076; **Supplementary Figure 2C**).

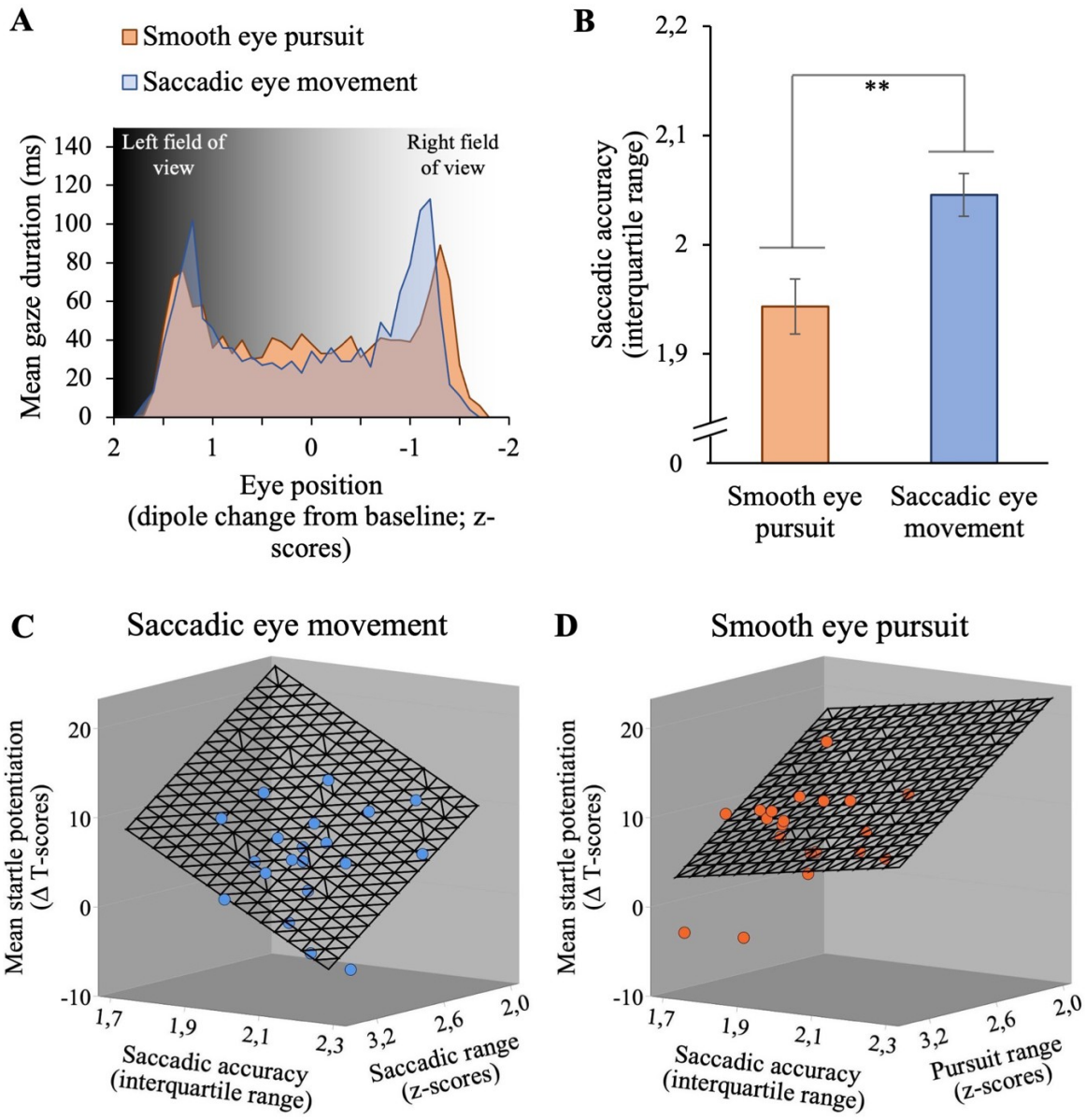


Figure 4. Higher saccadic accuracy and range is related to stronger extinction of fear potentiated startle

(A)**,** Distribution of mean gaze durations across eye positions during CS presentation during extinction.

(B)**,** Saccadic accuracy during CS presentation as indexed by the interquartile range of the individual distribution of mean gaze durations across eye positions. Error bars represent the standard error of the mean.

(C) and (D)**,** Scatter plots of mean standardized (T-scores) potentiation of the startle response (CS vs. ITI) as a function of saccadic accuracy (interquartile range) and range of eye movements in the saccadic eye movement condition (C) and in the smooth eye pursuit condition (D). The regression plane indicates the direction of the linear relationship between the variables.

Asterisks indicate statistical significance of between-group analyses or correlations, with * for $P < .05$, ** for $P < .01$, *** for $P < .001$.

3.1 Impaired short-term extinction recall and weak evidence for sustained defensive response inhibition after eye movement manipulation during extinction

After another twenty-four-hours consolidation period, the light was removed and we tested the short-term recall of extinction memory (short-term extinction recall; session 3) by presenting the CS without any administration of the US. All participants showed initial recovery of fear, indicated by initially increased shock expectancy (Recall, $F_{1,38} = 204.156$, $P < .001$; right panel of **Figure 2A**) and startle potentiation relative to the last trial of extinction training of the previous day (Recall x Potentiation, $F_{1,66.769} = 25.915$, $P < .001$; right panel of **Figure 2B**). Shock expectancies and startle potentiation decreased throughout the short-term extinction recall test (shock expectancy: Trial, $F_{15,38} = 46.034$, $P < .001$; startle: Trial, $F_{11,141.466} = 10.723$, $P < .001$; right panel of **Figures 2A, B, Supplementary Figure 1**),

indicating re-extinction of fear. Yet, startle potentiation was still evident throughout the entire short-term extinction recall test (Potentiation, $F_{1,284.387} = 318.912$, $P < .001$; right panel of **Figures 2B, 3B and C**), as was fear bradycardia (Time, $F_{13,75.739} = 5.182$, $P < .001$; right panel of **Figure 2C**), indicating robust defensive freezing.

