Fear memory uncovered: Prediction error as the key to memory plasticity

Sevenster, D.

Citation for published version (APA):
Sevenster, D. (2014). Fear memory uncovered: Prediction error as the key to memory plasticity.
Chapter 1

General Introduction
The notion that memories go through an initial labile, sensitive phase before being consolidated into stable long-term memory is well embedded in psychological and neurobiological models of memory (Dudai, 2004; McGaugh, 2000; Schafe & LeDoux, 2000). Exciting recent insights show that the retrieval of a consolidated memory can return it to a plastic state, and as a result the retrieved memory can be modified, strengthened, or even disrupted (Kindt, Soeter, & Vervliet, 2009; Nader, Schafe, & Le Doux, 2000). This phenomenon, referred to as reconsolidation, is considered crucial for the understanding of memory plasticity and, uniquely, offers the possibility to target the excessive emotional memory that comes with anxiety disorders. However, the circumstances under which memory reactivation effectively induces reconsolidation remain to be elucidated. Also, little is known about the adaptive function of the reconsolidation process. It has been suggested that the reconsolidation phase might enable the updating of a consolidated memory trace. Yet, experimental evidence that reconsolidation acts as an updating mechanism is scarce in the animal literature (Lee, 2008; Morris et al., 2006; Pedreira, Pérez-Cuesta, & Maldonado, 2004), and is even lacking for human fear memory. In the present dissertation, the functional role of reconsolidation is examined to gain further insights in the fundamentals of learning and memory, which ultimately will guide the development of reconsolidation-based interventions for anxiety disorders.

**Experimental model for fear learning and fear reduction**

Two events can become associated, when they occur closely together. If one event produces fear, the other event will bring fear as well. Pavlov, whose name is now practically a synonym for conditioned responding, formalized this process of associative learning. While his research mainly focused on classical conditioning of canine salivary reflexes, the principal is the same for conditioning of fear responses. During fear conditioning, an initially neutral or ambiguous conditioned stimulus (CS: e.g., a tone or picture) is paired with an aversive outcome or unconditioned stimulus (US: e.g., electric shock). As a result, the CS becomes associated with the US, and subsequent presentations of the CS elicit conditioned responses (CR) (Pavlov, 1927), such as increases in heart rate, stronger startle potentiation and freezing.

After its acquisition, the expression of learned fear can be diminished by means of extinction learning, the repeated presentation of the CS without the aversive consequence (Bouton & Bolles, 1979; Rescorla, 2001). While it can
successfully reduce conditioned fear responding, extinction is a fragile learning process. A large amount of studies show that in a number of situations fear responding can re-emerge. Examples of such situations are the unsignaled presentation of the aversive event (reinstatement), the mere passage of time (spontaneous recovery), or a context change after extinction (renewal). Based on these findings it is concluded that extinction does not erase the original fear memory. Instead, extinction learning involves the formation of an inhibitory memory trace that competes with the original fear memory but leaves the original fear memory intact (Bouton, 1993, 2002). The finding that long-term memory for fear extinction requires protein synthesis in the medial prefrontal cortex (mPFC) (Santini, Ge, Ren, Peña de Ortiz, & Quirk, 2004), a structure considered essential for extinction consolidation and extinction recall (Herry & Garcia, 2002; Morgan, Romanski, & LeDoux, 1993), supports the idea that during extinction, an additional memory trace is formed.

Extinction training is the dominant experimental model for current cognitive behavioural therapies, which are considered to be the most effective treatments for anxiety disorders (Craske, 1999). A central element in these therapies is the imaginary or in vivo exposure to the feared object. Similar to extinction learning, therapy does not rewrite the original fear memory but installs an additional inhibitory memory. Preservation of the fear memory explains the high rates of relapse after successful therapy (Craske, 1999). Pharmacological enhancement of the inhibitory memory trace has been proposed to improve the efficacy of exposure-based psychotherapies (Norberg, Krystal, & Tolin, 2008). However, such treatment would only strengthen the inhibitory extinction memory but still leave the original fear memory intact. To achieve more permanent effects of therapy, we should be able to eliminate the root of anxiety disorders.

