Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Trends Neurosci. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Trends Neurosci. 2016 May ; 39(5): 340–351. doi:10.1016/j.tins.2016.03.003.

New learning and unlearning: strangers or accomplices in threat memory attenuation?

Roger L. Clem[*] and **Daniela Schiller**

Departments of Neuroscience and Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY

Abstract

To achieve greatest efficacy, therapies for attenuating fear and anxiety should preclude the re-emergence of emotional responses. Of relevance to this aim, preclinical models of threat memory reduction are considered to engage one of two discrete neural processes: either establishment of a new behavioral response that competes with, and thereby temporarily interferes with expression of, an intact threat memory (new learning), or one which modifies and thereby disrupts an intact threat memory (unlearning). We contend that a strict dichotomy of new learning and unlearning does not provide a compelling explanation for current data. Instead, we suggest the evidence warrants consideration of alternative models that assume cooperation rather than competition between formation of new cellular traces and the modification of preexisting ones.

Keywords

Reconsolidation; extinction; engram; memory

The return of unlearning

Until year 2000, majority opinion held that memories are formed through a onetime process of consolidation, a form of synaptic plasticity requiring new protein synthesis and long-term molecular changes. This was thought to be the only “active” state in the lifetime of a memory. Once formed, memories were thought to be stored in an inactive state, from which they are then passively recalled [1].

Pavlovian, or classical, threat conditioning is the primary laboratory model of real life threat memories [2]. In this procedure, an innocuous stimulus, such as a light or a tone (conditioned stimulus, CS) reliably precedes a noxious stimulus, such as an air-puff to the eye or a mild electric foot shock (unconditioned stimulus, US). Several such pairings are typically sufficient for establishing a firm CS-US association, endowing the CS capability for triggering defensive responses (such as freezing) without the US. The traditional view of

[*]Correspondence: roger.clem@mssm.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

memory formation implied that the only way to attenuate such threat memories would be via new learning; such as via extinction, wherein the CS is repeatedly experienced in the absence of the US [2, 3]. Extinction was presumed to lead to the formation of a new CS-"no US" association via similar consolidation mechanisms [4]. However, around year 2000, several key studies provided compelling new evidence for a once neglected idea – that memories are not only active upon initiation but also when retrieved – promoting a paradigm shift in memory research [5, 6]. Under this scheme, retrieval returns memories to a labile state (termed reconsolidation [6, 7]) bearing great resemblance to consolidation in that it also requires synthesis of new proteins and molecular changes to once again confer long-term memory persistence. With the discovery of this post-retrieval re-storage state, new learning gave way to an alternative approach for threat memory attenuation: unlearning of the original association.

From a therapeutic standpoint, unlearning has the clear advantage that relapse to a defensive response is theoretically impossible following memory erasure. Despite behavioral evidence consistent with new learning and unlearning, predictions regarding the neural correlates of these processes remain ambiguous and incomplete. Here, we will discuss evidence in support of new learning and unlearning mechanisms, as well as their mutual exclusivity, from behavioral, cellular, molecular and systems levels of analysis.

Behavioral disambiguation of new learning from unlearning

Behavioral disambiguation of new learning from unlearning is typically judged against a set of circumstances under which the return of attenuated defensive responses occurs. Such situations are termed: spontaneous recovery, reinstatement, renewal and savings. If a memory is merely suppressed (i.e., via new learning) and not permanently disrupted it might spontaneously recover with the passage of time. A defensive response would then be produced when the CS (without the US) is next experienced. Similarly, a defensive response to the CS is said to be reinstated following extinction if it occurs when the CS is encountered shortly after the experience of stress (for example, exposure to a few unpaired USs). If a defensive response to a CS is generated when the context in which a CS is encountered differs from the one in which extinction occurred, the response is said to be renewed. Lastly, if new learning is responsible for the attenuation of a defensive response, then a previously conditioned CS should more readily enter a new association upon re-conditioning, a phenomenon termed savings. For example, following extinction, fewer conditioning trials can cause the CS to regain its threatening properties than are required for training with a naïve CS [3, 4, 8–10]. While memories attenuated through new learning are susceptible to at least one of the above challenges, those that have been unlearned are deemed incapable of reemergence in any way.

Putative unlearning protocols include pharmacological and behavioral interference with reconsolidation. In a typical experiment, memory is rendered labile by reactivation using a single presentation of the CS after the CS-US association has been fully consolidated. Following this reactivation, interference with reconsolidation is performed either pharmacologically (such as by administration of protein synthesis inhibitors, adrenergic antagonists, etc) or behaviorally (for example by way of extinction or other types of new

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

learning during reconsolidation). Because threat memory cannot persist without undergoing reconsolidation, it is predicted that these manipulations will confer a permanent disruption of defensive CS responses. Consistent with this prediction, various pharmacological agents and behavioral manipulations have been reported to induce amnesia and prevent the return of conditioned responses under a host of behavioral challenges [11–17].

If follows from the aforementioned behavioral criteria that new learning and unlearning are defined depending on whether memory recovery has been observed. However, this interpretation suffers from the complication that memory might still recover under conditions other than those tested; indeed, it is empirically impossible to provide behavioral proof that memory would *never* return. For example, a recent study of context conditioning in rats showed that conditioned freezing was prevented when tested two days after pharmacological blockade of reconsolidation, but a subsequent US reminder successfully recovered the memory and the rats returned to freezing in the conditioned context [18]. From a behavioral standpoint, memory recovery does not necessarily indicate that new learning was solely responsible for the attenuation of defensive responses to a CS. The recovery could stem, at least in part, from reversal of a partial unlearning process or even the relief of a retrieval impairment [19]. By the same token, some extinction protocols lead to persistent fear attenuation [20], and thus might engage an unlearning mechanism. Thus, behavioral criteria are only adequate to a certain extent for discriminating between new learning and unlearning. Specifically, they can identify conditions that lead to “more persistent” threat memory attenuation only in situations in which extinguished memories typically recover. Such observations are suggestive of unlearning, or could sometimes be interpreted as augmented new learning, but whether permanent memory modification had ensued would require another level of investigation. Could we aid behavioral disambiguation of new learning and unlearning by delineating their discrete neural substrates?