The type of eye movement manipulation during the previous day had no impact on any of the reported effects during the short-term extinction recall (shock expectancy: all $Fs \leq 1.490$, all $Ps \geq .158$; right panel of **Figures 2A, 3A**; startle potentiation: all $Fs \leq 1.536$, all $Ps \geq .125$; right panel of **Figures 2B, 3B and C, Supplementary Figure 1**). Likewise, no differences in overall risk assessments (Stimulation, $F_{1,38} = .829$, $P = .368$; right panel of **Figure 3A**), startle potentiation (Potentiation x Stimulation, $F_{1,284.387} = .246$, $P = .620$; right panel of **Figure 3B**) or fear bradycardia (Time x Stimulation, $F_{13,75.739} = 1.156$, $P = .328$; right panel of **Figure 3C**) were found between both groups during the short-term extinction recall test.

Remarkably, while shock expectancies were significantly lower in the short-term extinction recall test compared to the previous day (Phase, $F_{1,38} = 13.881$, $P < .001$; right panel of **Figure 3A**), startle potentiation and fear bradycardia were significantly stronger (startle: Phase x Potentiation, $F_{1,51.386} = 30.486$, $P < .001$; heart rate: Phase, $F_{1,38} = 13.782$, $P < .001$; right panel of **Figures 2B, C, 3B**). Potentiation scores even reached the level of the acquisition, as if participants never underwent extinction training (Phase, $F_{1,53.150} = 1.833$, $P = .182$; right panel of **Figure 3B**). Thus, the induction of eye movements during extinction generally seemed to hamper the recall of extinction memory on a defensive behavioral level, while declarative extinction memory could be recalled.

Following the short-term extinction recall test, three non-signaled electrical shocks were applied in the absence of the CS. As expected, we found significant reinstatement in all indicators of fear, that extinguished during a following extinction training (short-term return of fear test). We found no differences in reinstatement of fear and recall of extinction memory

with regard to the type of eye movements employed during the first extinction session (**Supplementary Figure 3** and **4;** for a detailed description of the results see **Supplementary Dataset 2**).

One week after session 3, we conducted a long-term extinction recall test, during which we repeated the procedure of session 3 (long-term extinction recall; session 4). Again, participants showed initial recovery of the fear response followed by a further extinction during the long-term extinction recall test. Neither initial fear recovery nor long-term extinction differed between eye movement conditions (**Supplementary Figure 1** and **3**). As during session 3, re-experience of the aversive event (reinstatement II) resulted in a significant reinstatement of fear in all dependent variables, which did not differ between eye movement conditions for shock expectancy ratings (**Supplementary Figure 3A**). Only for startle responses, reinstatement was attenuated in the saccadic eye movement condition (**Supplementary Figure 4**). During the following extinction training (long-term return of fear test), again, fear responses decreased and no modulatory effects of previous eye movements were found, aside from the already described less sensitized startles in the saccadic eye movement condition right after the reinstatement procedure (**Supplementary Figure 1** and **3**; for a detailed description of the results during session 4 see **Supplementary Dataset 2**).

4. Discussion

Eye movement desensitization and reprocessing therapy (EMDR) is a psychotherapeutic regimen for post-traumatic stress disorder, most renowned for manipulating eye movements during exposure towards fear-eliciting cues to achieve a reduction of affective distress (Landin-Romero et al., 2018; Shapiro, 1989). By showing that increased activity of the superior colliculus (SC) leads to an inhibition of the basolateral amygdala (BLA), recent rodent research provided a link between SC-controlled visual

attentional processing and the inhibition of BLA-controlled defensive responding, thus, suggesting a subcortical mechanism underlying this therapeutic approach (Amano et al., 2011; Baek et al., 2019; Herry et al., 2008; Leigh and Zee, 2015; Müller et al., 2005). The current study is a first endeavor to test such mechanism in humans by manipulating eye movements during fear extinction, the laboratory analog of exposure therapy (Hermans et al., 2006). Here, we hypothesized that saccadic eye movements would facilitate extinction of attentive immobility – a subcortically mediated defensive strategy during inevitable threat (Mobbs et al., 2020; Szeska et al., 2021) - but not cortically mediated shock expectancies (LeDoux, 1995; LeDoux and Pine, 2016). In order to test the predictions derived from the animal model, we compared fear extinction during such saccadic eye movements with fear extinction during smooth eye pursuits, as the former is known to recruit a larger population of SC neurons compared to the latter (Goossens and van Opstal, 2012; Leigh and Zee, 2015; Waitzman et al., 1988). Our results partly confirmed the animal model but also showed some important differences for human freezing.

The fear potentiated startle – a protective brain-stem reflex directly mediated by activation of the central amygdala in animals and humans (Davis, 2006; Hamm, 2015; Kuhn et al., 2019; Weike et al., 2005) – was significantly reduced, when saccadic eye movements were executed compared to smooth eye pursuits during extinction learning. Accordingly, fear bradycardia – a profound cardiac deceleration in the face of inevitable threat driven by amygdala projections (Hermans et al., 2013; Roelofs, 2017; Szeska et al., 2021) – was also attenuated in the saccadic eye movement condition during extinction. The inhibitory effects were, however, stronger for fear potentiated startle than for fear bradycardia, presumably due to its stronger link to subcortical circuits (Davis, 2006; Kuhn et al., 2019; Moratti and Keil, 2005). In fact, higher saccadic accuracy and range – linked to broader and more enduring activity in the SC (Goossens and van Opstal, 2012; Leigh and Zee, 2015; Waitzman et al., 1988) - was positively correlated with stronger startle inhibition. However, as saccadic

accuracy and range were assessed by the EOG, such correlation may only be considered preliminary and future eye tracking research may provide more thorough insights into the relationship between saccadic eye movements and low-level defensive responding. In contrast to low-level defensive responding, the type of eye movement manipulation had no impact on cortically-mediated cognitive risk assessments, which are closely related to reported fear during exposure therapy (Hollandt et al., 2020; LeDoux, 1995). Eye movement manipulation during exposure may therefore primarily inhibit subcortically-mediated indicators of fear activation during processing of the fear-eliciting stimuli.

Thus, our findings indicate that the suggested inhibitory SC-BLA pathway (Baek et al., 2019) may also exist in humans and that saccadic eye movements may hence lead to a more efficient attenuation of defensive responding compared to smooth eye pursuits, which are – despite initial instructions - commonly employed in EMDR practice (Shapiro, 1989; Stickgold, 2002). Yet, as saccadic eye movements have also been associated with stronger activation in the frontal eye fields relative to pursuits (O'Driscoll et al., 1998), and as frontal eye field activation has been found concomitantly with stronger coupling between the amygdala and prefrontal centers of emotion regulation (de Voogd et al., 2018), contributions of the prefrontal cortex to the inhibitory effects may not be ruled out. It is therefore possible that the stronger link of saccadic eye movements to the many beneficial effects of EMDR, including increased episodic memory retrieval (Christman et al., 2003; Landin-Romero et al., 2018), is driven by both mechanisms of action.