Retrieval-induced memory plasticity
In the 1960s and 1970s, animal studies on post-consolidation memory plasticity showed that a consolidated memory is less stable than previously assumed (Lewis, 1979; Misanin, Miller, & Lewis, 1968). These findings were somewhat pushed under the rug, since the field of learning and memory was dominated by consolidation dogma. Consolidation is the process by which newly learned information is transferred into a stable long-term memory (LTM). The classic consolidation hypothesis implies that a memory trace becomes resistant to manipulation once protein synthesis-dependent consolidation is over (Fig. 1A). This is supported by the
finding that infusions of a protein synthesis inhibitor in the amygdala — a site that is considered to be essential for fear learning (Davis & Lang, 2003; LeDoux, 2003) — produced long-term fear amnesia when administered immediately after fear learning but not after a 6h delay (Schafe & LeDoux, 2000).

![Diagram of memory states and stability over time.](image1)

**Figure 1.** Schematic models of memory states and stability over time. (A) According to traditional views, once memory is stabilized via consolidation mechanisms, it remains permanently. (B) Recent insights show that memory is not necessarily permanent but can become active upon retrieval of the memory. Based on the model of Dudai (2009).

The demonstration that memory is transiently dependent on protein synthesis for stability, not only after consolidation but also after its retrieval, marked the renaissance of the phenomenon of reconsolidation (Nader et al., 2000). In this landmark study, a consolidated fear memory was reactivated, followed by administration of a protein synthesis inhibitor. Noticeably, short-term fear memory was still intact 4 h after memory reactivation, but the long-term expression of fear (24 h) was erased (Nader et al., 2000). These findings strongly support the original suggestion that reactivation of a consolidated memory can return it to a labile state. This signalled a radical shift in the perspective on memory, which is presently appreciated as a dynamic rather than a static process (Fig 1B).

Preclinical studies in our own lab were the first to show that the expression of human fear memory can also be erased by pharmacological manipulation in
combination with memory reactivation. Propranolol, which is supposed to specifically block the noradrenergic receptors of the basolateral amygdala, administered before or after reactivation of a previously conditioned fear memory, erased the long-term fear expression. In contrast, the declarative memory or the factual knowledge of the “trauma” survived, suggesting that only the emotional component of the fear memory was affected (Kindt et al., 2009; Soeter & Kindt, 2010, 2011a, 2012). Importantly, it was shown that disruption of the fear memory requires active retrieval, as propranolol did not have any fear reducing effects in absence of a reminder cue. Later studies showed that the amnesic effects of propranolol were long lasting (i.e., up to a month after fear learning) and could also be accomplished in strong fear memories (Soeter & Kindt, 2010, 2011b). These findings are of tremendous importance for the development of reconsolidation-based treatments of anxiety disorders.

Note that in animal studies, the amnesic agents can directly inhibit protein synthesis by preventing the translation of mRNA into proteins (Fig. 2). Furthermore, the amnesic agent can be infused straight into the brain. In human research, neither the use of protein synthesis inhibitors nor local infusions are feasible. Contrary to protein synthesis inhibitors, propranolol does not directly disrupt the translation of mRNA in proteins. However, noradrenergic inhibition of the protein kinase A (PKA) signalling pathway does indirectly disturb new protein synthesis. The PKA pathway is involved in reconsolidation of fear memories, since fear memory reactivation followed by PKA inhibition in the amygdala disrupts the long-term fear expression (Tronson, Wiseman, Olausson, & Taylor, 2006). Noradrenergically innervated PKA signalling triggers the transcriptional activator cAMP response element-binding protein (CREB) (Purves et al., 2012), while CREB-driven transcription of DNA in mRNA results in the synthesis of new proteins (Fig. 2). Thus, indirect inhibition of the PKA pathway by noradrenergic blockade should ultimately affect new protein synthesis.

Constraints on reconsolidation

Although the phenomenon of reconsolidation has been demonstrated in various memory tasks across a variety of species (rats, mice, crabs, bees, humans), it was noted that memory reactivation does not automatically result in its destabilization. Instead, the induction of reconsolidation seemed to depend on specific parameters. Circumstances under which a memory that normally would become labile is resistant to manipulation are referred to as boundary conditions. Several
boundary conditions were proposed, such as age and strength of the memory (Suzuki et al., 2004), extinction consolidation (Bos, Beckers, & Kindt, 2012; Eisenberg, Kobilo, Berman, & Dudai, 2003; Lee, Everitt, & Thomas, 2004; Lee, Milton, & Everitt, 2006), and indirect memory reactivation (Débiec, Doyère, Nader, & LeDoux, 2006).