Differentiation of synaptic substrates of new learning and unlearning

Synaptic substrates of threat memory (Box 1) have been extensively investigated through electrophysiological recordings in brain slices, where it is possible to reveal the experience-dependent strengthening of synaptic connections as well as probe its cellular and biochemical basis. In a typical design, recordings are obtained from animals that have acquired a CS-US association through pairing an auditory stimulus (CS) with a mild electric foot-shock (US), and compared to those from various control conditions that involve stimulus exposure but do not result in an auditory threat association, such as unpaired presentations of the CS and US. For example, post-training recordings indicate that auditory threat conditioning strengthens excitatory synapses formed by thalamic and cortical afferents that convey sensory input to principal neurons of the lateral amygdala [21–29]. This is not due to mere context or stimulus exposure, but depends on CS-US associativity. Furthermore, the disruption of synaptic AMPA receptor incorporation in lateral amygdala neurons, which mediates thalamic synaptic strengthening, prevents the expression of conditioned defensive responses [30]. More recent work has also correlated threat conditioning with glutamatergic plasticity of excitatory and inhibitory neurons in the basal and central amygdala [31–33], where learning also impacts a large proportion of sampled neurons and stimulated pathways.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

These robust correlates make the amygdala an ideal testing ground for synaptic mechanisms of unlearning, since they constitute potential substrates for the reversal of memory encoding.

Box 1

Unlearning and the threat memory engram

Memory is nearly universally ascribed to a plastic reconfiguration of neural transmission, which sculpts new pathways linking sensory input to cognitive and emotional output. The mnemonic engram is defined as the physical substrates, such as molecular and structural synaptic changes, that potentiate a neural response to a stimulus and elicit a novel and reliable reaction. Associativity of synaptic plasticity as well as its calibration by learning-related neuromodulatory systems makes it an appealing mechanism for storage of threat-related information [94]. Evidence supports this notion by indicating overlapping requirements of classical conditioning and synaptic plasticity for a multitude of cellular biological processes in the amygdala and hippocampus [95], regions essential for acquisition and expression of threat associations. However, the clearest evidence that synaptic plasticity mediates threat learning is derived from elucidation and perturbation of the synaptic changes that are induced by threat conditioning, as detailed in this review.

A popular refinement of this account posits that memory traces are the specialization of a sparse, distributed ensemble of neurons. Often referred to as engram cells, these neurons are inferred to exhibit higher CS-evoked activity, and their functional or physical ablation renders threat memory irretrievable [93]. Importantly, there are two explanations for these properties with very different implications: these ensembles may be a physical substrate for threat memory encoding, or they may instead serve as a passive relay for retrieval of stored information from other neurons.

It follows from the above that if new pathways for memory can be created during learning, they must likewise be broken during unlearning in order to abolish encoded associations. In contrast, new learning is postulated to inhibit memory expression in specific situations without reversing the storage of previous information, thereby leaving a response subject to return under the appropriate conditions. The distinction between these forms of attenuation thus hinges on the persistence of a synaptic threat memory trace (Fig. 1).

Indeed, this form of unlearning has now been effectively modeled through optogenetic induction of thalamo-amygdala synaptic plasticity after threat conditioning [34]. In this study, the authors showed that reversal of synaptic potentiation abolishes conditioned responses, which can be subsequently recovered by optogenetic repotentiation. This supports the causal role of synaptic strength in memory encoding and the plausibility of bidirectional plasticity in mediating learning and unlearning. In contrast, the delivery of potentiating optogenetic stimulation fails to restore conditioned responses after extinction, consistent with the notion that extinction is a process distinct from unlearning in which new inhibitory CS associations override the potency of thalamic inputs. It therefore seems plausible that threat conditioning and extinction involve distinct cellular substrates, and that threat associations can be inhibited as well as unlearned.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Consistent with this thinking, electrophysiological studies suggest that some synaptic pathways are selectively modified during either threat conditioning or extinction, but not both, as might be expected for non-overlapping representations (Fig. 1). For example, extinction training has been associated with depression of medial prefrontal cortex (mPFC)-basal amygdala synapses [35] as well as potentiation of basal amygdala synapses onto amygdala intercalated cells (ITCs) [36], pathways that were unaffected by threat conditioning in these studies in interleaved control subjects. It is worth noting, however, that a more selective examination of inputs originating from prelimbic mPFC revealed a potentiation of synapses in basal amygdala principal neurons after threat conditioning [31], consistent with other evidence of close cooperation between mPFC and amygdala in threat memory encoding [37–39]. Within the basal amygdala, there exist discrete neuronal populations that fire preferentially during high and low defensive states of the animal resulting from threat conditioning and extinction [40], and that project selectively to hippocampal and prefrontal brain regions [40, 41]. It remains unknown whether experience-driven changes in synaptic efficacy play a role in these dynamics, but these populations represent a potential hardwired substrate for the storage of competing associations through new learning (Fig. 1), permitting extinction to be acquired without the reversal of prior CS-related plasticity.

In contrast to these extinction studies, manipulations targeting reconsolidation have been shown to alter transmission at synapses previously strengthened during threat conditioning (similar to Fig. 1). Pharmacological interference with reconsolidation is associated with reversal of excitatory synaptic potentiation in the lateral amygdala, as measured through *in vivo* recordings [42] as well as through postmortem brain slice physiology [43]. However, in the latter study, threat conditioning potentiated thalamic synapses by enhancing glutamate release, while reconsolidation blockade reduced the efficacy of AMPA receptor transmission. This appears to challenge the notion that reconsolidation is a process that, unlike extinction, acts exclusively and directly on threat memory substrates. In contrast, a more selective examination of hippocampal neurons activated by a conditioned context revealed enhancement of AMPA receptor currents, compared to adjacent non-activated cells, and reversal of this effect by reconsolidation blockade [44]. It therefore seems possible that only some forms of reconsolidation blockade or particular substrates will reflect unlearning at the cellular level. While the literature overwhelmingly attributes the efficacy of amnestic drug treatments to a disruption of memory restabilization, direct evidence of trace modification remains scant, and studies have mostly neglected potential *de novo* effects of these drugs.

Just as clarity is lacking about the mechanistic basis of pharmacological reconsolidation impairments, ambiguity exists regarding the capability of behavioral interventions to activate unlearning mechanisms. In particular, many have inferred that a cellular unlearning process could be sometimes engaged by extinction training based on the failure of defensive behavior to return after modified extinction protocols. For example, the partial NMDA receptor agonist D-cycloserine augments extinction and prevents spontaneous recovery [45–48], effects that are associated with reversal of threat-associated synaptic AMPA receptor enrichment [46]. Likewise, the application of a retrieval trial prior to extinction, a paradigm discussed in further detail below, can prevent threat memory spontaneous recovery, renewal

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and reinstatement [49, 50]. Retrieval-extinction is associated with removal of AMPA receptors from thalamic synapses in the lateral amygdala, leading to a long-lasting reversal of threat-related potentiation [21]. These effects are localized to the same cellular compartment and involve the same molecular substrates as threat-related synaptic strengthening, suggesting that it is possible to engage a molecular process that weakens memory traces using a purely behavioral manipulation. However, is the presence of such weakening sufficient to establish that threat associations have been unlearned and cannot return without relearning? We need only look to conventional extinction paradigms to show that such effects do not always exclude the possibility of relapse.