Contrary to the predictions from the animal model (Baek et al., 2019), we did not find a facilitation of extinction recall as a result of saccadic eye movements during extinction learning on day 3. Rather than sustaining defensive response inhibition, subcortically-mediated defensive responding was actually renewed in all participants as soon as eye movements were no longer manipulated and, for startle reflex potentiation, even reached levels of initial fear conditioning, as if an extinction training has never taken place.

Interestingly, successful recall of extinction memory was observed only for the expectancy ratings.

In contrast to rodent research, however, we tested the extinction recall in a context similar to the context of fear acquisition, which has shown to cause strong renewal of fear (Bouton, 2002), indicating that defensive response inhibition by eye movements does not withstand fear activation in previously threatening contexts. On the other hand, we manipulated eye movements *during* exposure towards a previously threat-predicting tone cue (i.e., during fear activation), which may have led to interference with the consolidation of extinction memory. Previous human research, indicating long-term inhibitory effects of eye movement manipulation, manipulated eye movements *after* fear activation (de Voogd et al., 2018) and may, thus, rather have disrupted the reconsolidation of fear memory, shown to persistently inhibit defensive responding (Kindt et al., 2009; Soeter and Kindt, 2010).

In fact, our data indicate that the manipulation of eye movements during extinction may be rather viewed as visual distraction, impeding the processing of threat-related information (see Roelofs, 2017), and hence may act as an avoidance strategy, which decreases defensive responses when executed (Lovibond et al., 2009; Vervliet and Indekeu, 2015). Importantly, such strategy may also prevent the consolidation and recall of extinction memory and, in turn, increase the probability of return of fear when no longer executed (Lovibond et al., 2009; Vervliet and Indekeu, 2015). On a neural level, cortical elaboration specifically involving the ventromedial prefrontal cortex has found to be required for such persistent extinction of fear, particularly for consolidation and recall of extinction memory (Milad and Quirk, 2012). Thus, the manipulation of eye movements may have resulted in a partial bypass of such cortical elaboration. Additional therapeutical or neuroscience-based strategies might therefore be necessary to achieve long-term fear attenuation, if eye movements are manipulated during exposure (Craske et al., 2014; Szeska et al., 2020).

Taken together, our findings suggest that the inhibitory pathway between the superior colliculus and the basolateral amygdala, by which visual attentional modulation may facilitate the extinction of defensive responding, might also exist in humans and, thus, act as one mechanism of action of EMDR. Nevertheless, any manipulation of eye movements during extinction may prevent a successful consolidation and, thus, recall of extinction memory. Increasing SC activity might, therefore, indeed boost exposure therapy in EMDR, but may come at the expense of sustained fear attenuation.

Acknowledgements

We thank Isabella Langen and Felicitas Güttner for their assistance during data assessment. Christoph Szeska was supported by the state excellence initiative MV “Digitization in Research”, funded by the European Union (ESF).

Conflict of Interest

The authors declare no competing interests.

References

Amano, T., Duvarci, S., Popa, D., Paré, D., 2011. The fear circuit revisited: contributions of the basal amygdala nuclei to conditioned fear. J. Neurosci. 31, 15481–9. https://doi.org/10.1523/JNEUROSCI.3410-11.2011

Baek, J., Lee, S., Cho, T., Kim, S.-W., Kim, M., Yoon, Y., Kim, K.K., Byun, J., Kim, S.J., Jeong, J., Shin, H.-S., 2019. Neural circuits underlying a psychotherapeutic regimen for fear disorders. Nature 566, 339–343. https://doi.org/10.1038/s41586-019-0931-y

Blumenthal, T.D., Cuthbert, B.N., Filion, D.L., Hackley, S., Lipp, O. V., Van Boxtel, A., 2005. Committee report: Guidelines for human startle eyeblink electromyographic studies. Psychophysiology 42, 1–15. https://doi.org/10.1111/j.1469-8986.2005.00271.x

Bouton, M.E., 2002. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. Biol. Psychiatry 52, 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Bradley, R., Greene, J., Russ, E., Dutra, L., Westen, D., 2005. A Multidimensional Meta-Analysis of Psychotherapy for PTSD. Am. J. Psychiatry 162, 214–227. https://doi.org/10.1176/appi.ajp.162.2.214

Buhr, K., Dugas, M.J., 2002. The intolerance of uncertainty scale: Psychometric properties of the English version. Behav. Res. Ther. 40, 931–945. https://doi.org/10.1016/S0005-7967(01)00092-4

Christman, S.D., Garvey, K.J., Propper, R.E., Phaneuf, K.A., 2003. Bilateral eye movements enhance the retrieval of episodic memories. Neuropsychology 17, 221–229. https://doi.org/10.1037/0894-4105.17.2.221

Craske, M.G., Treanor, M., Conway, C.C., Zbozinek, T., Vervliet, B., 2014. Maximizing exposure therapy: An inhibitory learning approach. Behav. Res. Ther. 58, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Davis, M., 2006. Neural systems involved in fear and anxiety measured with fear-potentiated startle. Am. Psychol. 61, 741–756. https://doi.org/10.1037/0003-066X.61.8.741

Davis, M., Walker, D.L., Miles, L., Grillon, C., 2010. Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. Neuropsychopharmacology 35, 105–35. https://doi.org/10.1038/npp.2009.109

de Voogd, L.D., Kanen, J.W., Neville, D.A., Roelofs, K., Fernández, G., Hermans, E.J., 2018. Eye-movement intervention enhances extinction via amygdala deactivation. J. Neurosci. 38, 8694–8706. https://doi.org/10.1523/JNEUROSCI.0703-18.2018

Duricki, D.A., Soleman, S., Moon, L.D.F., 2016. Analysis of longitudinal data from animals with missing values using SPSS. Nat. Protoc. 11, 1112–1129. https://doi.org/10.1038/nprot.2016.048

Gandhi, N.J., Keller, E.L., 1999. Comparison of saccades perturbed by stimulation of the rostral superior colliculus, the caudal superior colliculus, and the omnipause neuron region. J. Neurophysiol. 82, 3236–3253. https://doi.org/10.1152/jn.1999.82.6.3236

Globisch, J., Hamm, A.O., Schneider, R., Vaitl, D., 1993. A computer program for scoring reflex eyeblink and electrodermal responses written in PASCAL. Psychophysiology 30, 1, 30.

