

**Alma Mater Studiorum – Università di Bologna**

**DOTTORATO DI RICERCA IN**

**PSICOLOGIA**

Ciclo: XXXII

Settore Concorsuale: 11/E1

Settore Scientifico Disciplinare: M-PSI/02

**Functional and neural mechanisms of human fear conditioning:  
studies in healthy and brain-damaged individuals**

**Presentata da:**

Simone Battaglia

**Coordinatore Dottorato**

Prof. Monica Rubini

**Supervisore**

Prof. Giuseppe di Pellegrino

**Esame finale anno 2020**



# SUMMARY

<b>ABSTRACT .....</b>	<b>6</b>
<b>GENERAL INTRODUCTION .....</b>	<b>8</b>
<b>FEAR CONDITIONING .....</b>	<b>9</b>
ACQUISITION OF FEAR.....	10
EXTINCTION AND RETURN OF FEAR.....	12
SPONTANEOUS RECOVERY .....	14
CONTEXTUAL RENEWAL .....	15
REINSTATEMENT.....	17
MEMORY RECONSOLIDATION .....	19
<b>NEURAL SUBSTRATE OF FEAR CONDITIONING.....</b>	<b>22</b>
AMYGDALA.....	23
HIPPOCAMPUS.....	25
PREFRONTAL CORTEX .....	26
PSYCOPHYSIOLOGICAL MEASUREMENTS .....	31
<b>THESIS OUTLINE .....</b>	<b>34</b>
<b>STUDY 1: CONTEXT-DEPENDENT EXTINCTION OF THREAT MEMORIES: INFLUENCES OF HEALTHY AGING .....</b>	<b>38</b>
ABSTRACT.....	39
INTRODUCTION.....	40
METHODS.....	44
RESULTS.....	51

DISCUSSION .....	61
<b>STUDY 2: LESION TO THE VENTROMEDIAL PREFRONTAL CORTEX IMPAIRS THE ACQUISITION OF FEAR CONDITIONING IN HUMANS .....</b>	<b>69</b>
ABSTRACT .....	70
INTRODUCTION.....	71
METHODS.....	73
RESULTS.....	81
DISCUSSION .....	86
<b>STUDY 3: STATE-DEPENDENT TMS OVER PREFRONTAL CORTEX DISRUPTS MEMORY RECONSOLIDATION AND PREVENTS THE RETURN OF FEAR.....</b>	<b>91</b>
ABSTRACT .....	92
INTRODUCTION.....	93
METHODS.....	96
RESULTS.....	104
DISCUSSION .....	109
<b>STUDY 4: CHARACTERIZING CARDIAC AUTONOMIC SIGNATURE TO FEAR CONDITIONED STIMULI .....</b>	<b>114</b>
ABSTRACT .....	115
INTRODUCTION.....	116
METHODS.....	119
RESULTS.....	127
DISCUSSION .....	136

<b>GENERAL CONCLUSION</b> .....	<b>141</b>
<b>FINAL REMARKS</b> .....	<b>155</b>
<b>REFERENCES</b> .....	<b>156</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>199</b>

## **ABSTRACT**

Fear conditioning represents the learning process by which a stimulus, after repeated pairing with an aversive event, comes to evoke fear and becomes intrinsically aversive. This learning is essential to organisms throughout the animal kingdom and represents one of the most successful laboratory paradigms to reveal the psychological processes that govern the expression of emotional memory and explore its neurobiological underpinnings.

Although a large amount of research has been conducted on the behavioural or neural correlates of fear conditioning, some key questions remain unanswered. Accordingly, this thesis aims to respond to some unsolved theoretic and methodological issues, thus furthering our understanding of the neurofunctional basis of human fear conditioning both in healthy and brain-damaged individuals. Specifically, in this thesis, behavioural, psychophysiological, lesion and non-invasive brain stimulation studies were reported. Study 1 examined the influence of normal aging on context-dependent recall of extinction of fear conditioned stimulus. Results showed that older adults were less able to use contextual information to recall extinction memory and modulate the expression of defensive responses to threat in a context-dependent manner, despite their preserved ability to acquire and extinguish a conditioned response. This deficit may be linked to age-related changes in the neural structures underpinning context-dependent behaviour such as hippocampus and prefrontal cortex (PFC). Study 2 aimed to determine the causal role of the ventromedial PFC (vmPFC) in the acquisition of fear conditioning by systematically testing the effect of bilateral vmPFC brain-lesion. Results suggest that vmPFC is a crucial brain structure for fear conditioning in humans, impairing the ability to shape defensive anticipatory responses to the fear conditioned stimulus, but nevertheless sparing the ability to learn explicit contingencies regarding the conditioning. Study 3 aimed to interfere with the reconsolidation process of fear memory by the means of non-invasive brain stimulation (i.e. TMS) disrupting PFC neural activity. Results showed that interfering with

activity in both left and right PFC prevents the recall of fear, in contrast to control groups. These results suggest that non-invasive stimulation of PFC may attenuate the expression of learned fear, arguing in favour of a critical role of the PFC in the neural network underlying fear memory reconsolidation in humans. Finally, Study 4 aimed to investigate whether the parasympathetic – vagal – modulation of heart rate might reflect the anticipation of fearful, as compared to neutral, events during classical fear conditioning paradigm. Results indicate that fear conditioned stimuli elicit a strong and selective vagal response, supporting bradycardia during the acquisition of aversive conditioning.

Evidence reported in this PhD thesis might therefore provide key insights and deeper understanding of critical issues concerning the neurofunctional mechanisms underlying the acquisition, the extinction and the reconsolidation of fear memories in humans.

## **GENERAL INTRODUCTION**

Learning the relationships between aversive events and the environmental stimuli that predict such events is essential to the survival of organisms throughout the animal kingdom (Maren, 2001). Nearly a century ago, the advent of fear conditioning research in humans was marked by the historical experiment with little Albert, who was made to fear a rat by pairing it with a loud startling noise (Watson & Rayner, 1920). The experimental procedure of fear conditioning model derived from the more general conditioning model developed by Ivan Pavlov (1927), who initially studied appetitive conditioning processes.

Pavlovian fear conditioning is amongst the most successful laboratory paradigms in the history of experimental psychology. It was modelled after the appetitive conditioning procedure introduced by Pavlov in animals (1903/1928). The effect comes from the repeated pairing of an initially neutral stimulus (i.e. a tone) with a stimulus that is intrinsically aversive (i.e. a shock pulse). As a result, stimulus presentation typically comes to elicit a variety of psychophysiological reactions revealing fear (Lonsdorf et al., 2017). This simple procedure is an important paradigm for behavioural and cognitive sciences. Remarkably vast and deep understanding of fear itself and its related processes, such as learning mechanisms, memorization and retrieval, is the result of decades of scientific research conducted through the basic fear conditioning paradigm in both animal and humans (Beckers et al., 2012). This experimental paradigm has proven a tool of great use, not only in uncovering the psychological processes that govern the genesis and expression of fear and the functioning of emotional and general memory, but also in exploring the neurobiological underpinnings of emotion and learning (Craske et al., 2006; Fanselow and Poulos, 2005; Hartley and Phelps, 2010).

Alterations in fear conditioning regulation mechanisms may play a role in the development of anxiety related disorders such as panic disorder, specific phobias and PTSD (Rosen & Schulkin 1998, Wolpe 1981). Moreover, these altered mechanism are considered



critical in the pathogenesis and maintenance of pathological anxiety (Duits et al., 2015; Lissek et al., 2005a,b). Therefore, a complete understanding of the psychological and biological processes accountable for such disorders is of major importance, fear conditioning paradigm seems the privileged way to achieve this goal.

Given these premises, it is not surprising that in the last few decades we have seen an incredible surge in interest in the neurobiology of fear conditioning. Neural circuits underlying fear conditioning have been mapped, synaptic plasticity in these circuits has been identified, and biochemical and genetic manipulations are beginning to disentangle the molecular machinery responsible for the storage of fear memories (Maren, 2001; Kim & Jung, 2006).

## **FEAR CONDITIONING**

Fear conditioning represents the process by which a stimulus comes to evoke fear following its repeated pairing with an aversive event and becomes intrinsically aversive (Maren, 2001). Two main types of conditioning designs can be distinguished, which differ in the temporal relationship between stimulus and aversive event, hence in the *temporal contiguity*. In trace conditioning, a time interval ranging from for example 500 milliseconds to 10 seconds separates the presentation of the stimulus from the administration of aversive stimulus (Cheng et al., 2008; Knight et al., 2004). The expression ‘trace conditioning’ stems from the idea that a memory trace needs to bridge the gap between the stimulus and the aversive outcome to form an association, therefore working-memory processes are more strongly involved in trace conditioning. In contrast, in delayed conditioning the stimulus overlaps or is immediately followed by the aversive outcome (Sehlmeyer et al., 2009).

## ACQUISITION OF FEAR

A fear conditioning experiment commonly consists in a series of different experimental phases: habituation, fear acquisition, extinction and usually some specific protocols aimed to investigate return of fear. Habituation or familiarization phase in human fear conditioning precedes all the experimental manipulation that will be adopted. Habituation may have various roles: (1) it establishes a baseline of responses, which allows the determination and correction for possible pre-conditioning differences in each participant, (2) allowing to assess a decline in responding over the first number of trials (i.e. orienting responses), (3) ensure that participants understood the task, it may be useful to include a brief training phase for rating procedures (Lonsdorf et al., 2017). Once the habituation has been occurred acquisition may take place.

Acquisition of conditioned fear is achieved by presenting a neutral stimulus (NS, i.e. a tone) paired with an aversive event (US, i.e. electro-tactile stimulation), a procedure referred to as *fear acquisition*. As a result of this pairing, fear learning takes place, manifesting in the development of conditioned response (CR) to the NS that become a conditioned stimulus (CS+). Although CS+ response reflects the acquisition of the conditioned fear, it might valuable to use an additional CS as a control stimulus. This latter stimulus is also presented during the acquisition, but not paired with the US. To indicate that it was not paired, CS- is stated as opposed to the CS+, the conditioned stimulus that was actually paired with the US.

Conditioned responses consist of fear, orienting and defensive responses generated by the subject. Generally, the strength of the CR is also affected by the extrinsic (conditioned stimulus, i.e. tone) or intrinsic (natural characteristics of stimulus, i.e. spider or snake) salience of the CSs. The majority of fear conditioning studies rely on discrete exteroceptive stimuli, mostly visual CSs, such as pictures of differently coloured images, geometric shapes (Meulders et al., 2012; Vervliet et al., 2010a,b), human faces, or animals (Hermans et al., 2002), contexts or a combination of them (Milad et al., 2009). In addition, auditory, tactile, olfactory, and taste

CSs have been employed. Recently, proprioceptive CSs such as joystick arm movements (Meulders et al., 2011, 2013, 2015; Meulders and Vlaeyen, 2013) and interoceptive CSs (De Peuter et al., 2011) such as respiratory loads (Pappens et al., 2013), oesophageal balloon distension (Zaman et al., 2015), and inhalation of CO<sub>2</sub> enriched air (Acheson et al., 2007) have been applied in pain-related fear conditioning research.

US intensity, in particular in the case of electro-tactile stimulation, is often determined individually by assessing the participant's subjective evaluation in a procedure prior to fear acquisition. US salience (intensity) as well as CS salience, has an impact on the speed and duration of fear acquisition processes. While this distinction is clear for many US types (i.e. electro-tactile stimulation or aversive loud sound), emotional pictures are typically not commonly menacing and might lead to individual differences not related to the conditioning experimentation (Lonsdorf et al., 2017).

During fear acquisition, a CS can be paired with the US on every single trial (continuous reinforcement) or in a smaller number of trials only (partial reinforcement). *Reinforcement rate* refers to the probability of US occurrence in the presence of the CS. Although both continuous and partial reinforcement generally lead to fear acquisition, partial reinforcement rates is preferred in human (Dunsmoor et al., 2007; Haselgrove et al., 2004). Partial reinforcement protocols reduce response frequency (Flora and Pavlik, 1990; Huang et al., 1992) and CR amplitudes (Dunsmoor et al., 2007; Leonard, 1975) but produce more robust learning over the time (Gershman et al., 2015). It is important to highlight that a recent study comparing different reinforcement schedules concluded that partial followed by continuous reinforcement yields the strongest CRs during fear acquisition training (Grady et al., 2016).

## **EXTINCTION AND RETURN OF FEAR**

Behavioural flexibility has the same importance as acquisition of fear conditioning, as behaviour should no longer be guided by a stimulus that has lost its predictive value with respect to the related consequence.

*Extinction learning* is a well-known behavioural phenomenon that allows the organism to adapt its behaviour to a changing environment (Bouton et al., 2004). Extinction provides the leading theoretical framework and experimental model to describe how learned behaviours is reduced through the absence of anticipated and expected reinforcement (Dunsmoor et al., 2015). Indeed, extinction refers to the decrement in conditioned fear responses that occurs with repeated presentation of a conditioned fear stimulus that is unreinforced. In the past, extinction was regarded as a process of unlearning (CS+ and US association erased), but severe evidence suggest that extinction does not destroy the original learning, but instead generates new learning (for a review see Bouton, 2004). One of the first hypothesis, surprisingly common in models of learning (Rescorla and Wagner 1972; McClelland & Rumelhart 1985; McCloskey & Cohen 1989), was that extinction involves the destruction of what was originally learned, extinction was considered a form of learning erasing. However, much of the original learning survives extinction (Rescorla 2001; Bouton 2002; Myers and Davis 2002; Delamater 2004) and this hypothesis reduces its recognition over the time.

Return of fear (ROF) or extinguished behaviour is common following the passage of time (*spontaneous recovery*), when extinguished cues are encountered outside the extinction context (*contextual renewal*) and presentation of the unconditioned stimulus in absence of the conditioned stimulus (*reinstatement*; Bouton & King 1983; Rescorla & Heth 1975, Bouton, 2004; Vervliet et al., 2013). These effects provide support for the widely held view that extinction may be a new form of learning (new association CS+ and NoUS), and that conditioning and extinction memories may coexist in distinct neural circuits and be reactivated

independently based on environmental or situational factors (Milad & Quirk, 2012; Dunsmoor et al., 2015).

In laboratory experiments, when testing for return of fear, following acquisition and extinction training, the response that is measured depends on which of the two opposing and co-existing memory traces - original fear memory vs. extinction memory - is dominant. When the CR is weak or absent during testing, this is interpreted as dominance of the extinction memory trace and labelled as *extinction recall*. Conversely, when the CR is strong, dominance of the fear memory trace is assumed, which is referred to as return of fear or fear recall. Procedures that induce return of fear in the laboratory may serve as experimental models for clinical relapse, which affects a substantial percentage of patients (Craske, 1999).

Currently, theoretical views have been assumed: (1) extinction is one example of a retroactive inhibition phenomenon in which new learning inhibits old (2) extinction occurs because the omission of the US causes generalization decrement and violates the organism's expectation of the US and therefore initiates new learning (3) extinction as a context-dependent form of new inhibitory learning, and retrieval of the inhibitory memory (Bouton et al., 2004).

Like other forms of learning, the capacity to extinguish varies across the lifespan, and age-related changes in extinction reflect developmental changes in prefrontal-amygdala circuitry. For instance, studies have shown that extinction in pre-weanling rats violates the extinction (Kim & Richardson, 2010). Instead of potentiating inhibitory systems, early life extinction appears to erase fear memories from the amygdala (Kim & Richardson, 2008; Gogolla et al., 2009). During adolescence, extinction again becomes compromised, as twice the number of training trials are needed to learn extinction (Esmoris-Arranz et al., 2008). Finally, aged rats show impaired extinction learning (Kaczorowski et al., 2011). Developmental aspects of extinction learning have not yet been studied in humans, but it is known that older individuals show decreased awareness of CS-US contingencies that support conditioned

responses (LaBar et al 2004). These findings highlight the presence of windows of vulnerability with respect to extinction, but also provide a window of interest for therapeutic intervention.

Beyond interest in the basic mechanisms of learning and memory, renewed attention to extinction is due in large part to the clinical significance of extinction for the treatment of a variety of psychiatric disorders (Milad and Quirk, 2012; Vervliet et al., 2013; Lonsdorf et al., 2017). Specifically, extinction serves as the basis for exposure-based therapy, a primary treatment for anxiety disorders, addiction, trauma and stress related disorders (Powers et al., 2010). Experimental extinction is also considered within the National Institute of Mental Health's Research Domain Criteria as a scientific paradigm to provide objective neuro-behavioural measures of mental illness in the domain of Negative Affect (Dunsmoor et al., 2015). Advances in understanding of extinction across multiple fronts will translate to new, effective treatments for psychiatric conditions characterized by the inability to regulate pathological fear or anxiety (Dunsmoor et al., 2015).

## **SPONTANEOUS RECOVERY**

*Spontaneous recovery* refers to the return of CR as a function of time after successful extinction learning. According to Pavlov, evidence that the CR is preserved comes from the fact that it tends to return – spontaneously – over time. Pavlov (1927) considered spontaneous recovery to be a measure of the depth of the extinction process itself: “Extinction is measured, other conditions being equal, by the time taken for spontaneous restoration of the extinguished reflex to its original strength”.

Spontaneous recovery may be the effect that occurs when the CS is tested outside its temporal context. For example, a cue that is presented intermittently during the extinction can attenuate either spontaneous recovery if it is presented before the final test (Brooks and Bouton, 1993; Brooks 2000). Also, using a ‘gradual extinction’ – a paradigm in which some extinction

trials were reinforced with a US and the frequency of reinforced trials diminished throughout the extinction session, it has been found effective in preventing spontaneous recovery (Gershman et al., 2013). These results are consistent with the idea that rapid extinction is linked with more spontaneous recovery (Gershman and Hartley, 2015), and suggest that clinical protocol that aim to accelerate extinction might be counterproductive (Craske et al., 2008).

## **CONTEXTUAL RENEWAL**

*Context* is defined not only by spatial features, but also by temporal, interoceptive, cognitive, or social aspects of a given situation (Maren et al., 2013). Context can be considered another strong evidence for the persistence of the original fear memory (CS+ US association), in which the return of fear is specifically renewed (Dunsmoor et al., 2015). Context is theorized as a proxy for the expression of conditioning or extinction. Thus, following extinction, contextual information plays a critical role in determining whether the original fear memory or the new extinction memory controls fear expression (Bouton, 2004). Several versions of the contextual renewal effect have been studied. The most used paradigm is the ‘ABBA’, that is when the participant is conditioned in one context (‘A’), and then extinguished in a different context (‘B’); in the subsequent test phase, the extinction memory can only be expressed if the CS is presented within the extinction context (‘B’) or the conditioned response can only renewed if the CS is presented within the conditioning context (‘A’). In another paradigm’s version “ABC”, conditioning is conducted in a context A, extinction is conducted in a context B, and then testing phase is conducted in a third, “neutral-new” context (‘C’; Bouton and Bolles, 1979; Bouton, 2004). Moreover, conditioning and extinction can be both conducted in the same context ‘A’ and then the CS is tested in a second context ‘B’. Although ‘AAB’ paradigm may seems more ambiguous, there is evidence that conditioning fear is elicited (Bouton and Ricker, 1994). Classical models of fear conditioning (Rescorla and Wagner, 1972; Pearce and Hall,

1980) consider that the context is merely another CS that is presented in compound with the CS+ during the different phases of conditioning. Therefore context, enters in excitatory or inhibitory associations with the US. For example, using the ABA paradigm, context A might acquire excitatory associations with the US, and context B might acquire inhibitory associations. If this is the case, either context-association would cumulate with the CS+ and thus, produce the contextual renewal effect (Bouton et al., 2004).

Contextual renewal of conditioned fear responses appears to be supported by many varieties of contexts. For example, it has been observed in animals that when extinction is conducted within an interoceptive context, provided by benzodiazepine tranquilizers and diazepam, fear renewal it has observed only when rats were tested in the original non-drug state (Bouton et al. 1990). Also, Cunningham (1979) had reported similar results in experiment in which alcohol were administered.

The major theoretical basis of post-extinction return of fear effect is that proposed by Bouton (1993, 2004). Bouton considers extinction as a context-dependent form of new inhibitory learning, and retrieval of the inhibitory association interferes with expression of the excitatory memory. Moreover, Bouton considers that retrieval rarely survives after a context shift, so that extinction is encoded and processed to where it was learned. A key element in Bouton's theory of extinction is that new inhibitory learning makes the CS+ ambiguous because its presence signals either the presence or the absence of the US. Indeed, in a test phase, if the acquisition context is similar to the context in which extinction occurred, return of fear tends to the inhibitory 'CS+ noUS association' retrieval. If not, return of fear tends to elicit the 'CS+ US association', since this association was learned first and/or is more prominent. In other words, the extinction context acquires the ability to decrease the threshold at which the 'CS+ US association' is renewed.



Contextual renewal is particularly important in regulating the expression of responses to threat, in particular when such fear responses have been extinguished. The alteration in the mechanism underlie the ability to process contextual information, and flexibly adapt behaviour to situational changes, it may represent a fundamental mechanism that allow to survive. It has been suggested that PTSD patients fail to use flexibly contextual signals in order to regulate their behaviour (Jovanovic et al., 2012). Thus, the study of context has primary importance for the clinical domain, in translating the behavioural evidence to the clinic-based therapies in human psychopathology (Andreatta et al., 2015).

## **REINSTATEMENT**

The *reinstatement* effect is an experimental manipulation that consists in inducing return of fear following un-signaled and un-expected administration of US after successful extinction learning (Bouton and Bolles, 1979; Haaker et al., 2013; Rescorla & Heth, 1975). During extinction the “CS+ US association” is thought to be inhibited by formation of an “CS+ noUS association” that covers the original fear learning (Rescorla, 1979). Thus, after a US re-exposure, in a subsequent test phase, it restores the excitatory “CS+ US association” and consequently leading to a reinstate of the fear. Following reinstatement, the context may play a crucial role in CS discrimination and strength of the return of fear, and consequently, it is considered crucial in the theorisations of the reinstatement phenomenon (for an exhaustive review see Haaker et al., 2014). If fear acquisition and reinstatement manipulation occur in the same context, reinstatement should decrease contextual inhibition and increased attention to the CS+, leading to a reactivation of the “CS+ US association”. Contrary, if reinstatement occurred in the extinction context and the CS+ was tested in a context different from that, reinstatement should reinstate “CS+ noUS association” (Haaker et al., 2014).

This phenomenon can be explained by the fact that, fear acquisition produces an association between “CS+ US association”, because they are presented in simultaneous with US administration. Thus, US – alone – may cause a return of fear because they were encoded as part of the fear conditioning context (Baker et al., 1991; Bouton et al., 1993). Also, in many studies of reinstatement, testing phase is conducted 24h after the conditioning phase, in this case reinstatement producing similar results. (Norrhol, et al., 2006; Halladay et al., 2012).

Studies of the neural system underlie reinstatement in rodents observed a critical role of the amygdala and hippocampus (Frohardt et al., 2000). In humans, recently studies investigated the neural network with functional imaging, using a visceral pain US and a cue-conditioning paradigm found differential hemodynamic responses after reinstatement in the parahippocampus (Kattoor et al., 2013a) and the cerebellum (Kattoor et al., 2013b). Also, a study using “CS+ US association” and contextual conditioning found hippocampus activation to the contexts in which acquisition has been occurred. Furthermore, significant differential responses to the conditioned contexts were observed in the amygdala and the dmPFC after reinstatement. In particular, enhanced responses to CS+, were observed in the ACC/vmPFC, an area commonly implicated in fear inhibitory and monitoring processes (Lonsdorf et al., 2014a,b).

Currently, knowledge of experimental boundary conditions as well as biological or trait factors for reinstatement is very limited in humans. Further studies are needed to advance a more comprehensive understanding of this phenomenon. As a translational perspective, a better understanding the circumstances under which reinstatement occurs may offer a step toward the relapse as a clinical phenomenon and for the new developing of pharmacological or behavioural interventions to prevent the return of fear (Haaker et al., 2014).

## MEMORY RECONSOLIDATION

Citing the exact words of Kindt and colleagues 2009 ‘once emotional memory is established it appears to last forever’. Many studies demonstrated that even the most effective treatments only eliminate fear from responses, leaving the original fear memory intact, as demonstrated by the recurrent relapse after successful extinction (Bouton, 2002; Craske, 1999). However, studies have shown that fear memories can change when recalled, a phenomenon referred to as *reconsolidation* (Kindt et al., 2009). Reconsolidation is a process whereby previously consolidated memories can be reactivated and again made sensitive to mutate (Nader et al., 2000). Reconsolidation can be influenced by neurobiological manipulations during or shortly after the memory reactivation period (Tronson et al., 2007).

Considerable evidence in both animals and humans indicates that blockade of the process of reconsolidation by pharmacological manipulations produces amnesia for the original fear learning (Nader and Hardt, 2009, Sevenster et al., 2013). Among the most important studies in humans, Kindt and colleagues (2009) tested the first time the hypotheses that fear response can be weakened by disrupting the reconsolidation process of such memory and that disrupting should prevent the return of fear permanently. Reconsolidation disrupting was obtained by the administration of propranolol prior to memory reactivation and resulted in erasure of the fear response, an effect that has been found persisted over time (Soeter and Kindt, 2010). The erasure of the fear response could also have resulted from a more diffuse effect of the propranolol administration by reducing the fear aspects triggered by the aversive stimulus itself. However, authors argued that the propranolol manipulation specifically targeted the emotional expression of the memory at the same time leaving the declarative memory unchanged (Kindt et al., 2009; Soeter and Kindt, 2010). Moreover, authors suggest that propranolol selectively acts on the b-adrenergic receptors in the amygdala during emotional information processing in animals and humans. It may be hypothesized that beta-adrenergic blockade during

reconsolidation may selectively disrupt the protein synthesis of the amygdala, resulting in deconsolidation of the ‘CS+ US association’ while leaving the declarative knowledge in the hippocampus untouched (Phelps, 2004; McGaugh, 2004; van Stegeren 2005; Kindt et al., 2009). In addition, propranolol administration during memory reconsolidation resulted in selective erasure of the fear response to both the reactivated fear association and its category-related aspects. In particular, the memory reconsolidation effects following propranolol blockade were not found restricted to the reactivated fear CS+, but in its place generalized to those cues that were category related (Soeter and Kindt, 2011). The generalization of fear has been demonstrated to be dependent on the strength (intensity) of the ‘CS+ US association’ (Laxmi et al., 2003).

From these studies of memory reconsolidation in human, it is important to understand that the liability of a memory is to be dependent upon the ‘CS+ US association’ reactivation. It follows that, acquisition with partial reinforcement rate prevent that a single unreinforced reactivation trial would be sufficient to put the ‘CS+ US association’ in a sensitive state; if this is the case it should be necessary to use a sufficient number of reactivation trials to disrupt the reconsolidation process (Soeter and Kindt, 2010). Also, reconsolidation may only take place when memory reactivation involves an experience that engages new learning. Indeed, reconsolidation is triggered only when there is a new learning to take place during the specific reconsolidation time window. Reconsolidation might also be a considered as prediction error driven process, because associative learning requires prediction error (signalling discrepancy between actual and expected events) to create a new memory (Sevenster et al., 2013). Additionally, has been observed that using a CS reactivation trial or using a US reactivation trial (as a reinstatement) might also put the ‘CS+ US association’ labile and sensitive to disruption (Lonsdorf et al. 2014a).

A different but successful behavioural reconsolidation protocol has been described by Schiller and colleagues (2010). Authors demonstrated that the frequent presentations of unreinforced CS+ allowed for an updating of a more cognitive component of the emotional memory in humans. This procedure consists in an extinction protocol performed within the time window of reconsolidation. Extinction training within the reconsolidation window following reactivation was found to erase ‘CS+ US association’ leaving intact the declarative knowledge about the conditioning itself. As a consequence, fear responses were implicitly no longer expressed, the effect lasted at least a year and was selective only to reactivated fear memories.

Pharmacological manipulations affect reconsolidation process and as consequence in an incapacity to retrieve the fear conditioned memories, suggesting that they are erased or persistently weakened. Unfortunately, the use of pharmacological manipulations in humans can be always problematic. Obviously, a behavioural procedure will be preferred over pharmacological manipulations if providing similar effects. Change emotional memories has important implications for the treatment for anxiety disorders linked to traumatic memories, such as post-traumatic stress disorder (PTSD).

From an evolutionary perspective, it is functional to never forget the most important events in life, especially the negative ones. If emotional memory could be weakened or erased, then it might be possible to extirpate the root of many psychiatric disorders. For these reasons, reconsolidation phenomenon has important clinical implications; on one hand reconsolidation should not radically alter the functional reactions to potentially dangerous situations (US), but selectively weaken the maladaptive fear association ‘CS+ US association’. On the other hand, reconsolidation should not be specifically limited to the CS+ itself considering that generalization of fear is a main characteristic of anxiety disorders (Lissek et al., 2008). Indeed, reconsolidation it would be very useful to treat anxiety disorders when its effects spread to the category related of stimulus not previously associated (Soeter & Kindt, 2011). Currently, is a

matter of huge scientific interest identifying new flexibly and safely techniques in humans to target reconsolidation process (Schiller et al., 2010).

## **NEURAL SUBSTRATE OF FEAR CONDITIONING**

Fear conditioning is the most basic form of associative learning that has increased a considerable clinical relevance for the recent enhancing in the understanding of psychiatric disorders and thus, improving relative treatments. Modern neuroimaging techniques have significantly helped to provide the identification of anatomical structures and neural networks involved in human fear conditioning (Sehlmeyer et al., 2009).

Several studies aimed to investigate the anatomical contribution underlie fear conditioning across species and the whole evidence converge on that: amygdala is critical for the acquisition and the expression of conditioned fear (for review see LeDoux, 2000). Although the amygdala may be critical for the acquisition of extinction learning, the ventral medial prefrontal cortex (vmPFC) and hippocampus are other two key neural structure importantly implicated in this phenomenon, and together constitute the neural network of fear conditioning (Maren & Quirk, 2004). In particular, the amygdala stores both conditioning and extinction memories. The vmPFC integrates CS information with contextual information from the hippocampus in order to determine extinction retrieval. In the extinction context, the vmPFC inhibits amygdala projections, to reduce fear. Outside the extinction context, amygdala output is uninhibited (Quirk & Mueller, 2007). It is widely agreed that interactions between of these areas support the acquisition, storage, retrieval, expression and contextual modulation of fear conditioning (for review see Milad and Quirk, 2012). Consistent with studies in animal models, functional neuroimaging, lesion and morphology studies showed that extinction learning depends on the integrated functioning of this network and suggests that the neural mechanisms

supporting fear acquisition and extinction are phylogenetically conserved across species (Dunsmoor et al., 2017).

Since it has been discovered that amygdala and vmPFC are implicated in fear conditioning, a huge part of studies focused on psychiatric disorders, in particular, pervasive fear and/or anxiety, in which such brain areas are functional impaired. Positron emission tomography (PET) studies showed decreased prefrontal blood flow in PTSD patients (Semple et al., 1996; Bremner et al., 1999). Also, PTSD patients also showed reduced activation of vmPFC, as indicated by functional magnetic resonance imaging (fMRI), when recalling traumatic events (Shin et al., 1999). Thus, a better understanding of fear learning neural network may provide a solid ground to develop new specified and tailored treatment for psychiatric disorders.

## **AMYGDALA**

Surgical ad-hoc lesions, pharmacological drug administration and physiological evidence gained by both animal and human studies of the last century, provided a detailed model of the neural network underlie fear conditioning. Among all, amygdala became the core structure of the fear conditioning network when its involvement was discovered (LeDoux, 2000). Indeed, such brain area is now referred as follow: ‘amygdala as the locus of fear conditioning’ (Kim and Jung, 2006). One of the first reported evidence is that a lesion disrupts the acquisition and, thus the expression of conditioned fear responses (LeDoux et al., 1984, Hitchcock & Davis, 1986). Subsequent neuro-biological evidence reported that the association between CS+ and US is formed and expressed within different nuclei of the amygdala (Davis, 2000; Maren, 2005). The lateral nucleus of the amygdala (LA) is considered to be the core-site that encodes sensory inputs from both CS+ and US. In particular, with the presentation of the CS+, LA excites the central nucleus (CE), which is deputy to CR expression through projections to the

brainstem and thus, reach the whole peripheral nervous system. Also, it has been reported that LA indirectly projects to the CE through the basal nucleus and the intercalated cell masses (ITC). These pathways provide multiple potential circuits for gating fear expression (Dunsmoor et al., 2015).

Moreover, anatomical studies described the connections of the amygdala central nucleus with downstream structures implicated in the expression of fear conditioned responses, including the hypothalamus, periaqueductal gray, pons, and other brainstem regions (LeDoux et al., 1988, Romanski & LeDoux, 1993). Other studies described the inhibitory circuits within the amygdala that were found to be involved also in fear extinction, such as the lateral division of the central nucleus (Sun & Cassell, 1993), and inhibitory cells within the lateral and basolateral nuclei (Mahanty & Sah, 1998).

Consistent with the hypothesis that extinction results in new learning, not erasure of the original 'CS+ US association', a population of neurons in the LA have been identified in which the CS response is maintained despite a decrease in the expression of conditioned fear with extinction (Repa et al., 2001). This finding provides, again, evidence that the amygdala supports the maintenance of the original fear memory while at the same time allow extinction learning (Hartley and Phelps, 2010).

Early fMRI studies aimed to determine whether animals' neural models of fear conditioning might be overlapped and valid within the human brain. LaBar and colleagues (1998) and later Büchel and colleagues (1999), using a classical fear conditioning paradigm in healthy humans demonstrated increased amygdala functional activity in response to the CS associated with US, as compared to a neutral stimulus. In following fMRI studies using a large variety of CS and US, has been demonstrated beyond any doubt the crucial role of amygdala in fear conditioning in healthy human brain (for review see Dunsmoor et al., 2015; Milad & Quirk,



2012). Together these observations were critical and provided unequivocal evidence that suggest that amygdala functionality was preserved across species.

## **HIPPOCAMPUS**

The hippocampus plays an essential role in contextual learning (Bouton et al., 2006, Ji & Maren, 2007), as well as for the acquisition and the extinction of context conditioning (Radulovic & Tronson, 2010). In particular, the ventral hippocampus (vHPC) projects directly to both IL (PFC) and the BLA (amygdala) and it follows that it is in an anatomical position to modulate fear responses (Hugues & Garcia, 2007). It has been reported that, pre-conditioning hippocampal lesions selectively affects the acquisition of contextual memory (Phillips and LeDoux, 1992). Also, rodents with hippocampal lesions show impaired specific contextual renewal of the conditioned fear responses (Wilson et al., 1995). Finally, lesions to the nucleus accumbens (reached by hippocampal projections) disrupt contextual fear responses without affecting the explicit knowledge of conditioning itself (Riedel et al., 1997).

Extinction recall relies on contextual factors, suggesting a key role of the hippocampus in the retrieval extinction learning. It is important to highlight that, hippocampus involvement depends on the specific contextual renewal paradigm adopted (eg 'ABA' vs 'ABC'; Bouton et al, 2006). A clearer picture is emerged from studies using pharmacological inactivation of rodents' hippocampus. Inactivating it before extinction learning negatively affects the successful retrieval of extinction in later days, thus it has been noted an enhancement of conditioned fear responses compared to the controls (Corcoran et al., 2005). Also, the inactivation of the hippocampus before extinction recall phase prevented the renewal, thus it has been observed a reduction of conditioned fear responses as compared to the controls. (Corcoran and Maren, 2001; 2004; Hobin et al, 2006). Notably, converging evidence are

observed with inactivation of the mPFC (Sierra-Mercado et al, 2006), suggesting that the mPFC may be an important target of the hippocampus for contextual extinction recall (Hobin et al, 2003; Corcoran and Quirk, 2007).

Currently, hippocampus is considered essential to control the context-specific recall of extinction both indirectly through projections to the vmPFC, and directly through projections to the LA (for a review see Maren et al., 2013). Also, it has reported that different hippocampal subregions have been found implicated in different aspects of behaviour. In particular, the dorsal part for spatial-related behaviours and the ventral one for anxiety-related behaviours (for a review see Bannerman et al., 2004). Converging evidence suggest that hippocampus and different subregions are implicated in different characteristics of fear conditioning (Kin & Jung, 2006). Thus, the hippocampus appears to be essential for consolidation of extinction, especially in tasks such as inhibitory avoidance, which crucially require the hippocampus (Quirk & Mueller, 2007).

## **PREFRONTAL CORTEX**

One of the first evidence about the involvement of PFC in fear conditioning was demonstrated by Morgan et al. (1993) who found that rodents with ventral PFC lesions required more presentations of the CS to extinguish conditioned fear. In particular, authors reported that pre-training lesions of the ventral PFC had no effect on the acquisition of conditioned fear but impaired fear extinction in later days. Subsequently, it was found in electrophysiological studies that infralimbic cortex (IL) would be the functional homologous region of the vmPFC in humans (Quirk et al., 2000), since it inhibits the expression of conditioned fear during extinction through reciprocal connections with the amygdala. Milad and Quirk (2002) reported that IL neurons showed increased activity to the CS during extinction recall (Quirk et al., 2003) and

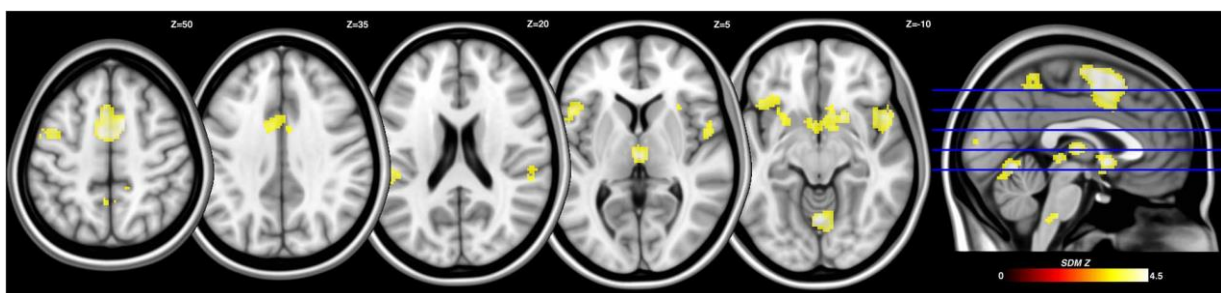
after surgical lesion, reduced conditioned response to a CS+ even before the extinction (Milad et al., 2004).

Phelps and colleagues (2004) conducted the first fMRI study to determine whether vmPFC in humans has the same functional properties as monkey or rodents IL in fear conditioning. Authors showed that human vmPFC increased its activation during the recall of extinction learning. Further studies, reported that, during extinction recall, the strength of vmPFC activation to an extinguished stimulus was positively correlated with the strength of extinction retention (Milad et al 2007). It follows that, the stronger the activation of the vmPFC, the better the ability to inhibit conditioned responding during extinction recall. Additionally, analysis of the thickness of the vmPFC and the dorsal anterior cingulate (dACC) was positively correlated with CR assessed by skin conductance responses (SCR) during the conditioning phase (Hartley et al., 2011). dACC involvement has been noted in previous studies of fear conditioning (Buchel et al., 1998; Phelps et al., 2004; Knight et al., 2004) but its involvement was not highlighted. However, dACC activation has been observed in response to both CS and US (Dunsmoor et al., 2008, Knight et al., 2010). Also, it has observed that when the participants anticipate the shock occurrence, dACC was also activated (Linnman et al. 2011). These evidences support the role of dACC in the expression of conditioned fear in humans, even if it is not deeply studied. Regarding other parts of the PFC, behavioural studies found no evidence for their involvement in fear conditioning. There are only few evidences about lesions in rats in the dorsal medial PFC, that enhance fear responses during acquisition but not extinction (Morgan and LeDoux, 1995) or mice (Vouimba et al., 2000).

Consistent to evidence of vmPFC functional alteration, it has reported that PTSD patients had normal ability to acquire fear conditioning and extinction, but the ability to recall extinction memory the following day were altered (Garfinkel et al., 2014). In particular, this deficit in extinction recall was associated with hypoactivation in the vmPFC and

hyperactivation in the dACC (Milad et al 2009). Similar observations have been reported in schizophrenic patients, in which vmPFC has been found functional impaired (Holt et al 2009). Taken together this evidence suggests that alteration in fear conditioning circuits might be transversal across many psychiatric disorders in humans (Insel et al 2010).

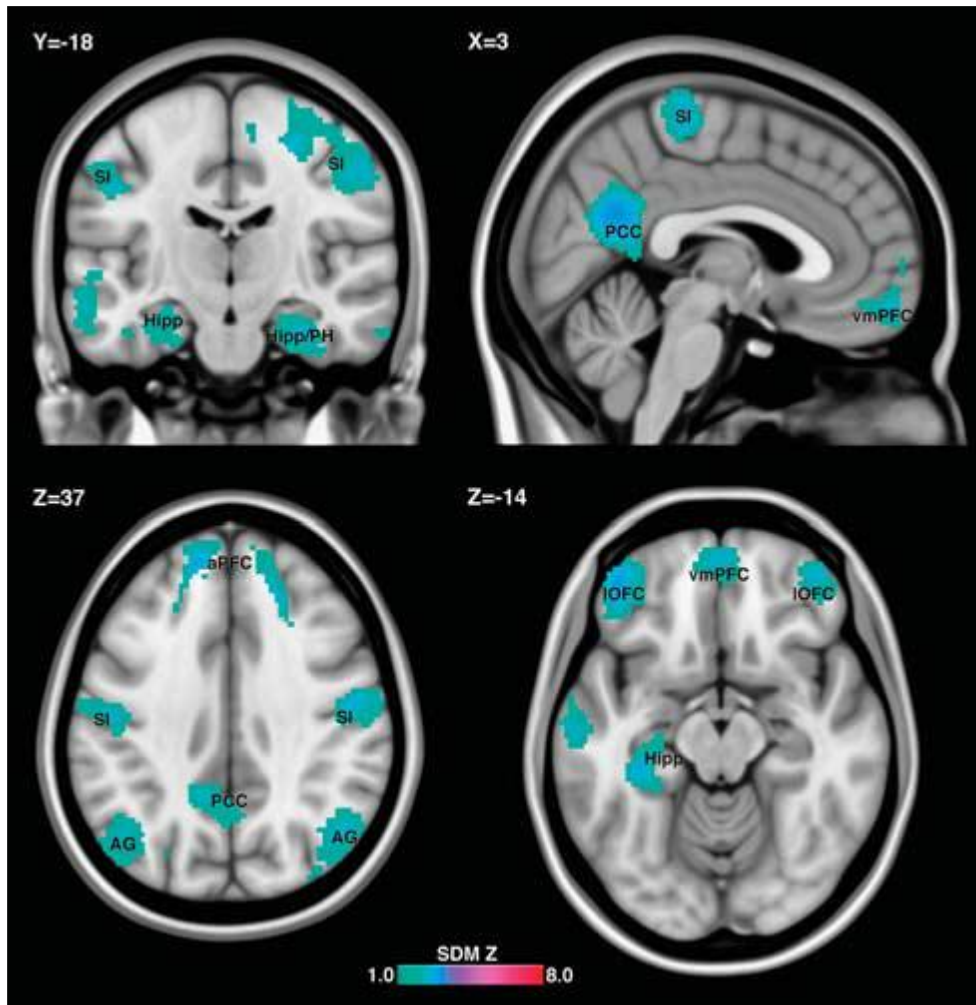
Although classical studies and interpretation agreed on the crucial role of amygdala in order to acquire fear conditioning, emerging evidence suggests that vmPFC may be also involved in the process underlie the acquisition fear. For example it has been reported that, vmPFC activity is initially suppressed by CS+ versus CS- during the acquisition of conditioning; such functional suppression gradually diminish over the course of extinction learning until there is no functional difference in response to both stimuli (Schiller et al., 2008; Schiller and Delgado, 2010). More recently, Fullana and colleagues (2016; 2018) reported in a reported in two meta-analysis studies of neuroimaging in humans, including a total of more than 1300 participants, evidence about the involvement of human vmPFC during the acquisition as well as the extinction. In fact, results showed that it was not possible to separate and thus identify specific brain network from acquisition of conditioning as compared to extinction (Figure 1). Moreover, results highlighted a prominent involvement of prefrontal cortex more than it emerges from individual studies. It follows that, the PFC and amygdala are both involved during acquisition and extinction.