Despite evidence that extinction introduces new synaptic modifications, and involves the recruitment of new circuitry, several lines of observation suggest that it can also involve trace modifications that spontaneously revert with the return of threat memory. One group has shown that a rather extensive, but nevertheless conventional, form of extinction leads to depotentiation of thalamo-amygdala synapses, as well as synaptic AMPA receptor removal [23, 51]. However, these effects do not preclude renewal of threat conditioning outside of the extinction context. More surprisingly, reinsertion of AMPA receptors occurs during renewal, despite the fact that these animals were never re-exposed to unconditioned stimuli [52]. A similar bidirectional regulation of inhibitory transmission has been observed in the lateral amygdala, in which high and low defensive states related to threat conditioning, extinction and renewal are correlated with internalization and re-accumulation of surface GABA receptors [24]. These electrophysiological and biochemical effects are further reminiscent of structural correlates of threat memory acquisition and extinction in the frontal association cortex, where the acquisition and attenuation of defensive responses are associated with retraction and subsequent reappearance of dendritic spines on persistently imaged neurons [53]. The reinstatement of memory after synapse destabilization is not exclusive to extinction, but also occurs during recovery of conditioned responses after reconsolidation blockade [54] as well as peptide-based amnestic treatments [54, 55]. Therefore, despite its potential contribution to fear reduction, trace modification may not be sufficient to block the return of memory. This suggests that memory can be maintained by secondary substrates in occult state, where its recovery may be subject to particular behavioral conditions or brain states.

While it is not currently possible to obtain a comprehensive map of plasticity in order to locate occult traces, a potential resolution to this impasse might be achieved by elucidating the role of so-called engram cells (Box 1). A recent elegant study demonstrated that putative hippocampal engram cells exhibit synaptic depotentiation following reconsolidation blockade, but memory can be recovered by bypassing synaptic inputs to directly activate these cells with optogenetic stimulation [44]. This suggests that a memory-related ensemble is not disintegrated by reconsolidation blockade, but retains properties conducive to memory recovery, such as unique connectivity. These residual properties may persist as stored information and support the reinstatement of active memory trace under conditions of relapse, a possibility that we consider below.

One important caveat to the above is that the rapid adaptation of threat-related ensembles with additional experience complicates their interpretation as a stable coherent engram. For

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

example, hippocampal place cells undergo remapping in response to both threat conditioning and extinction [56]. This adaptation is not consistent with discrete “threat” and “safety” ensembles, but rather a hybrid network in which firing persists in a subset of threat-related neurons after extinction, similar to unit recordings of extinction-related activity in the basal [40] and lateral amygdala [57]. It is possible that cells comprising an ensemble are tuned to specific experiential variables, such as valence [26], either as a result of their connectivity or plasticity. However, these populations may also be subject to continuous remodeling to make way for the storage of new information and the systems consolidation of previous associations [58–60]. While these properties are problematic for the conventional model of threat memory attenuation (Figure 1), we discuss below a mixed model in which their accommodation is possible (Figures 2–3).

Discrimination on the basis of biochemical mechanisms

If discerning a clear dichotomy in the synaptic expression mechanisms for new learning and unlearning proves inconclusive, perhaps these processes nevertheless engage a set of biochemical induction mechanisms that can serve as a signature of which process has occurred. Such a signature might take the form of a pharmacological induction requirement or a biochemical correlate, such as gene expression. In terms of its biochemical induction, reconsolidation is defined as a two-stage process in which memory is destabilized by retrieval and then later restabilized (or reconsolidated). Therefore, unlearning theoretically corresponds to the destabilization phase, at which point information is lost unless re-encoded into long-term memory, while requirements of the restabilization phase are irrelevant (and even antithetical) to unlearning. Thus, the ideal design for interrogating the unlearning component of reconsolidation entails the pre-retrieval administration of a test drug followed by a second agent, specifically an established reconsolidation blocker. A subsequent behavioral probe would then reveal whether the test drug had impaired trace destabilization, since this would prevent the loss of memory. It is worth noting that a significant pitfall of this approach is that interpretation becomes difficult when the drug in question also impairs memory retrieval [61], which is requisite for trace destabilization.

Application of the above approach has been limited, but the results of these studies indicate that all tested molecular mediators of extinction, including NMDA receptors [62], cannabinoid receptor 1 [62], L-type voltage-gated calcium channels [62] and the ubiquitin proteasome system [63, 64], are also requirements for memory loss via reconsolidation blockade [63, 65–67]. Some evidence suggests that a requirement for proteolysis in memory destabilization persists after memory retrieval, during the period of reconsolidation, and requires calcium/calmodulin-dependent kinase II (CaMKII) [68]. However, CaMKII is also requisite for extinction [69]. Thus, a reagent with the capacity to discriminate between extinction- and reconsolidation-based threat memory reduction, and their hypothetical basis in new learning versus unlearning, has yet to be identified.

Not only are reconsolidation and extinction considered to be discrete processes, but they are also postulated to exhibit mutual exclusivity, such that the inception of extinction precludes reconsolidation blockade [70]. Extinction paradigms leading to persistent threat memory attenuation therefore present an interesting test case for biochemical signatures of new

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

learning and unlearning, since they can presumably engage either extinction or reconsolidation, but not both. For example, in retrieval-extinction the introduction of an isolated CS trial shortly before extinction prevents the subsequent reemergence of threat memory during spontaneous recovery, renewal and reinstatement tests in both rodents and humans [21, 50, 71–80]; although for unknown reasons some studies have failed to replicate these effects (for in-depth reviews see [16, 81, 82]). Reports have favored the interpretation that the initial CS trial destabilizes threat memory and initiates a time-limited lability window during which extinction training can either interfere with memory restabilization or supply new safety-related information that can be integrated into the memory during its reconsolidation. However, a recent study argues against this interpretation, based on obtainment of similar behavioral effects when the order of retrieval and extinction are reversed [83].

Interestingly, many neural correlates of retrieval-extinction defy explanation exclusively as a reconsolidation- or extinction-based process. For example, compared to conventional extinction, retrieval-extinction is associated with diminished involvement of the human ventromedial prefrontal cortex (vmPFC), as measured by functional neuroimaging, suggesting a potential disengagement of pathways associated with new inhibitory CS learning [80]. In contrast, increased expression of protein kinase M ζ (PKMζ) [78] and zinc-finger 268 (Zif268) in rodent vmPFC [84] is consistent with enhanced extinction of appetitive memory in this paradigm. Most interestingly, increased expression of phosphorylated ribosomal protein S6 (rpS6P) is specifically observed in both the lateral amygdala and mPFC after retrieval-extinction, but not extinction or retrieval alone (conditions equivalent to extinction and reconsolidation controls) [84]. Among the potential upstream regulators of rpS6P is metabotropic glutamate receptor 1 (mGlu1) [85], which has been independently implicated in both persistent threat memory attenuation and synaptic weakening after retrieval-extinction [21]. While mGlu1 mediates the protein kinase-C (PKC)-dependent selective removal of GluA2-lacking calcium-permeable AMPA receptors (CP-AMPARs) from thalamo-amygdala synapses [22], both conventional extinction [23] and reconsolidation blockade [43, 66] have in contrast been associated with trafficking of amygdala GluA2-containing AMPARs. Available evidence therefore implies that retrieval-extinction is based on a third distinct mechanism, rather than extinction or reconsolidation as currently understood. This would appear plausible when viewing extinction after retrieval as an updating procedure, in which trace modification would have a signature distinct from reconsolidation. However, the assumption of an increasing number of distinct conditions that are all compatible with persistent threat memory attenuation further confounds the elucidation of a biochemical unlearning signature.