Goossens, H.H.L.M., van Opstal, A.J., 2012. Optimal control of saccades by spatial-temporal activity patterns in the monkey superior colliculus. PLoS Comput. Biol. 8. https://doi.org/10.1371/journal.pcbi.1002508

Graham, F.K., 1978. Constraints on Measuring Heart Rate and Period Sequentially Through Real and Cardiac Time. Psychophysiology 15, 492–495. https://doi.org/10.1111/j.1469-8986.1978.tb01422.x

Haaker, J., Maren, S., Andreatta, M., Merz, C.J., Richter, J., Richter, S.H., Meir Drexler, S., Lange, M.D., Jüngling, K., Nees, F., Seidenbecher, T., Fullana, M.A., Wotjak, C.T., Lonsdorf, T.B., 2019. Making translation work: Harmonizing cross-species methodology in the behavioural neuroscience of Pavlovian fear conditioning. Neurosci. Biobehav. Rev. 107, 329–345. https://doi.org/10.1016/j.neubiorev.2019.09.020

Hamm, A.O., 2020. Fear, anxiety, and their disorders from the perspective of psychophysiology. Psychophysiology 57, 1–14. https://doi.org/10.1111/psyp.13474

Hamm, A.O., 2015. Fear-Potentiated Startle, in: International Encyclopedia of the Social & Behavioral Sciences. Elsevier, pp. 860–867. https://doi.org/10.1016/B978-0-08-097086-8.55023-5

Hermans, D., Craske, M.G., Mineka, S., Lovibond, P.F., 2006. Extinction in Human Fear Conditioning. Biol. Psychiatry 60, 361–368. https://doi.org/10.1016/j.biopsych.2005.10.006

Hermans, E.J., Henckens, M.J.A.G., Roelofs, K., Fernández, G., 2013. Fear bradycardia and activation of the human periaqueductal grey. Neuroimage 66, 278–287. https://doi.org/10.1016/j.neuroimage.2012.10.063

Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., Lüthi, A., 2008. Switching on and off fear by distinct neuronal circuits. Nature 454, 600–606. https://doi.org/10.1038/nature07166

Hollandt, M., Wroblewski, A., Yang, Y., Ridderbusch, I.C., Kircher, T., Hamm, A.O., Straube, B., Richter, J., 2020. Facilitating translational science in anxiety disorders by adjusting extinction training in the laboratory to exposure-based therapy procedures. Transl. Psychiatry 10, 110. https://doi.org/10.1038/s41398-020-0786-x

Hoyer, J., Gloster, A.T., 2013. Psychologische Flexibilität messen: Der Fragebogen zu Akzeptanz und Handeln II. Verhaltenstherapie 23, 42–44. https://doi.org/10.1159/000347040

Kapp, B.S., Frysinger, R.C., Gallagher, M., Haselton, J.R., 1979. Amygdala central nucleus lesions: Effect on heart rate conditioning in the rabbit. Physiol. Behav. 23, 1109–1117. https://doi.org/10.1016/0031-9384(79)90304-4

Kindt, M., Soeter, M., Vervliet, B., 2009. Beyond extinction: erasing human fear responses and preventing the return of fear. Nat. Neurosci. 12, 256–258. https://doi.org/10.1038/nn.2271

Krauzlis, R.J., 2004. Recasting the Smooth Pursuit Eye Movement System. J. Neurophysiol. 91, 591–603. https://doi.org/10.1152/jn.00801.2003

Krauzlis, R.J., Dill, N., Fowler, G.A., 2012. Dissociation of pursuit target selection from saccade execution. Vision Res. 74, 72–79. https://doi.org/10.1016/j.visres.2012.09.006

Kuhn, M., Wendt, J., Sjouwerman, R., Büchel, C., Hamm, A., Lonsdorf, T.B., 2019. The Neurofunctional Basis of Affective Startle Modulation in Humans: Evidence From Combined Facial Electromyography and Functional Magnetic Resonance Imaging. Biol. Psychiatry 1–11. https://doi.org/10.1016/j.biopsych.2019.07.028

Landin-Romero, R., Moreno-Alcazar, A., Pagani, M., Amann, B.L., 2018. How does eye movement desensitization and reprocessing therapy work? A systematic review on suggested mechanisms of action. Front. Psychol. 9:1395. https://doi.org/10.3389/fpsyg.2018.01395

Leaton, R.N., Borszcz, G.S., 1985. Potentiated Startle. Its Relation to Freezing and Shock Intensity in Rats. J. Exp. Psychol. Anim. Behav. Process. 11, 421–428. https://doi.org/10.1037/0097-7403.11.3.421

LeDoux, J.E., 1995. Emotion: Clues from the Brain. Annu. Rev. Psychol. 46, 209–235. https://doi.org/10.1146/annurev.psych.46.1.209

LeDoux, J.E., Pine, D.S., 2016. Using neuroscience to help understand fear and anxiety: A two-system framework. Am. J. Psychiatry 173, 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Leigh, J.R., Zee, D.S., 2015. The neurology of eye movements, 5. ed. Oxford University Press, New York, New York.

Lovibond, P.F., Mitchell, C.J., Minard, E., Brady, A., Menzies, R.G., 2009. Safety behaviours preserve threat beliefs : Protection from extinction of human fear conditioning by an avoidance response. Behav. Res. Ther. 47, 716–720. https://doi.org/10.1016/j.brat.2009.04.013

Milad, M.R., Quirk, G.J., 2012. Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. Annu. Rev. Psychol. 63, 129–151. https://doi.org/10.1146/annurev.psych.121208.131631

Mobbs, D., Headley, D.B., Ding, W., Dayan, P., 2020. Space, Time, and Fear: Survival Computations along Defensive Circuits. Trends Cogn. Sci. 24, 228–241. https://doi.org/10.1016/j.tics.2019.12.016

Moratti, S., Keil, A., 2005. Cortical activation during Pavlovian fear conditioning depends on heart rate response patterns: An MEG study. Cogn. Brain Res. 25, 459–471. https://doi.org/10.1016/j.cogbrainres.2005.07.006