**Figure 2.** Molecular mechanisms of fear memory reconsolidation. Protein synthesis inhibitors directly disrupt the translation of mRNA in new proteins in the cell. Stimulation of the β-adrenergic receptors (β-AR) results in activation of protein kinase A (PKA) to activate transcription factors including cAMP response element-binding protein (CREB). CREB activation initiates transcription in mRNA, which ultimately results in new protein synthesis. Propranolol disrupts this downstream molecular signalling cascade.

**Memory updating as a key to memory erasure**

The idea that reconsolidation serves the adaptive function of memory updating suggests that these boundary conditions do not exist. As stated previously, it is hypothesized that the function of reconsolidation is to integrate relevant
information presented during memory retrieval in the existing memory trace. Hence, reconsolidation should only be induced upon presentation of significant novel information to maintain memory relevance in a dynamic ever-changing environment (Dudai, 2009; Lee, 2009). In theory, all memories should be able to reconsolidate as long as the retrieval procedure contains relevant information. Indeed, a preliminary study on crabs suggests that the absence of new learning during memory retrieval is a restraint on reconsolidation. The crabs were extensively fear conditioned on the first day (i.e., multiple CS-US trials). Administration of a protein synthesis inhibitor following memory reactivation on day 2 did not disrupt reconsolidation when the memory was reactivated with a reinforced reminder trial (i.e., a CS-US trial). However, memory reactivation with an unreinforced reminder trial (i.e., CS-noUS trial) resulted in long-term erasure of fearful responding (Pedreira et al., 2004). To omit presentation of the aversive consequence was in this case sufficient to destabilize the memory. These findings were interpreted as consistent with the hypothesis that the reminder trial should contain new information with regard to the original learning in order to update the memory.

Null results (absence of amnesia) do not necessarily imply that a given memory reactivation procedure is incapable of triggering reconsolidation. For example, in contrast to the findings by Pedreira et al. (2004), Lee (2008) showed that absence of reinforcement is not a boundary condition to reconsolidation per se. After a single CS-US acquisition trial, a reinforced reminder trial the next day increased subsequent memory strength. The strengthening of the previously formed memory trace relied on reconsolidation mechanisms. Thus, when a weakly trained animal receives a reinforced memory retrieval trial, reconsolidation is engaged (Lee, 2008). In contrast, a similar procedure does not result in memory updating in strongly trained animals (Pedreira et al., 2004). The interaction between the learning history and the reactivation procedure seems to be critical for memory updating.

It is important to note that new information during memory reactivation could not only result in effective destabilization, but could alternatively initiate the formation of a new memory trace. For example, prolonged or repeated memory retrieval triggers inhibitory learning (extinction learning), thereby preventing reconsolidation (Bos et al., 2012; Eisenberg et al., 2003; Lee et al., 2006). Thus, memory destabilization requires a certain amount of task-related novelty. This raises a serious problem for empirical falsifiability. How is the optimal amount of
new learning determined, such that memory destabilization but not additional
new learning is engaged by memory retrieval? Successful destabilization is
typically inferred from testing the sensitivity of a reconsolidation-disrupting agent
under a certain set of experimental parameters. If the procedure fails to reduce
the behavioural expression, it is concluded that under those conditions, updating
does not take place. Since the number of reactivation protocols is infinite, it can
never be ruled out that under slightly different reactivation conditions, memory
might be destabilized. It is therefore crucial that we provide a measure of memory
destabilization that is independent of the occurrence of reconsolidation itself. This
index should provide a guide as to whether reactivation engages memory
updating and renders the memory sensitive to disruption. From a clinical point of
view, the development of such an index is of great importance since it might
indicate whether a reactivation procedure has the potential to be effective or
not.