**Figure 1.** Neural correlates of fear conditioning versus extinction learning estimated by meta-analysis (Fullana et al., 2018). Results showed that it was not possible to separate specific brain activity from conditioning as compared to extinction.

Some explanation has been argued about this results, and authors explain that: extinction would be a form of implicit ‘emotional’ regulation (as reported by Schiller and Delgado, 2010), while during acquisition engages prefrontal cortical regions linked with more explicit ‘cognitive’ forms of emotion regulation (Delgado et al., 2008; Fullana et al., 2016). Also, behavioural studies suggest that cognitive-regulatory factors may be more involved in human fear extinction learning than conditioning (Lovibond, 2004). Other interpretation may be ground in earlier fMRI studies of fear conditioning (Milad et al., 2007; Schiller et al., 2008) in which authors have reported that vmPFC deactivation is possible to represent processing of the CS- as a non-threat cue, highlighting the inhibition of fear. Comparable ideas have also been supported in fear extinction studies using context, in which there are many evidences to suggest that vmPFC activity may be involved in the distinction between non-threatening and threatening contextual cues after conditioning. Although the precise role of vmPFC processing may vary across fear-learning paradigm, one idea is that the distinction between CS- and CS+ trigger a common neural substrate for the representation of reward value, in which CS-, as compared to CS+, has an intrinsic positive reward value (Schiller et al., 2010).

In conclusion, recently Dunsmoor and colleagues (2019) reported again that vmPFC appears more active during the presentation of CS- than CS+ (as Fullana et al., 2016; see Fig. 2), suggesting again its possible role in discriminating safety from threat. In conclusion, activity has also shown in the dmPFC, suggesting its possible role in the conscious appraisal of threat (Mechias et al., 2010).



**Figure 2.** Significant brain functional deactivation to the CS+ versus CS – determined by meta-analysis (Fullana et al., 2016). Results are displayed on the Montreal Neurological Institute. Abbreviations: AG, angular gyrus; aPFC, anterior prefrontal cortex; Hipp, hippocampus; IOFC, lateral orbitofrontal cortex; PCC, posterior cingulate cortex; PH, parahippocampal formation; SI, primary somatosensory cortex.

## PSYCOPHYSIOLOGICAL MEASUREMENTS

Human emotions are generally studied on different response levels: subjective verbal reports about fear experienced, behavioural, and physiological level as well as neurobiological changes (e.g. Bradley & Lang, 2000; Lonsdorf et al., 2017). For ethical and methodological reasons, the conditioned - fear - responses acquired in human are rarely strong enough to elicit a behavioural response such as flight (Löv et al., 2015).

Psychophysiological indices are the most commonly applied in human researches. They have the distinct advantage of not being biased by the participant itself and usually providing a direct comparison to animal research. The most commonly used physiological indices in human fear conditioning are Skin Conductance Responses (SCRs), Fear Potentiated Startle (FPS) reflex, Heart Rate (HR) and Pupillary Response (PS). However, such techniques needed methodological considerations before to be employed and should be specifically tailored to each fear conditioning research.

The most important (Steckle, 1933; Switzler, 1934) and still most employed psychophysiological index of conditioned fear responses is the electrodermal activity (EDA; Dawson et al., 2007). EDA may be measured as skin conductance response (SCR) or as skin conductance level (SCL). It is important to differentiate that, SCR refers to a phasic response to a stimulus, that can be computed as the difference between a pre-stimulus and the peak post-stimulus. On the other hand, SCL refers to the average levels during a specific time period and it is not related to a specific stimulus (Lykken and Venables, 1971). In fear conditioning research, CS+ onset elicits a stronger SCR (i.e. larger responses) as compared to the CS-. The application of SCL is mainly applied in context conditioning, where a larger level of electrodermal activity may be observed for the acquisition context (A) than for the extinction context (B; Lonsdorf et al., 2017). SCR are slow responses that reach their peak 0.5–5 s later the stimulus onset. Thus, experimental design needs to allow for acceptable temporal spacing

(between experimental stimuli) to allow a return to its baseline, thus avoiding the superimposed of stimuli. Indeed, a fast sequence of stimuli leads to superimposed SCR which suffer from distorted amplitudes and temporal characteristics (Boucsein et al., 2012).

The fear potentiated startle response is a defensive sequence of reflexes elicited by the occurring of a sensory event (Hunt et al., 1938). In humans, the most reliable component of the startle reflex is the startle eyeblink response (Blumenthal et al., 2005), usually assessed by the use of electromyography (EMG) over the orbicularis oculi muscle (Lonsdorf et al., 2017). The startle reflex is a probed response which is typically elicited by the 'startle probe', a brief burst of white noise with an instantaneous rise-time administered binaurally (Lissek et al., 2005a). However, the use of fear potentiated startle can be problematic and leading to methodological problem difficult to disentangle. In particular, the physical properties of the auditory startle probe, such as intensity, time, and bandwidth affect both startle amplitude and startle probability. Other issues could be that, the probing is in every trial, for both CS+ and CS- and this could affect participant's learning (Panayiotou et al., 2011).

The human pupillary response has also been described as a reliable measure for conditioning (Bitsios et al., 2004; Reinhard et al., 2006) and can be assessed by eye-tracking or pupillometry. Pupillary responses can be quantified in terms of pupil dilation to a mean baseline pre-stimulus. In contrast to slow SCRs, the pupillary response is fast and reflects a measure of psychological index both sympathetic and parasympathetic (Granholm and Steinhauer, 2004). Pupil response is really fast after stimulus onset (0.1 – 0.4s; Beatty and Lucero-Wagoner, 2000) and it should be employed in paradigms with short stimulus presentation and/or fast inter-interval stimulus (Lonsdorf et al., 2017).

Heart Rate (HR) as a measure of human fear conditioning has been employed in different recent studies (Lonsdorf et al., 2017). It is widely agreed that HR response to conditioned fear involves both deceleration and acceleration (Castegnetti et al., 2015). HR



decelerations reflects conditioned fear response, presumably indicating an orienting response to the CS+ presentation. The subsequent HR acceleration reflects a defensive response to the US predicted by the CS+ and, thus, reflecting fear learning (Hamm et al., 1993). Taken together, conditioned HR changes seem to reflect the stage of learning during CS processing (Lang et al., 1997). Compared to the past where HR has been extensively used for assessment of autonomic tone its use for investigating differences in emotional learning is relatively new. Recent studies have shown how HR may represent a useful tool to study fear conditioning (Liu et al., 2013; Pappens, et al., 2014; Castegnetti et al., 2015; Tzovara et al., 2018). Indeed, low Heart Rate Variability (HRV) has been associated with elevated contextual anxiety (Sevenster et al., 2015), higher levels of HRV at rest were associated with better extinction (Wendt et al., 2015; Pappens et al., 2014). Such studies suggest that HRV at rest may reflect the capacity of the high-level cognitive function to inhibit subcortical fear responses in the presence of safety or when former threat cues are presented in the absence of threat. Also, higher HRV is associated with a general ability to flexibly adapt to environmental demands (Lyonfields et al., 1995; Thayer et al., 2012) as well as specifically with more successful inhibition in the presence of emotional stimuli (Kryptos et al., 2011). Importantly, it has been shown that persons with low resting HRV have difficulty in adjusting their response to safety signals in a context where threat stimuli might occur (Melzig et al, 2009; Park et al., 2013; Ruiz-Padial et al., 2003; Ruiz-Padial & Thayer, 2014).

The use of such different and but complementary methodologies to study physiological responses of fear conditioning might serve as markers for maladaptive fear learning and contribute to the identification of individuals prone to the development of psychiatric disorder.

## THESIS OUTLINE

The aim of this thesis is to describe recent developments in understanding the neurobiological basis of human fear conditioning both in healthy and brain-damaged individuals. In the subsequent chapters the studies that I accomplished during my PhD will be reported in detail. In the subsequent paragraphs, the conducted studies will be briefly reported highlighting the experimental hypothesis and the most relevant results.

The study 1 aimed to examine the influence of normal aging on context-dependent recall of extinction to fear conditioned stimulus. Healthy young and old adults, for a total of 48 subjects, were tested in a multi-phase study over two days. We used a 2-day differential threat conditioning and extinction procedure to determine whether young and older adults differed in the contextual recall of conditioned responses. On the first day, conditioned stimuli were paired with an aversive electric shock in a context (danger) and then extinguished in a different context (safe). On the second day, the extinguished stimuli were presented to investigate both extinction recall (in safe context), and contextual fear renewal (in danger context). Results showed that young participants were able to use contextual information to adaptively guide their fear responses whereas older participants showed impaired modulation of the responses by contextual information.

The study 2 aimed to determine the causal role of the PFC in the acquisition of fear conditioning by systematically test the effect of a selective lesion of the vmPFC. In this study, participants were divided into three groups: 10 patients with a lesion to the vmPFC, 10 brain-damaged control patients with a lesion that did not involve the PFC or medial temporal lobe and 10 healthy control adults with no brain lesion. Results suggest that healthy controls and brain-damaged control patients had successful acquisition and extinction of threat conditioning. On the contrary, vmPFC patients were impaired in the acquisition of the conditioning. It is important to highlight that vmPFC patients were comparable CS-US awareness as the other two

groups. The results of the present study shed new light into the role of prefrontal cortex in acquisition of fear conditioning in humans. Unlike studies in animals and previous, anecdotal reports, the present results suggest that vmPFC is a crucial brain structure for fear conditioning in humans. Thus, damage to vmPFC in humans would impair the ability to shape defensive anticipatory responses to the fear conditioned stimulus but spare the ability to learn explicit contingencies regarding the conditioning.

The study 3 aimed to disrupt the reconsolidation process of acquired fear memory using a non-invasive brain stimulation protocol (i.e. TMS). The modification of emotional memories has been classically attempted with pharmacological or behavioural procedures. However, both approaches present limitations in terms of applicability and effectiveness in humans. To this end, 70 participants underwent a multi-session paradigm in three experimental days: on the first day participants acquire fear conditioning, on the second day participants reactivate the previously acquired fear memory and afterwards TMS was applied, and finally on the third day, participants recalled the fear memory, both before and after a reinstatement procedure. In particular, the experimental manipulation interfered with the memory reconsolidation process by applying repetitive TMS protocol over the right and left PFC or in other control cortical sites immediately after the memory reactivation in the second day. Results showed that interfering with activity in both left and right PFC prevents the recall of the fear, in contrast to other control groups. These results suggest that non-invasive stimulation of the PFC following memory reactivation may attenuate the expression of fear to a previously conditioned stimulus and argue in favour of a critical role of the PFC in the neural network that is underlie the reconsolidation in humans.

The study 4, a theoretical-methodological study on physiological measures have been carried out in the same framework of the previous studies. Here, it has been investigated whether the parasympathetic - vagal - modulation of heart rate might reflect the anticipation of

fearful as compared to safe outcomes during classical fear conditioning paradigm. To this aim, despite old and outdated analysis, it has proposed a new methodology to decode heart rate modulations. In particular, the presence of non-stationary mechanisms (i.e. transitory vagal responses) were assessed with a short-time Fourier analysis and applied to heart rate variability recorded throughout the task to catch specific component of heart rate (HR) that could be reflect conditioning. It has been hypothesized a different pattern of HR variability for CS+ as compared to CS- when participants ( $n = 50$ ) anticipated the fearful outcome administration. Results showed a significant cluster of power contribution from 0.15 to 0.30 Hz (i.e., high-frequency band), larger for the CS+ than the CS-, reflecting the vagal contribution and occurring at the time in which participants expected to receive the shock administration. These results indicate that the presentation of CS+ elicits a strong and selective vagal response compared to CS-, sustaining bradycardia during the acquisition phase. Thus, it implies that fear conditioning has occurred, revealing a specific biomarker of cardiac autonomic modulation in humans.

Evidence reported in this PhD thesis might provide key insights and deeper understanding of critical issues concerning both theoretical and methodological aspects underlying the acquisition, the extinction and the reconsolidation of fear memories in humans.



**STUDY 1:**  
**CONTEXT-DEPENDENT EXTINCTION OF THREAT MEMORIES:**  
**INFLUENCES OF HEALTHY AGING**

## **ABSTRACT**

Although a substantial progress has been made in recent years on understanding the processes mediating extinction of learned threat, little is known about the context-dependent extinction of threat memories in elderly individuals. We used a 2-day differential threat conditioning and extinction procedure to determine whether young and older adults differed in the contextual recall of conditioned responses after extinction. On Day 1, conditioned stimuli were paired with an aversive electric shock in a ‘danger’ context and then extinguished in a different ‘safe’ context. On Day 2, the extinguished stimulus was presented to assess extinction recall (safe context), and threat renewal (danger context). Physiological and verbal report measures of threat conditioning were collected throughout the experiment. Skin conductance response (SCR data revealed no significant differences between age groups during acquisition and extinction of threat conditioning on Day 1. On Day 2, however, older adults showed impaired recall of extinction memory, with increased SCR to the extinguished stimulus in the ‘safe’ context, and reduced ability to process context properly. In addition, there were no age group differences in fear ratings and contingency awareness, thus revealing that aging selectively impairs extinction memories as indexed by autonomic responses. These results reveal that aging affects the capacity to use context to modulate learned responses to threat, possibly due to changes in brain structures that enable context-dependent behaviour and are preferentially vulnerable during aging.

## INTRODUCTION

Extinction of threat memories is a phenomenon that allows animals and humans to adapt their behaviour to a changing environment. During extinction, repeated presentation of the conditioned stimulus (CS) alone after Pavlovian or classical, threat conditioning (CS–unconditional stimulus (US) pairings) causes attenuation of defensive responses<sup>1</sup> (Pavlov, 1927; see for recent review Dunsmoor et al., 2015). Several key studies (Bouton, 1993; Bouton, 2004) indicate that extinction does not involve permanent erasure (i.e., unlearning) of the original associative (i.e., CS-US) memory. Instead, there is converging evidence from animal (Quirk, 2002; Rescorla, 2001; Senn et al., 2014; Chhatwal, 2005) and human (Hobin et al., 2003; Kalisch et al., 2006) studies that the mechanisms supporting extinction entail new learning (i.e., CS-no US) that competes, and temporarily interferes, with the expression of the original conditioning trace. During this competition, contextual information appears to be a critical regulatory factor in determining whether the original threat memory or the new extinction memory should control defensive CS responses. For example, a renewal of responding is observed (Bouton & King, 1983; Rosas & Bouton, 1997; Bouton & Swartzentruber, 1986) when, after extinction in a context (Context B) different from the acquisition context (Context A), the CS is presented in the original acquisition context (Context A). This “ABA renewal effect” has been repeatedly demonstrated in both rats (Bouton & Bolles, 1979; Bouton & Ricker, 1994; Rauhut et al., 2001) and humans (Mineka et al., 1999; Mystkowski et al., 2002), and suggests that extinction involves just one more form of learning that is particularly context-dependent (for excellent comprehensive reviews on threat extinction and renewal, see Dunsmoor et al., 2015; Vervliet et al., 2013; Maren et al., 2013).

It is widely agreed that aging is accompanied by a cognitive decline in laboratory animals, as well as in humans (Park & Schwarz, 2000; Buckner, 2004; Hedden & Gabrieli, 2004). Declines in the ability to process contextual information, and flexibly adapt behaviour



to situational changes, may represent a fundamental mechanism of age-related cognitive alterations (Braver et al., 2001). Furthermore, considerable research in animals and humans reveals that contextual regulation of extinction memory requires coordinated activity of regions of prefrontal cortex, hippocampus, and amygdala (Maren, 2013). Of these, prefrontal cortex and hippocampus-dependent behaviours are preferentially vulnerable during aging, suggesting that impairments within these structures could underlie extinction deficits in advanced age (Salami, 2014; Van de Vijver, 2014). Although a substantial progress has been made in recent years on understanding the processes mediating extinction of learned threat (Dunsmoor, 2015; Milad & Quirk, 2012) the impact of healthy aging on the context-dependent extinction of threat memories has been relatively unexplored.

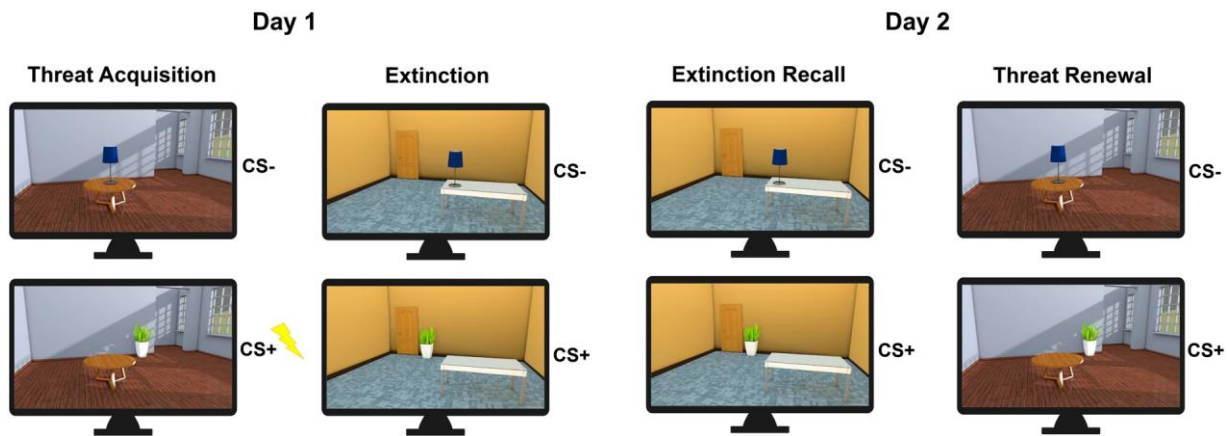
Previous deficits in the extinction of escape from spatial water maze have been reported in aged rats (Oliveira, 2007; Dere et al., 2005). However, these rats were also impaired at the initial acquisition of spatial water maze, thereby confounding clear assessment of how aging may specifically alter extinction. Recent studies have specifically demonstrated a decline of the capacity to extinguish in aged rats (Kaczorowski, 2012; Oler & Markus, 1998), and mice (Sanders, 2011), associated with difficulties in contextual regulation of extinction memory in older animals. Interestingly, age-related extinction deficits occurred in the absence of impairments in the initial acquisition and expression of defensive responses to threat stimuli, thus indicating that older animals have a selective difficulty using contextual information to modulate the expression of stimulus-response contingencies (Maren, 2013).

In humans, one prior study by LaBar and colleagues (LaBar, K. S. et al., 2004) examined the impact of aging on the acquisition and subsequent extinction of threat conditioning using a simple conditioning paradigm conducted within a single session. LaBar et al. (LaBar et al., 2004), reported no age-related reduction in threat conditioning and immediate extinction, provided that awareness of the CS–US contingency and arousal, assessed by unconditioned

responding, were taken into account. There is increasing evidence from the animal (Lebrón, 2004; Milad & Quirk, 2002; Quirk, 2000; Rhodes & Killcross, 2004; Santini, 2004), and human (Kalisch et al., 2006; Garfinkel et al., 2014; LaBar & Phelps, 2005) studies that within-session extinction (i.e., extinction conducted immediately after threat conditioning or short-term extinction) and between-session extinction recall (e.g., long-term extinction memory) involve different mechanisms and neurobiological substrates. To date, however, no prior study has directly examined age-related differences in delayed recall of extinction memory, and the contextual dependency of long-term extinction recall in young and older adults.

To test for context-dependent recall of extinction memory in aging, we used a 2-day differential threat conditioning and extinction procedure, modified from that previously described by Milad and colleagues (Garfinkel et al., 2014; Milad et al. 2005) (see Fig. 1). This protocol incorporates a temporal delay (24 hr) between extinction training and subsequent probing of extinction and threat memories, thus providing a more ecological test of long-term extinction memories in young and older adults. On Day 1, subjects received conditioning followed by extinction, with pictures of common objects as CSs, and electric shock as the US.

To manipulate context, we presented visual CSs embedded within pictures of two distinct rooms, such that, on Day 1, threat acquisition and extinction training were performed in contexts A and B, respectively. On Day 2, participants were presented with two additional phases: extinction recall, and threat renewal, in context B (extinction context) and context A (conditioning context), respectively. No US was delivered on Day 2. Physiological (skin conductance) and verbal report measures of threat conditioning were collected throughout the experiment.



**Figure 1. Stimuli and experimental design.** Threat acquisition and extinction were established on the first day (Day 1). Participants were threat conditioned in the danger context, in which the conditioned stimulus (CS+) was associated with a shock pulse on 60% of trials, while the CS- was not associated with any consequence. Extinction followed this phase, during which both CSs were presented within the safe context and none of them was associated with the shock pulse. Extinction recall and threat renewal were administered on the second day (Day 2). The recall of extinction was tested presenting the conditioned stimuli (CSs) within the safe context (in which extinction occurred on the first day). Subsequently, renewal of threat was tested presenting CSs within the danger context (in which the threat association was learned on the first day). On the second day, all CSs were presented in absence of the shock pulse.

Consistent with prior aging studies in humans and non-human mammals (Kaczorowski, 2012; Oler & Markus, 1998; Sanders, 2011), we hypothesized no significant age group differences during acquisition and extinction of threat conditioning on Day 1. However, we expected that, compared to young adults, older adults would show a selective deficit in contextual processing of extinction memory on Day 2. The results of the present study should thus yield insights into age-associated changes in the extinction of threat memories and the mechanisms that enable context-dependent behaviour.

## METHODS

**Participants.** A total of 48 right-handed healthy adults participated in the study. Participants were divided into two age groups: twenty-four young adults (12 female; mean age = 24.79 years, SD = 3.59 years; age range: 20–30 years; mean education = 14.45 years, SD = 2.32 years), and twenty-four older adults (12 female; mean age = 66.12 years, SD = 7.60; age range: 60–70 years; mean education = 13.33 years, SD = 2.41 years). The young group was composed of Bologna University students recruited through campus advertisements, whereas the old group was recruited through a referral from the Center for Studies and Research in Cognitive Neuroscience of Bologna University, where the study was conducted, or other referral sources. Prior to participation, subjects were screened to ensure that they had no history of neurological, psychiatric, or cardiovascular conditions. None of the participants were taking any medication affecting the central nervous system regularly. All participants had normal or corrected-to-normal vision. The two groups were matched for level of education ( $t(1,23) = 1.453$ ;  $p = 0.159$ ). It is widely known that anxiety and depression may affect SCR in classical conditioning<sup>42</sup>. To account for such variability, levels of anxiety and depression were measured by means of the State-Trait Anxiety Inventory (Spielberger et al., 1983), and the Hospital Anxiety and Depression Scale (Zigmond et al., 1983). The two groups did not show any significant difference in terms of anxiety (young group mean = 39.14, SD = 5.32 years; old group mean = 36.83 years, SD = 5.83 years;  $t(1,23) = 1.386$ ;  $p = 0.179$ ), and depression (young group mean = 4.68, SD = 1.89 years; old group mean = 5.37 years, SD = 2.97 years;  $t(1,23) = -0.939$ ;  $p = 0.357$ ). The study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki and was approved by the Ethics Committee of the Department of Psychology of the University of Bologna. All

participants gave informed written consent to participation after being informed about the procedure of the study.

**Neuropsychological assessment.** Young and older adults were given a series of standardized neuropsychological tests. The primary objective in performing these tests was to rule out the possibility that any older adult participants included in our sample were affected by age-associated cognitive deficits, rather than to assess differences between young and old groups. The battery included tests of abstract reasoning (Raven Progressive Matrices; Spinnler & Tognoni, 1987), verbal short-term and long-term memory (Verbal Span with disyllabic words, and Prose Recall; Spinnler, & Tognoni, 1987), selective attention (Attentional Matrices Test; Spinnler & Tognoni, 1987), and executive function (Weigl's Sorting Test; Spinnler, H. & Tognoni, G., 1987). Normative scores derived from a nationally representative sample of adults are available for each test. For all tests, participants' raw scores were converted into equivalent scores<sup>46</sup>, adjusted for age and years of education. Equivalent score is a 5-point scale, ranging from 0 to 4, with 0 = pathological performance, 1 = borderline performance, 2–4 = normal performance. The neuropsychological testing session was held one or two days before the experimental session, and only participants who were within normal ranges were asked to participate in the experiment. *Table 1* shows the means, standard deviations of the equivalent score on each test for young and older participants in the study.

Test	Equivalent Scores		t(23)	p
	Young	Old		
<b>Raven Progressive Matrices</b>	3.91 (±0.28)	3.79 (±0.58)	0.9	0.376
<b>Verbal Span</b>	3.12 (±0.94)	2.66 (±0.46)	1.141	0.265
<b>Prose Recall</b>	3.83 (±0.38)	3.70 (±0.46)	1.269	0.216
<b>Attentional Matrices Test</b>	3.33 (±0.70)	3.58 (±0.65)	-1.297	0.207
<b>Weigl's Sorting Test</b>	3.75 (±0.67)	3.45 (±0.77)	1.231	0.231

**Table 1.** Means, standard deviations (in brackets) and statistical comparison (t-test) between old and young participants of the equivalent scores on each test. The battery included tests of abstract reasoning (Raven Progressive Matrices), verbal short-term and long-term memory (Verbal Span, and Prose Recall), selective visual attention (Attentional Matrices Test), and executive function (Weigl's Sorting Test). No significant differences were found between young and older participants.

**Materials.** The experiment was implemented in Matlab R2016 (The MathWorks, Inc., Natick, Massachusetts, United States) software, and ran on a Windows-based PC (Lenovo ThinkCentre Desktop Computer). Stimuli were created with Blender (Blender Foundation, Amsterdam, Netherlands) and Cinema 4D R17 software (MAXON Computer GmbH, Friedrichsdorf, Germany), and were presented on a computer screen (screen size: 43 inches; resolution: 1920 × 1080; refresh rate: 60 Hz). Context scenes consisted of images of 2 different indoor scenes (i.e., a yellow-blue room, and a grey-red room), representing the acquisition ('danger') context and the extinction ('safe') context of threat associations, respectively. For half of the participants in each age group, the acquisition context and the extinction context were the yellow-blue room and the grey-red room, respectively. Context assignment was reversed for the other half so as to counterbalance across subjects which environment was associated with a shock. Conditioned stimuli (CSs) were images of two everyday common objects, a plant and a lamp, embedded within the context scenes (Garfinkel, S. N. et al., 2014; Milad, M. R. et al., 2005). For half of the participants in each group, the reinforced CS+ and the unreinforced CS- were the plant and the lamp, respectively, and vice versa for the other half. Neutral, rather than intrinsically emotional (i.e., spiders, snakes, or angry faces), stimuli were used as CSs, because conditioned responses to very salient CSs can be confounded by the ceiling effects of the respective outcome measures.

A mildly aversive electro-tactile stimulation served as unconditioned stimulus (US). The shock pulse was generated by a Digitimer Stimulator (Model DS7, Digitimer Ltd., UK) and delivered

to the participants' left inner wrist for 200 ms. The intensity of the stimulation was determined individually by assessing the participant's subjective evaluation in a standard work up procedure prior to threat acquisition. It was initially set at 0.5 mA and increased of 1 mA until participants reported it as a "highly annoying, but not painful" stimulation.

**SCR Recording.** The skin conductance response (SCR) was recorded with two Ag/AgCl electrodes (TSD203Model; Biopac Systems, USA), filled with isotonic hyposaturated conductant and attached to the distal phalanges of the second and the third finger of participants' left hand. A DC amplifier (Biopac EDA100C) was used while recording the SCR. A gain factor was 5  $\mu$ S/V and the low-pass filter was set at 10 Hz. The analog signal was then passed through a Biopac MP-150 digital converter at a 200 Hz rate. The signal was recorded with AcqKnowledge 3.9 (BIOPAC Systems, Inc., Goleta, California) and converted to microsiemens for offline analysis.

**Procedure.** The study was performed at the Center for Studies and Research in Cognitive Neuroscience of the University of Bologna, in Cesena, Italy. Participants were tested individually. They were comfortably seated in a silent and dimly lit room, and their position was centered relative to the computer screen, at 100 cm viewing distance. Electrodes for SCR recording, and for shock pulse administration were attached to the participant. The SCR was recorded continuously while participants completed the task and data were stored for offline analysis. Participants were asked to remain as quiet and still as possible during the task and to keep their attention at the center of the screen. After verifying that SCR was being properly recorded, the intensity of the shock pulse to be used as US was adjusted for each participant as described above. Finally, participants were informed that they had no effect on shock administration.

The experiment consisted in a modified version of a classical differential threat conditioning and extinction procedure (Garfinkel et al., 2014; Milad et al., 2005; Milad et al., 2009) (see Fig. 1). During the experiment, each trial consisted in the presentation of a context scene for 1 s, followed by one of the two CSs presented within the context scene for 4 s, and ending with the context scene still visible for 1 more second. The intertrial interval (ITI) was a white fixation cross on a black background, with a variable duration ranging from 11 to 16 s. The length of the ISI was chosen to avoid complete masking of conditioned SCRs by preceding unconditioned SCRs to the shock. The experimental protocol was administered over two separate days.

On Day 1, three different phases were presented: habituation, threat acquisition and threat extinction. At the beginning of the session, participants were informed that different images would be presented on the screen, and the task of the participant would be to carefully observe the images, as some of them might be paired with the electrical stimulation. The habituation phase included 4 trials, in which the CS+ and CS- (2 for each) were presented in random order either within the 'danger' context or the 'safe' context, to ensure the absence of any baseline differences within and between age groups in response to the CSs. Few habituation trials were used to avoid retardation of learning due to nonreinforced exposure to CS+ (the latent inhibition effect (Lubow, 1973)). The threat acquisition phase consisted of 20 CS+ and 20 CS- trials, all presented within the 'danger' context (yellow-blue room or grey-red room). One CS (plant or lamp) was associated with the administration of a shock pulse, resulting in the conditioned stimulus (CS+), while the other CS was never paired with any consequence, resulting in the neutral stimulus (CS-). In CS + trials, the US (shock) was administered 60% of times (12 out of 20 trials), 3.8 s after the CS+ onset, and co-terminated with the CS+. In CS- trials, the US was never administered. The trials were pseudo-randomly presented to participants such that no more of three identical CSs occurred in a row. During the extinction phase, which followed immediately, the CSs were presented within a distinct ('safe') context.



In this phase, participants learned that the CS+ was no longer followed by the US. Both CS+ and CS- stimuli were presented 20 times without the US. Characteristics of CSs, trial order, and ITI were as in the acquisition phase.

On Day 2, (24 hr after the extinction phase), two additional phases were presented: extinction recall, and threat renewal, during which the ability to selectively retrieve extinction memory as a function of context (safe vs. dangerous) was tested. Participants were told that the procedure for this second part of the experiment would be the same as on the previous day. During extinction recall, 10 CS+ (without the US) and 10 CS- were presented within the 'safe' context, where extinction learning previously occurred. During threat renewal, 10 CS+ (without the US) and 10 CS- were presented within the danger context, where the original threat conditioning was learned. Stimulus and ITI timings were identical on Days 1 and 2.

To assess the acquisition of a conditioned response to CSs, SCR was measured during all the experimental phases, and the responses related to CS+ were contrasted against those related to CS-. It has to be noted that shocks were delivered only in the acquisition phase of the first day and never delivered in all other phases of the experiment.

**SCR data analysis.** SCR data were offline analyzed using custom-made MATLAB scripts, and all statistical analyses were performed with STATISTICA (Dell Software, released September 2015, StatSoft STATISTICA for Windows, version 13.0, Round Rock, Texas, USA). Assumption of normal distribution of data was verified. Mixed-design analyses of variance (ANOVAs) were used to investigate differences within and between age groups. Post-hoc analyses were conducted with Newman-Keuls test and the significance threshold was  $p < 0.05$ . Data were extracted from the continuous signal and calculated for each trial as the peak-to-peak amplitude of the largest deflection during the 0.5 to 4.5 s time window after stimulus onset. The minimum response criterion was 0.02, and smaller responses were encoded as zero. SCR

following the US was analyzed to assess unconditioned responding, whereas SCR following the CS was analyzed to assess conditioned learning. Regarding SCR to the US, stimulus onset was represented by the time of shock administration; regarding SCR to CS, stimulus onset referred to the time of CS appearance.

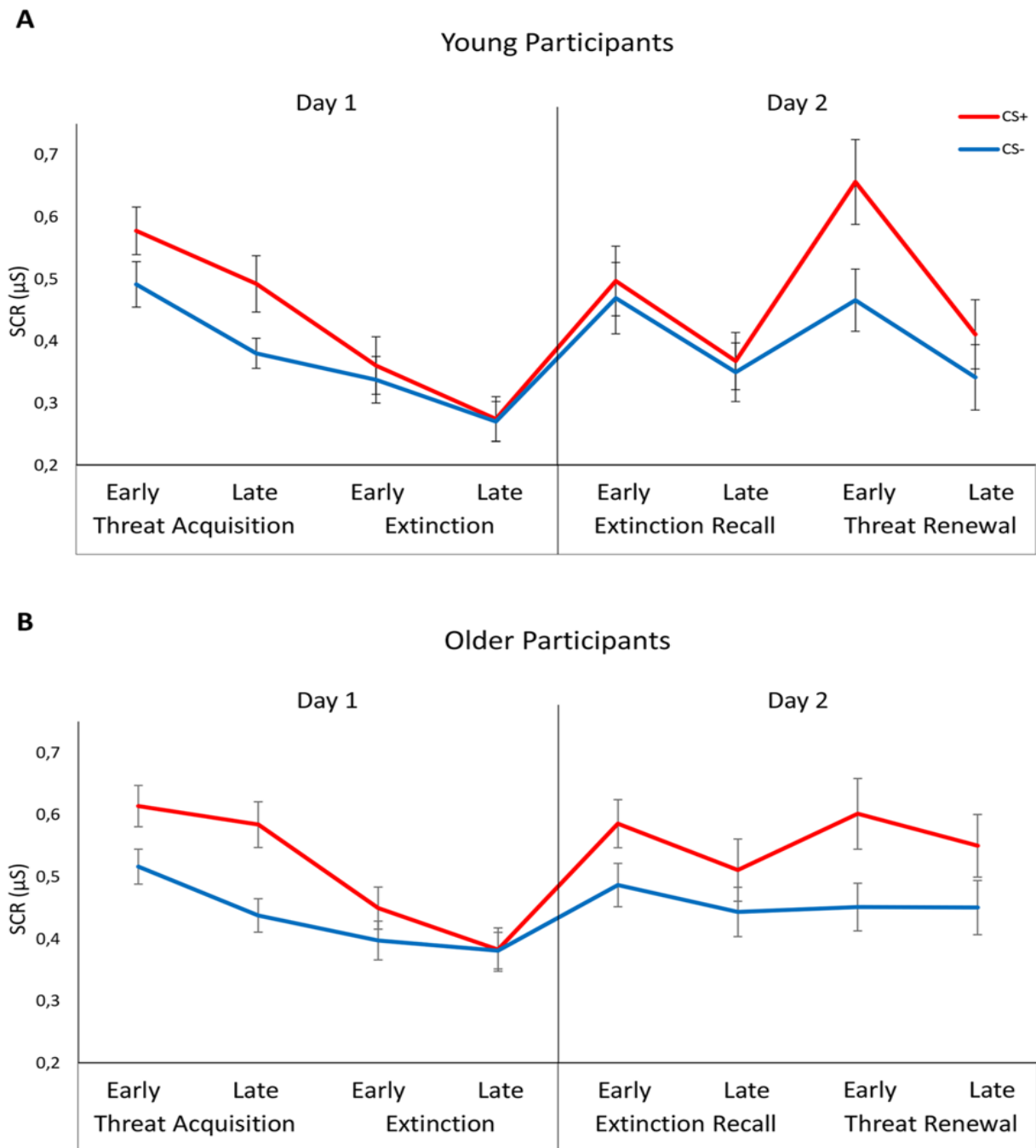
Raw SCR scores were square-root transformed to normalize the data distribution and scaled to each participant's mean square-root-transformed US response, to account for inter-individual variability (Schiller, et al., 2008). To reduce interindividual variability, raw scores were range corrected by dividing each individual score by the subject's mean SCR response to US (Lykken, 1972). This procedure can reduce error variance, thus increasing statistical power when comparing groups of participants. In this way, conditioned responses can be directly compared across groups without confounding baseline differences in skin conductance levels (LaBar et al., 2004). Because after range correction the resulting distribution was positively skewed, these data were then square-root transformed prior to statistical analyses (Siddle, et al., 1988). Regarding the response to the US, mean SCRs to the 12 shocks were analyzed. Concerning the response to the CS, SCR data were collapsed into "early" and "late" trial blocks of each phase (threat acquisition and threat extinction on Day 1; extinction recall and threat renewal on Day 2), as learning typically varies across time within each learning phase. On Day 1, to assess conditioned responses to the CS separated from unconditioned responses to the shocks themselves, only non-reinforced CS trials were analyzed. Learning-related changes were hypothesized to be found in the 'late acquisition' and 'late extinction' phases, as reported previously (Garfinkel et al., 2014; Milad et al., 2005).

## RESULTS

**US intensity and unconditioned responding.** One-way ANOVAs were used to evaluate differences in US intensity and mean SCR to the US. Results showed no difference in the intensity of shock pulses ( $F(1,46)=0.08$ ,  $p=0.928$ ) between young (mean = 7.49 mA, SD = 2.21 mA) and older (mean = 7.56 mA, SD = 2.62 mA) adults. Likewise, no difference between young (mean = 1.02  $\mu$ S, SD = 0.16) and old (mean = 0.97  $\mu$ S, SD = 0.14) group was found in the mean SCR in responses to US ( $F(1,46)=0.781$ ,  $p=0.381$ ). On average, therefore, the intensity of the electrical stimulation received by participants, the subjective quality of perception (“highly annoying, but not painful), as well as the physiological response to it (i.e., arousability) did not differ significantly between age groups.

**Habituation (Day 1).** To analyze habituation, a  $2 \times 2$  repeated measure ANOVA was performed on SCR, with Group (young/old) as a between-subject factor, and Stimulus (CS+/CS-), as a within-subject factor. Analysis showed no significant main effect of Group ( $F(1,46) = 0.23$ ,  $p = 0.632$ , partial  $\eta^2 = 0.02$ ), Stimulus ( $F(1, 46) = 0.304$ ,  $p = 0.58$ , partial  $\eta^2 = 0.01$ ), or Group by Stimulus interaction ( $F(1, 46) = 1.22$ ,  $p = 0.277$ , partial  $\eta^2 = 0.01$ ), thus revealing that at baseline there were neither within group nor between group differences in orienting responses to the CS+ and CS-.

**Threat acquisition and extinction (Day 1).** To analyze SCR data recorded in Day 1, a  $2 \times 2 \times 2$  repeated measure ANOVA with Group (young/old) as a between-subject factor, and Stimulus (CS+/CS-), and Block (early/late) as within-subject factors was carried out separately for each phase (threat acquisition and extinction; see Fig. 2).



**Figure 2. Skin conductance responses.** Graphs illustrate mean skin conductance responses (SCRs) to the conditioned (CS+) and neutral (CS-) stimuli during early and late blocks, in young (A) and older (B) participants on Day 1 (threat acquisition and extinction phase) and Day 2 (extinction recall and threat renewal phase). Data demonstrate no effect of aging on threat acquisition and extinction on Day 1. In contrast, only older participants failed to recall the previous extinction in the safe context on Day 2, while young participants specifically adapted their conditioned responses according to the context. Error bars represent standard error.

During threat acquisition, results showed a main effect of Stimulus ( $F(1,46) = 65.901$ ,  $p < 0.001$ ,  $\eta^2 = 0.58$ ), reflecting stronger responding to the CS+ (young group mean =  $0.52 \mu\text{S}$ ,  $\text{SD} = 0.17 \mu\text{S}$ ; old group mean =  $0.59 \mu\text{S}$ ,  $\text{SD} = 0.14 \mu\text{S}$ ) than to the CS- (young group mean =  $0.43 \mu\text{S}$ ,  $\text{SD} = 0.14 \mu\text{S}$ ; old group mean =  $0.47 \mu\text{S}$ ,  $\text{SD} = 0.12 \mu\text{S}$ ), and a main effect of Block ( $F(1,46) = 17.467$ ,  $p < 0.001$ ,  $\eta^2 = 0.27$ ), which reflected higher SCRs overall during late than early acquisition block. This result implies that that differential threat learning to the CS+ took place overall during the acquisition phase. Importantly, the analysis revealed neither a significant main effect of Group, nor interaction of Group with Stimulus or Block (all  $p$ s  $> 0.23$ ), thereby suggesting that conditioned learning took place equivalently in young and older participants.

During extinction, analysis revealed a significant Stimulus by Block interaction ( $F(1,46) = 7.17$ ,  $p = 0.010$ ,  $\eta^2 = 0.13$ ), but no significant main effect or interactions with the factor Group (all  $p$ s  $> 0.08$ ). Post-hoc analyses showed that participants had significantly stronger responses to CS+ than to CS- during early extinction (CS+: young group mean =  $0.36 \mu\text{S}$ ,  $\text{SD} = 0.22 \mu\text{S}$ ; old group mean =  $0.44 \mu\text{S}$ ,  $\text{SD} = 0.16 \mu\text{S}$ ; CS-: young group mean =  $0.33 \mu\text{S}$ ,  $\text{SD} = 0.18 \mu\text{S}$ ; old group mean =  $0.39 \mu\text{S}$ ,  $\text{SD} = 0.15 \mu\text{S}$ ), but SCR differences between CS+ (young group mean =  $0.27 \mu\text{S}$ ,  $\text{SD} = 0.17 \mu\text{S}$ ; old group mean =  $0.38 \mu\text{S}$ ,  $\text{SD} = 0.17 \mu\text{S}$ ) and CS- (young group mean =  $0.27 \mu\text{S}$ ,  $\text{SD} = 0.15 \mu\text{S}$ ; old group mean =  $0.38 \mu\text{S}$ ,  $\text{SD} = 0.14 \mu\text{S}$ ) disappeared for both groups during late extinction.

Thus, overall results showed equivalent responding of the two experimental groups across all three phases (i.e., habituation, threat acquisition, and extinction) of Day 1, prior to the extinction recall and threat renewal manipulations of Day 2.