Müller, J.R., Philiastides, M.G., Newsome, W.T., 2005. Microstimulation of the superior colliculus focuses attention without moving the eyes. Proc. Natl. Acad. Sci. U. S. A. 102, 524–529. https://doi.org/10.1073/pnas.0408311101

O'Driscoll, G.A., Strakowski, S.M., Alpert, N.M., Matthysse, S.W., Rauch, S.L., Levy, D.L., Holzman, P.S., 1998. Differences in cerebral activation during smooth pursuit and saccadic eye movements using positron-emission tomography. Biol. Psychiatry 44, 685–689. https://doi.org/10.1016/S0006-3223(98)00047-X

Roelofs, K., 2017. Freeze for action: neurobiological mechanisms in animal and human freezing. Philos. Trans. R. Soc. B Biol. Sci. 372, 20160206. https://doi.org/10.1098/rstb.2016.0206

SE Factory, 2020. EMDR Kit (Version 2.0) [Mobile application software] [WWW Document]. URL https://play.google.com/store/apps/details?id=nl.psy_zo.emdr_kit (accessed 5.20.19).

Shapiro, F., 1989. Eye movement desensitization: A new treatment for post-traumatic stress disorder. J. Behav. Ther. Exp. Psychiatry 20, 211–217. https://doi.org/10.1016/0005-7916(89)90025-6

Soeter, M., Kindt, M., 2010. Dissociating response systems: Erasing fear from memory. Neurobiol. Learn. Mem. 94, 30–41. https://doi.org/10.1016/j.nlm.2010.03.004

Spielberger, C.D., 2010. State-Trait Anxiety Inventory, in: The Corsini Encyclopedia of Psychology. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 3–4. https://doi.org/10.1002/9780470479216.corpsy0943

Stickgold, R., 2002. EMDR: A putative neurobiological mechanism of action. J. Clin. Psychol. 58, 61–75. https://doi.org/10.1002/jclp.1129

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A.O., 2021. Attentive immobility in the face of inevitable distal threat—Startle potentiation and fear bradycardia as an index of emotion and attention. Psychophysiology 1–17. https://doi.org/10.1111/psyp.13812

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A.O., 2020. Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training. Sci. Rep. 10, 1529. https://doi.org/10.1038/s41598-020-58412-w

Vervliet, B., Indekeu, E., 2015. Low-cost avoidance behaviors are resistant to fear extinction in humans. Front. Behav. Neurosci. 9, 1–12. https://doi.org/10.3389/fnbeh.2015.00351

Waitzman, D.M., Ma, T.P., Optican, L.M., Wurtz, R.H., 1988. Superior colliculus neurons provide the saccadic motor error signal. Exp. Brain Res. 72. https://doi.org/10.1007/BF00250610

Walker, P., Carrive, P., 2003. Role of ventrolateral periaqueductal gray neurons in the behavioral and cardiovascular responses to contextual conditioned fear and poststress recovery. Neuroscience 116, 897–912. https://doi.org/10.1016/S0306-4522(02)00744-3

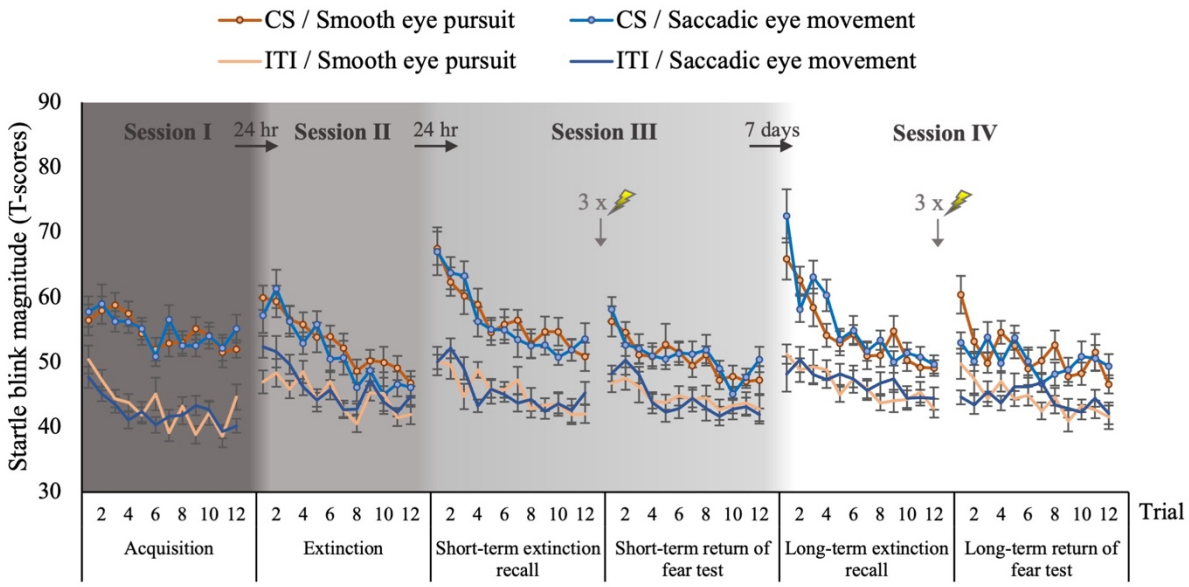
Weike, A.I., Hamm, A.O., Schupp, H.T., Runge, U., Schroeder, H.W.S., Kessler, C., 2005. Fear Conditioning following Unilateral Temporal Lobectomy: Dissociation of Conditioned Startle Potentiation and Autonomic Learning. J. Neurosci. 25, 11117–11124. https://doi.org/10.1523/JNEUROSCI.2032-05.2005

	Saccadic eye movement condition	Smooth eye pursuit condition	Statistics
N (female/male)	20 (11/9)	20 (14/6)	$X^2(1) = .960$, $p = .327$
Age (years)	21.25 (3.69)	22.50 (3.68)	$F(1,38) = 1.149$, $p = .291$, $\eta^2_p = .029$
Body-Mass-Index (kg/m^2)	21.78 (1.73)	21.75 (2.23)	$F(1,38) = .003$, $p = .956$, $\eta^2_p = .000$
Trait anxiety (STAI questionnaire sum score)	34.55 (5.79)	37.58 (6.31)	$F(1,37) = 2.436$, $p = .127$, $\eta^2_p = .062$
Psychological flexibility (FAH-II questionnaire sum score)	12.95 (5.69)	15.45 (5.84)	$F(1,38) = 1.879$, $p = .178$, $\eta^2_p = .047$
Intolerance of uncertainty (IUS-12 questionnaire mean)	2.35 (0.65)	2.55 (0.71)	$F(1,38) = .756$, $p = .390$, $\eta^2_p = .020$