Cognitive versus affective fear learning
Propranolol administered after memory reactivation reduces long-term startle
potentiation but leaves the skin conductance response (SCR) and threat
expectancy unaffected (Kindt et al., 2009; Soeter & Kindt, 2010, 2011b). These
findings indicate that while propranolol eliminates the emotional component of
fear one day after memory reactivation, declarative memory remains intact. In
other words, the emotional response that previously accompanied the encounter
with the feared object is not there anymore, and only the memory of the factual
events remains. The fear potentiated startle reflects the emotional component of
fear learning. Extensive work on the startle response indicates that it is an
amygdala-initiated reflex. Given that potentiation of the startle reflex is an index of
activation of the subcortical defensive system, the startle response is considered a
specific measure of fear (Davis, 1992, 2006). The cognitive component of fear
learning is reflected by US-expectancy ratings. Individuals learn under which
circumstances a certain stimulus predicts an aversive event. In conditioning
research, participants are instructed to rate to what amount they expect the CS to
be followed by the US, resulting in a trial-by-trial evolution of their expectancies.
This type of simple factual knowledge is known to rely on the hippocampal
complex (Hunsaker & Kesner, 2013). To what extent the skin conductance
response, a commonly used measure of conditioned responding, reflects the
emotional or cognitive component is currently under debate. Although less
intensively studied, it is known that SCR is due to fluctuations in sweat gland activity, resulting from the release of acetylcholine by the sympathetic nervous system (Boucsein, 1992; Wang, 1964).

The finding that the startle response and US-expectancy ratings can dissociate suggests that cognitive and emotional learning stem from multiple underlying systems (Baeyens, Eelen, & Crombez, 1995; Hamm & Weike, 2005). This is, however, opposed by theories advocating that both expectancy and emotional learning arise from a single common mechanism (Mitchell, Houwer, & Lovibond, 2009). In conclusion, there is a debate on (1) whether SCR reflects the cognitive or emotional component of fear learning and (2) whether these two components rely on a similar mechanism. The observation that propranolol only targets affective responding but not the factual knowledge of the events is very promising from a clinical point of view. Patients maintain the ability to recall the traumatic incident but are no longer haunted by events from the past. Therefore, it is important to fully understand the underlying mechanisms of the different response systems measured during fear conditioning.

Outline of the thesis

The aim of the current thesis is two-fold. First, since previous studies show a differential effect of propranolol on the cognitive and emotional component of fear learning, we further investigate this dissociation between response systems. Second, the central aim of the thesis is to investigate the role of reconsolidation in memory updating. If reconsolidation enables memories to be updated, its induction should be restricted to retrieval conditions that allow updating of the original memory. To address these issues, we conducted fear conditioning studies in healthy human participants. We measured both psychophysiological (startle response and SCR) and subjectively experienced conditioned responses.

This thesis comprises two sections. Before addressing the functional role of reconsolidation, we try to elucidate the dissociation between the cognitive and emotional component of fear learning in the first section (Chapters 2 and 3). To gain more insight into the mechanisms underlying fear learning is of special importance, since previous studies showed a dissociation in the effect of propranolol on the different conditioned responses involved in fear learning. In Chapter 2, we test whether fear acquisition could be observed in a differential paradigm with difficult-to-discriminate CSs. If the startle reflex is an automatic index of fear, it may not require the conscious acquisition of contingency
knowledge. In contrast, if SCR is a measure of anticipatory arousal, electrodermal conditioning should be observed only when participants are aware of the CS-US contingencies. Next, we investigate in Chapter 3 the effect of a simple verbal instruction that the CSs will no longer be followed by the aversive US one day after fear acquisition, on the cognitive and emotional components of fear memory. If the startle reflex is an automatic index of fear, this cognitive manipulation should reduce US-expectancy ratings but not the startle fear response. If SCR is indeed a measure of cognitive responding, it should mirror the US-expectancy ratings. We expect that manipulations of contingency awareness (Chapter 2 and 3) might reveal a dissociation between the cognitive and emotional components of fear learning.

The second section (Chapters 4, 5 and 6) focuses on the updating of human fear memory. In Chapter 4 we test whether absence of any new learning during memory retrieval prevents reconsolidation. When the retrieval conditions are perfectly predictable, new learning should not take place. Then, there should be no need for memory updating, leaving the fear memory trace unaltered. Next, in Chapter 5, we investigate whether new learning is required for reconsolidation. Reconsolidation should only occur when memory retrieval contains new learning, in order to update the memory. Finally, in Chapter 6, we investigate the transition from memory updating to new learning. Memory updating should be prevented by too much new learning.