**Extinction recall and threat renewal (Day 2).** To analyze SCR data collected in Day 2, a  $2 \times 2 \times 2$  repeated measure ANOVA with Group (young/old) as a between-subject factor, and Stimulus (CS+/CS-), and Block (early/late) as within-subject factors, was carried out separately for each phase (extinction recall and threat renewal; see Fig. 2).

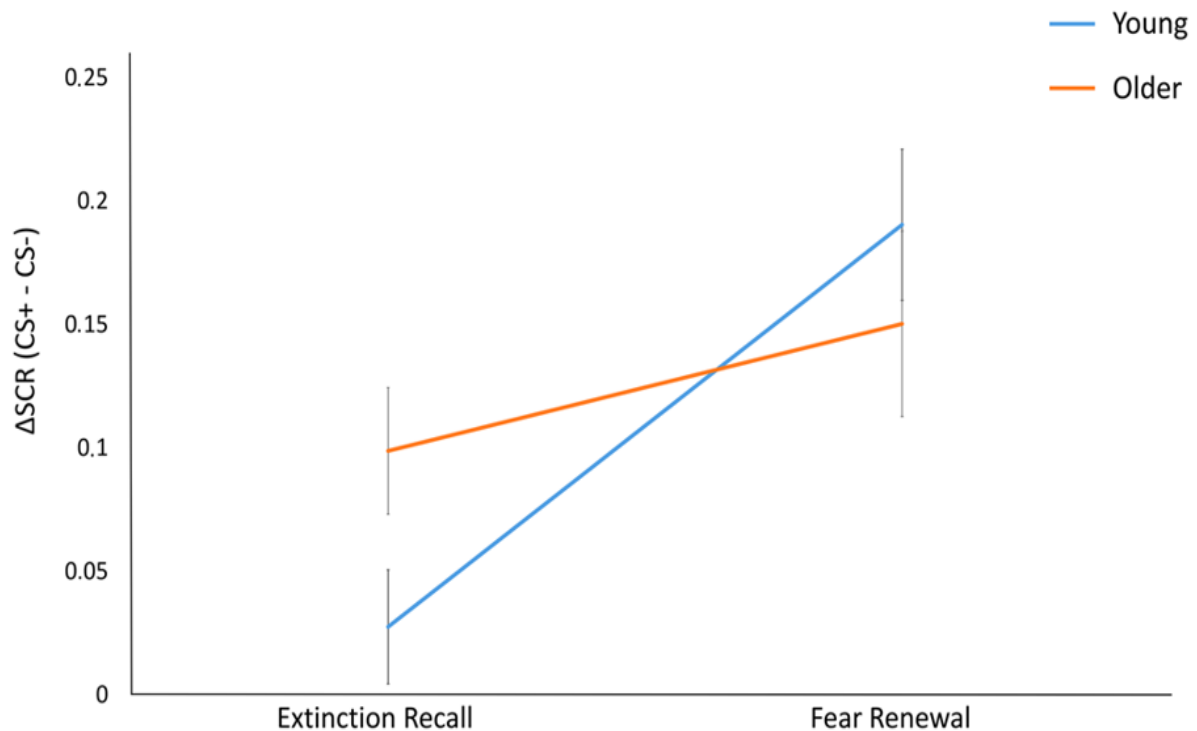
During extinction recall, analysis showed a main effect of Stimulus ( $F(1,46) = 13.512$ ,  $p = 0.001$ ,  $\eta^2 = 0.22$ ), which reflects elevated responses to the CS+ (young group mean =  $0.44 \mu\text{S}$ ,  $\text{SD} = 0.24 \mu\text{S}$ ; old group mean =  $0.54 \mu\text{S}$ ,  $\text{SD} = 0.21 \mu\text{S}$ ) relative to CS- (young group mean =  $0.41 \mu\text{S}$ ,  $\text{SD} = 0.22 \mu\text{S}$ ; old group mean =  $0.46 \mu\text{S}$ ,  $\text{SD} = 0.16 \mu\text{S}$ ), and a main effect of Block ( $F(1,46) = 14.368$ ,  $p = 0.001$ ,  $\eta^2 = 0.23$ ), due to a progressive decrease of conditioned SCRs during the extinction recall phase in both groups. Crucially, the analysis revealed a significant Group by Stimulus interaction ( $F(1,46) = 4.401$ ,  $p = 0.04$ ,  $\eta^2 = 0.087$ ). Follow-up Newman-Keuls tests showed different pattern of SCRs between groups. Specifically, no difference in SCR was found during extinction recall in the young group ( $p = 0.27$ ). In the old group, however, the SCR to CS+ was significantly higher than SCR to CS- ( $p < 0.001$ ), demonstrating a return of threat response to previously extinguished CS+. Other comparisons were not statistically significant ( $ps > 0.31$ ). Importantly, to control for the influence of depression and anxiety on extinction recall, we repeated the significant Group x Stimulus x Block analysis using ANCOVA with levels of depression and anxiety as additional covariates. The Group by Stimulus interaction remained statistically significant even after controlling for depression ( $F(1,45) = 3.981$ ,  $p = 0.05$ ,  $\eta^2 = 0.082$ ), and anxiety ( $F(1,45) = 4.798$ ,  $p = 0.03$ ,  $\eta^2 = 0.096$ ). Thus, aging was associated with impaired recall of extinction memory, both in the early and late portion of the phase.

During threat renewal, an ANOVA showed a significant main effect of Stimulus ( $F(1,46) = 38.551$ ,  $p < 0.001$ ,  $\eta^2 = 0.456$ ), indicating significantly greater SCRs associated to CS+ (young group mean =  $0.53 \mu\text{S}$ ,  $\text{SD} = 0.29 \mu\text{S}$ ; old group mean =  $0.57 \mu\text{S}$ ,  $\text{SD} = 0.25 \mu\text{S}$ )

than to CS<sup>-</sup> (young group mean = 0.40  $\mu$ S, SD = 0.24  $\mu$ S; old group mean = 0.45  $\mu$ S, SD = 0.18  $\mu$ S), in both groups. A main effect of Block ( $F(1,46) = 37.989, p < 0.001, \eta^2 = 0.452$ ), and a Group by Block interaction ( $F(1,46) = 21.535, p < 0.001, \eta^2 = 0.318$ ) were also found. This result reflected reduced SCRs overall during late than during early threat renewal in young ( $p < 0.001$ ), but not in older ( $p = 0.287$ ), adults. Importantly, neither a significant main effect of Group nor interaction of Group with Stimulus was found (all  $ps > 0.51$ ). Therefore, both groups showed differential SCR to CS<sup>+</sup> compared to CS<sup>-</sup> during renewal.

To directly assess context-dependent modulation of extinction memory in young and older participants, the differential threat response ( $\Delta$ SCR) was calculated by subtracting SCR to CS<sup>-</sup> from the SCR to CS<sup>+</sup>, both during early extinction recall and early threat renewal. Extinction recall analysis focused on the first block of trials ('early extinction recall') in order to avoid confounding extinction memory with new extinction learning taking place during the extinction recall phase itself (Milad et al., 2007). For the same reason and to be consistent, threat renewal also focused on the first block of trials ('early threat renewal'). An ANOVA, with Group (young/old) as a between-subject factor, and Phase (extinction recall/threat renewal) as within-subject factors, showed a main effect of Phase ( $F(1,46) = 22.108, p = 0.001, \eta^2 = 0.47$ ) and, more critically, a Phase by Group interaction ( $F(1,46) = 5.975, p = 0.018, \eta^2 = 0.11$ ). Follow-up Newman-Keuls tests revealed that the young adults showed normal context-sensitivity during extinction recall, with significantly lower  $\Delta$ SCR in the extinction (safe) compared with the acquisition (danger) context ( $p = 0.012$ ). In contrast, older adults did not demonstrate a significant effect of context on  $\Delta$ SCR on Day 2 ( $p = 0.12$ ). The Phase by Group interaction remained significant ( $p < 0.05$ ) even after adjusting for the influence of depression and anxiety levels as additional covariates, suggesting that impaired context-dependent modulation of threat and extinction memories were mediated by aging, and not by depression or anxiety. These results (Fig. 3) suggest that on Day 2 young participants adapted

their responses to threat based on the context in which the stimuli were presented. Differently, older participants did not recall extinction memory, responding specifically to CS+ regardless the context in which it was presented.

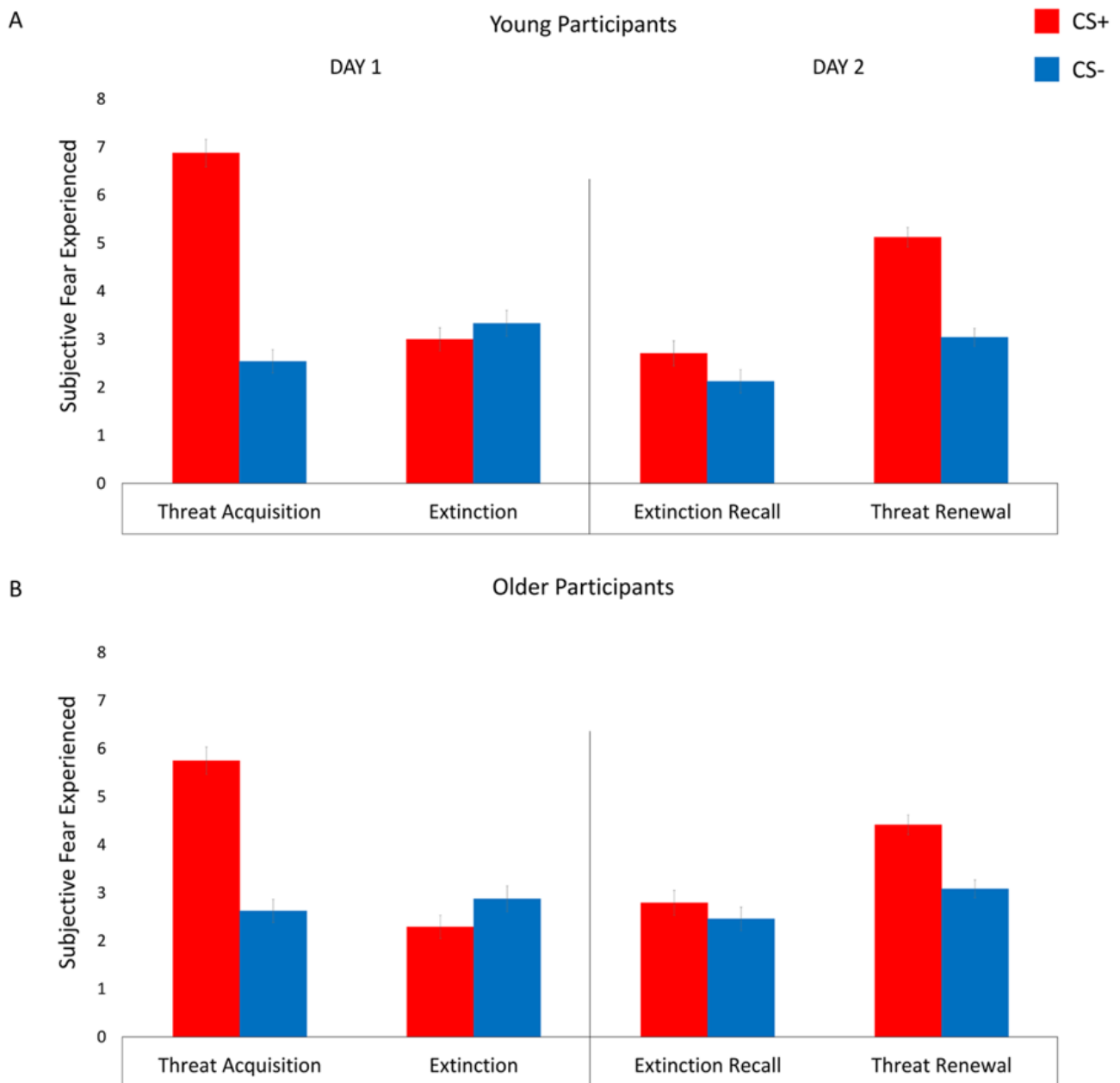


**Figure 3.**  $\Delta$ SCR (calculated by subtracting SCR to CS<sup>-</sup> from SCR to CS<sup>+</sup>) during early extinction recall and threat renewal (Day 2). While young participants adjusted their psychophysiological response based on the context, old participants show a similar activation regardless of the contextual information. Error bars represent standard error.

**Subjective fear ratings and contingency awareness.** A  $2 \times 2 \times 5$  repeated measure ANOVA with Group (young/old) as between-subject factor, Stimulus (CS<sup>+</sup>/CS<sup>-</sup>) and Phase (habituation/acquisition/extinction/extinction recall/threat renewal) as within-subject factors, was used to assess participants' fear ratings of conditioned stimuli in each experimental phase (Fig. 4). A significant Stimulus by Phase interaction ( $F(1,184) = 40.439, p < 0.001, \eta^2 = 0.49$ ) was found, indicating that self-report level of fear to CS<sup>+</sup> and CS<sup>-</sup> differed depending



on experimental phases. The Stimulus by Phase by Group interaction was not significant ( $F(1,184) = 1.081, p = 0.367, \eta^2 = 0.02$ ), indicating that the old group did not differ from the young group in the level of self-report fear to the conditioned stimuli during the experimental phases. Newman-Keuls test for the significant interaction showed that, in the habituation phase, self-report fear to CS+ (young group mean = 2.58, SD = 0.92; old group mean = 2.66, SD = 1.27) and CS- (young group mean = 2.87, SD = 1.22; old group mean = 2.79, SD = 1.41) were not significantly different ( $p = 0.846$ ). Instead, in the acquisition phase, self-report fear to the CS+ (young group mean = 6.87, SD = 1.39; old group mean = 5.75, SD = 1.42) was significant higher than fear to the CS- (young group mean = 2.54, SD = 1.17; old group mean = 2.62, SD = 1.46;  $p < 0.001$ ). During extinction, self-report fear to the CS+ (young group mean = 3, SD = 1.17; old group mean = 2.29, SD = 1.04) and CS- (young group mean = 3.33, SD = 1.30; old group mean = 2.87, SD = 1.19) were not significantly different ( $p = 0.967$ ), as well as in the extinction recall phase (CS+, young group mean = 2.70, SD = 1.26; old group mean = 2.79, SD = 0.93; CS-, young group mean = 2.12, SD = 1.19; old group mean = 2.45, SD = 0.88;  $p = 0.506$ ). Finally, during threat renewal, self-report fear to the CS+ (young group mean = 5.125, SD = 0.99; old group mean = 4.41, SD = 1.61) was significant higher compared to the CS- (young group mean = 3.04, SD = 0.90; old group mean = 3.08, SD = 1.24;  $p < 0.001$ ). Irrespective of the group, all participants correctly associated the context scenes with the administration of the electrical stimulation; moreover, 91% of young participants and 88% of older participants correctly paired the CSs with the corresponding outcome ( $p = 0.574$ ). Thus, both young and older participants were able to verbally express CS-US, as well as context-US, contingencies.

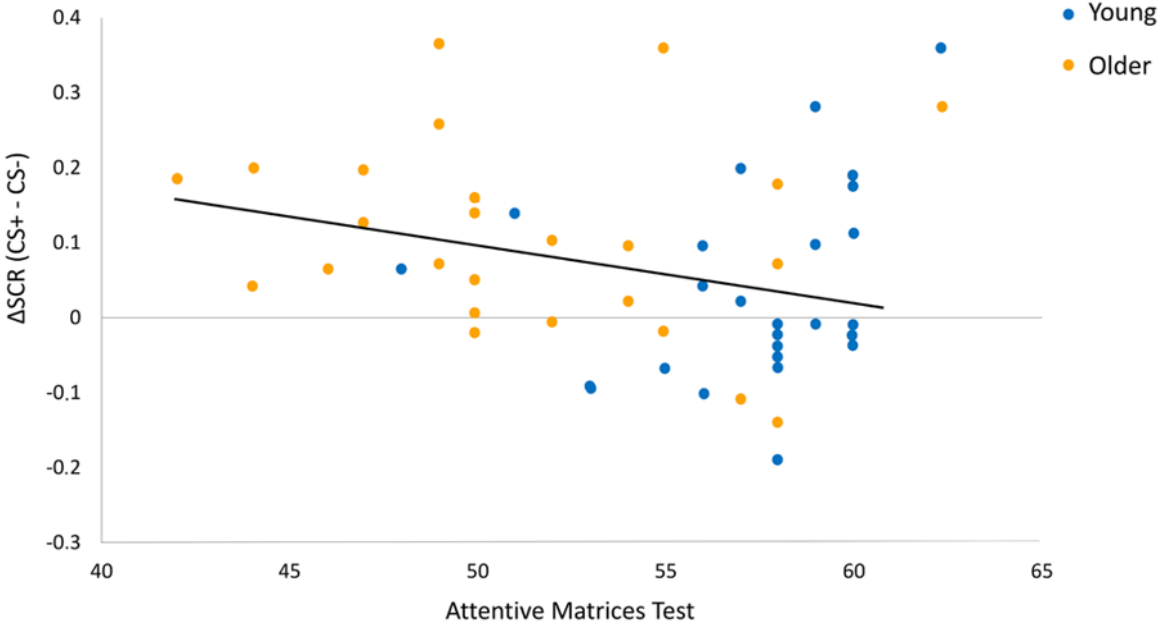


**Figure 4. Subjective fear ratings.** Graphs illustrate the level of self-reported fear to the conditioned stimuli during the experimental phases in young (A) and older (B) participants. Error bars represent standard error.

**Neuropsychological assessment.** Equivalent scores on each neuropsychological test were compared between young and older participants in the study and no significant differences were found (see Table 1).

To further test the impact of neuropsychological variables, four separate stepwise regression analysis (forward selection) were performed on each task phase (threat acquisition/extinction/extinction recall/threat renewal). The raw scores of all neuropsychological tests were used as regressors (namely, Raven Progressive Matrices, Verbal Span, Prose Recall, Attentional Matrices Test, and Weigl’s Sorting Test) and the differential conditioned response ( $\Delta$ SCR) was used as a dependent variable.

For the extinction phase (Day 1), the best model ( $F(1,46) = 15.93, p < 0.001, R^2 = 0.14$ ) reported a significant effect only of Weigl’s Sorting Test ( $\beta = -0.37, t = -2.68, p = 0.01$ ). For the extinction recall phase (Day 2), the best model ( $F(1,46) = 5.19, p = 0.03, R^2 = 0.11$ ) reported a significant effect only of the Attentional Matrices Test ( $\beta = -0.32, t = -2.27, p = 0.03$ ) (Fig. 5). No significant effects were found for acquisition and threat renewal phases.



**Figure 5. Impact of neuropsychological variables.** Regression analysis reported a significant influence of selective visual attention, as assessed by the Attentional Matrices Test, on  $\Delta$ SCR measured during early extinction recall phase.

**Influences between acquisition and recall of threat and extinction.** To test for a possible relation between the conditioned responses ( $\Delta$ SCR) at threat learning and retrieval, a correlation between threat acquisition (Day 1) and threat renewal (Day 2), and a correlation between extinction (Day 1) and extinction recall (Day 2), were calculated separately within each group. Furthermore, to test for a possible relation between the conditioned response ( $\Delta$ SCR) within each testing session, a correlation between threat acquisition and extinction (Day 1) and a correlation between threat renewal and extinction recall (Day 2), were calculated separately within each group. Pearson's correlation coefficient and one-tailed, Bonferroni-corrected p-value are reported.

Young participants showed a trend in the correlation between threat acquisition and threat renewal, but this resulted non-significant when Bonferroni-corrected ( $r = 0.40, p > 1$ ). No significant correlations were found between extinction and extinction recall ( $r = -0.02, p > 1$ ), between threat acquisition and extinction ( $r = 0.09, p > 1$ ), and between extinction recall and threat renewal ( $r = 0.006, p > 1$ ) in this group.

Older participants showed a significant positive correlation between threat acquisition and threat renewal ( $r = 0.48, p = 0.024$ ), and a significant positive correlation between extinction recall and threat renewal ( $r = 0.67, p < 0.01$ ). No significant correlations were found between extinction and extinction recall ( $r = 0.25, p = 0.42$ ), and between threat acquisition and extinction ( $r = 0.2, p = 0.32$ ) in this group.

Taken together, these results seem to indicate that similar processes may be involved in the acquisition and renewal of a threat in older and, possibly, in young participants. This second interpretation, however, has to be taken cautiously, as this trend is visible, but not significant when applying a Bonferroni correction. However, extinction recall and threat renewal clearly seem to involve similar processes in old, but not in young participants.

## DISCUSSION

Learning to disregard a stimulus that no longer predicts an aversive outcome, i.e., extinction, is critical for adaptive behaviour in a changing environment. Contextual information is particularly important in regulating the expression of responses to threat after these responses have been extinguished (Bouton, 1993). Declines in the ability to process contextual information may represent a fundamental mechanism of age-related cognitive changes (Braver et al., 2001). The present study was the first to examine the influence of normal aging on context-dependent recall of extinction of responses to threat. Healthy young and old adults were tested in a multi-phase study over two days (Garfinkel et al., 2014; Milad et al., 2007). During the first day, participants were threat conditioned to two visual stimuli (CS+ and CS-) within a specific (danger) visual context, and then underwent threat extinction within a different (safe) context. On the second day, the ability to selectively recall extinction memory within these two different contexts (danger and safe) was assessed.

Results showed that young participants were able to use contextual information to flexibly guide their learned responses to threat (as expressed by SCR), whereas older participants showed impaired modulation of the responses by contextual information. More specifically, on the first day, all participants were equally able to acquire and completely extinguish a threat conditioned response (i.e., higher SCR to CS+ as compared to CS- during threat acquisition, and equal SCR to CS+ and CS- during extinction). On the second day, young participants showed a context-dependent modulation of the autonomic responses, as higher SCR to CS+, compared to CS-, was observed in the danger context, but not in the safe context (Vansteenwegen et al., 2005). In stark contrast, older adults showed an impaired context-guided recall of extinction, with higher SCR to CS+, as compared to CS- in both danger and safe context (Fig. 2).

These results are consistent with the presence of either a specific extinction recall deficit, or a more general context-processing deficit. Our finding that differential responding to the CS+ versus the CS- increased from the safe (extinction) to the danger (renewal) context in the young but not in the older participants strongly suggests that aging is associated with a more general loss of context sensitivity in memory expression (Fig. 3). Moreover, on Day 2, there was a significant positive correlation between differential threat responses in the safe (extinction recall) and danger (threat renewal) context in the older, but not in the young, adult group. This further suggests that aging is associated with loss of contextual control of extinction, causing extinguished threat memories to inappropriately renew in any context. Interestingly, all participants were equally able to learn and explicitly report the association between conditioned stimuli, context scenes, and aversive US (i.e. contingency awareness), as well as rate how fearful each stimulus was in each context (i.e., affect ratings), thus revealing that aging specifically precludes recall of extinction memories as indexed by physiological responses (Figure. 4).

The present findings were not related to differences in global autonomic responsivity, as unconditioned responses to the shock were the same across both groups. Likewise, results were unlikely due to changes in trait anxiety, or depressive conditions, since we did not find differences in these control variables between young and old participants. Regarding neuropsychological performance, it is important to note that all participants performed within the normal range compared with age and education-adjusted norms, and that the groups did not differ on age and education adjusted scores (Table 1). Therefore, the impairment in context-dependent extinction recall in older participants was not related to age-related cognitive decline (American Psychiatric Association, 1994).

Taken together, these findings indicate that older adults were less able to use contextual information to recall extinction memory and modulate the expression of the defensive responses

to threat in a context-dependent manner, despite their preserved ability to acquire and extinguish a threat conditioned response.

Evidence of age-related changes in threat conditioning from non-human studies tend to report normal acquisition of simple forms of threat learning, but deficits in more complex aspects, such as acquisition and retention of contextual conditioning (Oler & Markus, 1998; Doyère et al., 2000; Houston & McNamara 1999; Stoehr & Wenk, 1995; Ohta et al., 2001). In line with the present findings, during tone threat conditioning, old mice exhibit a deficit in the use of context to modulate responses to threatening cues (Sanders, 2011). In particular, compared to young mice, aged mice showed low levels of threat responses regardless of the context, whereas young mice demonstrated context-dependent expression of renewal of responses (Sanders, 2011). Remarkably, both threat conditioning and immediate extinction were similar in the two groups.

In humans, LaBar and colleagues (LaBar et al., 2004) suggested an age-related impairment in threat conditioning as secondary to poor CS-US contingency awareness. More specifically, they found an age-related impairment in the expression of both threat conditioned responses and discriminative conditioning accounted for by a lack of awareness of the CS-US contingencies. Although awareness is neither necessary nor sufficient for normal conditioning learning (Lovibond, & Shanks, 2002; Bellebaum & Daum, 2004), it may play an important role in complex learning paradigms. Age effects may be at least partially due to a higher number of unaware subjects in old populations (Knuttnen et al., 2001), and it seems likely that old participants have more problems in recognizing the rule predicting US presentation during acquisition (Bellebaum & Daum 2004). The present results show that older participants had threat acquisition and explicit awareness of CS-US and context-US contingencies comparable to those of young participants. As such, results of the present study are in line with LaBar (LaBar et al., 2004) findings in showing no age-related reductions of threat learning and

extinction when contingency awareness is controlled. Thus, the failure in context-dependent extinction recall we observed in older participants does not seem to be due to a general learning deficit, or to a lack of contingency awareness and explicit knowledge acquired during the task.

In Pavlovian conditioning, the context is often referred to as an “occasion setter”, that is a modulating stimulus whose role is to disambiguate the current meaning of the conditioned stimulus (Bouton, 2004; Trask et al., 2017). Thus, in extinction procedures, context serves as an occasion setter that favours retrieval of the ‘safe’ CS–no US memory in the extinction context, and the ‘fearful’ CS–US memory in the acquisition (or any other) context (Holland, 1992; Schmajuk & Holland, 1998), which in turn inhibits and excites, respectively, the conditioned response (Todd et al., 2014). In older adults, the persistence of conditioned responses in the extinction context indicates an inability to correctly use contextual information to modulate responses to threat. In other words, in older participants, the context appears not able to operate as a gate that disambiguates the CS’s current relation with the US stimulus (Trask et al., 2017; Starosta et al., 2016). Current theorizing in cognitive aging offers a wide variety of accounts for performance decline in context processing and its utilization as occasion setter, including poor distribution of attentional resources (Braver et al., 2001; Hartley, 1992), reduction in working-memory capacity (Just & Carpenter, 1992; Van der Linden et al., 1999) and failure of inhibitory processes (Hasher & Zacks, 1988). These represent distinct but highly interdependent mechanisms that may influence each other (Spencer & Raz, 1995). Importantly, the present study found that the magnitude of the psychophysiological index of extinction recall was positively correlated with accuracy in the attentive matrices test (Fig. 5), a visual search task thought to index selective visual attention (Spinnler & Tognoni, 1987; Della Sala, 1992). That is, individual and age-related differences in selective attention performance predicted subsequent context-dependent recall of extinction memory. Thus, we tentatively suggest that age-related declines in the efficiency of selective attention, possible due to a age-related



reduction in available processing resources (Craik & Byrd, 1982), may lead to weak representation of contextual information and reduced ability to encode the appropriate CS-context relationship, thus promoting overgeneralization of threat responses to many contexts in older adults. These results are consistent with emerging theories that age-related declines in processing contextual information are attributable to poorer selective attention and/or greater inhibitory deficits in older adults (Powell et al., 2018). Additional research is certainly still warranted, however, that directly examines the relationship between selective attention and context dependency of extinction in young and older adults.

Although the present study did not directly investigate the neural substrates of threat conditioning and extinction in aging, deficit of context-guided recall of extinction may be linked to age-related changes in the neural structures underpinning context-dependent behaviour (Foster et al., 2012). Studies in animals support the view that a neural circuit that involves the hippocampus and medial prefrontal cortex is essential for contextual retrieval of threat and extinction memories (Maren, et al., 2013; Maren & Holmes, A., 2016). Consistent with this view, brain imaging studies in humans (Kalisch et al., 2006; Garfinkel et al., 2014) reported that the ventromedial prefrontal–hippocampal network is selectively involved in context-dependent regulation of extinction and threat memories. More specifically, during recall of extinction memory, the medial prefrontal cortex would act to inhibit the amygdala, preventing a response to threat, based on contextual information provided by the hippocampus (Hobin et al., 2003; Kalisch et al., 2006; Delamater, 2004). There is substantial evidence that a number of structural and physiological alterations preferentially influence the prefrontal cortex and medial temporal lobe in advanced aging, even in the absence of disease (Buckner, 2004; Bartzokis et al., 2001; Andrews-Hanna et al., 2007). These disruptive brain changes may underlie impairments in context-dependent extinction recall, as well as cause the decreased efficiency with which older adults use contextual information to determine when and where it

is appropriate to express fear. Additional research will be needed to clarify the underlying neuroanatomical mechanisms of extinction recall and context processing deficits in aging, providing important clues to the pathophysiology of these disorders. Moreover, such data could help to advance our understanding of the neural mechanisms underlying behavioural therapy, such as exposure therapy (Lovibond, 2004; Rauch et al., 2003) aimed at limiting pathological fear.

The results of this study are tempered by a number of limitations. First, the present study used mildly aversive electro-tactile stimulation as US. Since differences in threat learning and extinction may derive from differences in US reactivity, there is the need to replicate these results with a different type of US, for instance, aversive auditory stimuli, such loud noise or complex human scream. Second, extinction recall, and threat renewal were both tested at a single time-point after extinction learning (24 hr later, on Day 2). Future studies should also vary the interval between extinction training and recall/renewal testing, to determine whether aging may interfere with, or simply delay, the consolidation process of extinction memories. Third, we obtained one set of subjective measures following each phase of the study rather than continuous assessment. Online (i.e., trial-by-trial) measures could be used in future studies to provide a more accurate assessment of US expectancy and CS valence during learning and extinction. Note, however, that in older individuals the value of including ratings during the experimental learning phases should be carefully balanced against the possible impact of rating procedures on attention and executive resources, which in turn may affect the time course and strength of threat conditioning (Carter et al., 2003).

In conclusion, the present study documented the influence of normal aging on context-dependent recall of conditioned emotional responses. Contextual processing is especially vulnerable to advanced aging (Spencer & Raz, 1995; Braver et al., 2005; Fogelson, 2015). In line with this, the present data showed that (a) young and older participants were equally able

to acquire and extinguish an autonomic conditioned response to threat, and that (b) older participants failed to modulate such response based on a context-driven retrieval of threat memories, raising the possibility that their extinction recall deficit is a consequence of a more general impairment in using contextual information. This lack of flexible adaptation to contextual cues may play a role in the development of late-onset anxiety disorders (Le Roux et al., 2005), due to neural alterations that normally accompany healthy aging, particularly in the frontal and medial temporal lobes (Lenze et al., 2011). However, there is still a need for studies directly linking together the use of contextual information for flexible responses to threat, and age-related alterations of relevant neural structures underpinning aversive learning and memory processes.



**STUDY 2:**  
**LESION TO THE VENTROMEDIAL PREFRONTAL CORTEX**  
**IMPAIRS THE ACQUISITION OF FEAR CONDITIONING**  
**IN HUMANS**

## **ABSTRACT**

The role of the ventromedial prefrontal cortex (vmPFC) in Pavlovian fear conditioning has been largely attributed to extinction (Phelps et al., 2004) rather than the acquisition of conditioning. However, recent neuroimaging studies have questioned this view by showing the activation of vmPFC also during the acquisition of conditioning (Fullana et al., 2016). The lack of studies on the acquisition of fear conditioning with patients with vmPFC injury does not allow a complete view of this phenomenon. The only existing study with vmPFC patients, reports a preserved skin conductance response (SCR) to the presentation of an image previously associated with an aversive sound, indicating a preserved acquisition of conditioning (Bechara et al., 1999). This evidence must be taken with caution due to the limited sample size and the different etiology. The aim of the present study is to investigate the role of vmPFC in the acquisition of fear conditioning. Ten patients with specific vmPFC lesion, ten healthy participants and ten patients with control brain lesions underwent a classical fear conditioning paradigm, during which SCR was recorded. Unlike healthy participants and control patients, vmPFC patients showed no conditioned response during conditioning acquisition. This effect is not attributable either to differences in the intensity of the US, or to a lack of SCR response to the presentation of the US. Furthermore, vmPFC patients show a preserved explicit awareness of CS-US contingencies. Overall, these results demonstrated vmPFC injury seems to cause more pervasive difficulties than previously theorized. The preserved physiological response to the US and the preserved explicit awareness of the contingency between CS and US suggest that these processes are mediated by different networks from the one responsible for the conditioned psychophysiological responses, which do not require vmPFC. Finally, the deficit in the SCR conditioned response despite the preserved explicit CS-US contingency awareness suggests that declarative memory is not sufficient to produce conditioned responses.

## INTRODUCTION

In humans, the role of the ventromedial prefrontal cortex (vmPFC) in Pavlovian threat conditioning has been largely relegated to the extinction (Phelps et al., 2004) or reversal (Morris & Dolan, 2004) of previously acquired CS-US contingencies. The acquisition of threat conditioning has been mainly imputed to the amygdala (Bechara et al., 1999; LaBar et al., 1998; LaBar et al., 1995). However, recent neuroimaging evidence questions this view by showing activity in the vmPFC also during threat acquisition (Dunsmoor et al., 2019; Fullana et al., 2016), rising the hypothesis for a more prominent role of this region than previously thought during the early stages of conditioning.

The vmPFC has a crucial role in value and stimulus-outcome representation (Hiser & Koenigs, 2018; Schoenbaum et al., 2009; Schoenbaum et al., 2011) as well as model-based computations (Wilson et al., 2014). In fact, the acquisition of Pavlovian threat – fear – conditioning consists in updating the value of encountered stimuli based on changes in stimulus-outcome contingencies. Cognitive and computational models have traditionally described this updating in terms of model free mechanisms, in which prediction errors following the unexpected occurrence of unconditioned stimuli (US) drive the increment of predictions for conditioned stimuli (CS; Rescorla & Wagner, 1972; Sutton, 1988). Nevertheless, recent evidence shows the recruitment of model-based mechanisms to represent Pavlovian contingencies (Pauli et al., 2019; Prévost, et al., 2013), suggesting that also an abstract representation of the underlying structure of the Pavlovian contingencies is taken into account when computing predictions for CS, at least in humans.

The possibly neglected role of vmPFC in threat acquisition may have been reinforced also by a lack of neuropsychological studies assessing the consequences of a lesion to the vmPFC on the acquisition of threat conditioning. In fact, the only existing study found preserved conditioned skin conductance response (SCR) to the presentation of an image

previously associated with an aversive sound, despite patients with vmPFC lesion failing to show anticipatory SCR to negative (and positive) outcomes during a gambling task (Bechara et al., 1999). Nevertheless, this evidence should be taken cautiously because of the limited sample size (n=5) and the diverse aetiology of the lesion among patients. For example, one patient had a frontal cyst, which developed during childhood and was never removed, leaving room for the development of compensatory neural circuits. Additionally, the acquisition of conditioning was evaluated as an increase in SCR to a CS+ during acquisition as compared to habituation and extinction, lacking the comparison of SCR to a control stimulus (CS-). In fact, given that the vmPFC has been hypothesized to have a role in discriminating threat from safety (Fullana et al., 2016), the impairment in the acquisition of threat conditioning in patients with vmPFC lesion may consist in a failure in generating differential SCR between CS+ and CS-, rather than an overall failure in increasing SCR to a CS+ during acquisition.

The present study aims to reevaluate the role of the vmPFC in the acquisition of Pavlovian threat conditioning. Ten patients with a bilateral lesion to vmPFC, a group of healthy participants and of patients with a lesion outside PFC or medial temporal lobe completed a differential threat conditioning paradigm, while their SCR to CS+, CS- and US was recorded. Explicit awareness of CS-US contingencies was also assessed. Impaired conditioned responses during threat acquisition would indicate that, in humans, the vmPFC plays a causal role in this task.



## METHODS

**Participants.** Thirty right-handed adults participated in the study, equally divided into three groups, namely patients with lesion to the ventromedial prefrontal cortex (vmPFC), brain-damaged control patients (BDC) with a lesion that did not involve the PFC or medial temporal lobe, and healthy controls (HC, see Table 1 for demographic and clinical information). Groups were matched in terms of sex, education, illness chronicity, and neuropsychological assessment scores (see Table 1). BDC and HC differed in terms of age; therefore, this variable was used as covariate in all analyses (see Table 1).

Participants were recruited at the Center for Studies and Research in Cognitive Neuroscience of Bologna University, where the study was conducted. They had normal or corrected-to-normal vision and were naïve to the purposes of the study. All participants gave informed written consent to participation after being informed about the procedure of the study. The study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki and approved by the Bioethics Committee of the University of Bologna.

<b>Demographics</b>							
	<b>vmPFC</b> (n = 10)	<b>BDC</b> (n = 10)	<b>HC</b> (n = 10)	<b>F</b>	<b>df</b>	<b>p</b>	<b><math>\eta_p^2</math></b>
Age ( <i>yrs</i> )	58.90 ± 6.62	51.10 ± 14.37	67.90 ± 7.87	6.44	2, 16.9	0.01 <sub>+</sub>	0.33
Education ( <i>yrs</i> )	11.00 ± 4.19	13.30 ± 4.37	13.00 ± 2.45	1.09	2, 27	0.34	0.07
Chronicity ( <i>yrs</i> )	5.00 ± 4.74	2.80 ± 2.20	-	1.77	1, 12.7	0.21	0.09
Shock ( <i>mA</i> )	5.67 ± 2.58	7.45 ± 1.99	7.96 ± 2.94	0.81	2, 27	0.45	0.06
Sex ( <i>m/f</i> )	6/4	6/4	4/6	$X^2_{(2)} = 1,07$	2, 27	0.58	-

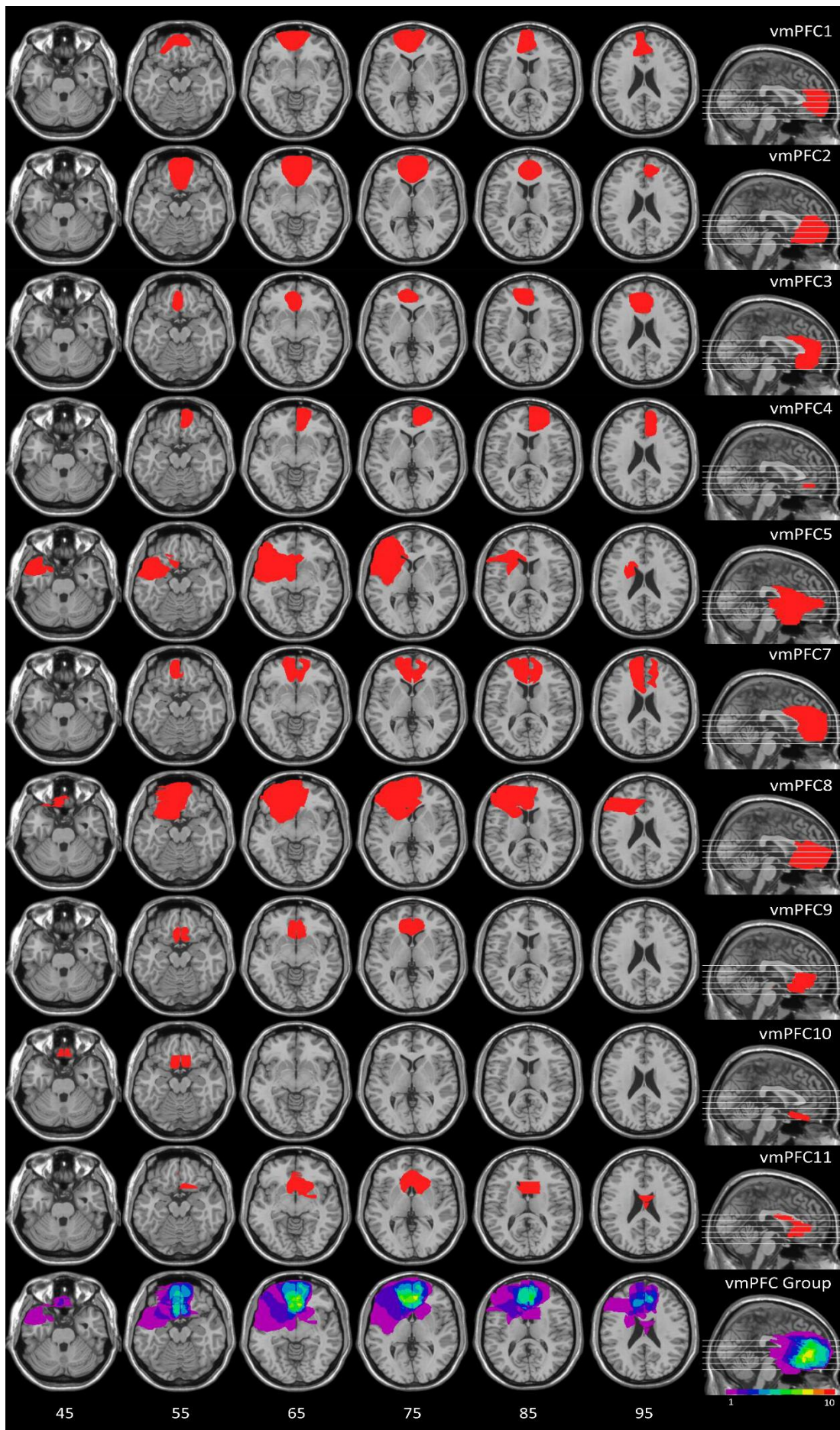
<b>Neuropsychological assessment</b>							
	<b>vmPFC</b> (n = 9)	<b>BDC</b> (n = 6)	-	<b>F</b>	<b>df</b>	<b>p</b>	<b><math>\eta_p^2</math></b>
Raven Progressive Matrices*	3.55 ± 0.72	3 ± 1.09	-	1.41	1, 13	0.25	0.09
Stroop Test*	2 ± 1.93	2.5 ± 1.04	-	0.41	1, 12.6	0.53	0.02
Towers of London#	100.66 ± 14.52	104 ± 13.56	-	0.19	1, 13	0.66	0.01
Digit Span*	3.4 ± 1.26	3.16 ± 1.60	-	0.11	1, 13	0.75	0.01
Phonemic Fluency*	3.11 ± 1.16	3.16 ± 1.32	-	0.01	1, 13	0.93	0.00 1
Semantic Fluency*	3.44 ± 0.72	3.66 ± 0.81	-	0.31	1, 13	0.59	0.02

**Table 1. Demographic characteristics and neuropsychological assessment for all groups.** All measures are reported as mean ± standard deviation, except for Sex, reported as frequency. vmPFC = ventromedial prefrontal cortex lesion group; BDC = brain-damaged control group; HC = healthy control group. \*Equivalent score. #Standard score. Post-hoc analysis reported a significant difference between BDC and HC  $p=0.003$  only, all other  $p>0.18$ .

**Lesion.** Lesion aetiologies for the vmPFC group were aneurysm of the anterior communicating artery (n=8), aneurysm of the anterior cerebral artery (n=1) and meningioma of the anterior cranial fossa, (n=1). Lesion aetiologies for the BDC group were Occipital Meningioma, Cerebellar Stroke, Vascular Encephalopathy, Occipital MAV, Brain Abscess, Posterior Artery Stroke, Basal Brain Abscess, Brain Anoxia, HSV Encephalopathy, and Occipital Stroke.