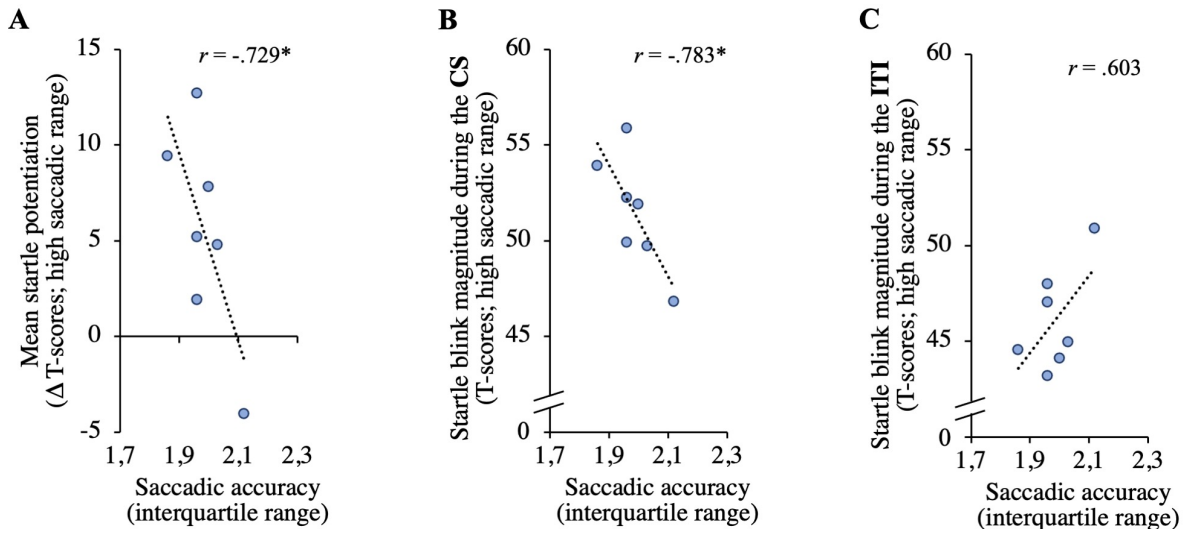
Supplementary Table 1. Means of demographics, BMI and personality trait scores for the saccadic eye movement and smooth eye pursuit condition with standard deviations in brackets.

	Acquisition (session 1)	Extinction (session 2)	Short-term extinction recall (session 3)	Short-term return of fear test (session 3)	Long-term extinction recall (session 4)	Long-term return of fear test (session 4)
Zero responses	3.7% (*M* = 0.88)	5.1% (*M* = 1.23)	6.2% (*M* = 1.48)	6.3% (*M* = 1.50)	8.2% (*M* = 1.98)	9.9% (*M* = 2.38)
Missing values	4.9% (*M* = 1.18)	3.7% (*M* = 0.88)	4.6% (*M* = 1.10)	6.2% (*M* = 1.48)	2.9% (*M* = 0.70)	5.3% (*M* = 1.28)

Supplementary Table 2. Proportion of zero responses and missing values in all scored startle responses for each experimental stage in percent, with the average number of trials scored as zero response or missing in brackets.



Supplementary Figure 1. Startle blink magnitudes during the experimental phases. Standardized blink magnitudes (T-scores) elicited during the CSs and inter-trial intervals during the acquisition (session 1), extinction (session 2), short-term extinction recall test (session 3), short-term return of fear test (session 3), long-term extinction recall test (session 4) and long-term return of fear test (session 4) for the smooth eye pursuit (orange [CS] and shaded orange [ITI] lines) and saccadic eye movement condition (blue [CS] and dark blue [ITI] lines). Error bars represent the standard error of the mean.



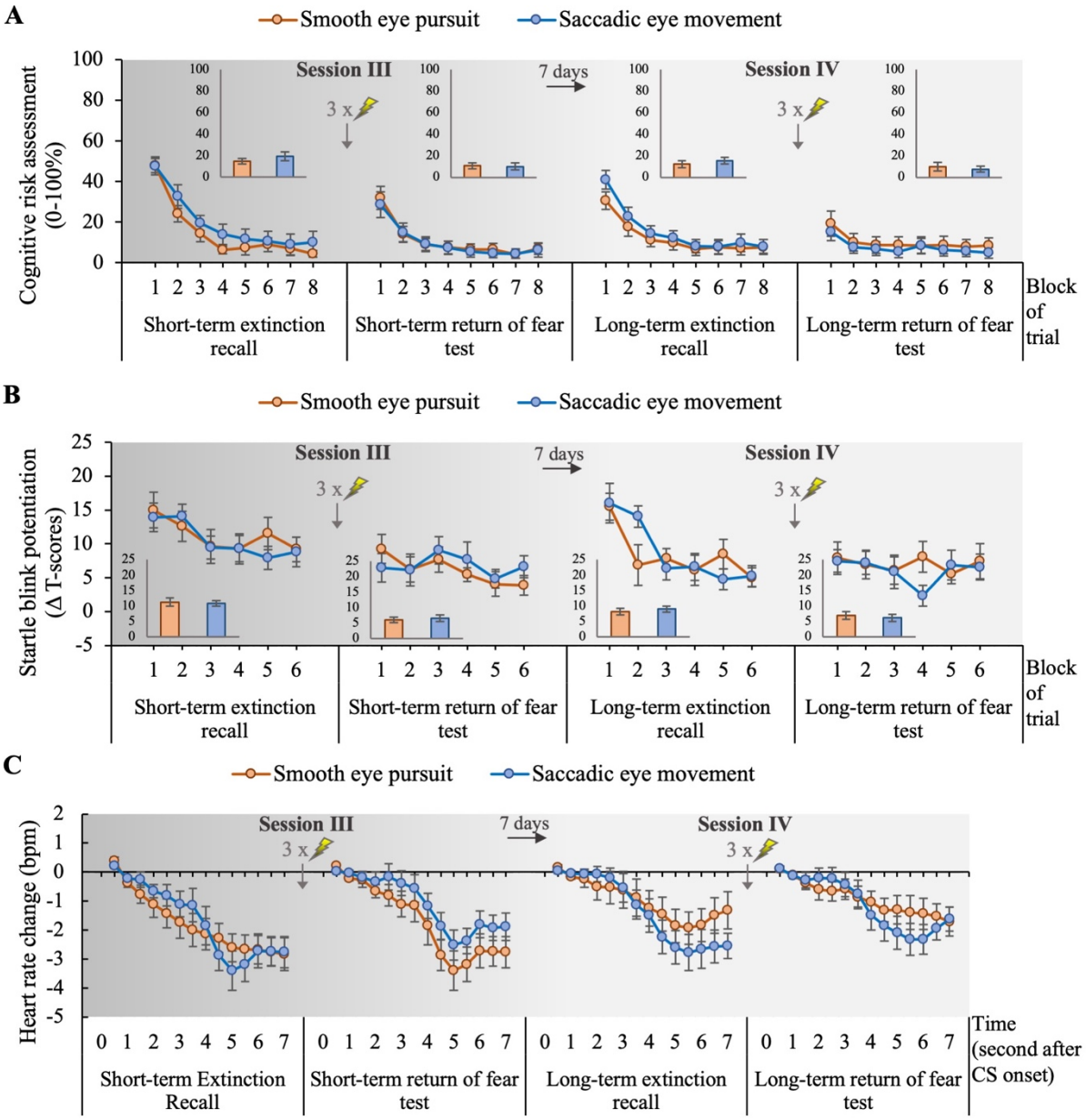
Supplementary Figure 2. The negative correlation between saccadic accuracy and amygdala-mediated behavioral defensive responding is most pronounced at high saccadic range