Considering the alternative: a mixed model

As was initially postulated on the basis of amnestic drug effects, substrates of memory are subject to destabilization. However, neither the presence of synaptic labilization nor associated biochemical and neural activity has proven useful in demarcating when a process of unlearning has ensued. Indeed, the return of threat memory can be observed even after the reversal of structural and synaptic memory traces, suggesting such memory is sustained by additional cellular substrates or extrinsic circuitry to be reactivated under the appropriate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

conditions. Both extensive overlap as well as heterogeneity in the outcomes and molecular mediators of different treatment paradigms pose major obstacles for reconciling their effects with the dichotomy of new learning and unlearning. It therefore appears that several issues will need to be clarified in order to understand the relationship between trace modification and the persistence of threat memory attenuation.

We suggest that neural representations of threat memories can exist in many configurations and that, likewise, many reconfigurations of these states are compatible with persistent threat memory attenuation. Long-term efficacy may depend not so much on whether therapy is based on new processes versus the reversal of old processes, but on the balance between the two or, more likely, their joint configuration. For example, susceptibility to relapse may require a residual trace of threat learning that is sufficient to support memory reactivation after its partial unlearning (Fig. 2). Ultimately, however, the long-term success of therapy depends on the interaction of attenuated threat memories with situational relapse triggers that prompt an abrupt return to a high defensive state. These triggers include altered context, stress, unconditioned stimulus exposure, and the mere passage of time. Perhaps the most influential model of relapse posits the context-dependent encoding of inhibitory CS associations during extinction, following which the absence of contextual cues leads to a failure of extinction retrieval when the CS is reencountered, due to a mismatch in the test conditions [86]. Some further suggest that the contextual dimension of this encoding encompasses both the exteroceptive and interoceptive state of the animal, thereby explaining the reemergence of threat memory under stress [87]. But how can behavioral responses return even after memory substrates have been modified by therapy?

New learning accounts of extinction stipulate competitive interactions between individually encoded threat and safety associations, implying that relapse biases this competition in favor of defensive responding by gating neural transmission upstream or downstream of the threat memory engram (Fig. 1). In contrast, putative unlearning paradigms like reconsolidation blockade could prevent this relapse in two ways: obliterating the engram or disconnecting it from its relevant inputs or outputs. As discussed above, new evidence suggests that while reconsolidation blockade weakens synaptic inputs to so-called engram cells, it preserves their capacity to evoke a threat response when directly activated [44]. This suggests that these threat-related ensembles exhibit a pre-existing or acquired pattern of connectivity that continues to support memory expression or the recruitment of other circuits where threat associations are stored. Such connectivity provides a basis for memory recovery even after synaptic depotentiation, and potentially the conversion of a once “silent” memory to a functional one through repotentiation (Fig. 3). In this way, the return of threat memory may be akin to a form of reconsolidation involving the reactivation-dependent strengthening of a very weak trace, consistent with reinstatement of memory-related molecular and structural modifications during relapse from extinction [23, 24, 51–53]. However, an important question is why do relapse triggers fail to restore these substrates after some manipulations, i.e. reconsolidation, but not others, i.e. extinction? A more sophisticated analysis of the dynamic interplay between various brain systems, neural ensembles (including putative engram cells), and synaptic pathways modulated by both the CS as well as relapse triggers would be instrumental in elucidating the underlying basis for these discrepancies.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In these studies, a clear rationale exists for emphasizing a brain network comprised of the amygdala, mPFC, and ventral hippocampus (vHC), which are considered to play important roles in situational relapse [40, 88–92] (see Outstanding Questions). One possibility is that various relapse triggers modulate a common circuit within this network to trigger the reemergence of defensive responses. While support exists for the gating influence of bulk connections between the amygdala, vHC and mPFC in CS-evoked freezing [90], the manipulation of specific projections will be required for a more precise blueprint of relapse circuits. The identification of a specific gating mechanism would provide a powerful tool for interrogating attenuated threat memories, since such a circuit could be brought under optogenetic control to elicit cellular processes leading to relapse, while anatomical and electrophysiological methods could reveal the logic of downstream connections with active neural ensembles. A role for such cells has been established in threat memory expression through the use of activity-based tagging and manipulation [93], but these methods should be extended to other ensembles whose activity correlates with extinction or relapse. In particular, it would be valuable to ascertain whether different ensembles compete at the cellular and behavioral level in threat memory encoding, as predicted in the conventional view (Fig. 1), and to determine how relapse circuits gate this competition.

Outstanding questions box

- **Do various behavioral relapse triggers modulate common or distinct circuits that lead to the reemergence of threat memory?** The causal influence of vHC and mPFC pathways in relapse can now be disentangled through the use of *in vivo* optogenetics, providing a basis for elaboration of downstream circuit and molecular processes that ensue upon the return of defensive responses. An even more powerful application of optogenetics would entail artificially eliciting relapse, and accompanying trace reinstatement, by stimulation or inhibition of projection-defined neurons or their target-specific axon terminals.
- **Do active neural ensembles constitute competing engrams in bidirectional regulation of threat memory?** Experimental strategies are needed to establish whether the increased firing of various cell populations in threat memory acquisition, attenuation and reemergence results from their expression of synaptic plasticity, rather than changes elsewhere in the network of excitatory and inhibitory neurons. This would be critical for localizing the neural substrates for memory storage and elucidating how they might compete with one another at the circuit and behavioral level.
- **Assuming the cooperation of new learning and unlearning mediates threat memory attenuation, how can we best aid therapeutic recovery?** Understanding the causal relationships of various neural populations to defensive responses, and how the circuitry of relapse gates their recruitment and plasticity, could reveal key targets for preventing the reemergence of a modified threat memory. This could

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

facilitate the development of more highly specific drugs as well as more effectively tailored behavioral interventions.

