Understanding the dynamic and destiny of memories

Lucas de Oliveira Alvares, Fabricio H. Do-Monte

ABSTRACT

Memory formation enables the retention of life experiences overtime. Based on previously acquired information, organisms can anticipate future events and adjust their behaviors to maximize survival. However, in an ever-changing environment, a memory needs to be malleable to maintain its relevance. In fact, substantial evidence suggests that a consolidated memory can become labile and susceptible to modifications after being reactivated, a process termed reconsolidation. When an extinction process takes place, a memory can also be temporarily inhibited by a second memory that carries information with opposite meaning. In addition, a memory can fade and lose its significance in a process known as forgetting. Thus, following retrieval, new life experiences can be integrated with the original memory trace to maintain its predictive value. In this review, we explore the determining factors that regulate the fate of a memory after its reactivation. We focus on three post-retrieval memory destinies (reconsolidation, extinction, and forgetting) and discuss recent rodent studies investigating the biological functions and neural mechanisms underlying each of these processes.

1. Introduction

Acquiring memories of our life experiences is vital to optimize our future decisions. Based on our memories we can adjust our behavior in response to stimuli that resemble prior experienced situations. For example, returning to your favorite restaurant or avoiding a bumpy road on the way home are situations in which our memories from past events directly guide our decisions. But how does our brain select and store the information that is necessary to appropriately shape our behaviors? Although we are constantly exposed to new environmental stimuli, most of the acquired information is lost in a few hours or days. Our brain has specialized mechanisms that act as filters to select which information should be retained as long-term memory. Because it is difficult to choose in real time (i.e., during learning) which piece of information is important to be kept, this filtering process occurs a posteriori, shortly after the experience has been acquired. There is a consensus that memories are unstable immediately after learning and can be easily modulated until they are consolidated into a more stable form. The time-dependent process of memory stabilization starts with an initial phase known as synaptic consolidation, which lasts for a few hours and allows for the encoded information to be stored as a memory trace. This phase is followed by an enduring phase named systems consolidation, which persists for days to weeks and is believed to be important for the maintenance and organization of different types of memories in distinct brain regions. While the ordinary process of memory stabilization has the biological function of selecting the memories that are potentially beneficial to be maintained in our brain, how it occurs and what sort of information is more prone to be retained as a long-lasting memory trace has been a major object of research in the last years.

Studies in both humans and laboratory animals have shown that memories containing emotional information are better retained than memories with neutral or trivial content. The underlying mechanisms of emotion seem to be a good candidate to modulate our memory filtering system by favoring the consolidation of emotionally-charged information. But, do memories remain in a steady form once this consolidation process is completed? Because we live in an ever-changing environment, an inflexible memory system would not be adaptive to changes and,
consequently, would not serve to properly guide our future behaviors. Instead, a system that permits our memories to be constantly updated in accordance with our needs seems to be evolutionarily more advantageous (Lee et al., 2017). Returning to our initial examples, a flexible memory system enables us to stop going to that favorite restaurant if the food doesn’t taste good anymore or resume taking that bumpy road on the way home after it gets repaved. Accumulating evidence from the last few decades suggests that, even after being consolidated, our memories can enter a plastic state that allows the incorporation of newly acquired information (Nader et al., 2000b; Wang et al., 2009). This updating process can be initiated when memories are retrieved, making them labile and susceptible to modifications. This post-retrieval mechanism seems to be critical to determine the destiny of our memories, as well as to maintain the efficiency of our memories in predicting future events (De Oliveira Alvares et al., 2013; Lee, 2009).

In this review, we explore the biological mechanisms that regulate the distinct fates of memories after retrieval. We focus mainly on studies using fear and reward conditioning paradigms in rodents, as they represent most of the memory literature that is currently available. First, we describe situations in which the environmental conditions during the retrieval phase resemble the original experience, in this way leading to a memory updating process called reconsolidation. Next, we discuss situations in which the current conditions during retrieval have a different meaning compared to the original experience, thereby causing an inhibition of the original memory trace in a process known as extinction. Last, we discuss situations in which memories fade away following the acquisition phase. This physiological process, often described as forgetting, occurs when a well-established memory is erased or becomes unable to be retrieved.