(A)**,** Scatter plot of mean standardized (T-scores) potentiation of the startle response (CS vs. ITI) during extinction as a function of saccadic accuracy in participants of the saccadic eye movement condition who executed saccades with high range (upper tertile).

(B)**,** Scatter plot of standardized (T-scores) CS startle responses during extinction as a function of saccadic accuracy in participants of the saccadic eye movement condition who executed saccades with high range (upper tertile).

(C)**,** Scatter plot of standardized (T-scores) ITI startle responses during extinction as a function of saccadic accuracy in participants of the saccadic eye movement condition who executed saccades with high range (upper tertile).

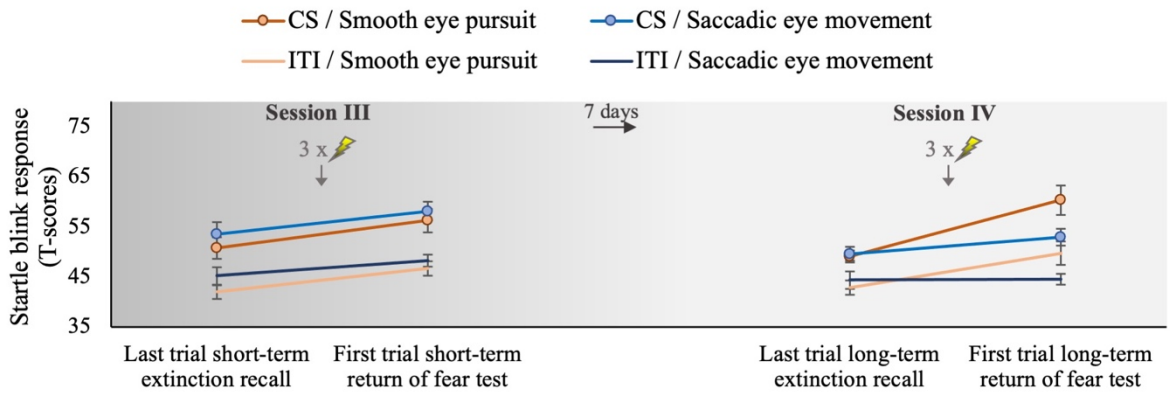
For all graphs: Orange graphs indicate data of the smooth eye pursuit condition, while blue graphs indicate data of the saccadic eye movement condition. Asterisks indicate statistical significance of between-group analyses or correlations, with * for $P < .05$, ** for $P < .01$, *** for $P < .001$.



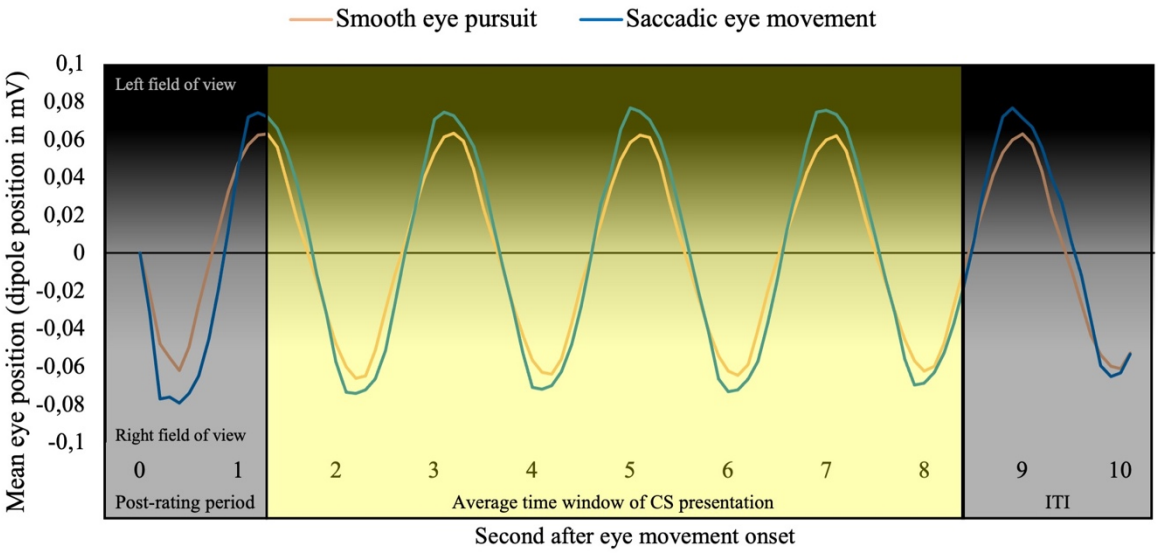
Supplementary Figure 3. Saccadic eye movements do not affect short- and long-term recall of extinction memory or return of fear

(A)**,** Cognitive risk assessment during the CS presentation throughout the short-term extinction recall test (session 3), short-term return of fear test (session 3), long-term extinction recall test (session 4) and long-term return of fear test (session 4), depicted in blocks of two averaged trials. Integrated bar charts represent mean cognitive risk assessment during the CS presentation across all trials during the respective experimental phases.