Certain outcomes of the above experiments would be consistent with a mixed mechanism for threat memory attenuation. For example, neurons activated by relapse may constitute a partial subset of those that participated in the original threat memory engram (Fig. 2). In this case, electrophysiological recordings should reveal preservation of synaptic potentiation in relapse-activated neurons and depotentiation of remaining cells. In the case of persistently attenuated memories, however, neither natural nor artificial stimulation of relapse circuitry would induce memory recovery, and a more complete depotentiation of engram cells would be observed. In contrast to this scenario, the establishment of a hidden engram poses a more difficult challenge for these forms of circuit-based dissection, in part because activity-based tagging would not discriminate these cells from the primary engram (Fig. 3). However, because a hidden engram should receive connections from relapse-related vHC and mPFC pathways, it is possible that the delineating the downstream substrates of these circuits will reveal potential candidates for the occult storage of memory. In contrast to primary engram cells, which possess potentiated sensory inputs, the hidden engram need not be directly activated by CS sensory relays, but should exhibit connections with the primary engram that can facilitate recovery of an attenuated threat memory. Experiments that silence these connections may not interfere with threat memory expression when CS inputs are in their potentiated state, but after threat memory attenuation such silencing would block the cellular events leading to repotentiation and trace reinstatement.

Concluding remarks

In conclusion, while the dichotomy of new learning and unlearning emerged as a theoretical solution to observations of temporary versus persistent threat memory attenuation, evidence for discrete neural mechanisms corresponding to these processes remains weak. On the one hand, the fact that memory substrates can be readily destabilized under many conditions implies that a potential basis for memory erasure is closer at hand than previously suspected. However, the pursuit of an endogenous routine for unlearning suffers from experimental limitations and ignores both the fluidity of threat memory modulation, in terms of its underlying mechanisms, as well as the dynamic contributions of situational relapse triggers to its long-term efficacy. Placed in context with these influences, the interplay between emotional states and environmental stimuli might be understood as aggregate transformations of a network, some of which confer a very low probability of threat memory recovery. Perhaps one day there will emerge a definitive test for irretrievable loss of threat associations. However, the lack of such a metric does not preclude the application of trace destabilization, along with other endogenous plasticity mechanisms, in the mitigation of relapse.

Acknowledgments

This work was funded by NIMH grant MH105414 and a Young Investigator Award from the Brain and Behavior Foundation to R.L.C. and by NIMH grant MH105515 and a Klingenstein-Simons Fellowship Award in the Neurosciences to D.S.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

1. McGaugh JL. Time-dependent processes in memory storage. Science. 1966; 153:1351–1358. [PubMed: 5917768]
2. Pavlov, IP. Conditioned Reflexes. Oxford University Press; 1927.
3. Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biological psychiatry. 2002; 52:976–986. [PubMed: 12437938]
4. Myers KM, Davis M. Mechanisms of fear extinction. Mol Psychiatry. 2007; 12:120–150. [PubMed: 17160066]
5. Przybyslawski J, Sara SJ. Reconsolidation of memory after its reactivation. Behav Brain Res. 1997; 84:241–246. [PubMed: 9079788]
6. Nader K, et al. The labile nature of consolidation theory. Nature reviews Neuroscience. 2000; 1:216–219. [PubMed: 11257912]
7. Spear NE. Retrieval of memory in animals. Psychol Rev. 1973; 80:163–194.
8. Bentz D, Schiller D. Threat processing: models and mechanisms. Wiley Interdiscip Rev Cogn Sci. 2015; 6:427–439. [PubMed: 26267313]
9. Bouton ME. Why behavior change is difficult to sustain. Prev Med. 2014; 68:29–36. [PubMed: 24937649]
10. Rescorla RA. Spontaneous recovery. Learn Mem. 2004; 11:501–509. [PubMed: 15466300]
11. Agren T. Human reconsolidation: a reactivation and update. Brain Res Bull. 2014; 105:70–82. [PubMed: 24397965]
12. Dudai Y. The restless engram: consolidations never end. Annu Rev Neurosci. 2012; 35:227–247. [PubMed: 22443508]
13. Nader K. Reconsolidation and the Dynamic Nature of Memory. Cold Spring Harb Perspect Biol. 2015:7.
14. Quirk GJ, et al. Erasing fear memories with extinction training. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2010; 30:14993–14997. [PubMed: 21068303]
15. Reichelt AC, Lee JL. Memory reconsolidation in aversive and appetitive settings. Frontiers in behavioral neuroscience. 2013; 7:118. [PubMed: 24058336]
16. Schiller D, Phelps EA. Does reconsolidation occur in humans? Frontiers in behavioral neuroscience. 2011; 5:24. [PubMed: 21629821]
17. Taylor JR, Torregrossa MM. Pharmacological disruption of maladaptive memory. Handb Exp Pharmacol. 2015; 228:381–415. [PubMed: 25977090]
18. Trent S, et al. Rescue of long-term memory after reconsolidation blockade. Nature communications. 2015; 6:7897.
19. Hardt O, et al. Storage or retrieval deficit: the yin and yang of amnesia. Learn Mem. 2009; 16:224–230. [PubMed: 19304892]
20. Dunsmoor JE, et al. Rethinking Extinction. Neuron. 2015; 88:47–63. [PubMed: 26447572]
21. Clem RL, Huganir RL. Calcium-permeable AMPA receptor dynamics mediate fear memory erasure. Science. 2010; 330:1108–1112. [PubMed: 21030604]
22. Clem RL, Huganir RL. Norepinephrine enhances a discrete form of long-term depression during fear memory storage. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2013; 33:11825–11832. [PubMed: 23864672]
23. Kim J, et al. Amygdala depotentiation and fear extinction. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:20955–20960. [PubMed: 18165656]
24. Lin HC, et al. GABAA receptor endocytosis in the basolateral amygdala is critical to the reinstatement of fear memory measured by fear-potentiated startle. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2011; 31:8851–8861. [PubMed: 21677169]
25. McKernan MG, Shinnick-Gallagher P. Fear conditioning induces a lasting potentiation of synaptic currents in vitro. Nature. 1997; 390:607–611. [PubMed: 9403689]
26. Namburi P, et al. A circuit mechanism for differentiating positive and negative associations. Nature. 2015; 520:675–678. [PubMed: 25925480]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