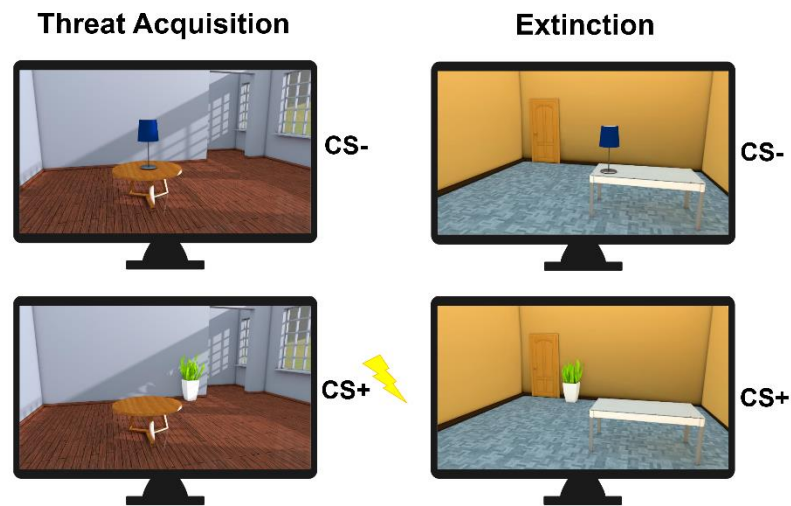
For each vmPFC patient, lesion extent and location were documented using the most recent clinical computerized tomography (CT; n=8) or magnetic resonance imaging (MRI; n=8). Lesions were traced directly on each slice of the normalized T1-weighted template MRI scan from the Montreal Neurological Institute (Holmes et al., 1998). This template is approximately oriented to match Talairach space and distributed with MRICro (Holmes et al.,

1998). MRIcro was used to estimate lesion volume (in cc) and generate brain lesion overlap images. Figure 1 shows the extent and overlap of brain lesions in vmPFC patients. In the vmPFC group, the lesion included Brodmann's areas (BA) 11 and 25 for all patients, and also BA 47, 10, 24, 32 for 9 patients. The region of maximal overlap occurred in BA 11 (M = 18.54 cc, SD=12.41), BA 10 (M = 9.51 cc, SD =8.07), BA 32 (M = 5.67 cc, SD =5.35), BA 25 (M = 3.48 cc, SD=3.12), BA 47 (M = 3.20 cc, SD =5.92).



**Figure 1. Extent and overlap of brain lesions.** The figure represents vmPFC patients' lesions projected on the axial slices of the standard Montreal Neurological Institute brain. The white horizontal lines on the sagittal view are the positions of the axial slices, and the white numbers under the axial views are the z-coordinates of each slice. The colour bar indicates the number of overlapping lesions. Maximal overlap occurs in BA 11, BA 10, BA 32, BA 25, BA 47.

**Pavlovian conditioning task.** The conditioned stimuli (CSs) were images of 2 different objects (i.e. a lamp and a plant), which appeared embedded in an indoor scene for 4s (Fig. 2, previously used in Battaglia et al., 2018). Their presentation was followed by a jittered 14-17s intertrial interval (ITI), consisting in a grey screen. Trial order was pseudo-randomized, such that no more than three identical CSs occurred in a row. To reduce initial orienting responses, the experiment began with a short habituation phase, which included two presentations of each CS (data not reported). Then acquisition included 40 trials (20 CS+, 20 CS-). Presentation of the threat conditioned stimulus (CS+, object assignment counterbalanced between subjects) co-terminated with the delivery of a 200ms aversive electrical shock (i.e. unconditioned stimulus, US) in 12 out of 20 trials (60% reinforcement rate). The shock was generated by a Digitimer Stimulator (Model DS7, Digitimer Ltd., UK) and delivered to the participants' left inner wrist through pre-gelled snapped electrodes. Presentation of the other object was never paired with the US, representing the within-subject control stimulus (CS-). Acquisition always started with a presentation of a CS- and of a reinforced CS+, in random order. Extinction included 40 trials (20 CS+, 20 CS-), during which CS presentation was never paired with the US.



**Figure 2. Schematic representation of the experimental design.** Participants performed a differential fear conditioning task in which threat acquisition and extinction were established. During threat acquisition, the conditioned stimulus (CS+) was associated with a shock pulse on 60% of trials, while the CS- was not associated with any consequence. Extinction followed this phase, during which both CSs were presented and none of them was associated with the shock pulse.

**Skin conductance response (SCR).** SCR was recorded continuously during the task from two Ag/AgCl electrodes (TSD203Model; Biopac Systems, USA) filled with isotonic hyposaturated conductant and attached to the distal phalanges of the second and the third finger of participants' left hand. The signal was amplified with a DC amplifier (Biopac EDA100C) with a gain factor of  $5\mu\text{S}/\text{V}$  and a low-pass filter of 10Hz. The analog signal was digitalized at 200Hz using a Biopac MP-150 digital converter and fed into AcqKnowledge 3.9 (BIOPAC Systems, Inc., Goleta, California) for offline analysis. The digitalized signal was analyzed using custom-made MATLAB scripts to obtain peak-to-peak SCR values, operationalized as the peak-to-peak amplitude of the largest deflection occurring between 0.5-4.5 s after event onset. The minimum response criterion was  $0.02\mu\text{S}$ , and smaller responses were encoded as zero. Raw SCR scores were square root transformed to normalize the data distribution (Siddle et al., 1988).

SCR following the US was analyzed to assess unconditioned responding, whereas SCR following the CS was analyzed to assess conditioned learning. Regarding SCR to the US, stimulus onset was represented by the time of shock administration; regarding SCR to CSs, stimulus onset referred to the time of CS appearance on the screen. To reduce interindividual variability, SCR was scaled to each participant's mean square-root-transformed US response (Schiller et al., 2008). Thus, raw scores were range corrected by dividing each individual score by the subject's mean SCR response to US (Lykken, 1972). This procedure can reduce error variance, thus increasing statistical power when comparing groups of participants. In this way, conditioned responses can be directly compared across groups without confounding baseline differences in skin conductance levels (LaBar et al., 2004). To assess conditioned responses to the CS+, only non-reinforced CS+ trials were analyzed (n=8). Regarding the response to the US, mean SCRs to the 12 shocks were analyzed.

**Explicit contingency awareness.** To assess the explicit awareness of CS-US contingencies, at the end of the conditioning task, participants were presented with one CS at the time and asked to indicate whether or not the stimulus was associated with a shock during the task. A score of one was given for a correct answer and of zero for an incorrect answer.

**Procedure.** Participants were seated comfortably in a silent and dimly lit room, and their position was centered relative to the computer screen, at 100cm viewing distance. Electrodes for SCR recording and for shock administration were attached to the participant. After verifying correct recording of SCR, shock intensity was adjusted using a standard workup procedure. It was initially set at 0.5mA and increased of 1mA until participants reported it as "highly annoying, but not painful". Participants were asked to remain as quiet and still as possible during the task and to keep their attention at the center of the screen. At the beginning of the

conditioning task, participants were informed that different visual stimuli would appear on the screen, and some of them might be paired with the shock. Participants received the instruction to carefully observe and learn which stimuli were paired with the shock. Participants were also informed that their behavior had no effect on shock delivery. At the end of the conditioning task, CS-US contingency awareness was tested.

**Neuropsychological assessment.** vmPFC and BDC patients were given a series of standardized neuropsychological tests. The selected neuropsychological battery included tests: Raven Progressive Matrices (Spinnler & Tognoni, 1987), Stroop Test (Caffarra et al., 2002), Tower of London (Kennedy et al., 2000), Digit Span (Orsini et al., 1987), Phonemic and Semantic Fluency (Novelli et al., 1986). Normative scores derived from a nationally representative sample of adults are available for each test. For Tower of London, patients' raw scores were converted in Standard scores (Culbertson & Zillmer, 1999). For all the other tests, patients' raw scores were converted into equivalent scores (Capitani & Laiacona, 1988), adjusted for age and years of education. Equivalent score is a 5-point scale, ranging from 0 to 4, with 0 = pathological performance, 1 = borderline performance, 2–4 = normal performance. T-tests in Table 1 showed no significant difference between vmPFC and BDC on any of the tests (all  $p > 0.25$ ).

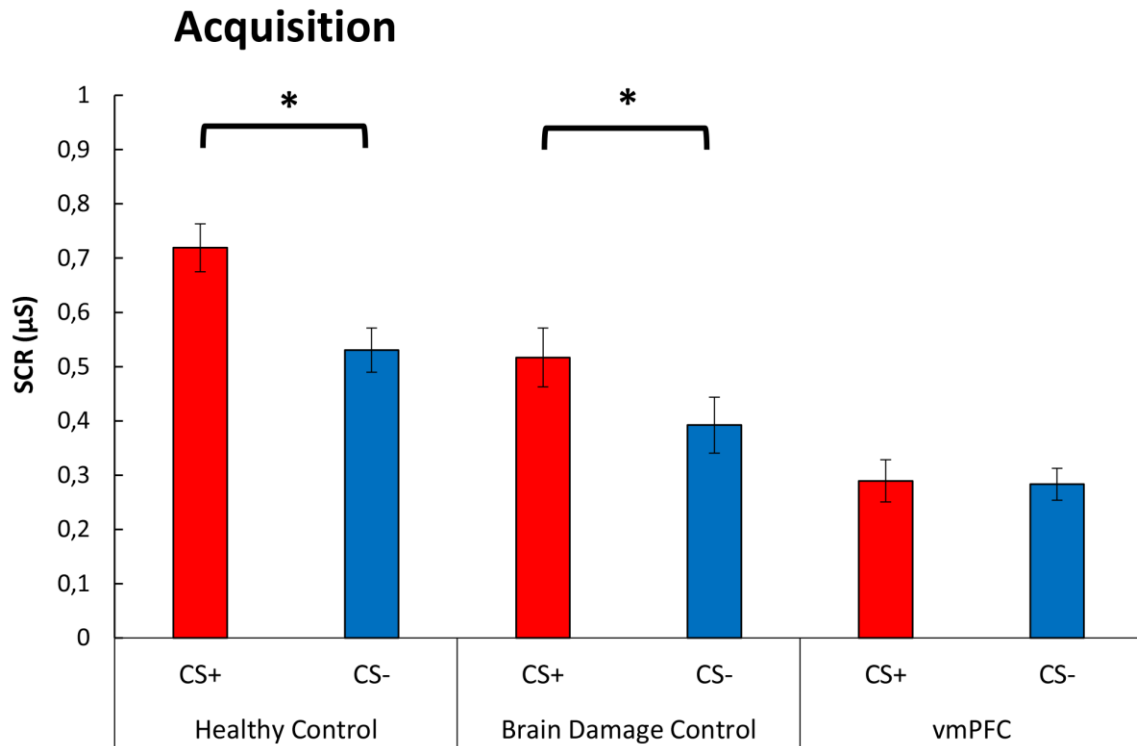
**Experimental design and statistical analysis.** Statistical analyses were performed using JASP (JASP Team 2019, Version 0.11.1). Assumptions for a correct use of parametric analyses were assessed. When homogeneity and sphericity assumptions were violated Welch or Greenhouse-Geissers corrections were applied, and corrected p value and degrees of freedom (df) reported. The significance threshold was  $p < 0.05$  and post-hoc analyses were conducted with Tukey HSD test.



Reported p-values are always intended as two-tailed. Since BDC and HC differed in terms of age (see Table 1), this variable was used as covariate in all ANOVAs. In all analyses the covariate age did not have any statistically significant influence (all  $p \geq 0.325$ ).

## RESULTS

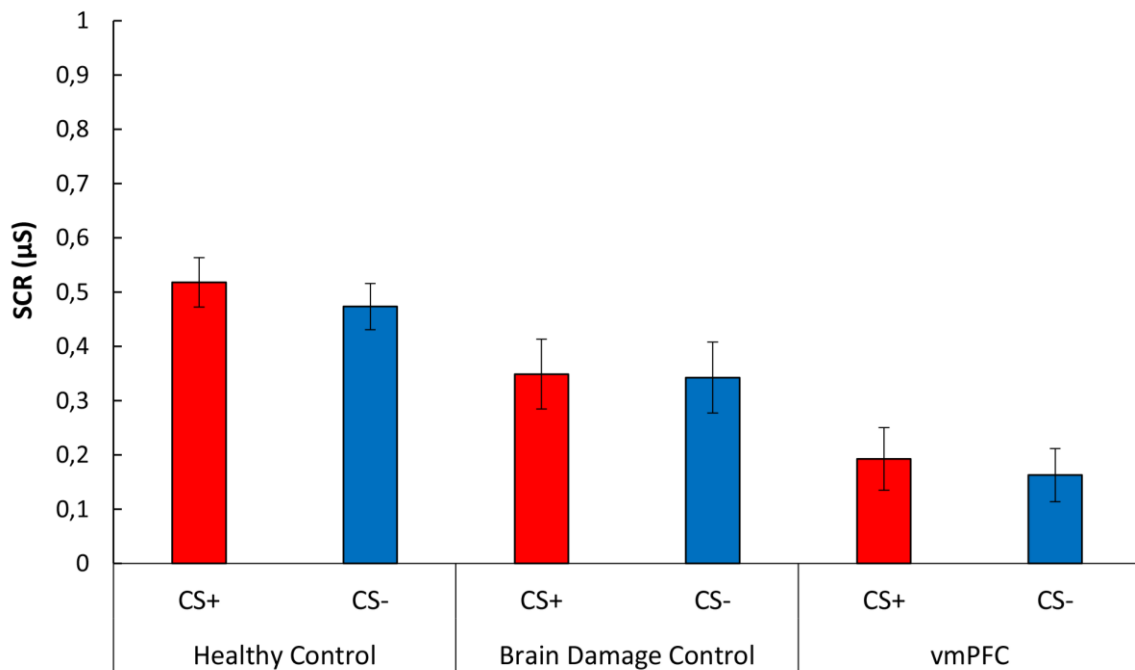
**Acquisition.** It has assessed whether lesion to vmPFC affects threat acquisition with a 2x3 mixed design ANCOVA (Stimulus: CS+/CS-; Group: vmPFC/BDC/ HC). There was no significant main effect of stimulus ( $p=0.399$ ). Instead, there was a significant main effect of group ( $F(2,26) = 14.74, p < 0.001, \eta p^2 = 0.53, 95\% \text{ CI } [.21, .73]$ ), qualified by a significant stimulus by group interaction ( $F(2,26) = 13.27, p < 0.001, \eta p^2 = 0.51, 95\% \text{ CI } [.18, .71]$ ). HC and BDC responded more to CS+ than CS- (Controls: CS+:  $M = 0.72\mu\text{S}$ , CS-:  $M = 0.53\mu\text{S}$ ;  $p < 0.001$ ; BDC CS+:  $M = 0.52\mu\text{S}$ ; CS-:  $M = 0.39\mu\text{S}$ ;  $p < 0.001$ ). In contrast, vmPFC did not show any significant difference in response between CS+ and CS- (CS+:  $M = 0.29\mu\text{S}$ ; CS-:  $M = 0.28\mu\text{S}$ ;  $p = 1.000$ ; Fig. 3). While HC and BDC patients showed successful acquisition of threat conditioning, vmPFC patients were impaired in the acquisition of threat conditioning (Fig. 5).



**Figure 3. Acquisition of threat conditioning.** Graphs illustrate mean skin conductance responses (SCRs) to the conditioned (CS+) and neutral (CS-) stimuli for each experimental group during the Acquisition phase. Data demonstrated that both HC and BDC groups showed successful acquisition of threat conditioning. In contrast, only vmPFC patients did not show any significant difference in response between CS+ and CS-. \* denotes significant comparisons ( $p < 0.05$ ) and error bars represent standard error.

**Extinction.** During extinction the 2x3 mixed design ANCOVA (Stimulus: CS+/CS-; Group: vmPFC/BDC/ HC) showed only a main effect of group ( $F(2, 26) = 8.22, p = 0.002, \eta^2 = 0.39, 95\% \text{ CI } [0.08, 0.62]$ ). VmPFC responded significantly less than HC (vmPFC:  $M = 0.17\mu\text{S}$ , Controls:  $M = 0.49\mu\text{S}$ ;  $p < 0.001$ ) but not than BDC (BDC:  $M = 0.34\mu\text{S}$ ;  $p = 0.060$ ; Fig 4). There was no significant difference between HC and BDC ( $p = 0.101$ ). No other main effects or interactions were significant (all  $p \geq 0.332$ ). Therefore, HC and BDC had successful extinction of threat conditioning (Fig. 5).

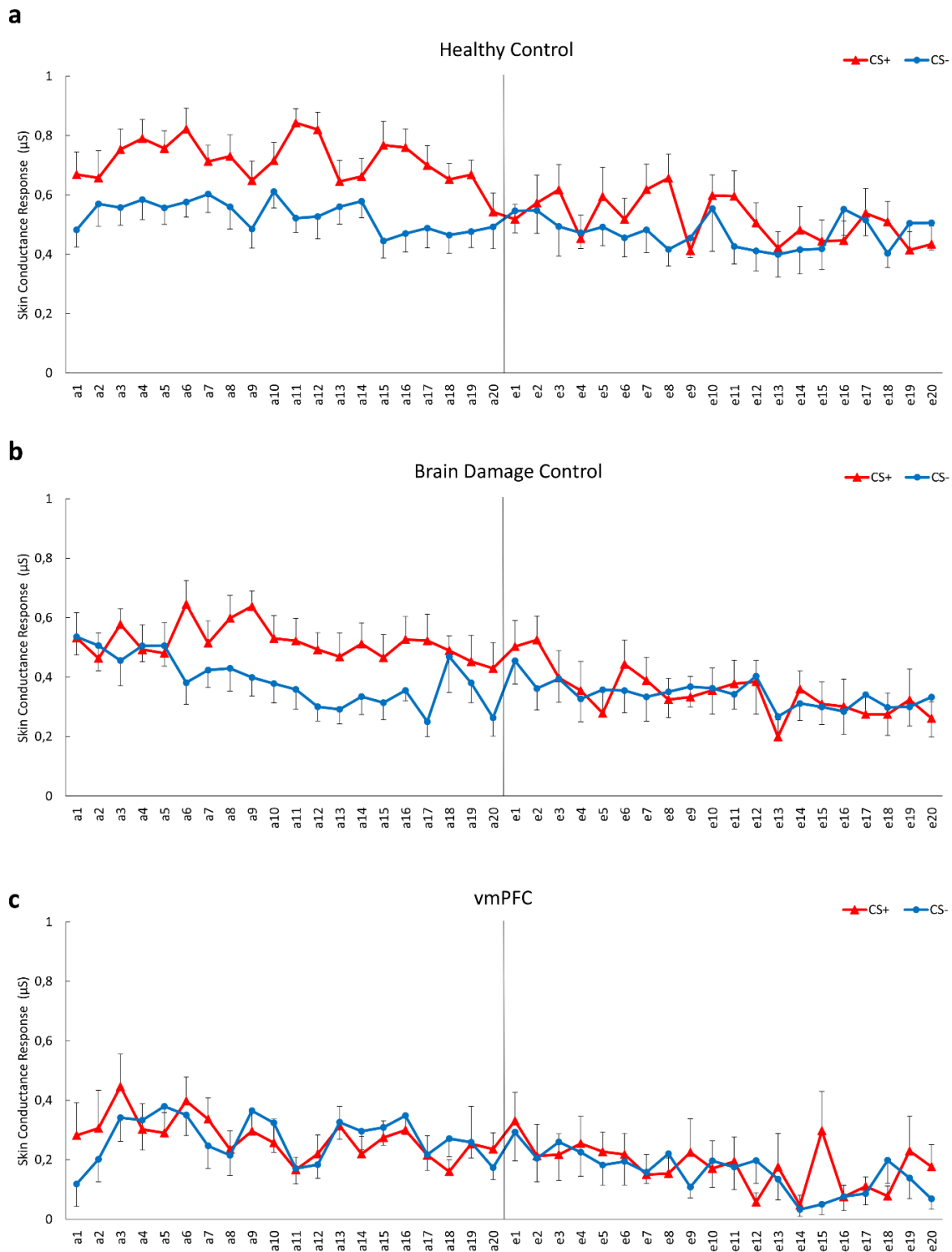
## Extinction



**Figure 4. Extinction of threat conditioning.** Graphs illustrate mean skin conductance responses (SCRs) to the conditioned (CS+) and neutral (CS-) stimuli for each experimental group during the Extinction phase. Data demonstrated that both, HC and BDC had successful extinction of threat conditioning. \* denotes significant comparisons ( $p < 0.05$ ) and error bars represent standard error.

**US intensity and SCR to the US.** We then tested the possibility that differences in threat acquisition between groups may have resulted from differences in intensity of delivered US and SCR to the US. The one-way ANCOVA on mean US intensity showed no significant effect of group ( $F(2, 26) = 1.17, p = 0.325, \eta^2 = 0.08, 95\% \text{ CI } [0.00, 0.31]$ ; vmPFC:  $M = 6.52\text{mA}$ , BDC:  $M = 7.44\text{mA}$ , HC:  $M = 7.95\text{mA}$ ). Similarly, the one-way ANCOVA on mean SCR to the US showed no significant effect of group ( $F(2, 26) = 1.71, p = 0.200, \eta^2 = 0.12, 95\% \text{ CI } [0.00, 0.35]$ ; vmPFC:  $M = 0.75\mu\text{S}$ , BDC:  $M = 1.05\mu\text{S}$ , HC:  $M = 1.01\mu\text{S}$ ). These results suggest that lesion to vmPFC does not significantly affect the unconditioned response.

**Explicit CS-US contingency awareness.** Finally, we tested whether groups differed in the awareness of CS-US contingency, as this may play a role in the acquisition of conditioned responses (Weike et al., 2005). The Chi-squared test showed no differences in the frequency of correct and wrong responses to either CS+ ( $\chi^2(2) = 0.001$ ,  $p = 1.000$ ) or CS- ( $\chi^2(2) = 0.952$ ,  $p = 0.621$ ) between the three groups.



**Figure 5. Trial by trial responses.** Graphs illustrate mean skin conductance responses (SCRs) to the conditioned (CS+) and neutral (CS-) stimuli for each experimental trial divided by group. In panel ‘a’ is depicted Healthy Control group, in panel ‘b’ Brain Damage Control group and in panel ‘c’ ventromedial prefrontal cortex group. Error bars represent standard errors.

## **DISCUSSION**

The aim of the present study was to reevaluate the role of vmPFC in the acquisition of Pavlovian threat conditioning. To this end, SCR to CS+, CS- and US and verbal reports of CS-US contingencies were recorded in a differential threat conditioning paradigm in ten patients with a bilateral lesion to vmPFC, a group of healthy participants and a group of patients with a lesion outside the PFC or the medial temporal lobe. Results show patients with a lesion to vmPFC fail to produce conditioned responses during threat acquisition, as evidenced by impaired anticipatory differential SCR between CS+ and CS- in vmPFC patients compared to healthy participants and control patients. In contrast, there was no evidence that lesion to vmPFC compromised unconditioned responses or declarative memory for CS-US contingencies. Groups did not differ significantly in the psychophysiological reactivity to the US itself nor in the scores for explicit CS-US contingency awareness.

Our results indicate that, in humans, the vmPFC is necessary since the early stages of threat conditioning, namely acquisition, in addition to extinction and reversal. Thus, a lesion to vmPFC results in more pervasive impairments than previously thought, occurring not only when facing a change in previously acquired contingencies, but already at the early stages of their acquisition. Additionally, the preserved unconditioned response and declarative memory of stimulus-outcome contingencies suggest that these two processes do not rely on a functional vmPFC and that they are mediated by different pathways than the one generating conditioned responses. Finally, preserved declarative memory does not appear sufficient to generate conditioned responses, as suggested by lack of conditioned responses despite awareness of stimulus-outcome contingencies.

Our main finding is that a lesion to vmPFC impairs conditioned physiological responses during the acquisition of threat conditioning. Given the role of vmPFC in value and stimulus-outcome representation (Hiser & Koenigs, 2018; Schoenbaum et al., 2009; Schoenbaum et al.,

2011), the vmPFC may be necessary to encode or learn the value of CS during acquisition. Although the sensory information about CS and US would still be able to converge into the amygdala, where stimulus-outcome associations are passively formed, the vmPFC may be necessary to turn this information into expectations regarding the value of future events (Roesch & Schoenbaum, 2006), thus enabling anticipatory responses. This idea seems also supported by recent evidence showing that Pavlovian learning is not passively driven by stimulus-reinforcement mechanisms but also involves model-based computations, occurring in vmPFC and amygdala, to create a detailed map of task space (Pauli et al., 2019; Prévost et al., 2013). Working in strict interplay with the amygdala, through their reciprocal connections (Krettek & Price, 1977; Cassell & Wright, 1986), the vmPFC may be causally involved in updating the value representation of the CS, in turn promoting the generation of the anticipatory physiological response usually observed during acquisition. This interpretation also suggests that in humans the vmPFC may have a more prominent role in threat conditioning than in rodents, where its lesion does not impair acquisition (Morgan et al., 1993; Morgan & LeDoux, 1995; Morrow et al., 1999; Quirk et al., 2000; Weible et al., 2000; Lebron et al., 2004; although see Fryszak & Neafsey, 1991, 1994). In fact, in rodents, the vmPFC appears necessary for the expression of conditioned responses rather than the initial encoding/learning of CS value. Inactivation of vmPFC prior to threat conditioning reduces conditioned responses during acquisition but does not impair responses at test on the following day (Corcoran & Quirk, 2007; Sierra-Mercado et al., 2006). In addition, recordings from prelimbic neurons show sustained conditioned response during acquisition that correlated with sustained freezing behavior (Burgos-Robles et al., 2009). Thus, it may be possible that the failure of patients with vmPFC lesion to produce conditioned responses indicates that the vmPFC is necessary for the expression of conditioned responses, rather than the encoding/learning of CS value. In particular, the vmPFC may receive amygdala inputs carrying information about CS value and

transform them into a sustained output that drives the expression of conditioned physiological responses.

Regardless of whether the vmPFC has a causal role in the acquisition of threat conditioning (i.e. encoding/learning CS value) or in its expression, the current results call for a reevaluation of the role of vmPFC in threat conditioning in humans, which warrants further research. For example, inducing a virtual lesion of vmPFC (e.g. through neurostimulation) that is temporally limited to the acquisition phase has the potential to clarify the role of this region during threat acquisition. The appearance of conditioned response in a subsequent recovery test would indicate that the vmPFC is necessary for the expression of conditioning. On the contrary, lack of such response would indicate the causal role of vmPFC in encoding/learning CS value. In addition, given the extended maturational trajectory of vmPFC over the lifespan (Gogtay et al, 2004; Lenroot & Giedd, 2006; Raznahan et al, 2014) and the age of our participants, it could be wondered whether or not a lesion to vmPFC at an earlier developmental stage also disrupts threat acquisition. Several studies have shown that the quality of threat conditioning acquisition changes over the course of development (Hartley & Lee, 2015), nevertheless the extent of vmPFC involvement in these changes remains unknown. The response pattern of vmPFC patients also seems to resemble that of patients with a bilateral lesion to the amygdala. In both cases, the lesion results in deficient conditioned but preserved unconditioned SCR during threat acquisition and declarative memory of CS-US contingencies. Nevertheless, differently than patients with unilateral amygdala lesion, where impaired conditioned response seems to partly depend on concurrent difficulties in declarative memory of CS-US contingencies (Weike et al., 2005), here, a bilateral lesion to vmPFC impaired conditioned response despite spared declarative memory.

The preserved unconditioned response is in line with evidence that inactivation of prelimbic cortex in rats does not abolish US response (Corcoran & Quirk, 2007) and adds to studies on



patients with amygdala lesions (Bechara et al., 1999, 1995) indicating that unconditioned physiological response is mediated by a different pathway than conditioned response, which does not require an intact vmPFC. In addition, the failure of generating conditioned responses despite preserved explicit awareness of stimulus-outcome contingencies confirms the idea that declarative memory is not sufficient to produce conditioned responses, as was the case of patients with amygdala lesions (LaBar et al., 1995). However, other studies have shown that in contrast with this, with a bilateral vmPFC lesion SCR was impaired despite preserved declarative memory of CS-US contingencies.



**STUDY 3:**  
**STATE-DEPENDENT TMS OVER PREFRONTAL CORTEX**  
**DISRUPTS MEMORY RECONSOLIDATION AND**  
**PREVENTS THE RETURN OF FEAR**

## **ABSTRACT**

Erasing maladaptive memories has been a challenge for years. A way to change emotional (e.g., fear) memories is to target the process of reconsolidation, during which a retrieved memory transiently returns to a labile state, amenable to modification (Bernard et al., 2012; Elsey et al., 2018). Disruption of human fear memories has been classically attempted with pharmacological (Kindt et al., 2014), or behavioral (e.g., extinction; Schiller et al., 2009) treatments which, however, do not clarify the underlying brain mechanism. To address this issue, here, we combined a fear conditioning paradigm with repetitive transcranial magnetic stimulation (rTMS) administered in a state-dependent manner. In a critical condition, we targeted the human dorsolateral prefrontal cortex (dlPFC) following presentation of a fear reminder that reactivated a fear memory acquired one day before. Twenty-four hours later, participants exhibited decreased physiological expression of fear, as shown by their skin conductance response (SCR). Similar reductions were observed when targeting the left and the right dlPFC. In striking contrast, no such decrease was observed in participants receiving either control rTMS (i.e., stimulation of a control site and sham stimulation), or dlPFC-rTMS without preceding reactivation of a fear memory (No-reminder), thus showing both the site-specificity and state-dependency of our rTMS intervention. Indeed, expression of fear was reduced only when dlPFC-rTMS was administered within the reconsolidation time-window (i.e., following memory reactivation). Moreover, only dlPFC-rTMS prevented subsequent return of fear after extinction training. These findings highlight the key role of the dlPFC in reconsolidation of fear memory, and suggest that rTMS can be safely used in humans to prevent the return of fear.

## INTRODUCTION

Since the dawn of psychology, psychologists and psychiatrists have tried to change undesired emotional memories. The result is that, even the most effective treatments, only eliminate fearful responding, leaving the original fear memory intact (Bouton, 2002). From an evolutionary perspective, it is extremely functional to never forget the most important events in life. However, the putative indelibility of emotional memory can also be maladaptive, such as in some trauma victims who suffer from dreadful memories and anxiety. In contrast to the traditional view of memory formation as a one-time process of consolidation (Squire and Davis, 1981), the reconsolidation hypothesis suggests that stored information is rendered labile after being retrieved and it raised the possibility of interfering with existing memories during a temporary window of lability. Thus, reconsolidation is defined as a process whereby previously consolidated memories can be reactivated and again make sensitive to mutate (Nader et al. 2000). Reconsolidation process can be influenced by neurobiological manipulations during or shortly after the memory reactivation period (Tronson & Taylor, 2007).

Recently, it was discovered that consolidated memories can re-enter unstable states when they are reactivated during retrieval or by a reminder cue and need to consolidate again in order to persist over longer periods of time (Nader and Hardt, 2009). Thus, the concept of 'reconsolidation' assumes that memories are not unchangeable consolidated once and for all. Indeed, memory reactivation may trigger reconsolidation process in a time-limited period. Kind and colleagues have suggested that permanent reduction of fear may be achieved through the pharmacological targeting the process of reconsolidation during which memories are activated into a labile state and can be re-stored in an altered form (Kindt et al., 2009). Pharmacological manipulations at this stage result in an inability to retrieve the memories at later times, suggesting that they are either erased or persistently inhibited (Nader and Hardt, 2009). Although these pharmacological manipulations are potentially useful for changing learned

fears, their use in humans can be problematic. Also, pharmacological effects have been not consistently replicated (Wood et al., 2015). Beside pharmacological interventions, various non-invasive methods have also been studied to interfere with fear memories. Schiller and colleagues (2010) used the reconsolidation window to behaviorally trigger reconsolidation process for fear memories with non-fearful information (i.e. extinction protocol; Schiller et al., 2010). Also, Burger and coworkers (2016) showed that tVNS (transcutaneous vagus nerve stimulation) accelerates explicit fear extinction, however, it did not lead to better retention of extinction memory 24 h later. Indeed, a common problem with most of these reconsolidation procedures is the lack of replicability and consistent long-term effects which are fundamental to treat psychiatric disorders. One potential way to disentangle these issues would be using non-invasive brain stimulation, as the Transcranial Magnetic Stimulation (TMS; Rossini et al., 2015), which would offer a safety and non-invasive tool to modulate brain activity in target regions known to have a critical role in the reconsolidation process, such as the dorsolateral prefrontal cortex (dlPFC) (Delgado et al., 2008). In particular, dlPFC has direct connection with the amygdala (Sladky et al., 2013), a key region for emotional learning process. Importantly, if emotional memory could be weakened or even erased, then we might be able to eliminate the root of many psychiatric disorders. Indeed, it would be particularly significant in pathologies in which intrusive traumatic memories significantly affect daily life, as in the case of Posttraumatic Stress Disorders (PTSD) or Generalized Anxiety Disorder (GAD).

Disrupt the neural activity of the dlPFC by means of the TMS during the reconsolidation window may seem to be the most promising tool to alter the state of consolidated maladaptive fear memories. To date, three studies targeting different processes have demonstrated that non-invasive brain stimulation over PFC can affect fear memories in humans. Van 't Wout and colleagues in 2016 stimulated the left dorsolateral prefrontal cortex during extinction learning, and reported enhanced subsequent extinction of conditioned fear. Asthana and coworkers

(2013) reported that the stimulation of the left DLPFC resulted in inhibition of fear memory consolidation. Previously, it has been shown that the stimulation of prefrontal cortices during the reconsolidation window resulted in enhancement of fear memories (Munjee et al., 2014). Although these studies have demonstrated the possibility to manipulate fear memories, none of these studies tested whether fear is persistently eliminated.

In the present study, using an experimental protocol of fear conditioning procedure and return of fear in different days (Schiller et al., 2010), will be tested whether the application of the rTMS over the right and left dlPFC during the reconsolidation window is able to impact the reconsolidation of the fear memory. In particular, the paradigm consists in a 3 days experiment: fear acquisition (day 1), memory reactivation and rTMS (day 2), and memory recall followed by extinction and reinstatement procedure (day 3). rTMS will be applied on day 2 in order to interfere with the reconsolidation process differently for each experimental group. It is expected that rTMS administered over both left and right dlPFC, would interfere the reconsolidation process of fear memory after memory reactivation.

## METHODS

**Participants.** Seventy healthy volunteers took part in the study. Participants were randomly assigned to one of five experimental groups: CtrlSham (14 participants, 8 females, mean age  $\pm$  SD:  $23.2 \pm 1.8$ ), CtrlOccipital (14 participants, 9 females,  $24.4 \pm 3.1$ ), No-reminder (14 participants, 11 females,  $21.6 \pm 2.0$ ), right-dIPFC (14 participants, 9 females,  $23.1 \pm 2.6$ ), and left-dIPFC (14 participants, 8 females,  $23.9 \pm 2.3$ ). All subjects were right-handed, had normal or corrected-to-normal visual acuity in both eyes, and were naïve about the purposes of the experiment. None of the participants had neurological, psychiatric or other medical problems, nor any contraindication to TMS (Rossi et al., 2009; Rossini et al., 2015). Participants provided written informed consent. The procedures were approved by the University of Bologna Bioethics Committee and were in accordance with the ethical standards of the 1964 Declaration of Helsinki (World Health Organisation, 2013). No discomfort or adverse effects of TMS were reported by participants or noticed by the experimenter. It is widely known that anxiety and depression may affect the skin conductance response (SCR) in classical fear conditioning (Duits et al., 2015). To account for such variability, levels of anxiety were measured by means of trait-anxiety scores using form Y2 of the State and Trait-Anxiety Inventory (STAI-Y2; Duits et al., 2015), whereas depression was assessed by means of the Hospital Anxiety and Depression Scale (Spielberger, 1983). A one-way ANOVA showed no significant effect of group on anxiety level ( $F_{4,65} = 1.71$ ,  $p = 0.16$ , CtrlSham, mean  $\pm$  SD:  $46.7 \pm 10.0$ ; CtrlOccipital,  $42.9 \pm 7.8$ ; No-reminder,  $45.6 \pm 9.0$ ; right-dIPFC,  $40.2 \pm 8.0$ ; left-dIPFC,  $39.6 \pm 11.1$ ), or depression ( $F_{4,65} = 1.43$ ,  $p = 0.23$ ; CtrlSham,  $3.1 \pm 2.2$ ; CtrlOccipital,  $2.9 \pm 1.4$ ; No-reminder,  $4.3 \pm 3.2$ ; right-dIPFC  $3.1 \pm 3.1$  and left-dIPFC,  $2.5 \pm 2.7$ ).



**Materials.** The study was implemented in Matlab R2016 software (The MathWorks, Inc., Natick, Massachusetts, United States) and stimuli presentation and shock administration were controlled by PsychToolbox (Zigmond & Snaith 1983), running on a Windows-based PC (Lenovo ThinkCentre Desktop Computer). Stimuli were created with Blender (Blender Foundation, Amsterdam, Netherlands) and Cinema 4D R17 software (MAXON Computer GmbH, Friedrichsdorf, Germany), and were presented on a computer screen (screen size: 43 inches; resolution: 1920 x 1080; refresh rate: 60Hz). Stimuli consisted of images of 2 different indoor scenes (i.e., a yellow-blue room and a grey-red room), representing the conditioned stimuli (CSs) of the study. Stimulus presentation and assignment to the experimental condition was counterbalanced across subjects, and the reinforced CS+ and the unreinforced CS- were counterbalanced, as well. The shock pulse was generated by a Digitimer Stimulator (Model DS7, Digitimer Ltd., UK) and delivered to the participants' left inner wrist for 200 ms. The intensity of the stimulation was set with a standard workup procedure. It was initially set at 0.5 mA and increased by 1 mA until participants reported it was highly annoying, but not painful. A one-way ANOVA on shock intensity showed no significant difference between groups ( $F_{4,65} = 1.96$ ,  $p = 0.11$ ,  $\eta^2 = 0.11$ ; CtrlSham, mean  $\pm$  SD:  $9.1 \pm 1.9$ ; CtrlOccipital,  $7.6 \pm 2.6$ ; No-reminder,  $8.8 \pm 1.2$ ; right-dIPFC,  $8.1 \pm 1.7$ ; left-dIPFC,  $9.2 \pm 1.5$ ).

**Procedure and experimental design.** The study was performed at the Center for Studies and Research in Cognitive Neuroscience at the University of Bologna campus in Cesena, Italy. Participants were tested individually. The procedure was the same for all participants. Participants were comfortably seated in a silent and dimly lit room, and their position was centered relative to the computer screen at a 100-cm viewing distance. Electrodes for SCR recording and for shock pulse administration were attached to the participant. The SCR was recorded continuously while participants completed the task, and data were stored for offline

analysis. Participants were asked to remain as quiet and still as possible during the task and to keep their attention on the center of the screen. After verifying that SCR was being properly recorded, the intensity of the shock pulse, to be used as the unconditioned stimulus (US), was adjusted for each participant as described above. Finally, participants were informed that they had no effect on shock administration.

The experiment used a differential fear conditioning paradigm. The testing protocol involved different phases administered over three consecutive days (Sevenster et al., 2013). During the experiment, regardless of the phase, each trial consisted of the presentation of the conditioned stimulus for 4 s. The interstimulus interval (ISI) was a grey blank screen, with a variable duration ranging from 14 to 17 s. The length of the ISI was chosen to avoid complete masking of conditioned SCRs by the unconditioned SCR to the shock in the preceding trial.

On day 1, two different phases were performed: habituation and fear acquisition. At the beginning of the session, participants were informed that different stimuli would be presented on the screen, and the participant's assignment would be to carefully observe the stimuli, as some of them might be paired with electrical stimulation. During the habituation phase, the CS+ and the CS- were presented 2 times each in a random order. To ensure the absence of baseline differences within and between groups in response to the CSs stimuli before conditioning, we performed a Group x Stimulus ANOVA on SCR data collected during habituation, which showed neither significant main effects, nor a significant interaction (all  $F > 0.14$ ; all  $p > 0.50$ ). The fear acquisition phase consisted of 16 CS+ and 16 CS- trials. One CS was associated with the administration of a shock pulse, resulting in the conditioned stimulus (CS+), while the other CS was never paired with any consequence, resulting in the neutral stimulus (CS-). In CS+ trials, the US (shock pulse) was administered 60% of the time (10 out of 16 trials), 3.8 s after the CS+ onset, and co-terminated with the CS+. In CS- trials, the US was never administered. The trials were pseudo-randomly presented to participants such that no more than two identical

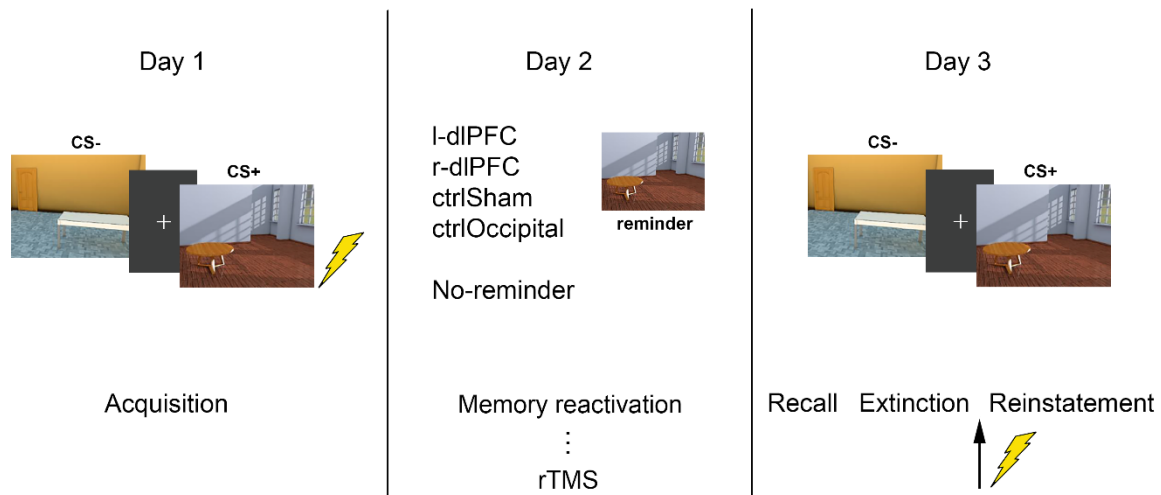
CSs occurred in a row. Participants received the instruction to press the spacebar when the CSs were presented in order to focus their attention on the screen and to learn which stimulus was followed by a shock pulse.

On day 2, 24 hours after the fear acquisition phase, fear memory reactivation was performed, except for the No-reminder group. Participants were told that the same stimuli would be presented, and they were explicitly instructed to remember what they had learned the day before (Sevenster et al., 2013). The memory was reactivated with two presentations of unreinforced (without US) conditioned stimuli (CS+). Based on previous findings showing that the reconsolidation process seems to begin between 3 and 10 min after memory reactivation (Monfils et al., 2009), subjects received rTMS (see details below) 10 min after reactivation by presentation of the reminder cues. For the No-reminder group, participants were tested in a different room with a different experimenter, and they underwent a single session of rTMS over left-dlPFC without any reactivation procedure.