(B)**,** Standardized (T-scores) fear potentiation of the startle response during the short-term extinction recall test (session 3), short-term return of fear test (session 3), long-term extinction recall test (session 4) and long-term return of fear test (session 4), depicted in blocks of two averaged probed trials. Integrated bar charts represent mean standardized fear potentiated startle across all probed trials during the respective experimental phases.
(C)**,** Heart rate change after CS onset during the short-term extinction recall test (session 3), short-term return of fear test (session 3), long-term extinction recall test (session 4) and long-term return of fear test (session 4), averaged across all trials, depicted in half-second bins.
For all graphs: Orange lines and bars represent data of the smooth eye pursuit condition, blue lines and bars represent data of the saccadic eye movement condition. Error bars represent the standard error of the mean.



Supplementary Figure 4. Saccadic eye movements do not inhibit short- but long-term re-sensitization of defensive responding after re-experiencing original aversive events. Standardized startle responses (T-scores) elicited during the last probed trial (CS vs. ITI) of the short-term extinction recall (session 3) and first probed trial (CS vs. ITI) of the short-term return of fear test are depicted in the left part of the panel. Standardized startle responses (T-scores) during the last probed trial (CS vs. ITI) of the long-term extinction recall (session 4) and first probed trial (CS vs. ITI) of the long-term return of fear test are depicted in the right part of the panel. Error bars represent the standard error of the mean. Orange graphs represent data of the smooth eye pursuit condition (orange [CS] and shaded orange [ITI] lines), and blue graphs represent data of the saccadic eye movement condition (blue [CS] and dark blue [ITI] lines)



Supplementary Figure 5. Averaged eye movements during extinction (session 2). Grand-averaged raw electrooculogram across all participants and all trials in the smooth eye pursuit (orange line) and saccadic eye movement condition (blue line) during session 2, corrected for individual different latencies in eye movement onset and drifts. The average time window of CS presentation is highlighted.

Appendix D: Publikationen / Wissenschaftliche Leistungen

Publikationen

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A. O. (2021). Attentive immobility in the face of inevitable distal threat – fear bradycardia and startle potentiation as an index of emotion and attention. *Psychophysiology, 58*(6), 1-17. doi: 10.1111/psyp.13812

Szeska, C., Richter, J., Wendt, J., Weymar, M., Hamm, A. O. (2020). Promoting long-term inhibition of human fear responses by non-invasive transcutaneous vagus nerve stimulation during extinction training. *Scientific reports, 10*, 1529. doi: 10.1038/s41598-020-58412-w

Vorträge

Innovative Ansätze zur Optimierung (psycho-)therapeutischer Verfahren mit Hilfe neuer Technologien (16. Nationale Branchenkonferenz Gesundheitswirtschaft; Rostock, 03.06.2021)

Non-invasive Vagusnervstimulation als potentielles Mittel der Begünstigung von Furchtextinktionslernen (11. Workshopkongress Klinische Psychologie und Psychotherapie / 37. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie der DGPs; Erlangen, 31.05.2019)

Non-invasive vagus nerve stimulation as a potential booster of fear extinction learning (9th World Congress of Behavioural and Cognitive Therapies; Berlin, 20.07.2019)

Poster

Neuroenhancement of fear reduction by stimulation of the brain's inhibitory pathways [11th European Meeting on Human Fear Conditioning (EMHFC); Würzburg, 07.05.2019]

About the discordance between physiological and subjective indicators of defensive reactivity in patients with anxiety and depressive disorders [10th European Meeting on Human Fear Conditioning (EMHFC); Hensol Castle (WAL), 17.04.2018]

About the discordance between physiological and subjective indicators of defensive reactivity in patients with anxiety and depressive disorders [Summer School of Emotional Learning and Memory in Health and Psychopathology; Leuven (BEL), 11.09.2017]

Danksagung

Eigenartigerweise gibt es kaum ein schöneres Gefühl als jenes, das entsteht, wenn man nach sorgfältiger Überprüfung feststellt, dass die eigene Theorie über die Wirklichkeit zumindest nicht gänzlich falsch ist. Ich möchte an dieser Stelle allen danken, die es ermöglicht haben, dass dieses Gefühl in den letzten fünf Jahren der Promotion fester Bestandteil meines Lebens war.

Mein tiefer Dank gilt hierbei meinem Betreuer Herrn Prof. Dr. Alfons Hamm. Für sein Vertrauen in mich und diese Arbeit. Für seine unentwegte fachliche Unterstützung. Für das fortlaufende Erinnern daran, dass wir bisher nur die Spitze des Eisbergs erkundet haben und die Expedition noch einiges bereit hält. Vor allem aber für die inspirierenden Gespräche und Diskussionen darüber, wie diese Welt wohl funktionieren mag.

Ganz besonders möchte ich mich zudem bei Herrn Dr. Jan Richter bedanken. Sein Einsatz war nicht weniger als der Initialfunke dieser Reise und seine bedingungslose Hilfe war mir in vielen Situationen ein Leuchtturm, der den richtigen Weg aufzeigte.

Herrn Dr. Heino Mohrmann und Sylvia Scholz danke ich nicht nur für die allzeit tatkräftige technische und organisatorische Unterstützung, sondern auch für die familiäre Arbeitsatmosphäre, ohne die das Arbeiten nicht einmal halb so schön gewesen wäre.

Ein großer Dank gilt darüber hinaus allen Kollegen und Freunden in Greifswald und in der Heimat. Danke, für jedes Mittagessen und für jeden Abend an wechselnden Küchentischen, am Hafen oder am Pavillon. Danke, für eure Unterstützung und dafür, dass ihr mich regelmäßig daran erinnert, zu entspannen.

Danke an Dich, lieber Leser oder liebe Leserin – dafür, dass Du diese Arbeit liest und so an diesen verrückten letzten fünf Jahren teilhast.

Zuletzt aber gilt mein größter Dank meiner Familie. Danke für eure bedingungslose Rückendeckung. Danke für euren beruhigenden Zuspruch. Danke für die Zuversicht.
Danke, dass ihr da seid.