27. Tsvetkov E, et al. Fear conditioning occludes LTP-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. Neuron. 2002; 34:289–300. [PubMed: 11970870]
28. Tye KM, et al. Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. Nature. 2008; 453:1253–1257. [PubMed: 18469802]
29. Zhou Y, et al. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nature neuroscience. 2009; 12:1438–1443. [PubMed: 19783993]
30. Rumpel S, et al. Postsynaptic receptor trafficking underlying a form of associative learning. Science. 2005; 308:83–88. [PubMed: 15746389]
31. Arruda-Carvalho M, Clem RL. Pathway-selective adjustment of prefrontal-amygdala transmission during fear encoding. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2014; 34:15601–15609. [PubMed: 25411488]
32. Li H, et al. Experience-dependent modification of a central amygdala fear circuit. Nature neuroscience. 2013; 16:332–339. [PubMed: 23354330]
33. Penzo MA, et al. The paraventricular thalamus controls a central amygdala fear circuit. Nature. 2015; 519:455–459. [PubMed: 25600269]
34. Nabavi S, et al. Engineering a memory with LTD and LTP. Nature. 2014; 511:348–352. [PubMed: 24896183]
35. Cho JH, et al. Synaptic encoding of fear extinction in mPFC-amygdala circuits. Neuron. 2013; 80:1491–1507. [PubMed: 24290204]
36. Amano T, et al. Synaptic correlates of fear extinction in the amygdala. Nature neuroscience. 2010; 13:489–494. [PubMed: 20208529]
37. Livneh U, Paz R. Amygdala-prefrontal synchronization underlies resistance to extinction of aversive memories. Neuron. 2012; 75:133–142. [PubMed: 22794267]
38. Likhtik E, et al. Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nature neuroscience. 2014; 17:106–113. [PubMed: 24241397]
39. Courtin J, et al. Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. Nature. 2014; 505:92–96. [PubMed: 24256726]
40. Herry C, et al. Switching on and off fear by distinct neuronal circuits. Nature. 2008; 454:600–606. [PubMed: 18615015]
41. Senn V, et al. Long-range connectivity defines behavioral specificity of amygdala neurons. Neuron. 2014; 81:428–437. [PubMed: 24462103]
42. Doyere V, et al. Synapse-specific reconsolidation of distinct fear memories in the lateral amygdala. Nature neuroscience. 2007; 10:414–416. [PubMed: 17351634]
43. Li Y, et al. Learning and reconsolidation implicate different synaptic mechanisms. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:4798–4803. [PubMed: 23487762]
44. Ryan TJ, et al. Memory. Engram cells retain memory under retrograde amnesia. Science. 2015; 348:1007–1013. [PubMed: 26023136]
45. Mickley GA, et al. Acute, but not chronic, exposure to d-cycloserine facilitates extinction and modulates spontaneous recovery of a conditioned taste aversion. Physiology & behavior. 2012; 105:417–427. [PubMed: 21925192]
46. Mao SC, et al. Extinction training in conjunction with a partial agonist of the glycine site on the NMDA receptor erases memory trace. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2006; 26:8892–8899. [PubMed: 16943544]
47. Walker DL, et al. Facilitation of conditioned fear extinction by systemic administration or intra-amygdala infusions of D-cycloserine as assessed with fear-potentiated startle in rats. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2002; 22:2343–2351. [PubMed: 11896173]
48. Ledgerwood L, et al. D-cycloserine and the facilitation of extinction of conditioned fear: consequences for reinstatement. Behavioral neuroscience. 2004; 118:505–513. [PubMed: 15174928]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

49. Monfils MH, et al. Extinction-reconsolidation boundaries: key to persistent attenuation of fear memories. Science. 2009; 324:951–955. [PubMed: 19342552]
50. Schiller D, et al. Preventing the return of fear in humans using reconsolidation update mechanisms. Nature. 2010; 463:49–53. [PubMed: 20010606]
51. Kim J, et al. Reactivation of fear memory renders consolidated amygdala synapses labile. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2010; 30:9631–9640. [PubMed: 20631192]
52. Lee S, et al. GluA1 phosphorylation at serine 831 in the lateral amygdala is required for fear renewal. Nature neuroscience. 2013; 16:1436–1444. [PubMed: 23974710]
53. Lai CS, et al. Opposite effects of fear conditioning and extinction on dendritic spine remodelling. Nature. 2012; 483:87–91. [PubMed: 22343895]
54. Chen S, et al. Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in Aplysia. eLife. 2014; 3:e03896. [PubMed: 25402831]
55. Parsons RG, Davis M. Temporary disruption of fear-potentiated startle following PKMzeta inhibition in the amygdala. Nature neuroscience. 2011; 14:295–296. [PubMed: 21258326]
56. Wang ME, et al. Extinction of Learned Fear Induces Hippocampal Place Cell Remapping. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2015; 35:9122–9136. [PubMed: 26085635]
57. An B, et al. Long-term neural correlates of reversible fear learning in the lateral amygdala. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2012; 32:16845–16856. [PubMed: 23175837]
58. Sacco T, Sacchetti B. Role of secondary sensory cortices in emotional memory storage and retrieval in rats. Science. 2010; 329:649–656. [PubMed: 20689011]
59. Frankland PW, et al. The involvement of the anterior cingulate cortex in remote contextual fear memory. Science. 2004; 304:881–883. [PubMed: 15131309]
60. Do-Monte FH, et al. A temporal shift in the circuits mediating retrieval of fear memory. Nature. 2015; 519:460–463. [PubMed: 25600268]
61. Lopez J, et al. Memory retrieval requires ongoing protein synthesis and NMDA receptor activity-mediated AMPA receptor trafficking. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2015; 35:2465–2475. [PubMed: 25673841]
62. Suzuki A, et al. Memory reconsolidation and extinction have distinct temporal and biochemical signatures. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2004; 24:4787–4795. [PubMed: 15152039]
63. Lee SH, et al. Synaptic protein degradation underlies destabilization of retrieved fear memory. Science. 2008; 319:1253–1256. [PubMed: 18258863]
64. Pick JE, et al. Neuronal expression of the ubiquitin E3 ligase APC/C-Cdh1 during development is required for long-term potentiation, behavioral flexibility, and extinction. Neurobiol Learn Mem. 2013; 100:25–31. [PubMed: 23238556]
65. Ben Mamou C, et al. NMDA receptors are critical for unleashing consolidated auditory fear memories. Nature neuroscience. 2006; 9:1237–1239. [PubMed: 16998481]
66. Hong I, et al. AMPA receptor exchange underlies transient memory destabilization on retrieval. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:8218–8223. [PubMed: 23630279]
67. Suzuki A, et al. Activation of LVGCCs and CB1 receptors required for destabilization of reactivated contextual fear memories. Learn Mem. 2008; 15:426–433. [PubMed: 18511694]
68. Jarome TJ, et al. CaMKII regulates proteasome phosphorylation and activity and promotes memory destabilization following retrieval. Neurobiol Learn Mem. 2016; 128:103–109. [PubMed: 26779588]
69. Kimura R, et al. Autophosphorylation of alphaCaMKII is differentially involved in new learning and unlearning mechanisms of memory extinction. Learn Mem. 2008; 15:837–843. [PubMed: 18984565]
70. Merlo E, et al. Reconsolidation and extinction are dissociable and mutually exclusive processes: behavioral and molecular evidence. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2014; 34:2422–2431. [PubMed: 24523532]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