To assess whether the unpleasantness of the stimulation could directly affect our results, at the end of the TMS session, participants were asked to provide subjective unpleasantness ratings of the sensations caused by the magnetic stimulation, using a 5-point Likert scale ranging from 1 (“not unpleasant at all”) to 5 (“extremely unpleasant”). A one-way ANOVA on unpleasantness ratings showed no significant effect of group ( $F_{4,65} = 1.49, p = 0.22$ ).

On day 3, fear memory recall and extinction-reinstatement took place 24h after memory reactivation. Participants were instructed that they would see the same two stimuli (CSs) from the first day. Importantly, the instructions did not reveal anything about the occurrence of the US. The fear memory recall phase consisted of 4 CS+ and 4 CS-, and the following extinction phase consisted of 12 CS+ and 12 CS- trials (the same that were presented during the fear acquisition phase), no longer followed by the US. After extinction learning, 3 unsignaled shocks (USs) were delivered to the wrist as a reinstatement procedure, followed by a fear memory

recall test. During this last phase, 4 CS+ and 4 CS- trials without any US were presented to participants. CSs characteristics, trial order, and ISI were the same in all experimental phases.



**Figure 1. Schematic representation of the experimental design and procedure for the fear memory reconsolidation experiment.** On separate days, participants performed a differential fear conditioning task. Images of two indoor scenes were used as conditioned stimuli (CS+ and CS-) presented in pseudorandom order. On day 1, during acquisition, the CS+ stimulus terminated with a shock (US, depicted as a lightning bolt) on 60% of the trials. On day 2, fear memory was reactivated with two presentations of unreinforced CS+ (reminder), except for the No-reminder group. Ten minutes after memory reactivation, participants received rTMS over either the dorsolateral prefrontal cortices (I-dIPFC and r-dIPFC), or a control site (CtrlOccipital), or placebo rTMS (CtrlSham). For the No-reminder group, rTMS was applied over the left dIPFC without memory reactivation. On day 3, participants were exposed to both the CSs without the US. After the extinction phase, and before reinstatement, participants received three unsignaled USs.

To assess conditioned responses to the CSs, SCR was measured during all the experimental phases, and the responses related to the CS+ were contrasted with those related to the CS-.

Moreover, at the end of day 1 and day 3 participants were asked to rate the expectancy of the US for the two CSs on a Visual Analogue Scale (VAS) ranging from 0 to 100.

Finally, to rule out the possibility that the observed effects of rTMS over the left and right dlPFC were simply due to a decline in higher-level cognitive processes such as working memory abilities, participants' working memory capacity (WMC) was assessed through the automated version of the operation span task (AOSPAN; Unsworth et al., 2005) at the end the experiment (day 3). A one-way ANOVA on WMC scores showed no significant effect of group ( $F_{4,65}=1.33$ ,  $p = 0.27$ ).

**Transcranial magnetic stimulation.** TMS was applied with a Magstim super rapid2 magnetic stimulator and a figure-of-eight coil with an outer winding diameter of 70mm (Magstim Company Limited, Whiteland, UK). After the memory reactivation phase on day 2, the individual resting motor threshold (rMT) of stimulation was established. The intersection of the coil was placed tangentially to the scalp with the handle pointing backward and laterally at a 45° angle away from the midline. Using a slightly suprathreshold stimulus intensity, the coil was moved to determine the optimal position from which maximal MEP amplitudes were elicited in the FDI muscle contralateral to the stimulated dlPFC cortex. For the CtrlSham and the CtrlOccipital stimulation groups, MEPs were elicited in the right FDI. The intensity of magnetic pulses was set at 110% of the resting motor threshold (rMT), defined as the minimal intensity of the stimulator output that produces MEPs with amplitudes of at least 50 $\mu$ V with 50% probability (Rossini et al., 1994). A one-way ANOVA on rMT intensity showed no significant effect of group ( $F_{4,65} = 0.63$ ,  $p = 0.64$ ; CtrlSham, mean  $\pm$  SD: 69.8  $\pm$  10.0; CtrlOccipital, 68.9  $\pm$  13.7; No-reminder, 66.3  $\pm$  13.0; right-dlPFC, 72.6  $\pm$  16.4; left-dlPFC, 72.9  $\pm$  10.5). After determination of each individual's rMT, rTMS was applied at 1 Hz for a total duration of 15 min (900 pulses), a protocol that has been shown to affect cortical excitability

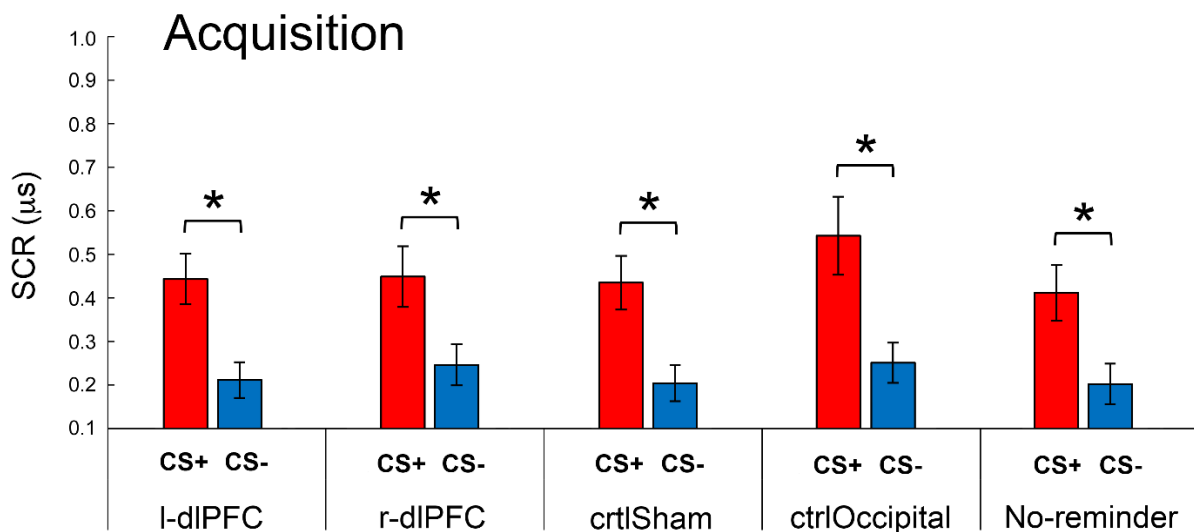
beyond the duration of the rTMS application itself (Chen et al., 1997). For stimulation of the left lateral PFC in both the left-dIPFC and the No-reminder groups, the TMS coil was placed over F3 using the international 10–20 electroencephalogram (EEG) system, while electrode F4 was chosen for the right lateral PFC, as in previous TMS studies (Sandrini et al., 2013; Rossi et al., 2001), corresponding to Brodmann area 9. The coil was held tangentially to the scalp with the handle positioned 45° with respect to the sagittal line. In the case of occipital cortex stimulation (CtrlOccipital group), the coil was placed positioned horizontally, with the coil handle pointing rightwards, over POz using the 10–20 EEG system (Jacobs et al., 2012). For sham stimulation, the coil was centered on CPZ and positioned perpendicular to the scalp surface. As shown by previous experiments (Lisanby et al., 2001; Sandrini et al., 2011), this procedure ensures that no effective magnetic stimulation reaches the brain during the sham condition, while keeping the subject’s feeling of coil–scalp contact and discharge noise similar to the real simulation.

**SCR Recording.** The skin conductance response (SCR) was recorded with two Ag/AgCl electrodes (TSD203Model; Biopac Systems, USA) filled with isotonic hyposaturated conductant and attached to the distal phalanges of the second and the third fingers of the participant’s left hand. A DC amplifier (Biopac EDA100C) was used while recording the SCR. The gain factor was 5 $\mu$ S/V and the low-pass filter was set at 10 Hz. The analog signal was then passed through a Biopac MP-150 digital converter at a 200-Hz rate. The signal was recorded with AcqKnowledge 3.9 (BIOPAC Systems, Inc., Goleta, California) and converted to microsiemens ( $\mu$ S) for offline analysis.

**SCR and subjective data analysis.** Data were analyzed offline using custom-made MATLAB scripts, and all statistical analyses were performed with STATISTICA (Dell Software, StatSoft STATISTICA, version 12.0, Round Rock, Texas, USA). Analysis of variance (ANOVA) was used to investigate differences within and between groups. Post-hoc analyses were conducted with Newman-Keuls test, and the significance threshold was  $p < 0.05$ . Moreover, effect size indices for main effects and interactions were computed using partial eta squared ( $\eta^2$ ), whereas Cohen's  $d$  values were computed for post-hoc comparisons (Cohen, 1977; Wolf, 1986). SCR data were extracted from the continuous signal and calculated for each trial as the peak-to-peak amplitude of the largest deflection during the 0.5 to 4.5 s time window after stimulus onset. The minimum response criterion was 0.02, and smaller responses were encoded as zero. Regarding SCR to CSs, stimulus onset referred to the time of the CS appearance on the screen. Regarding SCR to the US, stimulus onset was represented by the time of shock administration (3.8s after the onset of the CS). SCR following the CSs was analyzed to assess conditioned learning, whereas SCR following the US was analyzed to assess unconditioned responding. Raw SCR scores were square-root transformed to normalize the data distribution and scaled to each participant's mean square-root-transformed US response, to account for inter-individual variability (Schiller et al., 2009). SCRs were analyzed separately for each day. On day 1, to assess conditioned responses to the CS+, we separated CS+ from unconditioned responses to the shocks themselves. Hence, only non-reinforced CS+ trials were analyzed.

## RESULTS

**Day 1.** The analysis of the SCRs showed successful fear learning. That is, the stimulus (CS+ and CS-) by phase (early and late phase) interaction was significant ( $F_{1,65} = 15.15$ ,  $p = 0.0002$ ;  $\eta^2 = 0.19$ ). Follow-up tests revealed greater SCR to CS+ than to CS- trials during the early phase (mean SCR  $\pm$  SD for CS+:  $0.49 \mu\text{S} \pm 0.19 \mu\text{S}$ ; for CS-:  $0.36 \mu\text{S} \pm 0.21 \mu\text{S}$ ;  $p = 0.0001$ ;  $d = 0.85$ ) and the late phase of acquisition (CS+:  $0.45 \mu\text{S} \pm 0.25 \mu\text{S}$ ; CS-:  $0.22 \mu\text{S} \pm 0.16 \mu\text{S}$ ;  $p = 0.0001$ ;  $d = 1.36$ ) across all groups, and the difference between SCR to CS+ and CS- trials was greater in the late phase than in the early phase (early phase:  $0.13 \mu\text{S} \pm 0.02 \mu\text{S}$ ; late phase:  $0.23 \mu\text{S} \pm 0.02 \mu\text{S}$ ). Importantly, the analysis revealed neither a significant main effect of group, nor any interactions between group and either stimulus or phase (all  $p$ -values  $> 0.33$ ; see Figure 2).



**Figure 2. SCR during fear acquisition (day 1).** Data are represented as mean  $\pm$  SEM of the SCR amplitude recorded during acquisition (Day 1) in the five groups. \* denotes significant comparisons ( $p < 0.05$ ).



Likewise, CS-US contingency ratings – assessed on a 0-100 visual analog scale (VAS) at the beginning and the end of the session – revealed a significant stimulus x phase (pre- and post-fear acquisition) interaction ( $F_{1,65} = 82.57$ ,  $p < 0.0001$ ;  $\eta_p^2 = 0.56$ ). Follow-up tests showed that the CS+ elicited significantly larger shock-expectancy ratings than the CS- did after fear conditioning (mean ratings  $\pm$  SD for CS+:  $38.95 \pm 29.63$ ; CS-:  $6.31 \pm 12.31$ ;  $p < 0.001$ ), but not before fear conditioning (CS+:  $10.55 \pm 16.21$ ; CS-:  $10.42 \pm 16.76$ ;  $p = 0.96$ ). There were no significant differences between groups (see Table 1). Overall, these data demonstrate that fear learning took place equivalently across all groups of participants.

	<i>Day 1</i>		<i>Day 3</i>	
	CS+	CS-	CS+	CS-
<b>l-dIPFC</b>	37 $\pm$ 32	12 $\pm$ 19	20 $\pm$ 24	14 $\pm$ 21
<b>r-dIPFC</b>	30 $\pm$ 31	1 $\pm$ 2	13 $\pm$ 19	2 $\pm$ 4
<b>ctrlOccipital</b>	46 $\pm$ 27	7 $\pm$ 13	32 $\pm$ 31	1 $\pm$ 2
<b>ctrlSham</b>	44 $\pm$ 32	2 $\pm$ 4	17 $\pm$ 17	6 $\pm$ 11
<b>No-reminder</b>	36 $\pm$ 25	7 $\pm$ 10	22 $\pm$ 23	12 $\pm$ 10

**Table 1. CS-US contingency ratings.** Data are reported as mean  $\pm$  SD contingency ratings for the CS+ and the CS- stimulus assessed on 0-100 visual analog scale (VAS) at day 1 and day 3.

**Day 2.** On day 2 (fear memory reactivation and brain stimulation), in the four state-dependent rTMS groups, the CS+ was presented without the US to act as a reminder and reactivate the memory trace (Else et al., 2018, Kindt et al., 2014; Schiller et al., 2009). Afterwards, 1Hz rTMS was applied for 15 minutes to a specific brain region, according to the assigned group: l-dIPFC, r-dIPFC, ctrlOccipital, ctrlSham. For the additional control group (No-reminder), rTMS was administered to the left dIPFC without memory reactivation.

The four state-dependent groups expressed comparable levels of SCR during reactivation of fear memory (ctrlSham: mean SCR to two CS+ presentations  $\pm$  SD:  $0.71 \mu\text{S} \pm 0.44 \mu\text{S}$ ; ctrlOccipital:  $0.67 \mu\text{S} \pm 0.28 \mu\text{S}$ ; r-dlPFC:  $0.61 \mu\text{S} \pm 0.32 \mu\text{S}$  and l-dlPFC:  $0.75 \mu\text{S} \pm 0.49 \mu\text{S}$ ;  $F_{3,52} = 0.32$ ,  $p = 0.81$ ). In addition, fear memory was equally well consolidated in the four groups, as revealed by the absence of an interaction effect between group and phase ( $F_{3,52} = 0.58$ ,  $p = 0.63$ ). That is, there was no effect of group on SCR that differed between the last four acquisition trials (day 1) and the two reactivation trials (day 2). These data demonstrate that, before the reconsolidation manipulation, the conditioned response was equally expressed across groups.

**Day 3.** The analysis of SCRs revealed a significant interaction ( $F_{8,130} = 2.07$ ,  $p = 0.043$ ;  $\eta^2 = 0.11$ ) between group (r-dlPFC, l-dlPFC, ctrlSham, ctrlOccipital, and No-reminder), stimulus (CS+ and CS-) and phase (recall, extinction, and reinstatement).

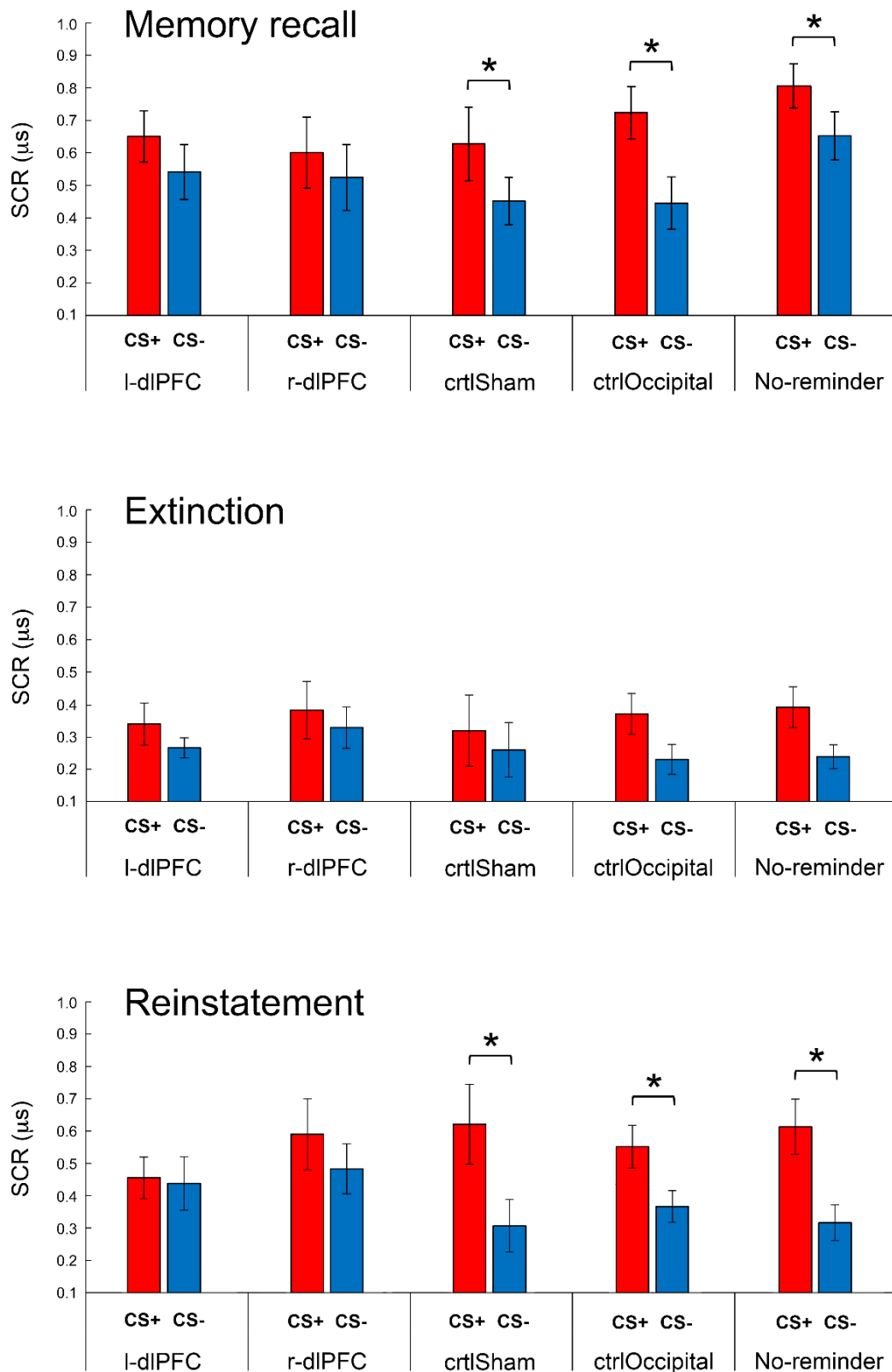
Specifically, in the memory recall test (48h after fear acquisition), the administration of rTMS over both right and left dlPFC significantly decreased SCR differences between CS+ and CS- (r-dlPFC, CS+:  $0.60 \mu\text{S} \pm 0.40 \mu\text{S}$ ; CS-:  $0.51 \mu\text{S} \pm 0.38 \mu\text{S}$ ;  $p = 0.41$ ; l-dlPFC, CS+:  $0.65 \mu\text{S} \pm 0.30 \mu\text{S}$ ; CS-:  $0.54 \mu\text{S} \pm 0.32 \mu\text{S}$ ;  $p = 0.33$ ). Conversely, the expression of fear memory remained stable in the three control groups, with SCRs to CS+ significantly larger than those to CS- (ctrlSham, CS+:  $0.63 \mu\text{S} \pm 0.42 \mu\text{S}$ ; CS-:  $0.46 \mu\text{S} \pm 0.27 \mu\text{S}$ ;  $p = 0.046$ ;  $d = 0.74$ ; ctrlOccipital, CS+:  $0.70 \mu\text{S} \pm 0.29 \mu\text{S}$ ; CS-:  $0.44 \mu\text{S} \pm 0.28 \mu\text{S}$ ;  $p < 0.001$ ;  $d = 0.99$ ; No-reminder, CS+:  $0.80 \mu\text{S} \pm 0.25 \mu\text{S}$ ; CS-:  $0.67 \mu\text{S} \pm 0.28 \mu\text{S}$ ;  $p = 0.02$ ;  $d = 0.92$ ).

In the extinction training phase, no significant SCR differences between CS+ and CS- were observed in any group (ctrlSham, CS+:  $0.32 \mu\text{S} \pm 0.41 \mu\text{S}$ ; CS-:  $0.26 \mu\text{S} \pm 0.31 \mu\text{S}$ ;  $p = 0.77$ ; ctrlOccipital, CS+:  $0.36 \mu\text{S} \pm 0.23 \mu\text{S}$ ; CS-:  $0.23 \mu\text{S} \pm 0.16 \mu\text{S}$ ;  $p = 0.25$ ; No-reminder, CS+:  $0.40 \mu\text{S} \pm 0.24 \mu\text{S}$ ; CS-:  $0.23 \mu\text{S} \pm 0.14 \mu\text{S}$ ;  $p = 0.06$ ; r-dlPFC, CS+:  $0.38 \mu\text{S} \pm 0.33 \mu\text{S}$ ; CS-:

0.33  $\mu\text{S} \pm 0.24 \mu\text{S}$ ;  $p = 0.83$ ; l-dIPFC, CS+: 0.34  $\mu\text{S} \pm 0.24 \mu\text{S}$ ; CS-: 0.27  $\mu\text{S} \pm 0.12 \mu\text{S}$ ;  $p = 0.70$ ; see Figure 3). Note that the differential fear response was already eliminated during the recall phase in the groups that received rTMS over the right and left dIPFC. This result ensures that the five groups were in a similar state of extinction. Namely, the conditioned fear response was equally reduced after the extinction training in all groups.

Exposure to the aversive stimulus (i.e., the shock) after fear extinction has been shown to reinstate the expression of the original fear memory in animals (Bouton, 2002) and humans (Norrholm et al., 2006). Accordingly, following fear memory reinstatement, we observed different SCRs between CS+ and CS- in the three control groups (ctrlSham, CS+: 0.62  $\mu\text{S} \pm 0.46 \mu\text{S}$ ; CS-: 0.31  $\mu\text{S} \pm 0.30 \mu\text{S}$ ;  $p = 0.00005$ ;  $d = 0.73$ ; ctrlOccipital, CS+: 0.52  $\mu\text{S} \pm 0.24 \mu\text{S}$ ; CS-: 0.36  $\mu\text{S} \pm 0.17 \mu\text{S}$ ;  $p = 0.056$ ;  $d = 0.67$ ; No-reminder, CS+: 0.61  $\mu\text{S} \pm 0.32 \mu\text{S}$ ; CS-: 0.32  $\mu\text{S} \pm 0.21 \mu\text{S}$ ;  $p = 0.00004$ ;  $d = 0.86$ ). Crucially, reinstatement was unsuccessful in the right and the left dIPFC groups, in which we observed no difference between CS+ and CS- (r-dIPFC, CS+: 0.59  $\mu\text{S} \pm 0.41 \mu\text{S}$ ; CS-: 0.48  $\mu\text{S} \pm 0.29 \mu\text{S}$ ;  $p = 0.20$ ; l-dIPFC, CS+: 0.46  $\mu\text{S} \pm 0.24 \mu\text{S}$ ; CS-: 0.44  $\mu\text{S} \pm 0.31 \mu\text{S}$ ;  $p = 0.94$ ).

Taken together, these data indicate that rTMS over the dIPFC (after fear memory reactivation) not only diminished fear expression at recall, but also prevented the return of fear following the reinstatement procedure. Interestingly, subjective US-expectancy, collected at the end of day 3, revealed a different pattern. The analysis showed significantly a higher shock-expectancy ( $F_{1,65} = 22.63$ ,  $p < 0.001$ ,  $\eta^2 = 0.26$ ) for CS+ (mean ratings  $\pm$  SD: 21.21  $\pm$  23.83) than CS- (7.09  $\pm$  13.06) that was equally present in all groups ( $F_{4,65} = 2.09$ ;  $p = 0.09$ ), thereby indicating no effect of dIPFC stimulation on participants' learned expectations of the unconditioned stimulus.



**Figure 3. SCR during memory recall, extinction, and reinstatement phases (day 3).** Data are represented as mean  $\pm$  SEM of the SCR amplitude recorded during memory recall, extinction, and reinstatement phases (Day 3) in the five groups. \* denotes significant comparisons ( $p < 0.05$ ).

## DISCUSSION

The neural substrates of fear memory reconsolidation in humans remain largely unknown. Here, to target brain processes implicated in fear memory reconsolidation, we administered rTMS to the dlPFC – a key area for learning and remembering events (Cabeza and Nyberg, 2000; Eichenbaum, 2017) – during the reconsolidation time-window of a previously acquired fear memory.

To ensure the state-dependent efficacy of the treatment, in four groups of healthy humans, we administered rTMS following a reminder of the fear memory able to trigger the reconsolidation process (Schiller et al., 2009; Sevenster et al., 2013; Merlo et al., 2014), and these groups were directly compared with an additional control group in which no reminder was used (No-reminder). In two of the groups receiving state-dependent rTMS, we targeted the dlPFC both in the left (l-dlPFC) and right (r-dlPFC) hemispheres; moreover, in the other two (control) groups, we stimulated the occipital cortex (ctrlOccipital) as an active control site, or administered sham stimulation (ctrlSham) to control for nonspecific effects of rTMS. Seventy participants were tested across three days, following established procedures to ensure acquisition, reconsolidation, extinction, and reinstatement of fear memory (Elsley et al., 2018, Kindt et al., 2014; Schiller et al., 2009). A physiological measure (i.e., SCR) and subjective reports (i.e., CS-US contingency ratings) of fear learning and memory were collected throughout the experiment as dependent measures.

On day 1, results showed that all experimental groups acquired fear conditioning, demonstrating that fear learning took place equivalently across all participants. On day 2, results revealed no differences of group on SCR for the CS+ reactivated trials. These data demonstrate that, before the reconsolidation manipulation, the conditioned response was equally expressed across groups. On day 3, both l-dlPFC and r-dlPFC groups exhibited decreased physiological expression of fear, indexed by their skin conductance response (SCR to CS+ similar to CS-),

both in memory recall and after an extinction-reinstatement phases. In striking contrast, no such decrease was observed in participants receiving either control rTMS (i.e., stimulation of a control site and sham stimulation), or dlPFC-rTMS without preceding reactivation of a fear memory (No-reminder), thus showing both the site-specificity and state-dependency of our rTMS intervention.

Several alternative explanations of the present findings can be discarded. First, the results cannot be explained by a general amnesic effect of brain stimulation, as the group receiving rTMS to the control brain area (occipital cortex) continued to express fear (higher SCR to CS+ compared to CS-) at recall and following reinstatement. Second, only the stimulation of the right and left dlPFC was causally associated with no fear response in both testing phases. Third, the evidence that participants persistently expressed fear (in terms of both psychophysiological reactions and subjective ratings) when the memory was not reactivated by presentation of the CS+ confirms that the dlPFC manipulation via rTMS was state-dependent, and specifically acted on the memory reconsolidation process (Else et al., 2018). These results, together with the absence of fear recovery following reinstatement (Barak and Ben Hamida, 2012), argue in favor of a direct modification of the original memory trace, rather than the formation of a new memory, as occurs in extinction (Bouton, 2002; Raij et al., 2018).

The present results confirm that the expression of fear, even if successfully extinguished, can be reinstated by reexposure to the threatening stimulus. Remarkably, we provide novel causal evidence that this return of fear can be prevented by reactivating the original memory trace and interfering with dlPFC activity. This highlights the critical role of the dlPFC in the modification of a previously acquired fear memory. It is widely accepted that the long-term consolidation of conditioned fear memories depends on plastic changes within the prefrontal cortex, which exhibits protein synthesis and degradation mechanisms similar to the classical activations found in hippocampus- and amygdala-dependent memory

consolidation (Gilmartin et al., 2014). However, the precise role of dorsal prefrontal regions in fear memories has been generally associated with the conscious appraisal of threat and the ongoing processing of the fear memory trace (i.e., working memory). Within this debate, the present results are in line with the idea that the prefrontal cortex plays a key role in the reconsolidation of memories (Gilmartin et al., 2014; Sandrini et al., 2013; Sandrini et al., 2014; Sandrini et al., 2015; Kitamura et al., 2017; Javadi and Cheng, 2013; Mungee et al., 2014; James et al., 2016). When reactivated, memories enter a transient and labile state that can result in the enhancement or weakening of that specific trace (Agren, 2014). Perturbation of the dlPFC during the reconsolidation time-window is likely to have altered prefrontal functional connections not only with the hippocampus – as already postulated for non-emotional memories (Sandrini et al., 2013; Sandrini et al., 2014) – but also with the amygdala, which is associated with the fearful component of the reactivated memory trace (Mungee et al., 2014). By interfering with normal brain activity during the consolidation time-window, the connections between frontal regions and the amygdala were weakened, thus resulting in decreased fear expression (Mungee et al., 2014).

Notably, dlPFC stimulation had no effect on the declarative memory about which conditioned stimulus had been paired with the shock, although this factual knowledge no longer accounted for reliable fear responses in those subjects. This finding therefore suggests that post-retrieval stimulation of the prefrontal cortices blocks the reconsolidation of the emotional component of the memory, while leaving the cognitive component of prior contingency learning unaffected.

It has to be acknowledged that the idea that brain stimulation can interfere with memory is not completely new. In 1968, two influential papers reported, in rodents, an elimination of the fear response by pairing a brief presentation of the conditioned stimulus with an electroconvulsive shock (Misanin et al., 1968; Schneider and Sherman, 1968). Even if

impressive, such an invasive approach could not be easily translated to humans. Crucially our study identified which neural regions should be the best target for interfering with the memory consolidation process, which represents a clinical priority. More recent non-invasive approaches to brain stimulation tried to tackle this issue (Raj et al., 2018; Sandrini et al., 2013; Sandrini et al., 2014; Javadi and Cheng, 2013; Murgee et al., 2014; Censor et al., 2010; Bernacer et al., 2013; Javadi and Walsh, 2012). However, none of the aforementioned studies aimed to reduce fear memories by interfering with the reconsolidation process. Moreover, they failed to investigate the critical role of the dlPFC in the reconsolidation process, and whether targeting the right or left dlPFC similarly impacts fear memory – a critical point in the design of clinical TMS protocols (Karsen et al., 2014). Finally, none of the existing non-invasive brain stimulation studies tested the strength of the neuromodulation by means of a reinstatement procedure.

To summarize, these results demonstrate that non-invasive stimulation of the prefrontal cortex following memory reactivation disrupts the expression of fear to a previously conditioned threatening stimulus, and argue in favor of a critical role of the dlPFC in the neural network that mediates the reconsolidation of conditioned fear memories in humans. These findings provide a step forward toward understanding the mechanisms underlying fear memory reconsolidation, and they have potential clinical implications for targeting emotional, maladaptive memories (Pennington and Fanselow, 2018). Uncovering the brain regions involved in the reconsolidation of emotional memories constitutes a challenging opportunity for non-invasive brain stimulation and reconsolidation-based interventions, which are increasingly applied to conditions like phobia, addiction, post-traumatic stress disorder, obsessive-compulsive disorder and many others (Schwabe et al., 2014).





**STUDY 4:**  
**CHARACTERIZING CARDIAC AUTONOMIC SIGNATURE**  
**TO FEAR CONDITIONED STIMULI**

## **ABSTRACT**

In humans, fear conditioning is often probed by measuring activity of the autonomic nervous system (ANS): via skin conductance responses (SCR; Critchley et al., 2000), or pupil size responses (Korn et al., 2016), by assessing the fear-potentiated startle (Brown et al., 1951; Khemka et al., 2016), or measures of heart rate variability (HRV; Paulus et al., 2016). HRV refers to fluctuations of the lengths of time between consecutive heartbeats, or interbeat intervals. It is generated by the sinoatrial node of the heart but antagonistically modulated by the sympathetic and parasympathetic (vagal) branches of the ANS. According to the neurovisceral integration (NVI) model (Thayer and Lane 2000), the functioning of prefrontal-subcortical inhibitory circuits critical for self-regulation is linked with the heart via the vagus nerve that provides inhibitory inputs to the heart. Although several earlier studies have argued that the vagus nerve may play a crucial role in fear conditioning due to the related bradycardia observed to the conditioned stimulus, there is scarce evidence that proves the vagus nerve involvement. To test whether vagus nerve plays a crucial role in fear conditioning, healthy subjects were divided into two experiments were tested in a classical paradigm of fear conditioning, healthy subjects were divided into two experiments were tested in a classical paradigm of fear conditioning (acquisition and extinction) and skin conductance responses and heart rate were recorded. We developed a series of analysis specifically to characterize autonomic modulations of heart rate modulations by using spectral analysis approaches, quantifying the vagus nerve involvement to the fear conditioned stimuli. Results demonstrated that during the acquisition phase, a significant cluster of spectral power reflecting vagus nerve contribution, that is occurring around the time in which participants expected the shock administration. The results of the present study provide the first direct evidence that systematically investigate and quantified its involvement in the human fear conditioning framework.

## INTRODUCTION

Learning to respond to stimuli or circumstances that predict impending danger is a highly adaptive function for animals and humans alike. From an evolutionary perspective, learned fear serves to activate defensive responses in anticipation of harm, thus minimizing the impact of noxious challenge (Ploghaus et al., 2003). In the laboratory, a paradigm most often used to study this process is Pavlovian fear conditioning, wherein an initially neutral stimulus (the conditioned stimuli or CS), is paired with a noxious stimulus (usually a mild electric shock, the unconditioned stimulus or US). As a result, the CS acquires the ability to elicit various behavioural and physiological conditioned fear responses when later presented alone. Fear conditioning is widely held to be a model for pathogenesis of phobia or anxiety disorders.

In humans, fear conditioning is often probed by measuring activity of the autonomic nervous system (ANS), for instance via skin conductance responses (SCR; Critchley et al., 2000), or pupil size responses (Korn et al., 2016), or by assessing the fear-potentiated startle (Brown et al., 1951; Khemka et al., 2016). Moreover, measures of heart rate variability (HRV), a noninvasive marker of autonomic control, are increasingly being employed to clarify the relationship between psychological and physiological processes (Castegnetti et al., 2016; Paulus et al., 2016).

HRV refers to fluctuations of the lengths of time between consecutive heartbeats, or interbeat intervals. It is generated by the sinoatrial node of the heart, but antagonistically modulated by the sympathetic and parasympathetic (vagal) branches of the ANS. Moreover, higher neural networks can exert a flexible control over the ANS, evoking reciprocal (i.e., increase in activity of one branch is associated with decreased activity in the other), or independent changes of the sympathetic and parasympathetic nervous systems (Berntson et al. 1991; 1993; Koizumi & Kollai 1992; Tessa et al., 2019).

Fear conditioned stimuli (CS) typically elicits transient heart rate deceleration that reach its nadir at some point in time near the US (Obrist et al., 1965; Hugdahl, 1995; Castegnetti et al., 2016). This anticipatory fear bradycardia is generally believed to reflect almost exclusive vagal control. However, given the antagonistic effects of the sympathetic and parasympathetic branches, a bradycardic response to a stimulus could arise, for example, from either increased vagal activation, or sympathetic withdrawal, or even from vagal and sympathetic coactivation, in which the vagal effects exceed those due to the sympathetic nervous system. Measure of heart rate *per se*, therefore, may not provide an accurate rendition of the underlying autonomic mechanisms.

To gain deeper insight into the physiological underpinnings of fear learning, HRV was analysed here as a quantitative index of the interplay between sympathetic and parasympathetic influences on cardiac activity. Owing to the difference in their latencies of action (i.e., vagal effects unfold faster than sympathetic effects), the periodic oscillations in heart rate produced by the two autonomic branches occur at different frequency (Akselrod et al., 1981). This serves as the basis for frequency-domain techniques, such as spectral analysis, to distinguish between the frequency-specific contribution of the sympathetic and parasympathetic systems to HRV at any given time. Spectral analysis allows the intensity of the HRV spectral components [i.e., the high-frequency band (HF), low-frequency band (LF), and very low frequency band (VLF)] to be determined. The HF component is believed to be mediated primarily by cardiac parasympathetic outflows, and thus may provide a direct index of vagal activity; whereas the LF is commonly viewed as a product of both sympathetic and parasympathetic activity (Malliani et al., 1991; Pumprla et al., 2002; Task Force, 1996).

Conventional analysis of HRV requires an observation window in the range of 2-3 minutes and some level of stationarity during this period (Task Force, 1996; Acharya, 2006). It is, therefore, inadequate to capture transient components in heart rate due to interaction of the

autonomic nervous system, as those occurring after the CS onset in typical fear learning paradigm. In the present study, two approaches were used to perform a frequency-domain analysis of heart rate: short-time Fourier transform (STFT), and instantaneous spectral estimates extracted from a point-process modelling algorithm (Barbieri et al., 2005). The STFT method was used to obtain the time-frequency structure of the expected alteration of the autonomic regulation. Because this method is known to introduce temporal and spectral spread of the components the point-process modelling algorithm was used in order to obtain unbiased spectral estimates at the cost of a reduced statistical sensitivity.

In two separate experiments, participants underwent a delay fear conditioning procedure in which a visual stimulus (CS+) was probabilistically paired with an electric shock as US, while a different stimulus (CS-) was never paired with a US. Fear acquisition was followed by an extinction phase during which both CSs were presented in the absence of the US. We tested whether the spectral components of the HRV, as a non-invasive marker of sympathetic and parasympathetic mechanisms, can dissociate between conditioned and neutral stimuli related to fear learning. To this end, we combined the electrocardiogram (ECG) signal recording with an established psychophysiological measure of fear conditioning, that is the SCR. The results of the present study should thus provide unique insights into the psychophysiological metrics of fear learning, both in healthy individuals and patients suffering from anxiety disorders.

## **METHODS**

**Participants.** A total of 50 healthy individuals took part in two independent experiments. Twenty-eight subjects (17 female; mean age  $\pm$  SD = 23.25  $\pm$  2.32 years; mean education  $\pm$  SD = 15.11  $\pm$  1.99 years) participated in Experiment 1, and a different sample of 22 subjects (12 female; mean age  $\pm$  SD = 24.51  $\pm$  3.52 years; mean education  $\pm$  SD = 14.22  $\pm$  2.13 years) in Experiment 2. Prospective participants were recruited from the student population of the University of Bologna using campus advertisements. All subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), had normal or corrected-to-normal visual acuity in both eyes, and were naive to the purposes of the experiment. Individuals who reported a history of psychiatric care, neurological disease, cardiovascular conditions or substance abuse were excluded, as were individuals currently treated with any medication known to affect the central nervous system. Trait anxiety and depression were measured, given evidence for their relationship with fear learning (Otto et al., 2007; Prenoveau et al., 2011). Trait anxiety was assessed using the State-Trait Anxiety Inventory (Spielberg et al., 1983), which possesses good reliability and validity. Depression and anxiety symptomatology was assessed with the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), which has moderate to high convergent validity. All participants gave informed written consent to participation after being informed about the procedures of the study. All procedures were conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki and were approved by the Ethics Committee of the Department of Psychology of the University of Bologna.

**Apparatus, stimuli and task.** All experiments were implemented in MATLAB environment (version R2018b; The MathWorks, Inc., Natick, Massachusetts, USA), and ran on a Windows-based PC (Lenovo ThinkCentre Desktop Computer). A delay fear conditioning task with partial

reinforcement was used in both Experiment 1 and 2. The task consisted of habituation, Acquisition and Extinction Phases presented continuously. Unconditioned stimulus (US) was a 200-ms train of electrical square pulses (individual pulse width of 0.2 ms), generated by a constant-current stimulator (DS7A, Digitimer Ltd., UK), and applied via two surface electrodes fixated on the inner side of the participants' left wrist. The intensity of the electrical stimulus was determined individually by assessing the participant's subjective evaluation in a standard work-up procedure prior to the experimental task. The current was initially set at 0.5 mA and increased in steps of 1 mA until participants reported it as a "highly annoying, but not painful" stimulation. Conditioned stimuli (CSs) consisted of two visual stimuli created with Blender (Blender Foundation, Amsterdam, Netherlands), and Cinema 4D R17 software (MAXON Computer GmbH, Friedrichsdorf, Germany), and presented on a computer screen.

In Experiment 1, CSs were images of two different indoor environments (i.e., a yellow-blue room, and a grey-red room) that covered the entire screen. In experiment 2, CSs consisted of images of two common objects (i.e., a plant and a lamp) embedded within an indoor scene. Neutral, rather than intrinsically emotional (i.e., spiders, snakes, or angry faces), visual stimuli were used as CSs, because conditioned responses to very salient CSs can be confounded by the ceiling effects of the respective outcome measures (reference). The type of stimuli associated to the CS+ and CS- was counterbalanced across participants. During the experiments, each trial consisted in the presentation of one CS in the centre of a computer screen for 4 s, followed by an inter-trial interval (ITI) of variable duration, from 13 to 16 s in Experiment 1, and from 14 to 17 s in Experiment 2, during which the screen turned completely grey and empty.

The habituation phase included 4 trials (2 for each CS), during which the CSs were presented without reinforcement. Few habituation trials were used to avoid retardation of learning due to non-reinforced exposure to CS+ (the latent inhibition effect; Lubow, 1973). Habituation trials were not analyzed. During the Acquisition Phase, one CS was designated as



CS+, and was associated with the US 60% of times (12 out of 20 trials), while the unreinforced stimulus (CS-) was a different visual stimulus not associated with any consequence. The Acquisition Phase consisted of 20 CS+ and 20 CS- trials. In CS + trials, the US was administered 3.8 s after the CS+ onset, and co-terminated with the CS+, 0.2 s later. During the Extinction Phase, both CS+ and CS- stimuli were presented 20 times without the US. In both experiments, the trials were pseudo-randomly presented to participants such that no more of three identical CSs occurred in a row.

**Procedure.** The study was performed at the Centre for Studies and Research in Cognitive Neuroscience of the University of Bologna, in Cesena, Italy. Participants were tested individually. They were comfortably seated in a silent and dimly lit room in front of a computer screen, (size: 43 inches; resolution: 1920 × 1080 pixels; refresh rate: 60 Hz), at ~75 cm viewing distance. Once seated, the experimental procedure was explained and written informed consent was obtained from participants. After verifying that signals were being properly acquired by the instruments, and the intensity of the electrical stimulation was set, the experiment began. Participants were instructed that different images would be presented on the screen, and that they would have to carefully observe the screen, as some of the displayed stimuli might be paired with electrical stimulation. During experiments, visual and electrical stimuli were automatically administered by the task presentation system implemented in MATLAB environment, while ECG and SCR signals were recorded continuously.

**Physiological signal recordings.** In both experiments, signals were recorded with a Biopac MP-150 system at 200 Hz sampling rate and fed into AcqKnowledge 3.9 software (BIOPAC Systems, Inc., Goleta, California, USA) for offline analysis. The SCR was acquired with two Ag/AgCl electrodes (TSD203 model; BIOPAC Systems), filled with isotonic hyposaturated

conductant gel, and attached to the distal phalanges of the second and third finger of the participant left hand. A Biopac EDA100C module was used as to amplify SCR signal (gain switch set to 5  $\mu$ S/V, low pass to 35 Hz, high pass to DC). The ECG was acquired with three Ag/AgCl electrodes (EL503 model; BIOPAC Systems), filled with isotonic hyposaturated conductant gel. Electrodes were positioned in a modified bipolar lead I configuration, with the positive electrode placed on the participant's left wrist, the negative electrode on the right wrist, and the ground electrode attached just underneath the right clavicle. A Biopac ECG100C module was used to amplify ECG signals (gain switch set to 500, low pass to 35 Hz, high pass to 0.05 Hz).

**SCR data processing and statistical analysis.** SCR data were analysed offline in a MATLAB environment, and all statistical analyses were performed with STATISTICA (StatSoft, v. 13.0, Round Rock, Texas, USA). SCR following the CS was analysed to assess conditioned responses, whereas SCR following the US was analysed to assess unconditioned responses. The onset was represented respectively by the time of stimulus presentation and electrical shock administration. Each trial was extracted from the entire SCR signal and peak-to-peak value was calculated as the amplitude of the largest deflection during the 0.5 to 4.5 s time window after stimulus onset. The minimum response criterion was 0.02, and smaller responses were encoded as zero. Then, SCR peak-to-peak values were square-root transformed and scaled to each participant's average square-root of US responses (Schiller et al., 2008; Battaglia et al., 2018), to reduce interindividual variability and increasing statistical power. Finally, SCR values were collapsed into "early" and "late" responses of each phase - Acquisition and Extinction - as learning typically varies across time (Milad et al., 2005). Learning-related changes in SCR were hypothesized to be found in the 'late' phase of both Acquisition and Extinction, as previously reported (Milad et al., 2005). Normality of data distributions were verified with Shapiro-Wilk

tests. Mixed-design analyses of variance (ANOVAs) were used to investigate differences within experimental phases and post-hoc analyses were conducted with Scheffe test.

**Identification of QRS complex peaks.** ECG processing and analyses were performed in a MATLAB environment. Identification of QRS complex peaks from the ECG was performed automatically by a sample-based envelope detector of the demeaned ECG signal. The value of the envelope is updated ( $y_{env}$ ) with the value of the ECG signal as long as the amplitude is increasing. When the subsequent sample is smaller, the ECG amplitude is stored in  $y_0$  and time instant in  $t_0$ ;  $y_{env}$  is subsequently updated by a decay function modulated by the square-root of the time passed from  $t_0$  in seconds and a positive coefficient ( $r = 0.65$ ) to tune the rate of decay.

$$d = r \cdot \sqrt{t - t_0}$$
$$y_{env} = (1 - d) \cdot y_0$$

Local maxima of ECG signal ( $y_{max}$ ) detected as changes from positive to negative slope are flagged as QRS complex peaks when  $y_{max}$  is within a tolerance ( $p = 80\%$ ) of the envelope ( $p \cdot y_{env} < y_{max} < y_{env}$ ) and the time from the last QRS complex peak is greater than 0.45 s (equivalent of a heart rate inferior to 134 beats per minute). Plots of subject's ECG and resulting inter-beat intervals (IBI) sequence are presented to the operator and identified peaks appear on the ECG plot to allow for quality check and interactive corrections. A trained operator was instructed to inspect the resulting sequences, manually correct for misidentifications and regularize singular ectopic events by means of linear interpolation (Nabil et al., 2015). Furthermore, the operator reconstructed the IBI sequence during the administration of US and reported that none of the enrolled participants presented multiple consecutive ectopic events. To account for the time delay between SA node depolarization and apical depolarization,

detected as the QRS complex peak, IBI sequences were anticipated by 0.20 s in respect to the trials' trigger timings (Malmivuo & Plonsev, 1995).

**Signal processing of heart rate dynamics.** RR wave sequence was obtained by homogeneous resampling of the IBI sequence at 10 Hz, allowing a recalculation of the trials' trigger timings with a temporal accuracy of  $\pm 0.05$  s. Three different interpolation methods were tested: zero-order hold, linear and spline. Tonic component of heart rate was removed from the RR wave sequence subtracting the moving-median computed over the past T seconds to preserve causality as opposed to a centered moving metric; the median was chosen as a more robust estimator than the mean for skewed distributions and in presence of outliers (Hedges et al., 2003). The time window length T was chosen to be 15 seconds to encompass the entire duration of the RR wave response as observed under similar experimental conditions (Castegnetti et al., 2016). The Short-time Fourier transform (STFT) was computed using a centered Hamming window of length T, at time periods of 0.10 s, with interpolated spectral resolution of 0.01 Hz for components from 0 Hz to 0.50 Hz.

**Point-process modelling of heart rate dynamics.** The sequence of systolic peaks timing was passed to an autoregressive (AR) point-process modelling algorithm (Barbieri et al., 2005) to compute instantaneous estimates of heart rate variability defined in the time and frequency domains, with regression order  $K = 16$ , local likelihood interval  $W = 120$  s, forget factor  $F = 0.99$  and updating interval  $\Delta = 5$  ms. All parameters were determined after preliminary goodness-of-fit analysis of the data with evaluation of Kolmogorov-Smirnov statistic. This approach models the stochastic nature of heartbeat generation considering a physiologically plausible, history-dependent, inverse-Gaussian process of ventricular repolarization (Barbieri et al., 2005). This allows us to obtain an instantaneous RR wave mean estimate ( $\mu$ ) at a very

fine timescale, which requires no interpolation between the arrival times of two beats. Moreover, we can use the instantaneous K coefficients of the AR model to compute the distribution of spectral powers (Mainardi 2008). HRV indices were computed integrating spectral powers within the four frequency ranges of interest (VLF, LF, HFinf and HFsup). Mu and HRV indices were down-sampled at 10Hz as a data compression solution without loss of relevant information and from the Mu sequence was subtracted the moving-median component in the same fashion as per the RR wave.

**Statistical analysis of STFT and HRV indices.** Each spectral component of the STFT, and the HRV indices, were normalized using a moving modified z-score (Iglewicz & Hoaglin 1993), where the median and the median of absolute distances (MAD) were calculated over the past T seconds to preserve causality. For each spectral component, the numerator of the moving modified z-score plays the role of removing the uninformative trends (tonic component) allowing to focus the analysis on transitory oscillations (phasic component), while the denominator is used to scale for variability allowing for both inter-component and inter-subject comparability. Single trial responses of RR wave, STFT, Mu and HRV indices were considered in a time window spanning 3 s before and 15 s after the CS onset and analyzing only 'late' phase of both acquisition and extinction. Based on trials' collections, non-parametric Mann–Whitney U test (TMW) was performed for each spectral component at each timepoint, to compare CS+/US against CS- and CS+ against CS-. The number of spectral components N was set at 50 for the STFT analysis and 4 for the HRV indices analysis. Since at each timepoint spectral power tends to appear in distinct clusters (i.e., if one spectral component is significant it is likely that adjacent components will be as well) we cannot assume independence of the statistics computed along the N spectral components; nevertheless such dependence resemble that of gene paths which come in "relatively small, disjoint groups" (Storey, 2003) therefore

under this assumption positive false discovery rate (pFDR; Storey, 2003) correction was computed at each timepoint. Convergence of pFDR is reliably achieved when computed on 1000 tests (Storey, 2003), therefore a comparable amount of spectral components should be estimated but such an approach would serve little practical purpose and imply superfluous computational burden. Instead of constraining the number of spectral components  $N$  to the correction requirements we exploited the generation of  $B$  bootstrap replicates of the TMW statistics to obtain  $B$  times  $N$  p-values at each timepoint; the  $B$  resampling of data were created randomizing trials with replacement (pooling was not involved and collections numerosity was preserved). With this approach once the number of spectral components  $N$  has been defined the sufficient amount of  $B$  bootstrap replicates can be calculated as 1000 over  $N$ ; in this study  $B = 20$  for the STFT analysis and  $B = 250$  for the HRV indices analysis. The resulting corrected p-values were approximated by the median computed over the  $B$  values associated to each of the  $N$  spectral components (Bhattacharya & Habtzghi 2002).

## RESULTS

Primary our results aimed to investigating, separately for Experiment 1 and 2, if participants successfully learn the association between stimuli and the relative outcome using skin conductance responses (SCR), a well-established psychophysiological measure of fear conditioning. Once fear acquisition and extinction are assessed in all participants our main goal is to characterize autonomic signatures of heart rate modulation during fear conditioning with combined datasets.

**SCR Results.** SCR data recorded in experiment 1 are analysed using a  $4 \times 2$  repeated measure ANOVA with Phase (Acquisition Early/Acquisition Late/Extinction Early/Extinction Late) and Stimulus (CS+/CS-) as within-subject factors was carried out. Results show a main effect of Phase ( $F(3,81) = 6.899, p < 0.001, \eta^2 = 0.20$ ) which reflected differences of SCRs in the different phases of the experimental task and a main effect of Stimulus ( $F(1,27) = 48.856, p < 0.001, \eta^2 = 0.64$ ). Crucially, analysis reveals a significant Phase x Stimulus interaction ( $F(3,81) = 6.975, p < 0.001, \eta^2 = 0.20$ ). Follow-up post hoc analysis shows different pattern of SCRs between phases. Specifically, during both early (CS+ mean  $\pm$  SD =  $0.53 \pm 0.20$ ; CS- mean  $\pm$  SD =  $0.38 \pm 0.20$ ) and late (CS+ mean  $\pm$  SD =  $0.49 \pm 0.27$ ; CS- mean  $\pm$  SD =  $0.24 \pm 0.16$ ) phase of acquisition, SCR to CS+ is higher than to CS- (all  $p < 0.001$ ), suggesting a successful acquisition of learning. During the early phase of extinction, the previous conditioned response to CS+ (mean  $\pm$  SD =  $0.56 \pm 0.35$ ) is higher than to CS- (mean  $\pm$  SD =  $0.40 \pm 0.27$ ;  $p < 0.001$ ), due to the strength of the acquired conditioning and the few extinction trials presented; indeed, in late phase of extinction, no difference in SCR was found between CS+ (mean  $\pm$  SD =  $0.32 \pm 0.34$ ) and CS- (mean  $\pm$  SD =  $0.23 \pm 0.25$ ;  $p = 0.12$ ), suggesting that extinction has been occurred. Together these results demonstrate that participants showed anticipatory responses to fear-conditioned stimulus in acquisition and extinguished it in the late phases of experimental

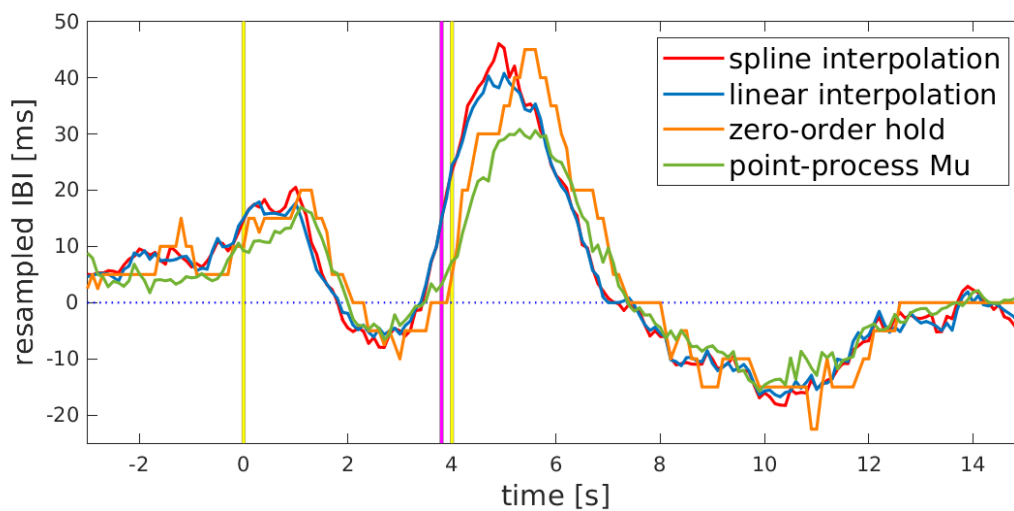
task. The SCR recorded in experiment 2 are analyzed with a 4 x 2 repeated measure ANOVA with Phase (Acquisition Early/Acquisition Late/Extinction Early/Extinction Late) and Stimulus (CS+/CS-) as within-subject factors. The statistical analysis reveals a main effect of Phase ( $F(3,63) = 25.003, p < 0.001, \eta^2 = 0.54$ ), a main effect of Stimulus ( $F(1,21) = 18.559, p < 0.001, \eta^2 = 0.47$ ) and crucially a significant Phase X Stimulus interaction ( $F(3,63) = 6.221, p < 0.001, \eta^2 = 0.23$ ). Follow-up post hoc analysis shows that during late phase of acquisition SCR to CS+ (mean  $\pm$  SD =  $0.52 \pm 0.22$ ) is higher than to CS- (mean  $\pm$  SD =  $0.41 \pm 0.12; p < 0.001$ ), but not in early phase (CS+ mean  $\pm$  SD =  $0.58 \pm 0.19$ ; CS- mean  $\pm$  SD =  $0.50 \pm 0.18; p = 0.08$ ). Also, during both early (CS+ mean  $\pm$  SD =  $0.38 \pm 0.23$ ; CS- mean  $\pm$  SD =  $0.35 \pm 0.19$ ) and late (CS+ mean  $\pm$  SD =  $0.28 \pm 0.18$ ; CS- mean  $\pm$  SD =  $0.28 \pm 0.16$ ) no difference in SCR is found between CS+ and CS- (all  $p > 0.96$ ). These results demonstrate that participants were able to acquire fear learning during the Acquisition Phase and extinguish in the subsequent phase of the task. Overall, SCR results of both experiment 1 and 2, demonstrate that all participants were able to successfully acquire fear conditioning as demonstrated by higher responses to CS+ than CS- in Acquisition Phase. Although some differences between experiments have been shown, all participants were able to acquire and extinguish fear-conditioning in late phases of the experimental task.

**HRV Results.** Cardiac signals of both experiments are analysed together in order to highlight different responses in late Acquisition Phase for CS+ as compared to CS-, and thus extracting autonomic cardiac signatures of fear conditioning.

**RR interpolation.** Three methods of homogenous resampling of the IBI sequence (spline, linear, zero-order hold) are compared to the point-process Mu using the median responses to CS+ during Acquisition Phase as reference, on data from the two experiments combined (see



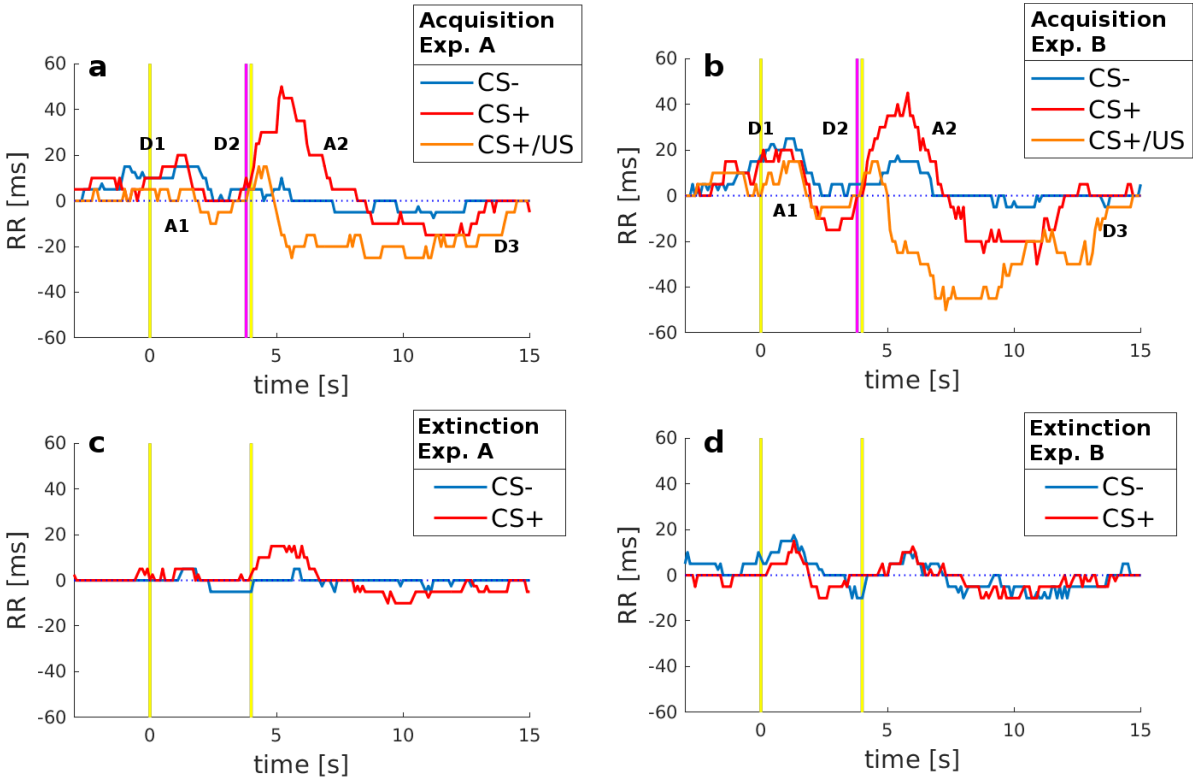
Fig. 1). Pairwise cross-correlation function among responses is computed to extract time lag of the maximum evidencing for spline and linear methods a temporal anticipation of 0.4 s compared to zero-order hold and 0.3 s compared to point-process Mu. No temporal shift results directly between zero-order hold and point-process Mu. To appropriately integrate results from STFT and point-process model subsequent analyses are conducted on RR wave obtained from zero-order hold interpolation of the IBI sequences.



**Figure 1.** Cardiac IBI sequence response to CS+ during Acquisition Phase was used as reference to compare different methods of homogeneous resampling a point-process Mu. Time lag of the cross-correlation function maximum between all interpolation methods and point-process Mu was computed evidencing that spline and linear methods involve a temporal anticipation (0.4-0.3 s) in comparison to the use of zero-order hold and point-process Mu. No temporal shift resulted directly between zero-order hold and point-process Mu. In the graph, yellow vertical lines represent respectively the onset and the offset of stimuli. Pink vertical line represents the shock pulses occurrence.

**Heart Period Responses.** The grand medians of the RR relative to CS-, CS+ and CS+/US of the two experiments are represented in Fig. 2. Heart rate modulations can be decomposed into their deceleration (positive slope of RR) and acceleration (negative slope of RR) components.

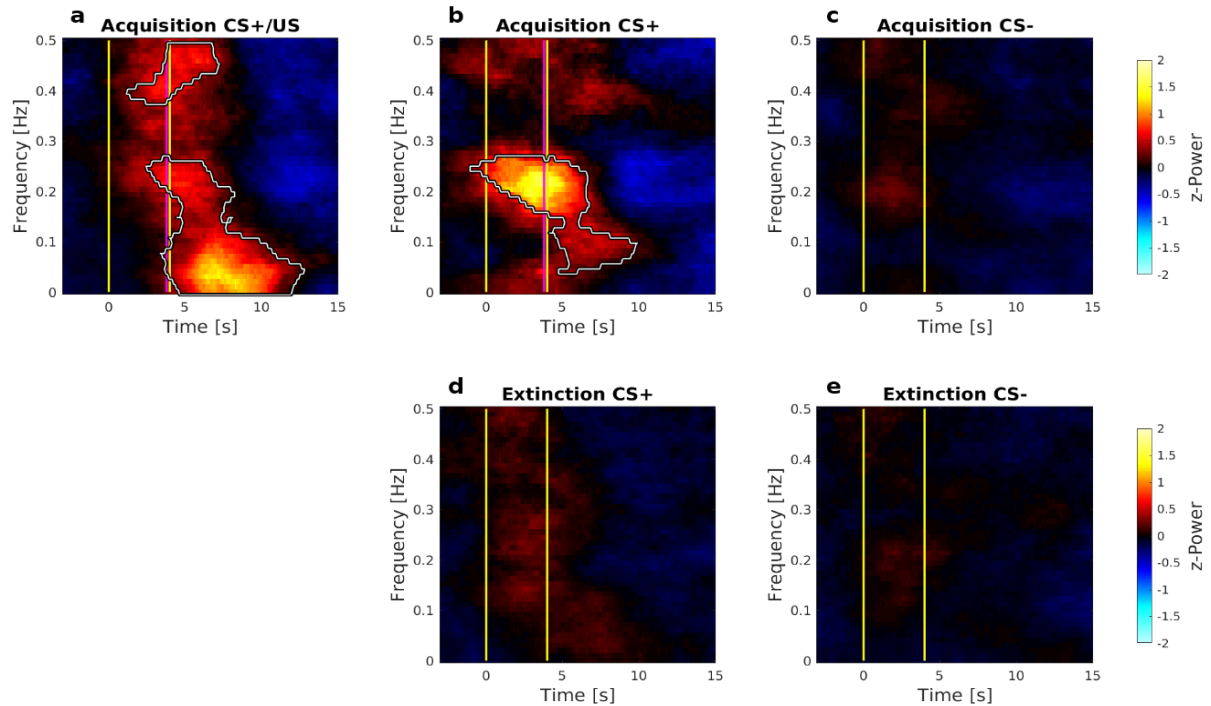
In line with similar experiments of classical conditioning (Bohlin and Kjelberg, 1979; Castegnetti et al., 2016), we observe an initial deceleration (D1 slope, Fig.2-a,b), followed by an acceleration (A1 slope, Fig.2-a,b) and another deceleration (D2 slope, Fig.2-a,b). In our study we extend the time window to include the observation of two more components: an acceleration (A2 slope, Fig.2-a,b) and a final deceleration to baseline (D3 slope, Fig.2-a,b). In particular, it is observed that the A2 slope for the CS+/US in both experiments, is steeper and temporally anticipated than its counterpart to the CS+. Overall, these responses show that during Acquisition Phase, RR to CS+ shows qualitatively larger dynamics than RR to CS- (Fig.2-a,b). On the other hand, during the Extinction Phase CS+ and CS- show a seemingly equivalence in heart rate modulation, and this effect is noticeable for both experiments (Fig.2-c,d).



**Figure 2.** RR wave responses to stimuli can be decomposed into their components of heart rate deceleration (positive slope of RR) and acceleration (negative slope of RR) respectively labelled with capital D and A. During the Acquisition Phase (panel a and b) responses showed

consistent similarities between both experiments, presenting the same deceleration and acceleration components and comparable response amplitude. Responses to CS+/US present a steeper and anticipated A2 slope subsequent to the shock administration, than its counterpart to the CS. During the Extinction Phase CS+ and CS- show a seemingly equivalence in heart rate modulation, noticeable for both experiments. In the graphs, yellow vertical lines represent respectively the onset and the offset of stimuli. Pink vertical lines represent the shock pulses occurrence.

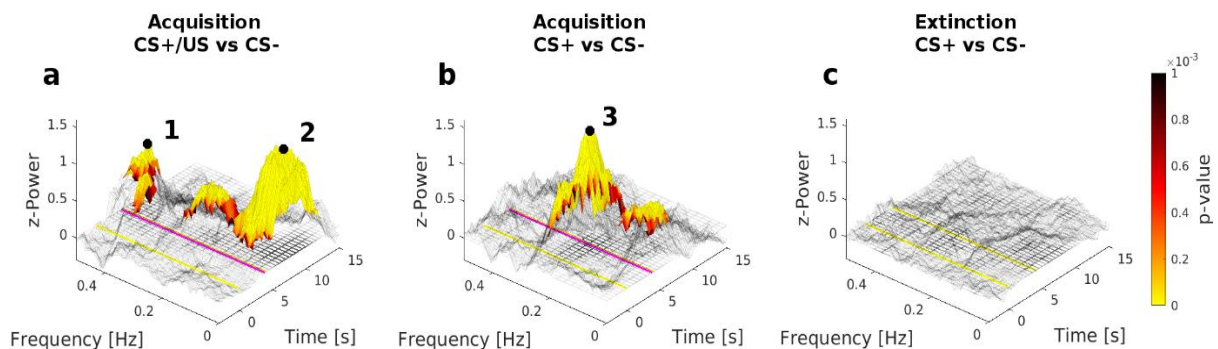
**Frequency analysis.** CSs responses of the normalized power (z-Power) of STFT for combined experiments are depicted in Fig. 3. Separate statistical analyses of CS+/US and CS+ against CS- ( $p < 0.001$ ) are conducted in both Acquisition and Extinction Phases and statistically significant results are highlighted with a coloured white boundary (Fig. 3-a,b). Comparison analysis between CS+ and CS- during Acquisition Phase, shows a significant cluster of power contribution from 0.05 Hz to 0.30 Hz, with a higher concentration of power at 0.21 Hz, that is occurring at 3.90 s, approximately the time in which participants expect the shock administration (Tab. 1; Fig. 3-b). Regarding the differences between CS+/US and CS-, analysis shows a significant cluster of power contribution from 0 Hz to 0.30 Hz, with a higher concentration of power below 0.15 Hz subsequent to the shock administration (Fig. 3-a). Furthermore, the same analysis shows the presence of another significant cluster above 0.40 Hz around the time of shock administration (Fig. 3-a). Finally, analyses between CS+ and CS- during the Extinction Phase shows no significant differences (Fig. 3-d,e).



**Figure 3.** Normalized power of STFT responses of Acquisition and Extinction Phases where the statistical differences ( $p < 0.001$ ) of CS+/US and CS+ against CS- are represented as the areas delimited by white boundaries. For a spectral power cluster in the range 0-0.30 Hz larger power contribution appears below 0.15 Hz in response to CS+/US (panel a) larger power contribution appears below 0.15 Hz, while in response to CS+ (panel b) and appears above 0.15 Hz, in response to CS+ (panel b) against CS- (panel c). For the same analysis a cluster above 0.40 Hz in response to CS+/US (panel a) around the timing of stimulus administration also results as statistically significant. For the Extinction Phase no differences were found between CS+ (panel d) and CS- (panel e) responses. Yellow vertical lines represent respectively the onset and the offset of stimuli. Pink vertical lines represent the shock pulses occurrence.

Maps of p-value (Fig. 4) are used to evaluate the distribution of the significance level in the resulting clusters ( $p < 0.001$ ). These are super-imposed over a grey-scale mesh representing the difference of STFT z-Power between the analysed responses: CS+/US and CS+ as compared to CS- in the Acquisition Phase (Fig. 4 a-b), and CS+ as compared to CS- in the Extinction Phase (Fig. 4 c). For each resulting cluster, maxima in difference between analyzed responses are

identified in the graph (Fig. 4) in order to extract STFT z-Power values and significance level of the comparison (Tab. 1). As a whole these results indicate that the presentation of CS+ elicits a strong response in the range of 0.15 - 0.30 Hz as compared to CS- (Fig. 4-b; Tab. 1-cluster index 3) sustaining bradycardia (Fig. 4-a,b; D2 to CS+). This imply that fear conditioning has occurred during the Acquisition Phase and revealing the cardiac autonomic signature of fear conditioning in humans. Differently, in response to CS+/US the spectral content presents two well distinct clusters (Fig. 4-a): one which mediates the rapid heart rate acceleration A2 to CS+/US (Fig. 2-a,b; Tab. 1-cluster index 1) and the other convergence of the responses to baseline (Tab. 1-cluster 2), namely D3 to CS+/US (Fig. 2-a,b). Moreover, results indicate no difference in the spectral content of CS+ as compared to CS- during Extinction Phase (Fig. 4-c), reflecting extinction of fear conditioning.

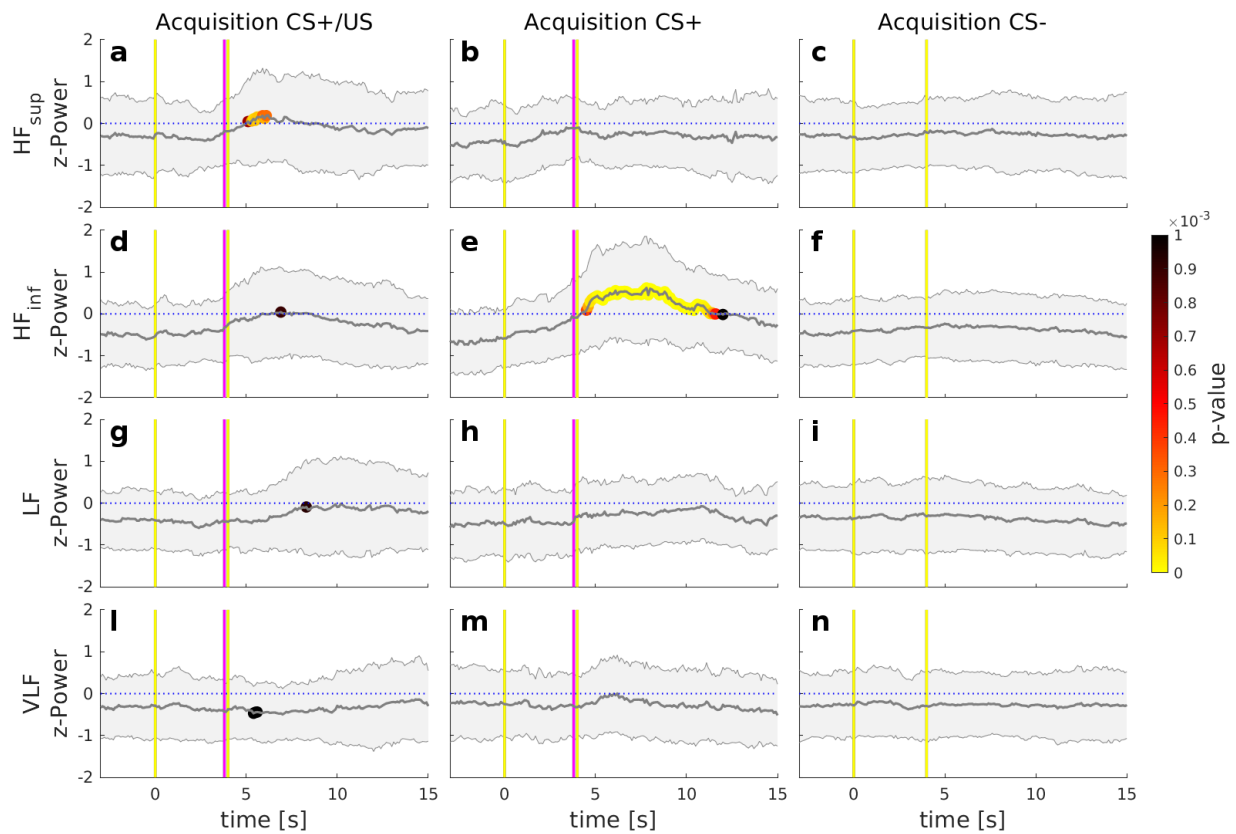


**Figure 4.** Maps of p-value super-imposed over a grey-scale mesh representing the difference of STFT z-Power between the analysed responses of Acquisition and Extinction Phases. For each resulting cluster of maxima in difference between analysed responses were identified in the graph. Values of STFT z-Power at identified times and frequencies are reported in Table 1 for each analysed response including the significance level of the difference. In graphs, yellow lines represent respectively the onset and the offset of stimuli. Pink lines represent the shock pulses occurrence.

Cluster index	Time [s]	Frequency [Hz]	Stimulus	STFT median [z-Power]	STFT MAD [z-Power]	p-value
1	5.80	0.45	CS+/US	0.79	2.79	3.5*10 <sup>-7</sup>
			CS-	-0.04	2.15	
2	7.20	0.04	CS+/US	1.29	7.68	2.1*10 <sup>-12</sup>
			CS-	-0.18	2.48	
3	3.90	0.21	CS+	1.66	4.05	6.4*10 <sup>-12</sup>
			CS-	0.07	2.58	

**Table 1.** Time and Frequency of STFT z-Power maxima in difference between analysed responses were identified for each significant cluster ( $p < 0.001$ ). Median and median of absolute distances (MAD) values of STFT z-Power are reported separately for the compared responses indicated is the Stimulus column. The column p-value contains the significance level of the comparison at the identified maximum difference.

**Point-process modelling of heart rate dynamics.** To investigate unbiased spectral powers in light of the STFT results a point-process modelling algorithm of cardiac dynamics is involved to extract instantaneous spectral power indices of heart rate variability (HRV indices). These are calculated in four frequency ranges: very low frequencies (VLF) [0.003 0.03) Hz, low frequencies (LF) [0.03 0.15) Hz, inferior range of high frequencies (HF<sub>inf</sub>) [0.15 0.30) Hz, superior range of high frequencies (HF<sub>sup</sub>) [0.30 0.45) Hz. Stimuli responses of the normalized HRV indices for combined experiments are depicted in Fig. 5. Separate statistical analyses of CS+/US and CS+ against CS- ( $p < 0.001$ ) are conducted. During the Acquisition Phase, responses to the CS+/US compared to CS- results in significantly increased HF<sub>sup</sub>, HF<sub>inf</sub> (Fig. 5-a,d) and reduced VLF around 5 s (Fig. 5-l). Also, a significant increase results in LF around 8 s (Fig. 5-g). Regarding responses to the CS+ compared to CS-, results shows significantly increased HF<sub>inf</sub> around 4s sustained up to around 12 s (Fig. 5-e). No other significant differences were found in the other frequency ranges (Fig. 5, b,h,m). Finally, CS+ and CS- comparison during the Extinction Phase shows no significant differences in any range of frequencies investigated.



**Figure 5.** Normalized HRV indices median and median of absolute distances (MAD) responses for Acquisition Phase with highlighted spots representation of p-value at timepoints of significant ( $p < 0.001$ ) difference between stimuli and CS-. Significant increase of HF<sub>sup</sub>, HF<sub>inf</sub> (panel a, d) and reduced VLF (panel l) is associated to CS+/US subsequent to the shock administration. These are followed by an increase in LF (panel g). In response to the CS+ increased HF<sub>inf</sub> (panel e) around 4s is sustained up to around 12 s. In graphs yellow vertical lines represent respectively the onset and the offset of stimuli. Pink vertical lines represent the shock pulses occurrence.

## DISCUSSION

In humans, fear conditioning is usually assessed by skin conductance responses (SCR; Critchley et al., 2000) and fear potentiated startle (FPS; Brown et al., 1951). However, both SCR and FPS present some methodological and practical limitations. Thus, SCR is subject to a fast habituation decay, while FPS requires the presentation of US during both CS+ and CS- leading to a possible interference in the learning process (Castegnetti et al., 2016). Although previous research has extensively used HRV to assess sustained autonomic tone, its use as a tool to investigate more phasic changes, such as conditioned responses in emotional learning, is relatively new. Recent studies have shown how HRV may represent an important technique to study fear conditioning (Liu et al., 2013; Wendt et al., 2015; Pappens, et al., 2014; Tzovara et al., 2018). Specifically, it has been reported that fear conditioned stimuli (CS+) typically elicit transient heart rate deceleration, and this anticipatory fear bradycardia is generally believed to reflect vagal control (Castegnetti et al., 2016). The use of different and complementary methodologies to study physiological responses of fear conditioning may serve as markers for maladaptive fear learning and contribute to the identification of individuals prone to the development of psychiatric disorder (Sevenster et al., 2015). Indeed, an important difference between HRV and SCR is that HRV, as analyzed in the present study, is almost exclusively modulated by the parasympathetic nervous system (PNS; Berntson et al., 2007), while SCR is under almost exclusive control of the sympathetic nervous system (SNS; Boucsein, 2012). Assessing both measures at once could provide a sensitive methodological approach to evaluate the contribution of both sympathetic and parasympathetic autonomic learning.

The present study aimed to gain deeper insight into the autonomic cardiac underpinnings of fear learning in humans. In particular, HRV was analysed as a quantitative index of the interplay between sympathetic and parasympathetic influences on cardiac activity using frequency-domain techniques. Spectral analysis allows to study the intensity of the HRV



spectral components, with high frequency band [0.15 0.45 Hz) considered to be related primarily to cardiac parasympathetic outflows, and thus provide a direct index of vagal activity.

To this aim, we tested whether the spectral components of HRV, as a non-invasive marker of sympathetic and parasympathetic mechanisms, can dissociate between conditioned and neutral stimuli to fear learning. In two separate experiments, healthy participants underwent a fear conditioning acquisition procedure in which a visual stimulus (CS+) was paired with an electric shock as US, while a different and neutral stimulus (CS-) was never paired with a US. Fear acquisition was followed by an extinction phase during which both CSs were presented in the absence of the US.

Results revealed a specific pattern of heart rate modulations during the acquisition of fear conditioning. More specifically, RR wave responses to stimuli were decomposed into their components of heart rate during the acquisition phase, which allowed us to observe a well-known triphasic response (see Bohlin & Kjellberg, 1979), consisting of: (a) an initial deceleration, (b) followed by an acceleration, (c) and a further late deceleration around the time point in which US was expected. More importantly for the present purposes, the RR to CS+ showed larger dynamics than RR to CS-. Since the responses to conditioned and neutral stimuli begin to differ at about 1 s before the expected US onset, the late deceleration component appears to be due to the CS+ presentation rather than the US, thereby reflecting the anticipatory response to the fear conditioned stimulus. During the extinction phase, CS+ and CS- showed equivalence RR modulation of heart rate. Overall, these findings appear consistent with recent HR data of fear learning in humans (Castegnetti et al., 2016).

Crucially, subsequent frequency analysis of HRV during the acquisition phase revealed a higher concentration of power in the high frequency (HF) band, significantly larger after the CS+ than the CS- presentation, approximately around the time point in which participants expected the shock administration. These results indicate that the presentation of CS+,

compared to CS-, elicits a strong anticipatory response in the range of HF, indicating a specific vagal contribution to fear conditioning. Finally, spectral analyses revealed no significant difference between CS+ and CS- during the extinction phase.

These results suggest that fear conditioning has occurred during the acquisition phase of the experiment, and directly demonstrate the involvement of the vagus nerve on cardiac activity modulation during fear conditioning in humans. Overall, it appears that heart rate fluctuations (i.e., HRV) provides a powerful and robust biomarker of fear learning, particularly when explored with a power spectrum analysis approach. As previously noted (Castegnetti et al., 2016), this could be of particular importance in neuroimaging experiments, since MRI scanners are standardly equipped with a ECG to record cardiac activity, while equipment for recording SCR is less commonly available.

Importantly, these results may be in line with the neurovisceral integration (NVI) model (Thayer and Lane 2000; 2002), which suggests an extensive anatomical overlap between the distributed network of brain areas composing the central autonomic network (CAN), and the neural circuit critically involved in fear conditioning and emotional learning in humans. The structures of the CAN include regions in the prefrontal cortex, the insula, the amygdala, nuclei of the hypothalamus, and several brain stem nuclei. The CAN supports regulated emotional responding by flexibly adjusting physiological arousal in accordance with changing situational demands (Thayer & Lane, 2000; Thayer et al., 2009; Park et al., 2013). Through the sympathetic and parasympathetic – vagal – branches of the autonomic nervous system (ANS), the CAN directly regulates heart rate. Thus, HRV reflects the moment-to-moment output of the CAN and, by proxy, an individual's capacity to generate regulated physiological responses in the context of emotional learning (Thayer & Lane, 2000; Thayer & Siegle, 2002). Indeed, phasic HRV enhancement and parasympathetically dominated heart rate deceleration, as observed here when the US is expected, has been associated with emotional regulation (Park and Thayer,

2014), and ventral aspect of the medial prefrontal cortex (mPFC) activation (see Roelofs, 2017). This suggests that high frequency fluctuations in heart rate may prepare subject for an impending threat, an idea in keeping with our result that a higher concentration of power in the high frequency (HF) peaked approximately around the time the US was expected. However, additional datasets with more diverse SOAs between CS and US are needed to unambiguously confirm this result.

To sum up, the aim of the present study was twofold: first, to contribute to this field by investigating autonomic signatures, as biomarkers, of fear conditioning using spectral analysis approaches and as result, the second aim was to quantify the vagal contribution to learned fear. We report that HRV is particularly sensitive to changes in ANS activity (i.e., changes in both sympathetic and parasympathetic nervous system) during fear acquisition and extinction. High parasympathetic activity, which is characterized by an increase in the HF and a decrease in the LF was found in response to CS+, as compared to CS. HRV may be associated with the activity of a network of neural structures, the CAN, which are dynamically organized in response to environmental challenges. Indeed, neuroimaging studies suggest that HRV may be linked to reduced threat perception, mediated by cortical regions (e.g., mPFC) involved in the appraisal of threatening situations. Although, several earlier studies have argued that the vagus nerve may play a crucial role in fear conditioning, the results of the present study provide the first direct evidence that systematically investigate and more importantly quantified its involvement in the human fear conditioning framework.



## GENERAL CONCLUSION

The main goal of the present thesis was to enrich the knowledge and understanding of the functional and neural mechanisms underlying fear conditioning in humans, by the means of studying its psychophysiological correlates in both healthy and brain-damaged individuals. A series of studies have been detailed reported in previous chapters and several results uncovered new important evidence that might provide bases to shed light on existing theories of fear learning and extinction as well as develop treatments for a variety of neurological and psychiatric disorders. In the next paragraphs, studies, results and their consequent implications will be discussed.

**The influence of normal aging on context-dependent recall of extinction.** The Study 1 was the first in human fear conditioning literature aiming to examine the influence of normal aging on context-dependent recall of extinction. Healthy young and old adults were tested in a multi-phase study over two days (Garfinkel et al., 2014; Milad et al., 2007). During the first day, participants were fear conditioned to two visual stimuli (CS+ and CS-) within a specific (danger) visual context, and then extinguished the previous conditioning within a different (safe) context. On the second day, the ability to selectively recall extinction memory within these two different contexts (danger and safe) was assessed. In particular, the extinguished stimuli were presented to investigate both extinction recall (in safe context), and contextual fear renewal (in danger context). Results showed that young participants were able to use contextual information to flexibly guide their learned responses to threat (as expressed by SCR), whereas older participants showed impaired modulation of the responses by contextual information. In particular, on the first day, all participants were equally able to acquire and completely extinguish the conditioned response. On the second day, young participants showed a context-dependent fear response for CS+, as compared to CS-, in the danger context, but not in the safe

context (Vansteenwegen et al., 2005). In contrast, older adults showed an impaired context-guided recall of extinction, with higher SCR to CS+, as compared to CS-, in both danger and safe context. These results are consistent with the presence of either a specific extinction recall deficit, or a more general context-processing deficit. These finding that differential responding to the CS+ versus the CS- increased from the safe (extinction) to the danger (renewal) context in the young but not in the older participants strongly suggests that aging is associated with a more general loss of context sensitivity in memory expression. Moreover, on day 2, a positive correlation between differential threat responses in the safe (extinction recall) and danger (threat renewal) context was found in the older, but not in the young group. This suggests that aging is associated with loss of contextual control of extinction, causing extinguished memories to inappropriately renew in any - irrelevant - context. Taken together, these findings indicate that older adults were less able to use contextual information to recall extinction memory and modulate the expression of the defensive responses to threat in a context-dependent manner, despite their preserved ability to acquire and extinguish the conditioned response.