71. Graff J, et al. Epigenetic Priming of Memory Updating during Reconsolidation to Attenuate Remote Fear Memories. Cell. 2014; 156:261–276. [PubMed: 24439381]
72. Oyarzun JP, et al. Updating fearful memories with extinction training during reconsolidation: a human study using auditory aversive stimuli. PloS one. 2012; 7:e38849. [PubMed: 22768048]
73. Flavell CR, et al. Behavioural memory reconsolidation of food and fear memories. Nature communications. 2011; 2:504.
74. Rao-Ruiz P, et al. Retrieval-specific endocytosis of GluA2-AMPARs underlies adaptive reconsolidation of contextual fear. Nature neuroscience. 2011; 14:1302–1308. [PubMed: 21909089]
75. Ma X, et al. Post-retrieval extinction training enhances or hinders the extinction of morphine-induced conditioned place preference in rats dependent on the retrieval-extinction interval. Psychopharmacology. 2012; 221:19–26. [PubMed: 22048130]
76. Olshavsky ME, et al. Updating appetitive memory during reconsolidation window: critical role of cue-directed behavior and amygdala central nucleus. Frontiers in behavioral neuroscience. 2013; 7:186. [PubMed: 24367304]
77. Sartor, GC.; Aston-Jones, G. Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology. 2013. Post-Retrieval Extinction Attenuates Cocaine Memories.
78. Xue YX, et al. A memory retrieval-extinction procedure to prevent drug craving and relapse. Science. 2012; 336:241–245. [PubMed: 22499948]
79. Agren T, et al. Disruption of reconsolidation erases a fear memory trace in the human amygdala. Science. 2012; 337:1550–1552. [PubMed: 22997340]
80. Schiller D, et al. Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:20040–20045. [PubMed: 24277809]
81. Auber A, et al. Post-retrieval extinction as reconsolidation interference: methodological issues or boundary conditions? Psychopharmacology. 2013; 226:631–647. [PubMed: 23404065]
82. Kredlow MA, et al. Harnessing reconsolidation to weaken fear and appetitive memories: A meta-analysis of post-retrieval extinction effects. Psychol Bull. 2016; 142:314–336. [PubMed: 26689086]
83. Baker KD, et al. Memory retrieval before or after extinction reduces recovery of fear in adolescent rats. Learn Mem. 2013; 20:467–473. [PubMed: 23950194]
84. Tedesco V, et al. Extinction, applied after retrieval of auditory fear memory, selectively increases zinc-finger protein 268 and phosphorylated ribosomal protein S6 expression in prefrontal cortex and lateral amygdala. Neurobiol Learn Mem. 2014; 115:78–85. [PubMed: 25196703]
85. Antion MD, et al. Removal of S6K1 and S6K2 leads to divergent alterations in learning, memory, and synaptic plasticity. Learn Mem. 2008; 15:29–38. [PubMed: 18174371]
86. Maren S, et al. The contextual brain: implications for fear conditioning, extinction and psychopathology. Nature reviews Neuroscience. 2013; 14:417–428. [PubMed: 23635870]
87. Bouton ME, et al. Contextual and temporal modulation of extinction: behavioral and biological mechanisms. Biological psychiatry. 2006; 60:352–360. [PubMed: 16616731]
88. Jin J, Maren S. Fear renewal preferentially activates ventral hippocampal neurons projecting to both amygdala and prefrontal cortex in rats. Sci Rep. 2015; 5:8388. [PubMed: 25669753]
89. Orsini CA, et al. Hippocampal and prefrontal projections to the basal amygdala mediate contextual regulation of fear after extinction. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2011; 31:17269–17277. [PubMed: 22114293]
90. Sotres-Bayon F, et al. Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. Neuron. 2012; 76:804–812. [PubMed: 23177964]
91. Burgos-Robles A, et al. Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2009; 29:8474–8482. [PubMed: 19571138]
92. Ji J, Maren S. Electrolytic lesions of the dorsal hippocampus disrupt renewal of conditional fear after extinction. Learn Mem. 2005; 12:270–276. [PubMed: 15930505]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

93. Kandel ER, et al. The molecular and systems biology of memory. Cell. 2014; 157:163–186. [PubMed: 24679534]

94. Herry C, Johansen JP. Encoding of fear learning and memory in distributed neuronal circuits. Nature neuroscience. 2014; 17:1644–1654. [PubMed: 25413091]

95. Johansen JP, et al. Molecular mechanisms of fear learning and memory. Cell. 2011; 147:509–524. [PubMed: 22036561]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Trends box

- Laboratory protocols for attenuating learned defensive reactions are thought to function selectively via one of the following mechanisms: by forming a new association that inhibits threat memory expression (new learning), or by permanently disrupting threat memory encoding (unlearning).
- The dichotomy of new learning and unlearning provides a theoretical explanation for return of fear after some protocols, but not others. This is highly relevant to alleviating relapse after clinical therapy for fear-related disorders.
- Evidence for discrete neural mechanisms corresponding to new learning and unlearning remains weak: defensive behavioral responses, synaptic plasticity, and associated biochemical activity fail to clearly demarcate which process has been engaged. We therefore consider the possibility of a “mixed mechanism” and its potential implications.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

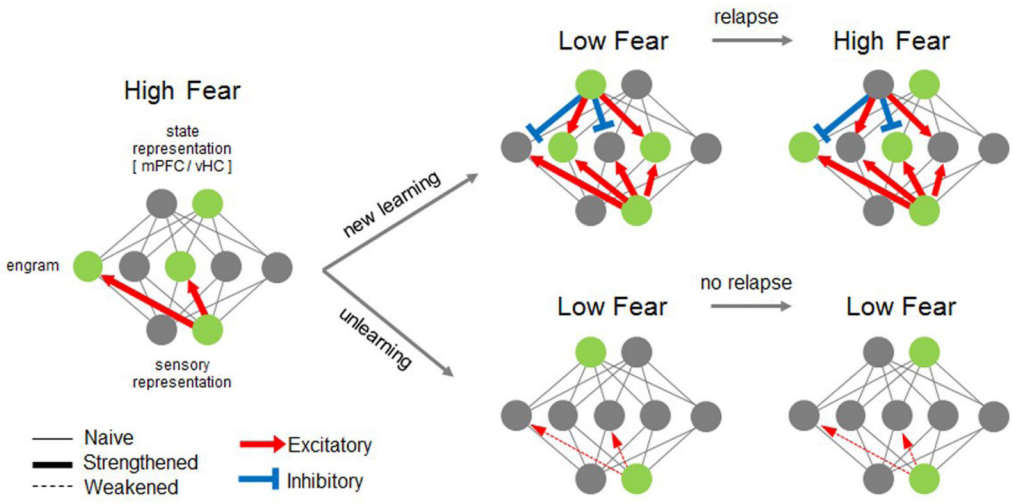

Figure 1.