Current theorizing in human cognitive aging offers a wide variety of accounts for performance decline in context processing and its utilization as occasion setter, including poor distribution of attentional resources (Braver et al., 2001; Hartley, 1992), reduction in working-memory capacity (Just & Carpenter, 1992; Van der Linden et al., 1999) and failure of inhibitory processes (Hasher & Zacks, 1988). These represent distinct but highly interdependent mechanisms that may influence each other (Spencer & Raz, 1995). Although the present study did not directly investigate the neural substrates of conditioning and extinction in aging, the deficit of context-guided recall of extinction may be linked to age-related changes in the neural structures underpinning context-dependent behaviour (Foster et al., 2012). Consistent with this view, neuroimaging studies in humans (Kalisch et al., 2006; Garfinkel et al., 2014) reported that ventromedial prefrontal cortex and hippocampus, as neural network, is selectively involved

in context-dependent regulation of extinction and threat memories. More specifically, during recall of extinction memory, the medial prefrontal cortex would act to inhibit the amygdala, preventing a response to a threat, based on contextual information provided by the hippocampus (Hobin et al., 2003; Kalisch et al., 2006; Delamater, 2004). Furthermore, there is substantial evidence that several structural and physiological alterations preferentially influence the prefrontal cortex and medial temporal lobe in advanced aging, even in the absence of disease (Buckner, 2004; Bartzokis et al., 2001; Andrews-Hanna et al., 2007). These disruptive brain changes may underlie impairments in context-dependent extinction recall, as well as cause the decreased efficiency with which older adults use contextual information to determine when and where it is appropriate to express fear.

Some considerations and questions could be made. Well-grounded literature assumes that human prefrontal cortex regulates the activity of amygdala after fear conditioning as in animals, but scarce human evidence supports such theorization. Also, it is possible to consider that there are functional differences within the prefrontal cortex between humans and animals. At this point, it is possible to wonder: what is the role of the prefrontal cortex in human fear conditioning? Possible answers may be provided by Study 2 and discussed in the next paragraph.

**Rethinking human prefrontal cortex: evidence from brain-damaged patients.** The role of the ventromedial prefrontal cortex (vmPFC) in human fear conditioning has been largely attributed to the extinction learning (Phelps et al., 2004) rather than the acquisition of conditioning. However, recent neuroimaging studies have questioned this view by showing the activation of vmPFC also during the acquisition of conditioning, rising the hypothesis for a critical role of this brain region than previously theorized during the early stages of conditioning (Fullana et al., 2016; Dunsmoor et al., 2019). The ignored role of vmPFC in the acquisition of

fear conditioning may be also by a lack of neuropsychological studies assessing the consequences of a lesion to the vmPFC in humans. The only existing study found preserved conditioned SCR to the presentation of a visual stimulus previously associated with an aversive sound, despite patients with vmPFC lesion failing to show anticipatory SCR to negative during a gambling task (Bechara et al., 1999). Thus, Study 2 aimed to reevaluate the role of vmPFC in the acquisition of fear conditioning in humans. A group of ten patients with a bilateral lesion to vmPFC, a group of ten patients with a lesion outside PFC or medial temporal lobe and a group of ten healthy participants completed a differential fear conditioning paradigm, while their SCR to CS+, CS- and US was recorded. Results showed patients with a lesion to vmPFC failed to produce conditioned responses during the acquisition of conditioning, as evidenced by impaired anticipatory differential SCR between CS+ and CS- in both healthy participants and brain damage control patients. In contrast, there was no evidence that lesion to vmPFC compromised unconditioned responses or declarative memory for CS-US contingencies.

Study 2 provides the first evidence establishing that human vmPFC is necessary for the expression of conditioned physiological responses (i.e. SCR) during the acquisition of fear conditioning, indicating that, vmPFC is necessary since the early stages of conditioning. Thus, a lesion to vmPFC results in more pervasive impairments than previously reported (Phelps et al., 2004). However, the preserved unconditioned response and declarative memory of stimulus-outcome contingencies suggest that these two processes do not rely on a functional vmPFC and that they are mediated by different pathways than the one generating conditioned responses. Also, preserved declarative memory does not appear sufficient to generate conditioned responses, as suggested by the lack of conditioned responses despite awareness of stimulus-outcome contingencies. Furthermore, the preserved unconditioned response is consistent with evidence that inactivation of prelimbic cortex in rats does not abolish US response (Corcoran & Quirk, 2007) indicating that unconditioned physiological response is



mediated by a different pathway than conditioned response, which does not require an intact vmPFC.

Current theorizations on the vmPFC suggest that it has a crucial role in the value and stimulus-outcome representation (Hiser & Koenigs, 2018; Schoenbaum et al., 2009; Schoenbaum et al., 2011) as well as model-based computations (Wilson et al., 2014). Given the role of vmPFC in value and stimulus-outcome representation (Hiser & Koenigs, 2018; Schoenbaum et al., 2009; Schoenbaum et al., 2011), vmPFC may be necessary to encode or learn the value of CS during acquisition. Although the sensory information about CS and US would still be able to converge into the amygdala, where stimulus-outcome associations are passively formed, the vmPFC may be necessary to turn this information into expectations regarding the value of future events (Roesch & Schoenbaum, 2006) and thus, enabling anticipatory responses. This interpretation also suggests that in humans, vmPFC may have a more prominent role in fear conditioning than in rodents, where its lesion does not impair acquisition (Morgan et al., 1993; Morgan & LeDoux, 1995; Morrow et al., 1999; Quirk et al., 2000). Thus, it may be possible that the failure of patients with vmPFC lesion to produce conditioned responses indicates that the vmPFC is necessary for the expression of conditioned responses, rather than the encoding/learning of CS value. Further theoretical consideration could be made about the functional role of prefrontal cortex in humans.

Although Study 1 did not directly investigate the neural substrates of conditioning in ageing, it demonstrated a specific deficit of context-guided recall of extinction which may be linked to functional prefrontal cortex alteration due to age-related changes. Indeed, has been found considerable evidence that prefrontal cortex and medial temporal lobe are functionally altered in advanced ageing, even in the absence of neurological disease. Furthermore, Study 2 highlighted the crucial role of ventromedial prefrontal cortex in humans in acquiring fear conditioning. In particular, Study 2 highlighted that vmPFC lesion did not alter the ability to

acquire declarative knowledge about which conditioned stimulus was paired with the unconditioned stimulus (i.e., contingency awareness) revealing a dissociation between how vmPFC patients anticipate the impending threat and how they explicitly report the experimental contingencies. Importantly, Study 2 results cannot be explained based on autonomic response deficit (i.e. SCR impaired) after neurological damage. All prefrontal cortex patients showed normal SCRs whenever the US was administered together with the CS.

Given the psychophysiological and behavioural results of Study 1 and the brain-damaged patients' evidence of Study 2, it is a must to ask if human prefrontal cortex may unveil other neurobiological functional differences in fear conditioning from the animal kingdom. In the subsequent paragraph, Study 3 will be discussed aiming to answer about the role of human prefrontal cortex in reconsolidation of fear memory.

**Beyond fear: erasing human fear responses by disrupting prefrontal cortex.** From an evolutionary perspective, it is extremely functional to never forget the most important events in life. However, the putative indelibility of emotional memory can also be maladaptive, such as in some traumatized individuals who may suffer from dreadful memories and anxiety. In contrast to the traditional view of memory consolidation (Squire and Davis, 1981), the reconsolidation hypothesis suggests that stored information is rendered labile after being retrieved and it raised the possibility of interfering with existing memories during a temporary window of lability. Considerable evidence in both animals and humans indicates that blockade of the process of reconsolidation by pharmacological manipulations produces amnesia for the original fear learning (Nader and Hardt, 2009, Sevenster et al., 2013). Kindt and colleagues (2009) demonstrated that fear response can be weakened by disrupting the reconsolidation process of fear memory and that disrupting should prevent the return of fear. Although pharmacological manipulations are potentially useful for changing learned fears, their use in

humans can be problematic. Also, pharmacological effects have been not consistently replicated (Wood et al., 2015). Beside pharmacological interventions, non-invasive methods have also been studied to interfere with fear memories. Schiller and colleagues (2010) used an extinction protocol within the reconsolidation window to behaviourally trigger reconsolidation for fear memories. However, a common problem with most of the reported reconsolidation procedures is the lack of replicability and consistent long-term effects. If emotional memory could be weakened or even erased, then we might be able to eliminate the root of many psychiatric disorders. Indeed, it would be particularly significant in pathologies in which intrusive traumatic memories significantly affect daily life, as in the case of Posttraumatic Stress Disorders (PTSD) or Generalized Anxiety Disorder (GAD).

Although severe evidence has been reported within the reconsolidation framework in the last decade, the neural substrates of fear memory reconsolidation in humans remain largely unknown. One potential way to unveil the neural mechanism underlie reconsolidation would be using non-invasive brain stimulation, as the Transcranial Magnetic Stimulation (TMS). Delgado and colleagues (2008) suggested that prefrontal cortex could be also implicated in reconsolidation process because has direct connections with the amygdala (Sladky et al., 2013), which is a key region for emotional learning. Recent evidence suggested that prefrontal cortex may be implicated in the reconsolidation process (Munsee et al., 2014). It has been reported that stimulating the dorsolateral prefrontal cortex during extinction learning, produce an enhancing of extinction (Van 't Wout et al., 2016), or inhibition of fear memory consolidation (Asthana et al., 2013).

In Study 3, using an experimental protocol of fear conditioning procedure and return of fear in different days (Schiller et al., 2010), it has been tested whether the application of the rTMS over the right and left dlPFC during the reconsolidation window is able to impact the reconsolidation of the fear memory. Thus, it was expected that TMS administered over dlPFC,

would interfere with the reconsolidation process of fear memory after memory reactivation. The experimental paradigm consisted in a 3 days experiment: fear acquisition (day 1), memory reactivation and rTMS (day 2), and memory recall followed by extinction and reinstatement procedure (day 3). To ensure the state-dependent efficacy of the experimental protocol, in four groups of healthy humans, we administered rTMS following a reminder of the fear memory able to trigger the reconsolidation process (Schiller et al., 2009; Sevenster et al., 2013; Merlo et al., 2014), and these groups were directly compared with an additional control group in which no reminder was used (No-reminder). Results showed on day 1, that all experimental groups acquired fear conditioning, demonstrating that fear learning took place equivalently across all participants. On day 2, results revealed no differences of the group on SCR for the reactivated trials. These data demonstrate that, before the reconsolidation manipulation, the conditioned response was equally expressed across groups. On day 3, both l-dIPFC and r- dIPFC groups showed decreased physiological expression of fear, indexed by SCR (CS+ similar to CS-), both in memory recall and after extinction-reinstatement phases. On contrary, no decrease was observed in participants receiving either control rTMS (i.e., stimulation of a control site and sham stimulation), or dIPFC-rTMS without preceding reactivation of fear memory (No-reminder), thus showing both the site-specificity and state-dependency of our rTMS intervention. These results showed that participants persistently expressed fear - in terms of both psychophysiological reactions and subjective ratings - when the memory was not reactivated by the presentation of the CS+ confirms that the dIPFC manipulation via rTMS was state-dependent, and specifically acted on the memory reconsolidation process (Elsley et al., 2018). These results, together with the absence of fear recovery following reinstatement (Barak & Ben Hamida, 2012), argue for direct modification of the original memory trace, rather than the formation of a new memory, as occurs in extinction (Bouton, 2002; Raji et al., 2018).

Study 3 provides strong evidence supporting the following hypothesis: when reactivated, memories enter a transient and labile state that can result in the enhancement or weakening of that specific trace (Agren, 2014). Also, disrupting the dlPFC during the reconsolidation time-window is likely to have altered prefrontal functional connections not only with the hippocampus but also with the amygdala, which is associated with the fearful component of the reactivated memory trace (Munsee et al., 2014). This suggests that interfering with normal brain activity during the consolidation time-window, the connections between frontal regions and the amygdala were weakened, thus resulting in decreased fear expression (Munsee et al., 2014).

Remarkably, Study 3 provide causal evidence that the return of fear can be prevented by reactivating the original memory trace by disrupting dlPFC activity by the mean of state-dependent TMS. These findings provide a step forward understanding the mechanisms underlying fear memory reconsolidation, and they have potential clinical implications for targeting emotional, maladaptive memories (Pennington et al., 2018). Study 3 results demonstrate that non-invasive stimulation of the prefrontal cortex following memory reactivation disrupts the expression of fear to a previously conditioned threatening stimulus, and highlight the critical role of the PFC within the neural network that mediates the reconsolidation of conditioned fear memories in humans. Finally, this study unveiled an important crucial role of the PFC in the modification of a previously acquired fear memory and thus, preventing the return of fear in humans.

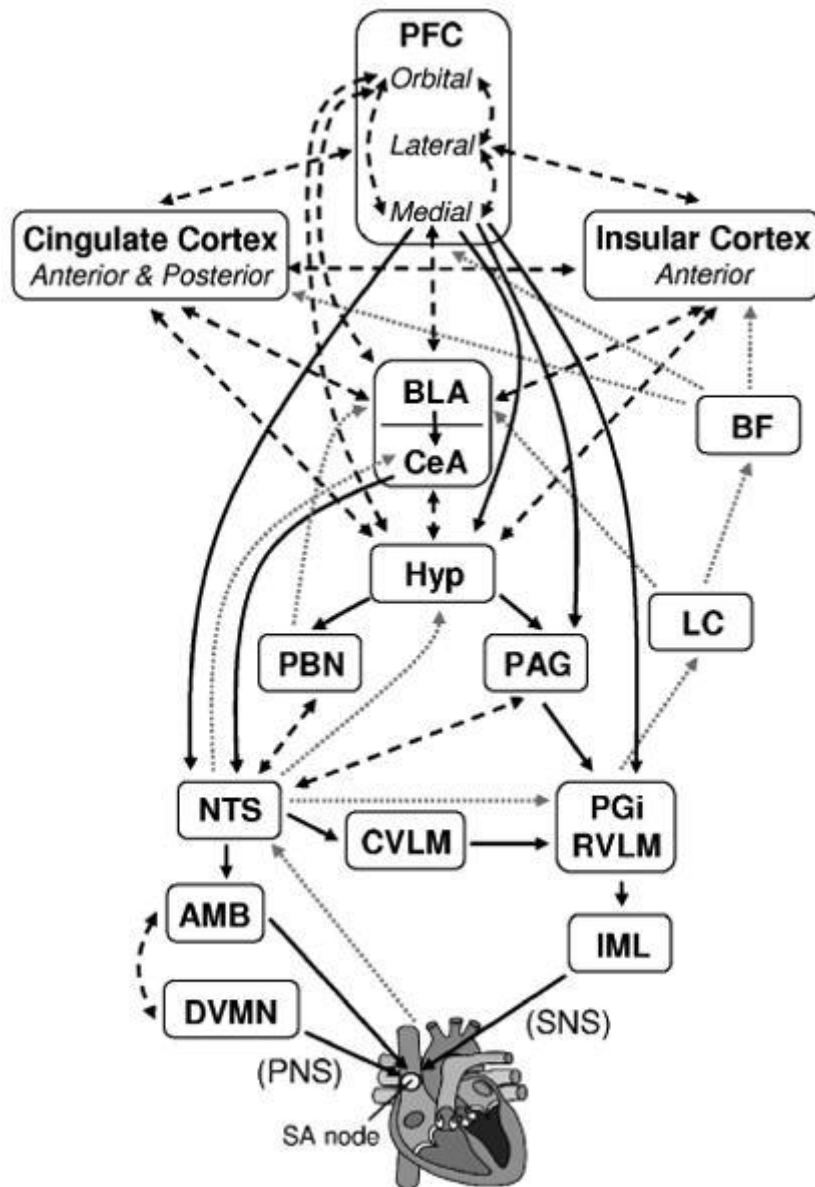
**Vagally mediated heart rate dynamics reflect the encoding of fearful memory.** Learning to respond to specific cues or environmental circumstances that predict impending threat is a highly adaptive function for animals and humans as well (Bouton et al., 2002). Fear conditioning is the most used laboratory paradigm to study this phenomenon producing both behavioural and physiological conditioned responses (LeDoux, 2000). The study of fear

conditioning is widely held to be a model for psychiatric disorders in humans. Thus, the development of different and complementary methodologies to study physiological responses of fear conditioning might serve as markers for maladaptive fear learning and contribute to the identification of individuals prone to the development of psychiatric disorders (Sevenster et al., 2015).

Study 4 intended to contribute to the human fear conditioning framework by investigating autonomic signatures, and characterizing as biomarkers, using new methodological approach to quantify the vagus nerve involvement in heart rate modulations during fear conditioning paradigm. In Study 4, participants underwent a delay fear conditioning procedure in which a visual stimulus (CS+) was paired with an electric shock as US, while a different stimulus (CS-) was never paired with a US. Fear acquisition was followed by an extinction phase during which both CSs were presented in the absence of the US. It has tested whether the spectral components of the HRV, as a non-invasive marker of sympathetic and parasympathetic mechanisms, can dissociate between conditioned and neutral stimuli related to fear learning. Results revealed a specific pattern of heart rate dynamics during the acquisition of fear conditioning. Primarily, consistent with precedent studies (Bohlin & Kjellberg, 1979; Castegnetti et al., 2016) it has observed well-known triphasic RR responses of heart rate during the acquisition phase. In particular, such responses were found larger to CS+ as compared to CS-, reflecting the acquisition of fear conditioning and a reduction during the extinction phase. More importantly, spectral analysis of HRV during the acquisition phase revealed a higher concentration of power in the high frequency (HF) band, significantly larger after the CS+ than the CS-, approximately around the time in which participants expected the shock administration. These results indicate that the presentation of CS+ elicits a strong anticipatory response in the range of HF, indicating a specific vagal contribution encoding fear conditioned stimuli.

Overall, Study 4 results demonstrated the involvement of the vagus nerve on cardiac activity modulation during fear conditioning in humans. In particular, HRV was assessed as a quantitative index of the interplay between sympathetic and parasympathetic influences on cardiac activity using frequency-domain techniques. The spectral analysis allowed to study the intensity of the HRV spectral components, and the high-frequency band has found a significant cluster to CS+, as compared to CS-. It is important to consider that, High-Frequency band [0.15-0.45 Hz) is considered to be related primarily to cardiac parasympathetic outflows, and thus provide a direct index of vagal activity (Task Force, 1996).

Notably, Study 4 results may be consistent with the neurovisceral integration (NVI) model (Thayer and Lane 2000; 2002), which suggests an extensive anatomical overlap between the distributed network of brain areas composing the central autonomic network (CAN), and the neural circuit critically involved in fear conditioning and emotional learning in humans. The structures of the CAN include: anterior cingulate cortex (ACC), anterior (AI) and posterior (PI) insula, ventromedial prefrontal cortex (VMPFC), orbitofrontal cortex (OFC), amygdala, nucleus of the stria terminalis (NST), hypothalamus, periaqueductal gray (PAG), parabrachial nucleus (PBN), nucleus of the solitary tract (NTS), nucleus ambiguus (NA), dorsal motor nucleus of the vagus (DMNV), noradrenergic locus coeruleus (LC), the rostral (RVLM) and caudal (CVLM) ventrolateral medulla (see Diagram. 1; Ellis and Thayer, 2010; Park and Thayer, 2014).



**Diagram 1. Neural structures involved in the control of heart rate.** Solid black arrows indicate efferent pathways to the heart, including right vagus nerve (PNS) and stellate ganglion (SNS) inputs to the SA node. Dotted gray arrows indicate afferent pathways to medullary structures via aortic baroreceptor signals carried through the vagus. Dashed black arrows indicate bidirectional connections. AMB: nucleus ambiguus; BF: basal forebrain; BLA: basolateral amygdala; CeA: central nucleus of the amygdala; CVLM: caudal ventrolateral medullary neurons; DVMN: dorsal vagal motor nuclei; Hyp: hypothalamus (lateral and paraventricular); IML: intermediolateral cell column of the spinal cord LC: locus coeruleus; NTS: nucleus of the solitary tract; PAG: periaqueductal gray; PBN: parabrachial nuclei; PFC: prefrontal cortex; PGi: nucleus paragigantocellularis; RVLM: rostral ventrolateral medullary neurons. Diagram is reproduced from Ellis and Thayer (2010).



The model suggests the existence of a reciprocal inhibitory cortico-subcortical brain circuit, and in particular highlight, the role of the prefrontal cortex (PFC) in the modulation of subcortical cardio-acceleratory circuits via an inhibitory pathway that is associated with vagal function and it can be indexed by heart rate variability (Luque-Casado et al., 2016).

Under normal circumstances, prefrontal cortex identifies safety cues from the environment and exerts its inhibitory control over sympatho-excitatory subcortical circuits (Heatherton and Wagner, 2011). On contrary, in threatening and uncertain situations, prefrontal inhibitory regulation diminishes and sympatho-excitatory subcortical circuits eliciting fear responses (Park et al., 2013). Indeed, the role of the PFC in exerting inhibitory control over subcortical brain structures is crucial for modulating vagally mediated HRV (Wendt et al., 2015). These pathways have been discussed in detail by Thayer & Lane (2009) and Thayer and colleagues (2012), suggesting that the dorsal and the ventral surface of the prefrontal cortex, are involved in threat responses, and modulate amygdala activity via GABAergic intercalated cells. The output of the amygdala via the nucleus of the solitary tract (NTS) impacts the output of the vagal motor neurons in the medulla through a network of interneurons connecting the NTS with the nucleus ambiguus (NA) and the dorsal motor nucleus of the vagus (DVM). The effect is that sympatho-excitatory circuits in the medulla are tonically inhibited by the vmPFC. Consistently, disruption of prefrontal-subcortical circuits has been associated with a wide range of psychopathologies, including depression (Davidson et al., 2002; Johnstone et al., 2007), anxiety (Kim & Whalen, 2009), schizophrenia (Callicott et al., 2003; Lewis et al., 2005), and addictive behaviour (for a review, Li & Sinha, 2008). Furthermore, the heart is dually innervated by the sympathetic and parasympathetic branches of the autonomic nervous system, which innervate the heart through both stellate ganglia and the vagus nerve. The integrated effects of these different signal pathways, converge into the sino-atrial (SA) node of the heart, determining HRV. Thus, the heart's beat-to-beat variation is the result of the interplay of

sympathetic and parasympathetic activity and HRV reflects the moment-to-moment output of the CAN and, by proxy, an individual's capacity to generate regulated physiological responses in the framework of emotional learning (Thayer & Lane, 2000; Thayer & Siegle, 2002; Lane et al., 2009).

According to the NVI model and brain areas composing the CAN, the functioning of prefrontal-subcortical inhibitory circuits critical for self-regulation is linked with the heart via the vagus nerve that provides inhibitory inputs to the heart (Levy, 1971; Benarroch, 1993; Ellis and Thayer, 2010). Although several studies have argued that the vagus nerve may play a crucial role in fear conditioning, Study 4 results provide the first direct evidence that systematically investigates, and more importantly, quantified its involvement in the human fear conditioning. Study 4 provides a powerful and robust biomarker of fear learning, particularly when assessed with a power spectrum analysis approach and thus, unique insights into the psychophysiological measurement of fear memory. Finally, Study 4 results may provide a solid ground to develop new specific diagnosis protocol and treatment monitoring for several psychiatric disorders.

## **FINAL REMARKS**

In sum, the present PhD thesis contributed to extending the current literature on the description and the understanding of emotional, fear learning in humans. Evidence reported in this thesis might provide key insights and deeper understanding of critical issues concerning the acquisition, the extinction and the reconsolidation of fear memories. Also, it may help to solve methodological issues concerning the assessment of heart rate as psychophysiological index in fear conditioning framework, providing a new reliable biomarker in humans.

In particular, in Study 1 I showed that healthy aging processes have no impact on conditioned learning but may affect the capacity to use contextual information to modulate fear learned responses to an impending threat.

In Study 2, I reported that patients with selective bilateral damage to the ventromedial prefrontal cortex (vmPFC) maintain intact explicit knowledge of the conditioning, but fail to elicit conditioned to visual stimuli predicting threat, thus establishing that the vmPFC is necessary for the expression of conditioned fear.

In Study 3, I found that transcranial magnetic stimulation administered over the dorsolateral PFC, following memory reactivation, can erase the expression of autonomic responses of fear memory, and prevent the return of fear. Thus, demonstrating that PFC is necessary for the post-retrieval modification of learned fear memories.

In Study 4, using a classical fear conditioning experimental procedure, I studied the heart rate modulations to conditioned stimuli and developed new algorithm to analyse transitory heart rate variation. Finally, I was able to characterize the autonomic cardiac signatures, as biomarkers, of fear conditioned stimuli.

## REFERENCES

- Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability: a review. *Medical and biological engineering and computing*, 44(12), 1031-1051.
- Acheson, D. T., Forsyth, J. P., Prenoveau, J. M., & Bouton, M. E. (2007). Interoceptive fear conditioning as a learning model of panic disorder: An experimental evaluation using 20% CO<sub>2</sub>-enriched air in a non-clinical sample. *Behaviour Research and Therapy*, 45(10), 2280-2294.
- Agren, T. (2014). Human reconsolidation: A reactivation and update. *Brain research bulletin*, 105, 70-82.
- Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Berger, A. C., & Cohen, R. J. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *science*, 213(4504), 220-222.
- Andreatta, M., Glotzbach-Schoon, E., Mühlberger, A., Schulz, S. M., Wiemer, J., & Pauli, P. (2015). Initial and sustained brain responses to contextual conditioned anxiety in humans. *Cortex*, 63, 352-363.
- Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron*, 56(5), 924-935.
- Asthana, M., Nueckel, K., Mühlberger, A., Neueder, D., Polak, T., Domschke, K., ... & Herrmann, M. J. (2013). Effects of transcranial direct current stimulation on consolidation of fear memory. *Frontiers in psychiatry*, 4(107), 1-7.
- Baker, A. G., Steinwald, H., & Bouton, M. E. (1991). Contextual conditioning and reinstatement of extinguished instrumental responding. *The Quarterly Journal of Experimental Psychology*, 43(2), 199-218.

- Bannerman, D. M., Rawlins, J. N. P., McHugh, S. B., Deacon, R. M. J., Yee, B. K., Bast, T., ... & Feldon, J. (2004). Regional dissociations within the hippocampus—memory and anxiety. *Neuroscience & biobehavioral reviews*, 28(3), 273-283.
- Barak, S., & Hamida, S. B. (2012). Memory erasure, enhanced extinction and disrupted reconsolidation. *Journal of Neuroscience*, 32(7), 2250-2251.
- Barbieri, R., Matten, E. C., Alabi, A. A., & Brown, E. N. (2005). A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology*, 288(1), H424-H435.
- Bartzokis, G., Beckson, M., Lu, P. H., Nuechterlein, K. H., Edwards, N., & Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Archives of general psychiatry*, 58(5), 461-465.
- Battaglia, S., Garofalo, S., & di Pellegrino, G. (2018). Context-dependent extinction of threat memories: influences of healthy aging. *Scientific reports*, 8(1), 12592.
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology*, 2(142-162).
- Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of neuroscience*, 19(13), 5473-5481.
- Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, 269(5227), 1115-1118.
- Beckers, T., Krypotos, A. M., Boddez, Y., Effting, M., & Kindt, M. (2013). What's wrong with fear conditioning?. *Biological psychology*, 92(1), 90-96.
- Bellebaum, C. & Daum, I. (2004). Fear, conditioning, and aging: Theoretical comment on LaBar et al. (2004). *Behavioral Neuroscience*. 118, 1137–1139.

- Benarroch, E.E. (1993). The central autonomic network: functional organization, dysfunction, and perspective. In *Mayo Clinic Proceedings*, 68(10), 988-1001.
- Bernacer, J., Corlett, P. R., Ramachandra, P., McFarlane, B., Turner, D. C., Clark, L., ... & Murray, G. K. (2013). Methamphetamine-induced disruption of frontostriatal reward learning signals: relation to psychotic symptoms. *American Journal of Psychiatry*, 170(11), 1326-1334.
- Berntson G.G., Quigley K. S., Lozano D., Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology*, 2007, (pp. 182-210). New York, NY, US: Cambridge University Press.
- Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: the modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological review*, 98(4), 459-487.
- Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Cardiac psychophysiology and autonomic space in humans: empirical perspectives and conceptual implications. *Psychological bulletin*, 114(2), 296-322.
- Besnard, A., Caboche, J., & Laroche, S. (2012). Reconsolidation of memory: a decade of debate. *Progress in neurobiology*, 99(1), 61-80.
- Bhattacharya, B., & Habtzghi, D. (2002). Median of the p value under the alternative hypothesis. *The American Statistician*, 56(3), 202-206.
- Bitsios, P., Szabadi, E., & Bradshaw, C. M. (2004). The fear-inhibited light reflex: importance of the anticipation of an aversive event. *International Journal of Psychophysiology*, 52(1), 87-95.
- 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.

- Bohlin G., & Kjellberg A. (1979) Orienting activity in two-stimulus paradigms as reflected in heart rate. In: H.D. Kimmel, E.H. van Olst, & J.F. Orlebeke (Eds), *The orienting reflex in humans* (pp. 169-197), New Jersey: Lawrence Erlbaum Associates.
- Boucsein W., Fowles D.C., Grimnes S., Ben-Shakhar G., Roth W.T., Dawson M.E., Filion D.L. (2012). Publication recommendations for electrodermal measurements, *Psychophysiology*, 49:1017-1034.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological bulletin*, 114(1), 80-99.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological psychiatry*, 52(10), 976-986.
- Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & memory*, 11(5), 485-494.
- Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and motivation*, 10(4), 445-466.
- Bouton, M. E., & Brooks, D. C. (1993). Time and context effects on performance in a Pavlovian discrimination reversal. *Journal of Experimental Psychology: Animal Behavior Processes*, 19(2), 165-179.
- Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 248-265.
- Bouton, M. E., & Ricker, S. T. (1994). Renewal of extinguished responding in a second context. *Animal Learning & Behavior*, 22(3), 317-324.
- Bouton, M. E., & Swartzentruber, D. (1986). Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 12(4), 333-350.

- Bouton, M. E., Kenney, F. A., & Rosengard, C. (1990). State-dependent fear extinction with two benzodiazepine tranquilizers. *Behavioral neuroscience*, 104(1), 44-55.
- Bouton, M. E., Westbrook, R. F., Corcoran, K. A., & Maren, S. (2006). Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biological psychiatry*, 60(4), 352-360.
- Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. *Cognitive neuroscience of emotion*, 25, 49-59.
- Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433-436.
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., ... & Jagust, W. J. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, 130(4), 746-763.
- Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., & Barch, D. M. (2005). Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. *Psychology and aging*, 20(1), 33-46.
- Bremner, J. D., Staib, L. H., Kaloupek, D., Southwick, S. M., Soufer, R., & Charney, D. S. (1999). Neural correlates of exposure to traumatic pictures and sound in Vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study. *Biological psychiatry*, 45(7), 806-816.
- Brooks, D. C. (2000). Recent and remote extinction cues reduce spontaneous recovery. *The Quarterly Journal of Experimental Psychology: Section B*, 53(1), 25-58.
- 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.



- Büchel, C., Dolan, R. J., Armony, J. L., & Friston, K. J. (1999). Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 19(24), 10869-10876.
- Büchel, C., Morris, J., Dolan, R. J., & Friston, K. J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron*, 20(5), 947-957.
- Buckner, R. L. (2004). Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. *Neuron*, 44(1), 195-208.
- 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.
- Burgos-Robles, A., Vidal-Gonzalez, I., & Quirk, G. J. (2009). Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. *Journal of Neuroscience*, 29(26), 8474-8482.
- Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: functional neuroimaging evidence. *Current opinion in neurology*, 13(4), 415-421.
- Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). A short version of the Stroop test: normative data in an Italian population sample. *Nuova Rivista di Neurologia*, 12(4), 111-115.
- Callicott, J. H., Egan, M. F., Mattay, V. S., Bertolino, A., Bone, A. D., Verchinski, B., & Weinberger, D. R. (2003). Abnormal fMRI response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. *American Journal of Psychiatry*, 160(4), 709-719.
- Capitani, E., & Laiacona, M. (1988). Aging and psychometric diagnosis of intellectual impairment: some considerations on test scores and their use. *Developmental Neuropsychology*, 4(4), 325-330.

- Carter, R. M., Hofstötter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences*, 100(3), 1399-1404.
- Cassell, M. D., & Wright, D. J. (1986). Topography of projections from the medial prefrontal cortex to the amygdala in the rat. *Brain research bulletin*, 17(3), 321-333.
- Castegnetti, G., Tzovara, A., Staib, M., Paulus, P. C., Hofer, N., & Bach, D. R. (2016). Modeling fear-conditioned bradycardia in humans. *Psychophysiology*, 53(6), 930-939.
- Censor, N., Dimyan, M. A., & Cohen, L. G. (2010). Modification of existing human motor memories is enabled by primary cortical processing during memory reactivation. *Current Biology*, 20(17), 1545-1549.
- Chen, R. M. M. F., Classen, J., Gerloff, C., Celnik, P., Wassermann, E. M., Hallett, M., & Cohen, L. G. (1997). Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology*, 48(5), 1398-1403.
- Cheng, D. T., Disterhoft, J. F., Power, J. M., Ellis, D. A., & Desmond, J. E. (2008). Neural substrates underlying human delay and trace eyeblink conditioning. *Proceedings of the National Academy of Sciences*, 105(23), 8108-8113.
- Chhatwal, J. P., Myers, K. M., Ressler, K. J., & Davis, M. (2005). Regulation of gephyrin and GABAA receptor binding within the amygdala after fear acquisition and extinction. *Journal of Neuroscience*, 25(2), 502-506.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.
- Corcoran, K. A., & Maren, S. (2001). Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *Journal of Neuroscience*, 21(5), 1720-1726.
- Corcoran, K. A., & Maren, S. (2004). Factors regulating the effects of hippocampal inactivation on renewal of conditional fear after extinction. *Learning & Memory*, 11(5), 598-603.

- Corcoran, K. A., & Quirk, G. J. (2007). Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *Journal of Neuroscience*, 27(4), 840-844.
- Corcoran, K. A., & Quirk, G. J. (2007). Recalling safety: cooperative functions of the ventromedial prefrontal cortex and the hippocampus in extinction. *CNS spectrums*, 12(3), 200-206.
- Corcoran, K. A., Desmond, T. J., Frey, K. A., & Maren, S. (2005). Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *Journal of Neuroscience*, 25(39), 8978-8987.
- Craik, F. I., & Byrd, M. (1982). Aging and cognitive deficits. In *Aging and cognitive processes* (pp. 191-211). Springer, Boston, MA.
- Craske, M. G. (1999). *Anxiety disorders: psychological approaches to theory and treatment*. Westview Press.
- Craske, M. G., Hermans, D. E., & Vansteenwegen, D. E. (2006). *Fear and learning: From basic processes to clinical implications*. American Psychological Association.
- Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour research and therapy*, 46(1), 5-27.
- Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron*, 33(4), 653-663.
- Culbertson, W. C., & Zillmer, E. A. (1999). *The Tower of London, Drexel University, research version: examiner's manual*. North Tonawanda: Multi-Health Systems.
- Cunningham, C. L. (1979). Alcohol as a cue for extinction: State dependency produced by conditioned inhibition. *Animal Learning & Behavior*, 7(1), 45-52.

- Davidson, R. J., Pizzagalli, D., Nitschke, J. B., & Putnam, K. (2002). Depression: perspectives from affective neuroscience. *Annual review of psychology*, 53(1), 545-574.
- Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. *The amygdala*, 2, 213-287.
- Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of psychophysiology*, 2, 200-223.
- De Peuterl, S., Van Diestl, I., Vansteenwegen, D., Van den Berghl, O., & Vlaeyenl, J. W. (2011). Understanding fear of pain in chronic pain: interoceptive fear conditioning as a novel approach. *European journal of pain*, 15(9), 889-894.
- Delamater, A. R. (2004). Experimental extinction in Pavlovian conditioning: behavioural and neuroscience perspectives. *Quarterly Journal of Experimental Psychology Section B*, 57(2), 97-132.
- Delgado, M.R., Nearing, K. I., LeDoux, J. E., & Phelps, E. A. (2008). Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron*, 59(5), 829-838.
- Della Sala, S., Laiacona, M., Spinnler, H., & Ubezio, C. (1992). A cancellation test: its reliability in assessing attentional deficits in Alzheimer's disease. *Psychological medicine*, 22(4), 885-901.
- Dere, E., Schulz, D., de Souza Silva, M. A., Jocham, G., Kart, E., & Huston, J. P. (2005). Aged and adult rats compared in acquisition and extinction of escape from the water maze: focus on individual differences. *Behavioral neuroscience*, 119(1), 127-144.
- Doyère, V., Gisquet-Verrier, P., De Marsanich, B., & Ammassari-Teule, M. (2000). Age-related modifications of contextual information processing in rats: role of emotional reactivity, arousal and testing procedure. *Behavioural brain research*, 114(1-2), 153-165.

- Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... & Baas, J. M. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and anxiety*, 32(4), 239-253.
- Dunsmoor, J. E., & Paz, R. (2015). Fear generalization and anxiety: behavioral and neural mechanisms. *Biological psychiatry*, 78(5), 336-343.
- Dunsmoor, J. E., Bandettini, P. A., & Knight, D. C. (2007). Impact of continuous versus intermittent CS-UCS pairing on human brain activation during Pavlovian fear conditioning. *Behavioral neuroscience*, 121(4), 635-642.
- Dunsmoor, J. E., Bandettini, P. A., & Knight, D. C. (2008). Neural correlates of unconditioned response diminution during Pavlovian conditioning. *Neuroimage*, 40(2), 811-817.
- Dunsmoor, J. E., Kroes, M. C., Li, J., Daw, N. D., Simpson, H. B., & Phelps, E. A. (2019). Role of human ventromedial prefrontal cortex in learning and recall of enhanced extinction. *Journal of Neuroscience*, 39(17), 3264-3276.
- Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking extinction. *Neuron*, 88(1), 47-63.
- Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, 18(9), 547-558.
- Ellis, R. J., & Thayer, J. F. (2010). Music and autonomic nervous system (dys) function. *Music Perception: An Interdisciplinary Journal*, 27(4), 317-326.
- Elsey, J. W., Van Ast, V. A., & Kindt, M. (2018). Human memory reconsolidation: A guiding framework and critical review of the evidence. *Psychological bulletin*, 144(8), 797-848.
- Esmorís-Arranz, F. J., Méndez, C., & Spear, N. E. (2008). Contextual fear conditioning differs for infant, adolescent, and adult rats. *Behavioural processes*, 78(3), 340-350.

- Fanselow, M.S., & Poulos, A. M. (2005). The neuroscience of mammalian associative learning. *Annu. Rev. Psychol.*, 56, 207-234.
- Flora, S.R., & Pavlik, W. B. (1990). Conventional and reversed partial reinforcement effects in human operant responding. *Bulletin of the Psychonomic Society*, 28(5), 429-432.
- Fogelson, N. (2015). Neural correlates of local contextual processing across stimulus modalities and patient populations. *Neuroscience & Biobehavioral Reviews*, 52, 207-220.
- Foster, T. C., DeFazio, R. A., & Bizon, J. L. (2012). Characterizing cognitive aging of spatial and contextual memory in animal models. *Frontiers in Aging Neuroscience*, 4(12), 1-18.
- Frohardt, R. J., Guarraci, F. A., & Bouton, M. E. (2000). The effects of neurotoxic hippocampal lesions on two effects of context after fear extinction. *Behavioral neuroscience*, 114(2), 227-240.
- Frysztak, R. J., & Neafsey, E. J. (1991). The effect of medial frontal cortex lesions on respiration, “freezing,” and ultrasonic vocalizations during conditioned emotional responses in rats. *Cerebral Cortex*, 1(5), 418-425.
- Frysztak, R. J., & Neafsey, E. J. (1994). The effect of medial frontal cortex lesions on cardiovascular conditioned emotional responses in the rat. *Brain research*, 643(1-2), 181-193.
- Fullana, M. A., Albajes-Eizagirre, A., Soriano-Mas, C., Vervliet, B., Cardoner, N., Benet, O., ... & Harrison, B. J. (2018). Fear extinction in the human brain: a meta-analysis of fMRI studies in healthy participants. *Neuroscience & Biobehavioral Reviews*, 88, 16-25.
- Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Àvila-Parcet, A., & Radua, J. (2016). Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry*, 21(4), 500-508.

- Garfinkel, S. N., Abelson, J. L., King, A. P., Sripada, R. K., Wang, X., Gaines, L. M., & Liberzon, I. (2014). Impaired contextual modulation of memories in PTSD: an fMRI and psychophysiological study of extinction retention and fear renewal. *Journal of Neuroscience*, 34(40), 13435-13443.
- Gershman, S. J., & Hartley, C. A. (2015). Individual differences in learning predict the return of fear. *Learning & behavior*, 43(3), 243-250.
- Gershman, S. J., Jones, C. E., Norman, K. A., Monfils, M. H., & Niv, Y. (2013). Gradual extinction prevents the return of fear: implications for the discovery of state. *Frontiers in behavioral neuroscience*, 7(164), 1-6.
- Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, 5, 43-50.
- Gilmartin, M. R., Balderston, N. L., & Helmstetter, F. J. (2014). Prefrontal cortical regulation of fear learning. *Trends in neurosciences*, 37(8), 455-464.
- Gogolla, N., Caroni, P., Lüthi, A., & Herry, C. (2009). Perineuronal nets protect fear memories from erasure. *Science*, 325(5945), 1258-1261.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., ... & Rapoport, J. L. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, 101(21), 8174-8179.
- Grady, A. K., Bowen, K. H., Hyde, A. T., Totsch, S. K., & Knight, D. C. (2016). Effect of continuous and partial reinforcement on the acquisition and extinction of human conditioned fear. *Behavioral neuroscience*, 130(1), 36-43.
- Granholm, E. E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, 52: 1 – 6.

- 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.
- Haaker, J., Lonsdorf, T. B., Thanellou, A., & Kalisch, R. (2013). Multimodal assessment of long-term memory recall and reinstatement in a combined cue and context fear conditioning and extinction paradigm in humans. *PloS one*, 8(10), e76179.
- Halladay, L. R., Zelikowsky, M., Blair, H. T., & Fanselow, M. S. (2012). Reinstatement of extinguished fear by an unextinguished conditional stimulus. *Frontiers in behavioral neuroscience*, 6(18), 1-7.
- 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.
- Hartley, A. A., Kieley, J., & Mckenzie, C. R. (1992). Allocation of visual attention in younger and older adults. *Perception & Psychophysics*, 52(2), 175-185.
- Hartley, C. A., & Lee, F. S. (2015). Sensitive periods in affective development: Nonlinear maturation of fear learning. *Neuropsychopharmacology*, 40(1), 50-60.
- Hartley, C. A., & Phelps, E. A. (2010). Changing fear: the neurocircuitry of emotion regulation. *Neuropsychopharmacology*, 35(1), 136-146.
- Hartley, C. A., Fischl, B., & Phelps, E. A. (2011). Brain structure correlates of individual differences in the acquisition and inhibition of conditioned fear. *Cerebral cortex*, 21(9), 1954-1962.
- Haselgrove, M., Aydin, A., & Pearce, J. M. (2004). A partial reinforcement extinction effect despite equal rates of reinforcement during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 30(3), 240-250.



- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In *Psychology of learning and motivation* (Vol. 22, pp. 193-225). Academic Press.
- Heatherton, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in cognitive sciences*, 15(3), 132-139.
- Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature reviews neuroscience*, 5(2), 87-96.
- Hedges, S. B., & Shah, P. (2003). Comparison of mode estimation methods and application in molecular clock analysis. *BMC bioinformatics*, 4(1), 31.
- Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., & Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: Affective priming as an indirect and unobtrusive measure of conditioned stimulus valence. *Behaviour Research and Therapy*, 40(3), 217-234.
- Hiser, J., & Koenigs, M. (2018). The multifaceted role of the ventromedial prefrontal cortex in emotion, decision making, social cognition, and psychopathology. *Biological Psychiatry*, 83(8), 638-647.
- Hitchcock, J., & Davis, M. (1986). Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. *Behavioral neuroscience*, 100(1), 11-22.
- Hobin, J. A., Goosens, K. A., & Maren, S. (2003). Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. *Journal of Neuroscience*, 23(23), 8410-8416.
- Hobin, J. A., Ji, J., & Maren, S. (2006). Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. *Hippocampus*, 16(2), 174-182.

- Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. In *Psychology of Learning and Motivation* (Vol. 28, pp. 69-125). Academic press.
- Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of computer assisted tomography*, 22(2), 324-333.
- Holt, D. J., Lebron-Milad, K., Milad, M. R., Rauch, S. L., Pitman, R. K., Orr, S. P., ... & Goff, D. C. (2009). Extinction memory is impaired in schizophrenia. *Biological psychiatry*, 65(6), 455-463.
- Houston, A. I., & McNamara, J. M. (1999). *Models of adaptive behaviour: an approach based on state*. Cambridge University Press.
- Huang, I. N., Krukar, J. D., & Miles, S. P. (1992). Effects of reinforcement schedules on rats' choice behavior in extinction. *The Journal of general psychology*, 119(2), 201-211.
- Hugdahl, K. (1995). *Psychophysiology: The mind-body perspective*. Harvard University Press.
- Hugues, S., & Garcia, R. (2007). Reorganization of learning-associated prefrontal synaptic plasticity between the recall of recent and remote fear extinction memory. *Learning & memory*, 14(8), 520-524.
- Hunt, W. A., Strauss, H., & Landis, C. (1938). The startle pattern in epileptic patients. *Psychiatric Quarterly*, 12(2), 375-382.
- Iglewicz, B. & Hoaglin, D. (1993) *The ASQC Basic References in Quality Control: Statistical Techniques*. In: Mykytka, E.F., Eds., *How to Detect and Handle Outliers*, ASQC Quality Press, Milwaukee, Vol. 16.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *167*: 748 – 751.

- James, E. L., Lau-Zhu, A., Tickle, H., Horsch, A., & Holmes, E. A. (2016). Playing the computer game Tetris prior to viewing traumatic film material and subsequent intrusive memories: Examining proactive interference. *Journal of behavior therapy and experimental psychiatry*, 53, 25-33.
- Javadi, A. H., & Cheng, P. (2013). Transcranial direct current stimulation (tDCS) enhances reconsolidation of long-term memory. *Brain stimulation*, 6(4), 668-674.
- Javadi, A. H., & Walsh, V. (2012). Transcranial direct current stimulation (tDCS) of the left dorsolateral prefrontal cortex modulates declarative memory. *Brain stimulation*, 5(3), 231-241.
- Ji, J., & Maren, S. (2007). Hippocampal involvement in contextual modulation of fear extinction. *Hippocampus*, 17(9), 749-758.
- Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *Journal of Neuroscience*, 27(33), 8877-8884.
- Jovanovic, T., Kazama, A., Bachevalier, J., & Davis, M. (2012). Impaired safety signal learning may be a biomarker of PTSD. *Neuropharmacology*, 62(2), 695-704.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, 99(1), 122-149.
- Kaczorowski, C. C., Davis, S. J., & Moyer Jr, J. R. (2012). Aging redistributes medial prefrontal neuronal excitability and impedes extinction of trace fear conditioning. *Neurobiology of aging*, 33(8), 1744-1757.
- Kaczorowski, C. C., Sametsky, E., Shah, S., Vassar, R., & Disterhoft, J. F. (2011). Mechanisms underlying basal and learning-related intrinsic excitability in a mouse model of Alzheimer's disease. *Neurobiology of aging*, 32(8), 1452-1465.

- Kalisch, R., Korenfeld, E., Stephan, K. E., Weiskopf, N., Seymour, B., & Dolan, R. J. (2006). Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. *Journal of Neuroscience*, 26(37), 9503-9511.
- Karsen, E. F., Watts, B. V., & Holtzheimer, P. E. (2014). Review of the effectiveness of transcranial magnetic stimulation for post-traumatic stress disorder. *Brain Stimulation*, 7(2), 151-157.
- Kattoor, J., Gizewski, E. R., Kotsis, V., Benson, S., Gramsch, C., Theysohn, N., ... & Elsenbruch, S. (2013). Fear conditioning in an abdominal pain model: neural responses during associative learning and extinction in healthy subjects. *PloS one*, 8(2), e51149.
- Kattoor, J., Thürling, M., Gizewski, E. R., Forsting, M., Timmann, D., & Elsenbruch, S. (2014). Cerebellar contributions to different phases of visceral aversive extinction learning. *The Cerebellum*, 13(1), 1-8.
- Kennedy, C. H., Zillmer, E. A., & Culbertson, W. C. (2000). Assessing executive function in the MR population: an investigation of the TOL-Dx and MWCST. *Archives of Clinical Neuropsychology*, 15(8), 725-725.
- Khemka, S., Tzovara, A., Gerster, S., Quednow, B. B., & Bach, D. R. (2017). Modeling startle eyeblink electromyogram to assess fear learning. *Psychophysiology*, 54(2), 204-214.
- Kim, J. H., & Richardson, R. (2008). The effect of temporary amygdala inactivation on extinction and reextinction of fear in the developing rat: unlearning as a potential mechanism for extinction early in development. *Journal of Neuroscience*, 28(6), 1282-1290.
- Kim, J. H., & Richardson, R. (2010). New findings on extinction of conditioned fear early in development: theoretical and clinical implications. *Biological psychiatry*, 67(4), 297-303.

- Kim, J. J., & Jung, M. W. (2006). Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neuroscience & Biobehavioral Reviews*, 30(2), 188-202.
- Kim, M. J., & Whalen, P. J. (2009). The structural integrity of an amygdala–prefrontal pathway predicts trait anxiety. *Journal of Neuroscience*, 29(37), 11614-11618.
- Kindt, M., Soeter, M., & Sevenster, D. (2014). Disrupting reconsolidation of fear memory in humans by a noradrenergic  $\beta$ -blocker. *JoVE (Journal of Visualized Experiments)*, (94), e52151.
- Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature neuroscience*, 12(3), 256-258.
- Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., ... & Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science*, 356(6333), 73-78.
- Knight, D. C., Cheng, D. T., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (2004). Neural substrates mediating human delay and trace fear conditioning. *Journal of Neuroscience*, 24(1), 218-228.
- 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, & Behavioral Neuroscience*, 4(3), 317-325.
- Knight, D. C., Waters, N. S., King, M. K., & Bandettini, P. A. (2010). Learning-related diminution of unconditioned SCR and fMRI signal responses. *Neuroimage*, 49(1), 843-848.
- Knuttinen, M. G., Power, J. M., Preston, A. R., & Disterhoft, J. F. (2001). Awareness in classical differential eyeblink conditioning in young and aging humans. *Behavioral neuroscience*, 115(4), 747-757.

- Koizumi, K., & Kollai, M. (1992). Multiple modes of operation of cardiac autonomic control: development of the ideas from Cannon and Brooks to the present. *Journal of the autonomic nervous system*, 41(1-2), 19-29.
- Korn, C. W., Staib, M., Tzovara, A., Castegnetti, G., & Bach, D. R. (2017). A pupil size response model to assess fear learning. *Psychophysiology*, 54(3), 330-343.
- Krettek, J. E., & Price, J. L. (1977). Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *Journal of Comparative Neurology*, 172(4), 687-722.
- Kryptos, A. M., Jahfari, S., van Ast, V. A., Kindt, M., & Forstmann, B. U. (2011). Individual differences in heart rate variability predict the degree of slowing during response inhibition and initiation in the presence of emotional stimuli. *Frontiers in psychology*, 2(278), 1-8.
- LaBar, K. S., & Phelps, E. A. (2005). Reinstatement of conditioned fear in humans is context dependent and impaired in amnesia. *Behavioral neuroscience*, 119(3), 677-686.
- LaBar, K. S., Cook, C. A., Torpey, D. C., & Welsh-Bohmer, K. A. (2004). Impact of healthy aging on awareness and fear conditioning. *Behavioral Neuroscience*, 118(5), 905-915.
- LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron*, 20(5), 937-945.
- LaBar, K. S., LeDoux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *Journal of neuroscience*, 15(10), 6846-6855.
- Lane, R. D., McRae, K., Reiman, E. M., Chen, K., Ahern, G. L., & Thayer, J. F. (2009). Neural correlates of heart rate variability during emotion. *Neuroimage*, 44(1), 213-222.

- 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). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Laxmi, T. R., Stork, O., & Pape, H. C. (2003). Generalisation of conditioned fear and its behavioural expression in mice. *Behavioural brain research*, 145(1-2), 89-98.
- Le Roux, H., Gatz, M., & Wetherell, J. L. (2005). Age at onset of generalized anxiety disorder in older adults. *The American journal of geriatric psychiatry*, 13(1), 23-30.
- Lebrón, K., Milad, M. R., & Quirk, G. J. (2004). Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. *Learning & memory*, 11(5), 544-548.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience*, 23(1), 155-184.
- LeDoux, J. E., Iwata, J., Cicchetti, P. R. D. J., & Reis, D. J. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *Journal of Neuroscience*, 8(7), 2517-2529.
- LeDoux, J. E., Sakaguchi, A.K., & Reis, D. J. (1984). Subcortical efferent projections of the medial geniculate nucleus mediate emotional responses conditioned to acoustic stimuli. *Journal of Neuroscience*, 4(3), 683-698.
- Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neuroscience & biobehavioral reviews*, 30(6), 718-729.
- Lenze, E. J., & Wetherell, J. L. (2011). A lifespan view of anxiety disorders. *Dialogues in clinical neuroscience*, 13(4), 381-399.

- Leonard, D. W. (1975). Partial reinforcement effects in classical aversive conditioning in rabbits and human beings. *Journal of comparative and physiological psychology*, 88(2), 596-608.
- Levy, M. N. (1971). Brief reviews: sympathetic-parasympathetic interactions in the heart. *Circulation research*, 29(5), 437-445.
- Lewis, D. A., Hashimoto, T., & Volk, D. W. (2005). Cortical inhibitory neurons and schizophrenia. *Nature Reviews Neuroscience*, 6(4), 312-324.
- Li, C. S. R., & Sinha, R. (2008). Inhibitory control and emotional stress regulation: Neuroimaging evidence for frontal–limbic dysfunction in psycho-stimulant addiction. *Neuroscience & Biobehavioral Reviews*, 32(3), 581-597.
- Linnman, C., Rougemont-Bücking, A., Beucke, J. C., Zeffiro, T. A., & Milad, M. R. (2011). Unconditioned responses and functional fear networks in human classical conditioning. *Behavioural brain research*, 221(1), 237-245.
- Lisanby, S. H., Gutman, D., Luber, B., Schroeder, C., & Sackeim, H. A. (2001). Sham TMS: intracerebral measurement of the induced electrical field and the induction of motor-evoked potentials. *Biological psychiatry*, 49(5), 460-463.
- Lissek, S., Baas, J. M., Pine, D. S., Orme, K., Dvir, S., Nugent, M., ... & Grillon, C. (2005a). Airpuff startle probes: an efficacious and less aversive alternative to white-noise. *Biological Psychology*, 68(3), 283-297.
- Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., & Grillon, C. (2008). Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behaviour research and therapy*, 46(5), 678-687.



- Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005b). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour research and therapy*, 43(11), 1391-1424.
- Liu, J., Wei, W., Kuang, H., Zhao, F., & Tsien, J. Z. (2013). Changes in heart rate variability are associated with expression of short-term and long-term contextual and cued fear memories. *PLoS One*, 8(5), e63590.
- Lonsdorf, T. B., Haaker, J., & Kalisch, R. (2014a). Long-term expression of human contextual fear and extinction memories involves amygdala, hippocampus and ventromedial prefrontal cortex: a reinstatement study in two independent samples. *Social cognitive and affective neuroscience*, 9(12), 1973-1983.
- Lonsdorf, T. B., Haaker, J., Fadai, T., & Kalisch, R. (2014b). No evidence for enhanced extinction memory consolidation through noradrenergic reuptake inhibition—delayed memory test and reinstatement in human fMRI. *Psychopharmacology*, 231(9), 1949-1962.
- Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., ... & Drexler, S. M. (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 & Biobehavioral Reviews*, 77, 247-285.
- Lovibond, P. F. (2004). Cognitive processes in extinction. *Learning & Memory*, 11(5), 495-500.
- Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, 28(1), 3-26.
- 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.

- Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin*, 79, 398–407.
- Luque-Casado, A., Perakakis, P., Ciria, L. F., & Sanabria, D. (2016). Transient autonomic responses during sustained attention in high and low fit young adults. *Scientific reports*, 6, 27556.
- Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, 8(5), 656-672.
- Lykken, D.T. (1972). Range correction applied to heart rate and to GSR data. *Psychophysiology*, 9(3), 373-379.
- Lyonfields, J. D., Borkovec, T. D., & Thayer, J. F. (1995). Vagal tone in generalized anxiety disorder and the effects of aversive imagery and worrisome thinking. *Behavior Therapy*, 26(3), 457-466.
- Mahanty, N. K., & Sah, P. (1998). Calcium-permeable AMPA receptors mediate long-term potentiation in interneurons in the amygdala. *Nature*, 394(6694), 683-687.
- Mainardi, L. T. (2008). On the quantification of heart rate variability spectral parameters using time–frequency and time-varying methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1887), 255-275.
- Malliani, A., Pagani, M., Lombardi, F., & Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84(2), 482-492.
- Malmivuo, P., Malmivuo, J., & Plonsey, R. (1995). *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA.
- Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual review of neuroscience*, 24(1), 897-931.
- Maren, S. (2005). Building and burying fear memories in the brain. *The Neuroscientist*, 11(1), 89-99.

- Maren, S., & Holmes, A. (2016). Stress and fear extinction. *Neuropsychopharmacology*, 41(1), 58-79.
- Maren, S., & Quirk, G. J. (2004). Neuronal signalling of fear memory. *Nature Reviews Neuroscience*, 5(11), 844-852.
- 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.
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114(2), 159-188.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation* (Vol. 24, pp. 109-165). Academic Press.
- McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.*, 27, 1-28.
- Mechias, M. L., Etkin, A., & Kalisch, R. (2010). A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *Neuroimage*, 49(2), 1760-1768.
- Melzig, C. A., Weike, A. I., Hamm, A. O., & Thayer, J. F. (2009). Individual differences in fear-potentiated startle as a function of resting heart rate variability: implications for panic disorder. *International Journal of Psychophysiology*, 71(2), 109-117.
- Merlo, E., Milton, A. L., Goozée, Z. Y., Theobald, D. E., & Everitt, B. J. (2014). Reconsolidation and extinction are dissociable and mutually exclusive processes: behavioral and molecular evidence. *Journal of Neuroscience*, 34(7), 2422-2431.

- Meulders, A., & Vlaeyen, J. W. (2013). The acquisition and generalization of cued and contextual pain-related fear: an experimental study using a voluntary movement paradigm. *Pain*, 154(2), 272-282.
- Meulders, A., Jans, A., & Vlaeyen, J. W. (2015). Differences in pain-related fear acquisition and generalization: an experimental study comparing patients with fibromyalgia and healthy controls. *Pain*, 156(1), 108-122.
- Meulders, A., Vandebroek, N., Vervliet, B., & Vlaeyen, J. W. (2013). Generalization gradients in cued and contextual pain-related fear: an experimental study in healthy participants. *Frontiers in human neuroscience*, 7(345), 1-12.
- Meulders, A., Vansteenwegen, D., & Vlaeyen, J. W. (2011). The acquisition of fear of movement-related pain and associative learning: a novel pain-relevant human fear conditioning paradigm. *Pain*, 152(11), 2460-2469.
- Meulders, A., Vervliet, B., Fonteyne, R., Baeyens, F., Hermans, D., & Vansteenwegen, D. (2012). Preexposure to (un) predictable shock modulates discriminative fear learning between cue and context: An investigation of the interaction between fear and anxiety. *International Journal of Psychophysiology*, 84(2), 180-187.
- Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*, 420(6911), 70-74.
- Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Annual review of psychology*, 63, 129-151.
- Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., ... & Rauch, S. L. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. *Biological psychiatry*, 66(12), 1075-1082.

- Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proceedings of the National Academy of Sciences*, 102(30), 10706-10711.
- Milad, M. R., Vidal-Gonzalez, I., & Quirk, G. J. (2004). Electrical stimulation of medial prefrontal cortex reduces conditioned fear in a temporally specific manner. *Behavioral neuroscience*, 118(2), 389-395.
- Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological psychiatry*, 62(5), 446-454.
- Mineka, S., Mystkowski, J. L., Hladek, D., & Rodriguez, B. I. (1999). The effects of changing contexts on return of fear following exposure therapy for spider fear. *Journal of consulting and clinical psychology*, 67(4), 599-604.
- Misanin, J. R., Miller, R. R., & Lewis, D. J. (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science*, 160(3827), 554-555.
- Monfils, M. H., Cowansage, K. K., Klann, E., & LeDoux, J. E. (2009). Extinction-reconsolidation boundaries: key to persistent attenuation of fear memories. *science*, 324(5929), 951-955.
- Morgan, M. A., & LeDoux, J. E. (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behavioral neuroscience*, 109(4), 681-688.
- Morgan, M. A., Romanski, L. M., & LeDoux, J. E. (1993). Extinction of emotional learning: contribution of medial prefrontal cortex. *Neuroscience letters*, 163(1), 109-113.
- Morris, J. S., & Dolan, R. J. (2004). Dissociable amygdala and orbitofrontal responses during reversal fear conditioning. *Neuroimage*, 22(1), 372-380.

- Mungee, A., Kazzer, P., Feeser, M., Nitsche, M. A., Schiller, D., & Bajbouj, M. (2014). Transcranial direct current stimulation of the prefrontal cortex: a means to modulate fear memories. *Neuroreport*, 25(7), 480-484.
- Myers, K. M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron*, 36(4), 567-584.
- Mystkowski, J. L., Craske, M. G., & Echiverri, A. M. (2002). Treatment context and return of fear in spider phobia. *Behavior Therapy*, 33(3), 399-416.
- Nabil, D., & Reguig, F. B. (2015). Ectopic beats detection and correction methods: A review. *Biomedical Signal Processing and Control*, 18, 228-244.
- Nader, K., & Hardt, O. (2009). A single standard for memory: the case for reconsolidation. *Nature Reviews Neuroscience*, 10(3), 224-234.
- Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406(6797), 722-726.
- Norrholm, S. D., Jovanovic, T., Vervliet, B., Myers, K. M., Davis, M., Rothbaum, B. O., & Duncan, E. J. (2006). Conditioned fear extinction and reinstatement in a human fear-potentiated startle paradigm. *Learning & Memory*, 13(6), 681-685.
- Novelli, G., Papagno, C., Capitani, E., Laiacona, M., Cappa, S. F., & Vallar, G. (1986). Three clinical tests for the assessment of verbal long-term memory function: norms from 320 normal subjects. *Archivio di Psicologia, Neurologia e Psichiatria*, 47(2), 278-296.
- Obrist, P. A., Wood, D. M., & Perez-Reyes, M. (1965). Heart rate during conditioning in humans: Effects of UCS intensity, vagal blockade, and adrenergic block of vasomotor activity. *Journal of Experimental Psychology*, 70(1), 32-42.
- Ohta, A., Akiguchi, I., Seriu, N., Ohnishi, K., Yagi, H., Higuchi, K., & Hosokawa, M. (2001). Deterioration in learning and memory of fear conditioning in response to context in aged SAMP8 mice. *Neurobiology of aging*, 22(3), 479-484.

- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Oler, J. A., & Markus, E. J. (1998). Age-related deficits on the radial maze and in fear conditioning: hippocampal processing and consolidation. *Hippocampus*, 8(4), 402-415.
- Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *Journal of cognitive neuroscience*, 19(12), 1994-2004.
- Orsini, A., Grossi, D., Capitani, E., Laiacona, M., Papagno, C., & Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *The Italian Journal of Neurological Sciences*, 8(6), 537-548.
- Otto, M. W., Church, T. S., Craft, L. L., Greer, T. L., Smits, J. A., & Trivedi, M. H. (2007). Exercise for mood and anxiety disorders. *Primary care companion to the Journal of clinical psychiatry*, 9(4), 287-294.
- Panayiotou, G., van Oyen Witvliet, C., Robinson, J. D., & Vrana, S. R. (2011). A startling absence of emotion effects: Active attention to the startle probe as a motor task cue appears to eliminate modulation of the startle reflex by valence and arousal. *Biological psychology*, 87(2), 226-233.
- Pappens, M., Schroijen, M., Sütterlin, S., Smets, E., Van den Bergh, O., Thayer, J. F., & Van Diest, I. (2014). Resting heart rate variability predicts safety learning and fear extinction in an interoceptive fear conditioning paradigm. *PloS one*, 9(9), e105054.
- Pappens, M., Van den Bergh, O., Vansteenwegen, D., Ceunen, E., De Peuter, S., & Van Diest, I. (2013). Learning to fear obstructed breathing: comparing interoceptive and exteroceptive cues. *Biological psychology*, 92(1), 36-42.
- Park, D. C. & Schwarz, N. (2000) *Cognitive aging: a primer*. Psychology Press.

- Park, G., & Thayer, J. F. (2014). From the heart to the mind: cardiac vagal tone modulates top-down and bottom-up visual perception and attention to emotional stimuli. *Frontiers in psychology*, 5(278),1-8.
- Park, G., Van Bavel, J. J., Vasey, M. W., & Thayer, J. F. (2013). Cardiac vagal tone predicts attentional engagement to and disengagement from fearful faces. *Emotion*, 13(4), 645-656.
- Pauli, W. M., Gentile, G., Collette, S., Tyszka, J. M., & O’Doherty, J. P. (2019). Evidence for model-based encoding of Pavlovian contingencies in the human brain. *Nature communications*, 10(1), 1099.
- Paulus, P. C., Castagnetti, G., & Bach, D. R. (2016). Modeling event-related heart period responses. *Psychophysiology*, 53(6), 837-846.
- Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Translated and edited by Anrep, GV (Oxford University Press, London, 1927).
- Pavlov, I.P. (1903/1928). *Experimental psychology and psychopathology in animals* Herald of the Military Medical Academy, vol. 7, pp. 109-121. Reprinted in Gantt, W.H. Edition, 1928. *Lectures on Conditioned Reflexes*, vol. 1. Lawrence & Wishart, London, pp. 47–60.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, 87(6), 532-552.
- Pennington, Z. T., & Fanselow, M. S. (2018). Indirect Targeting of Subsuperficial Brain Structures With Transcranial Magnetic Stimulation Reveals a Promising Way Forward in the Treatment of Fear. *Biological psychiatry*, 84(2), 80-81.



- Phelps, E. A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current opinion in neurobiology*, 14(2), 198-202.
- Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron*, 43(6), 897-905.
- Phillips, R. G., & LeDoux, J. E. (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral neuroscience*, 106(2), 274-285.
- Ploghaus, A., Becerra, L., Borras, C., & Borsook, D. (2003). Neural circuitry underlying pain modulation: expectation, hypnosis, placebo. *Trends in cognitive sciences*, 7(5), 197-200.
- Powell, P. S., Strunk, J., James, T., Polyn, S. M., & Duarte, A. (2018). Decoding selective attention to context memory: An aging study. *Neuroimage*, 181, 95-107.
- Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical psychology review*, 30(6), 635-641.
- Prenoveau, J. M., Craske, M. G., Liao, B., & Ornitz, E. M. (2013). Human fear conditioning and extinction: Timing is everything... or is it?. *Biological psychology*, 92(1), 59-68.
- Prévost, C., McNamee, D., Jessup, R. K., Bossaerts, P., & O'Doherty, J. P. (2013). Evidence for model-based computations in the human amygdala during Pavlovian conditioning. *PLoS computational biology*, 9(2), e1002918.
- Pumprla, J., Howorka, K., Groves, D., Chester, M., & Nolan, J. (2002). Functional assessment of heart rate variability: physiological basis and practical applications. *International journal of cardiology*, 84(1), 1-14.

- Quirk, G. J. (2002). Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. *Learning & memory*, 9(6), 402-407.
- Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, 33(1), 56-72.
- Quirk, G. J., Likhtik, E., Pelletier, J. G., & Paré, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience*, 23(25), 8800-8807.
- Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *Journal of Neuroscience*, 20(16), 6225-6231.
- Radulovic, J., & Tronson, N. C. (2010). Molecular specificity of multiple hippocampal processes governing fear extinction. *Reviews in the neurosciences*, 21(1), 1-18.
- Raji, T., Nummenmaa, A., Marin, M. F., Porter, D., Furtak, S., Setsompop, K., & Milad, M. R. (2018). Prefrontal cortex stimulation enhances fear extinction memory in humans. *Biological psychiatry*, 84(2), 129-137.
- Rauch, S. L., Shin, L. M., Segal, E., Pitman, R. K., Carson, M. A., McMullin, K., ... & Makris, N. (2003). Selectively reduced regional cortical volumes in post-traumatic stress disorder. *Neuroreport*, 14(7), 913-916.
- Rauhut, A. S., Thomas, B. L., & Ayres, J. J. (2001). Treatments that weaken Pavlovian conditioned fear and thwart its renewal in rats: Implications for treating human phobias. *Journal of Experimental Psychology: Animal Behavior Processes*, 27(2), 99-114.
- Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., ... & Giedd, J. N. (2014). Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proceedings of the National Academy of Sciences*, 111(4), 1592-1597.

- Reinhard, G., Lachnit, H., & König, S. (2006). Tracking stimulus processing in Pavlovian pupillary conditioning. *Psychophysiology*, 43(1), 73-83.
- Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., & LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nature Neuroscience*, 4(7), 724-731.
- Rescorla, R. A. (1979). Conditioned inhibition and extinction. *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski*, 83-110.
- Rescorla, R. A. (2001). Experimental extinction. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 119-154). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Rescorla, R.A. (2001). Retraining of extinguished Pavlovian stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 27(2), 115-124.
- Rescorla, R.A., & Heth, C.D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 1(1), 88-96.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64-99.
- Rhodes, S. E., & Killcross, S. (2004). Lesions of rat infralimbic cortex enhance recovery and reinstatement of an appetitive Pavlovian response. *Learning & Memory*, 11(5), 611-616.
- Riedel, G., Harrington, N. R., Hall, G., & Macphail, E. M. (1997). Nucleus accumbens lesions impair context, but not cue, conditioning in rats. *Neuroreport*, 8(11), 2477-2481.
- 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.

- Roesch, M., & Schoenbaum, G. (2006). From associations to expectancies: orbitofrontal cortex as gateway between the limbic system and representational memory. *The orbitofrontal cortex*, 199-235.
- Romanski, L. M., & LeDoux, J. E. (1993). Information cascade from primary auditory cortex to the amygdala: corticocortical and corticoamygdaloid projections of temporal cortex in the rat. *Cerebral Cortex*, 3(6), 515-532.
- Rosas, J. M., & Bouton, M. E. (1997). Renewal of a conditioned taste aversion upon return to the conditioning context after extinction in another one. *Learning and motivation*, 28(2), 216-229.
- Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological review*, 105(2), 325-350.
- Rossi, S., Cappa, S. F., Babiloni, C., Pasqualetti, P., Miniussi, C., Carducci, F., ... & Rossini, P. M. (2001). Prefrontal cortex in long-term memory: an “interference” approach using magnetic stimulation. *Nature neuroscience*, 4(9), 948-952.
- Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical neurophysiology*, 120(12), 2008-2039.
- Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., ... & De Noordhout, A. M. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography and clinical neurophysiology*, 91(2), 79-92.
- Rossini, P. M., Burke, D., Chen, R., Cohen, L. G., Daskalakis, Z., Di Iorio, R., ... & Hallett, M. (2015). Non-invasive electrical and magnetic stimulation of the brain, spinal cord,

- roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an IFCN Committee. *Clinical Neurophysiology*, 126(6), 1071-1107.
- Ruiz-Padial, E., & Thayer, J. F. (2014). Resting heart rate variability and the startle reflex to briefly presented affective pictures. *International Journal of Psychophysiology*, 94(3), 329-335.
  - Ruiz-Padial, E., Sollers III, J. J., Vila, J., & Thayer, J. F. (2003). The rhythm of the heart in the blink of an eye: Emotion-modulated startle magnitude covaries with heart rate variability. *Psychophysiology*, 40(2), 306-313.
  - Salami, A., Pudas, S., & Nyberg, L. (2014). Elevated hippocampal resting-state connectivity underlies deficient neurocognitive function in aging. *Proceedings of the National Academy of Sciences*, 111(49), 17654-17659.
  - Sanders, M. J. (2011). Context processing in aging: older mice are impaired in renewal of extinguished fear. *Experimental aging research*, 37(5), 572-594.
  - Sandrini, M., Brambilla, M., Manenti, R., Rosini, S., Cohen, L. G., & Cotelli, M. (2014). Noninvasive stimulation of prefrontal cortex strengthens existing episodic memories and reduces forgetting in the elderly. *Frontiers in aging neuroscience*, 6(289), 1-9.
  - Sandrini, M., Censor, N., Mishoe, J., & Cohen, L. G. (2013). Causal role of prefrontal cortex in strengthening of episodic memories through reconsolidation. *Current Biology*, 23(21), 2181-2184.
  - Sandrini, M., Cohen, L. G., & Censor, N. (2015). Modulating reconsolidation: a link to causal systems-level dynamics of human memories. *Trends in cognitive sciences*, 19(8), 475-482.

- Sandrini, M., Umiltà, C., & Rusconi, E. (2011). The use of transcranial magnetic stimulation in cognitive neuroscience: a new synthesis of methodological issues. *Neuroscience & Biobehavioral Reviews*, 35(3), 516-536.
- Santini, E., Ge, H., Ren, K., de Ortiz, S. P., & Quirk, G. J. (2004). Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. *Journal of Neuroscience*, 24(25), 5704-5710.
- Schiller, D., & Delgado, M. R. (2010). Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in cognitive sciences*, 14(6), 268-276.
- Schiller, D., Cain, C. K., Curley, N. G., Schwartz, J. S., Stern, S. A., LeDoux, J. E., & Phelps, E. A. (2008). Evidence for recovery of fear following immediate extinction in rats and humans. *Learning & Memory*, 15(6), 394-402.
- Schiller, D., Levy, I., Niv, Y., LeDoux, J. E., & Phelps, E. A. (2008). From fear to safety and back: reversal of fear in the human brain. *Journal of Neuroscience*, 28(45), 11517-11525.
- Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, 463(7277), 49-53.
- Schmajuk, N. A., & Holland, P. C. (1998). Occasion setting: Associative learning and cognition in animals (pp. xxi-440). American Psychological Association.
- Schneider, A. M., & Sherman, W. (1968). Amnesia: a function of the temporal relation of footshock to electroconvulsive shock. *Science*, 159(3811), 219-221.
- Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, 10(12), 885-892.

- Schoenbaum, G., Takahashi, Y., Liu, T. L., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value?. *Annals of the New York Academy of Sciences*, 1239, 87-99.
- Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of human memory: Brain mechanisms and clinical relevance. *Biological psychiatry*, 76(4), 274-280.
- Sehlmeier, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., & Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PloS one*, 4(6), e5865.
- Semple, W. E., Goyer, P. F., McCormick, R., Compton-Toth, B., Morris, E., Donovan, B., ... & Leisure, G. (1996). Attention and regional cerebral blood flow in posttraumatic stress disorder patients with substance abuse histories. *Psychiatry Research: Neuroimaging*, 67(1), 17-28.
- Senn, V., Wolff, S. B., Herry, C., Grenier, F., Ehrlich, I., Gründemann, J., ... & Lüthi, A. (2014). Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron*, 81(2), 428-437.
- Sevenster, D., Beckers, T., & Kindt, M. (2013). Prediction error governs pharmacologically induced amnesia for learned fear. *Science*, 339(6121), 830-833.
- Sevenster, D., Hamm, A., Beckers, T., & Kindt, M. (2015). Heart rate pattern and resting heart rate variability mediate individual differences in contextual anxiety and conditioned responses. *International Journal of Psychophysiology*, 98(3), 567-576.
- Shin, L. M., McNally, R. J., Kosslyn, S. M., Thompson, W. L., Rauch, S. L., Alpert, N. M., ... & Pitman, R. K. (1999). Regional cerebral blood flow during script-driven imagery in childhood sexual abuse-related PTSD: a PET investigation. *American Journal of Psychiatry*, 156(4), 575-584.

- Siddle, D. A., Bond, N. W., & Packer, J. S. (1988). Comparator theories of habituation: A comment on Mackintosh's analysis.
- Sierra-Mercado Jr, D., Corcoran, K. A., Lebrón-Milad, K., & Quirk, G. J. (2006). Inactivation of the ventromedial prefrontal cortex reduces expression of conditioned fear and impairs subsequent recall of extinction. *European Journal of Neuroscience*, 24(6), 1751-1758.
- Sladky, R., Höflich, A., Küblböck, M., Kraus, C., Baldinger, P., Moser, E., ... & Windischberger, C. (2013). Disrupted effective connectivity between the amygdala and orbitofrontal cortex in social anxiety disorder during emotion discrimination revealed by dynamic causal modeling for fMRI. *Cerebral Cortex*, 25(4), 895-903.
- Soeter, M., & Kindt, M. (2010). Dissociating response systems: erasing fear from memory. *Neurobiology of learning and memory*, 94(1), 30-41.
- Soeter, M., & Kindt, M. (2011). Disrupting reconsolidation: pharmacological and behavioral manipulations. *Learning & Memory*, 18(6), 357-366.
- Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychology and aging*, 10(4), 527-539.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory* (Palo Alto, CA, Consulting Psychologists Press). Inc.
- Spinnler, H. (1987). Standardizzazione e taratura italiana di test neuropsicologici. *The Italian Journal of Neurological Sciences*, 6, 21-120.
- Squire, L. R., & Davis, H. P. (1981). The pharmacology of memory: A neurobiological perspective. *Annual Review of Pharmacology and Toxicology*, 21(1), 323-356.



- Starosta, S., Uengoer, M., Bartetzko, I., Lucke, S., Güntürkün, O., & Stüttgen, M. C. (2016). Context specificity of both acquisition and extinction of a Pavlovian conditioned response. *Learning & Memory*, 23(11), 639-643.
- Steckle, L. C. (1933). A trace conditioning of the galvanic reflex. *The Journal of General Psychology*, 9(2), 475-480.
- Stoehr, J. D., & Wenk, G. L. (1995). Effects of age and lesions of the nucleus basalis on contextual fear conditioning. *Psychobiology*, 23(3), 173-177.
- Storey, J. D. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics*, 31(6), 2013-2035.
- Sun, N., & Cassell, M. D. (1993). Intrinsic GABAergic neurons in the rat central extended amygdala. *Journal of Comparative Neurology*, 330(3), 381-404.
- Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning*, 3(1), 9-44.
- Switzer, S. C. (1934). Anticipatory and inhibitory characteristics of delayed conditioned reactions. *Journal of Experimental Psychology*, 17(5), 603.
- Task Force of the European Society of Cardiology, the North American Society of Pacing, and Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*, 93: 1043 – 1065.
- Tessa, C., Toschi, N., Orsolini, S., Valenza, G., Lucetti, C., Barbieri, R., & Diciotti, S. (2019). Central modulation of parasympathetic outflow is impaired in de novo Parkinson's disease patients. *PloS one*, 14(1), e0210324.
- Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of affective disorders*, 61(3), 201-216.
- Thayer, J. F., & Lane, R. D. (2002). Perseverative thinking and health: Neurovisceral concomitants. *Psychology and Health*, 17(5), 685-695.

- Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, 33(2), 81-88.
- Thayer, J. F., & Siegle, G. J. (2002). Neurovisceral integration in cardiac and emotional regulation. *IEEE Engineering in Medicine and Biology Magazine*, 21(4), 24-29.
- Thayer, J. F., Åhs, F., Fredrikson, M., Sollers III, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36(2), 747-756.
- Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine*, 37(2), 141-153.
- Todd, T. P., Vurbic, D., & Bouton, M. E. (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiology of learning and memory*, 108, 52-64.
- Trask, S., Shipman, M. L., Green, J. T., & Bouton, M. E. (2017). Inactivation of the prelimbic cortex attenuates context-dependent operant responding. *Journal of Neuroscience*, 37(9), 2317-2324.
- Tronson, N. C., & Taylor, J. R. (2007). Molecular mechanisms of memory reconsolidation. *Nature Reviews Neuroscience*, 8(4), 262-275.
- Tzovara, A., Korn, C. W., & Bach, D. R. (2018). Human Pavlovian fear conditioning conforms to probabilistic learning. *PLoS computational biology*, 14(8), e1006243.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*, 37(3), 498-505.

- Van de Vijver, I., Cohen, M. X., & Ridderinkhof, K. R. (2014). Aging affects medial but not anterior frontal learning-related theta oscillations. *Neurobiology of Aging*, 35(3), 692-704.
- Van der Linden, M., Hupet, M., Feyereisen, P., Schelstraete, M. A., Bestgen, Y., Bruyer, R., ... & Seron, X. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology, and Cognition*, 6(1), 32-55.
- Van Stegeren, A. H., Goekoop, R., Everaerd, W., Scheltens, P., Barkhof, F., Kuijjer, J. P., & Rombouts, S. A. (2005). Noradrenaline mediates amygdala activation in men and women during encoding of emotional material. *Neuroimage*, 24(3), 898-909.
- Vansteenwegen, D., Hermans, D., Vervliet, B., Francken, G., Beckers, T., Baeyens, F., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a return to the original acquisition context. *Behaviour research and therapy*, 43(3), 323-336.
- Van't Wout, M., Mariano, T. Y., Garnaat, S. L., Reddy, M. K., Rasmussen, S. A., & Greenberg, B. D. (2016). Can transcranial direct current stimulation augment extinction of conditioned fear?. *Brain stimulation*, 9(4), 529-536.
- Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2013). Extinction, generalization, and return of fear: a critical review of renewal research in humans. *Biological psychology*, 92(1), 51-58.
- Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: state of the art. *Annual review of clinical psychology*, 9, 215-248.
- Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010a). Fear generalization in humans: impact of verbal instructions. *Behaviour research and therapy*, 48(1), 38-43.

- Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010b). Fear generalization in humans: impact of prior non-fearful experiences. *Behaviour research and therapy*, 48(11), 1078-1084.
- Vouimba, R. M., Garcia, R., Baudry, M., & Thompson, R. F. (2000). Potentiation of conditioned freezing following dorsomedial prefrontal cortex lesions does not interfere with fear reduction in mice. *Behavioral neuroscience*, 114(4), 720-724.
- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of experimental psychology*, 3(1), 1-14.
- Weible, A. P., McEchron, M. D., & Disterhoft, J. F. (2000). Cortical involvement in acquisition and extinction of trace eyeblink conditioning. *Behavioral neuroscience*, 114(6), 1058-1067.
- Weike, A. I., Hamm, A. O., Schupp, H. T., Runge, U., Schroeder, H. W., & Kessler, C. (2005). Fear conditioning following unilateral temporal lobectomy: dissociation of conditioned startle potentiation and autonomic learning. *Journal of Neuroscience*, 25(48), 11117-11124.
- Wendt, J., Neubert, J., Koenig, J., Thayer, J. F., & Hamm, A. O. (2015). Resting heart rate variability is associated with inhibition of conditioned fear. *Psychophysiology*, 52(9), 1161-1166.
- Wilson, A., Brooks, D. C., & Bouton, M. E. (1995). The role of the rat hippocampal system in several effects of context in extinction. *Behavioral neuroscience*, 109(5), 828-836.
- Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2), 267-279.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis (Vol. 59)*. Sage.

- Wolpe, J. (1981). The dichotomy between classical conditioned and cognitively learned anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 12(1), 35-42.
- Wood, N. E., Rosasco, M. L., Suris, A. M., Spring, J. D., Marin, M. F., Lasko, N. B., ... & Pitman, R. K. (2015). Pharmacological blockade of memory reconsolidation in posttraumatic stress disorder: three negative psychophysiological studies. *Psychiatry research*, 225(1-2), 31-39.
- Zaman, J., Weltens, N., Ly, H. G., Struyf, D., Vlaeyen, J. W., Van den Bergh, O., ... & Van Diest, I. (2016). Influence of interoceptive fear learning on visceral perception. *Psychosomatic medicine*, 78(2), 248-258.
- Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica*, 67(6), 361-370.



## ACKNOWLEDGEMENTS

Undertaking this PhD course has been a truly life-changing experience and it would not have been possible without the support and guidance that I received from many people that have been crucial during this long path.

I would like to first express my gratitude to my supervisor, prof. Giuseppe di Pellegrino for all the support and encouragement that he was able to give me. His wisdom and guidance traced my path since I was an undergraduate student attending my first Neuroscience course. His passion ignites my deepest ambition and interest in this field. After almost 10 years, without his supervision and constant feedback, this PhD degree would not have been achievable.

I would like to truly thank Dr. Sara Borgomaneri for her theoretical, practical and human advice that she gave to me during these years and for all the support she provided, leading me here today. She encouraged me during my worst moments, pushing me to pursue everything in which I believe. Also, she was always present when I needed her, she has been the best colleague, co-desk worker, and friend that this PhD could have offered me. I am really grateful for what she has been with me.

Also, I am thankful to all my collaborator students, interns and colleagues that worked with me during these years, for the support received from them. Among all, I would like to mention Francesco Padovani, Stefano Orsolini, Francesco Tortora, Cristina Roperti and Matteo Mazzà, for their remarkable contribution in this shared path. I truly hope that the period spent together will be helpful for their future to achieving their goals, as it was for me.

Many thanks to the Centre and Studies of Cognitive Neuroscience of the Department of Psychology of the University of Bologna for the support that I received.

This PhD would not have been possible also without the support of my family and my friends. I will always be grateful for what my father Alessandro and my mother Anna gave to me, for all their love and encouragement. It is well above my best ability to write something commensurate to their commitment. I can never thank my parents enough for having raised me believing in every step that I made and supporting me in all my pursuits. Furthermore, I must mention my grandparents Simone and Enza, who have been close to me for a lifetime, supporting and taking care of me. Also, I want to thank my grandfather Francesco for giving me his ingeniousness and his passion to discover the world.

I would like to thank Giulia Petrillo for holding me in spite of everything. I thank her for all for the love, the encourage, and to be a patient and faithful partner that supported me during all the stages of this PhD. I could not have done it without her on my side.

Also, I would like to thank other relatives and friends that always scaffolded my life to allow me to achieve this life-goal. In particular, Maria Rita Battaglia, Gabriele Prestifilippo, Gianfranco Catania, Elena Catania, Andrea Curcuruto and many others with whom I shared joy and rage.

Finally, I would like to thank the hundreds of participants that I tested during these years, without their kind help this would not have been possible.

Simone Battaglia,

October 31<sup>th</sup>, 2019