The conventional view of discrete plasticity processes leading to temporary and persistent threat memory attenuation via new learning and unlearning. The threat memory engram is depicted as a set of active nodes (green) that correspond to single neurons or ensembles with potentiated CS-evoked firing. Two types of inputs contribute to engram cell firing that are modulated by the following neural representations: sensory regions encoding basic attributes of the CS (sensory representation) and multimodal regions that represent the exteroceptive and interoceptive state of the animal (state representation) [87]. The exteroceptive state at the time of stimulus presentation includes the spatial and temporal context (both external variables), while the internal state is defined as the emotional and cognitive state of the animal (for example, the presence of stress- or drug-induced brain states). In most accounts, both forms of state representation are ascribed to the hippocampus and medial prefrontal cortex [86]. High fear develops when strengthening of excitatory inputs occurs within the sensory pathway. When fear is attenuated through new learning, such as extinction, a new engram is encoded (green nodes) and becomes linked to a new state representation through excitatory plasticity (denoted by thicker lines). Conversely, increased inhibition develops onto the threat-encoding neurons. Consequently, when the CS is encountered in this new configuration, defensive responses diminish. However, relapse occurs when the animal re-encounters the CS in a different state, such as in the context in which the original learning took place. Because the encoding of both threat- and safety-related ensembles are preserved, alternation between high and low fear can occur without any additional plasticity and is governed according to which state the animal detects at the time of test. In contrast to new learning, unlearning through protocols such as reconsolidation blockade is considered to destabilize and reverse potentiation of engram cell pathways. Resulting erasure of information precludes the return of threat memory, regardless of the state in which the CS is encountered.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

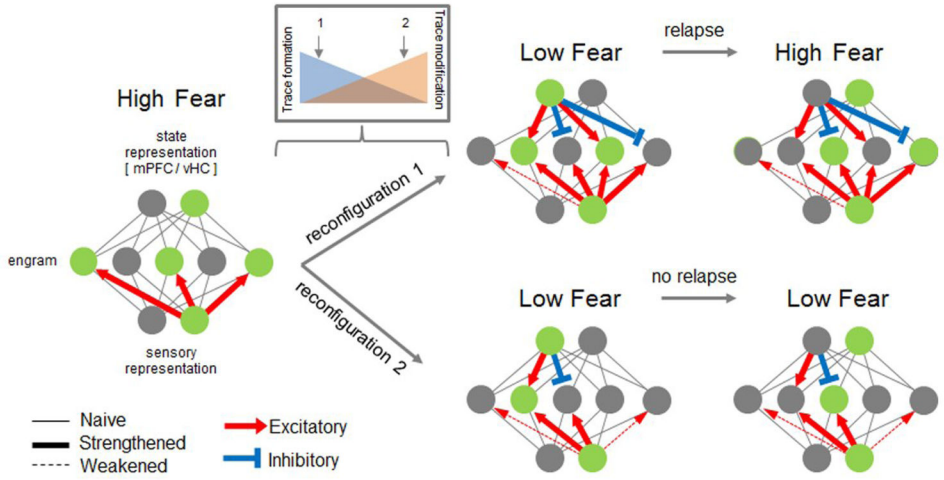

Figure 2.

Mixed mechanism model of divergent outcomes in threat memory attenuation. Inconsistent with the dichotomy of new learning and unlearning, some reports indicate that cellular trace modifications such as synaptic depotentiation (dashed arrows) can occur during the transition from high to low fear under many conditions, including extinction [21, 23, 24, 46, 51, 52]. One explanation for differences in susceptibility to relapse might involve the relative balance between new learning and unlearning. A predominance of the new trace formation (reconfiguration 1) could temporarily render a threat response attenuated while preserving an adequate substrate for threat memory reactivation (green nodes). In contrast, a predominance of original trace modification during attenuation (reconfiguration 2) could prevent relapse by leaving an insufficient residual trace of threat conditioning. While the outcomes depicted relate to a system with only two exteroceptive/interoceptive states, it is possible that additional state-specificity of relapse could be determined by the joint configuration of newly-formed and modified traces.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

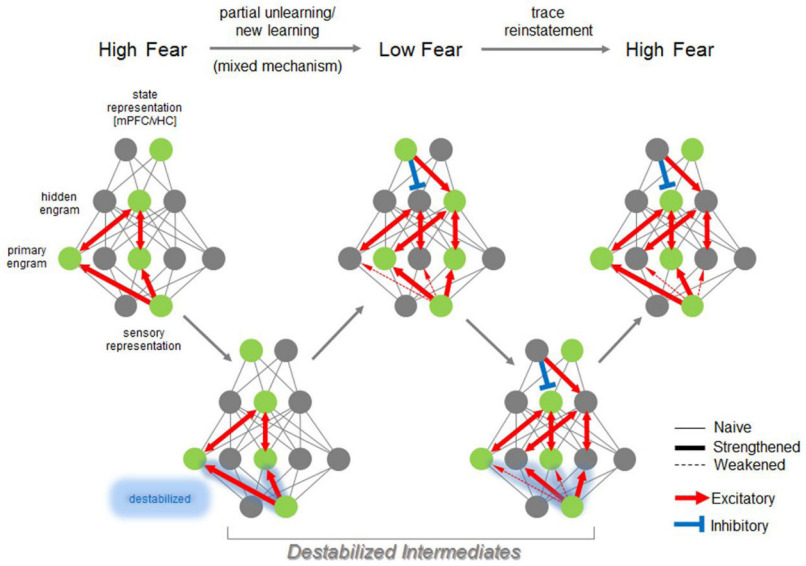

Figure 3.

Model for reinstatement of a weak memory trace following reactivation by a hidden engram. Despite the reversal of threat-related plasticity, fear sometimes paradoxically returns and this can be accompanied by reinstatement of its synaptic substrates [24, 52, 53]. The conditions under which this relapse is observed requires that some residual trace of previous learning is retained in the brain, since animals were not re-exposed to the US and therefore had no opportunity to relearn the threat association. One possibility is a hidden engram (depicted as a nested layer) where such a trace can persist in an inactive state, while retaining strengthened connections to other CS-responsive engram cells, here termed the primary engram. The hidden engram might also receive CS input, but for the purpose of simplicity this has not been depicted. As in the conventional model of relapse (Fig. 1), the return of high fear is elicited by the detection of a new intero/exteroceptive state that has not been associated with fear attenuation. However, in contrast to unlearning or new learning, the recruitment of a hidden engram is sufficient, together with CS inputs, to reactivate primary engram cells due to the retention of synaptic strengthening in their connectivity with the hidden engram (bidirectional arrows). This reactivation might destabilize CS connections and permit their repotentiation through Hebbian synaptic strengthening or reconsolidation-based mechanisms.