2. Basic phases of memory formation

Before discussing the post-retrieval destiny of a memory, it is important to revisit the distinct phases of memory. For didactic purposes, we separate memory into three different phases that progress over time. The first phase is called acquisition or encoding and refers to the stage in which the information is initially received and processed in the brain, by changing neuronal excitability and the strength of connections between neurons (i.e., synaptic plasticity). The second phase, called consolidation, is the dynamic process by which the newly encoded information is gradually stabilized and retained as a memory trace. And the third phase, called retrieval, is the phase in which the previously acquired information is accessed and re-experienced by the subject.

Most of our knowledge about the different phases of memory formation comes from animal models of associative learning (Dickinson, 2012; Hawkins and Byrne, 2015; Pearce and Bouton, 2001). Such models have enabled us to systematically distinguish these three memory phases. This differentiation has been made possible by manipulating the activity of specific brain regions or neurotransmission system before each specific phase and using the animal’s response to associated cues as a measure of memory. In contrast, a systematic differentiation between the distinct mechanisms that follow memory retrieval has been more difficult because we have just begun to understand the molecular and temporal processes that separate reconsolidation, extinction, and forgetting, and how each one of these processes contributes to the destiny of a memory (Fig. 1).

3. Memory reconsolidation: mechanisms and biological functions

The mainstream assumption regarding memory formation was based on the dogma that, once consolidated, memories would become stable and protected from further modifications. However, pioneer studies by the middle of the last century started to challenge this view by showing that consolidated memories can undergo modifications after being retrieved (Lewis, 1979; Lewis et al., 1968). Subsequent studies a few decades later demonstrated that reactivation of a well-consolidated memory can trigger cellular events that are critical for memory persistence (Nader et al., 2000b; Przybyslawski and Sara, 1997). The observation that consolidated memories can enter a labile state and become editable upon retrieval has given origin to the term memory reconsolidation. In the following years, hundreds of studies demonstrated that memory reconsolidation is a ubiquitous process that happens in different species – from worms and fishes to birds and mammals including humans – and can be studied by using distinct memory paradigms (for a review see: Besnard et al., 2012; Dudai, 2012; Haubrich

---

**Fig. 1.** Schematic of memory destinies following retrieval. During the initial phase of memory formation, new information is acquired and consolidated as a long-term memory. Subsequent retrieval of the consolidated memory may activate three distinct processes: i) reconsolidation is triggered after a short period of memory reactivation and a mild degree of mismatching between the original memory and the retrieval session. Reconsolidation involves an initial destabilization of the memory trace followed by memory updating through the strengthening of the original memory and the incorporation of new information; ii) extinction is triggered after a long period of memory reactivation and a high degree of mismatching between the original memory and the retrieval session. Extinction involves the formation of a new memory that inhibits the original memory trace; and iii) forgetting results from a deficit in the retrieval of the original memory. Forgetting may serve as a filter to remove unnecessary information.
and Nader, 2018; Taujanskaite et al., 2020).

The discovery that memories can be manipulated by interfering with the mechanisms of reconsolidation has generated great excitement and interest in the field because pathological memories are in the heart of many psychiatric disorders including post-traumatic stress disorder (PTSD), phobias and drug addiction (Beckers and Kindt, 2017a; Elsey et al., 2018). Because of the potential clinical relevance of disrupting reconsolidation processes to modulate emotional memories in humans, reconsolidation rapidly became one of the most studied topics in the memory field (Monfils and Holmes, 2018; Phelps and Hofmann, 2019).

While many studies have focused on investigating reconsolidation processes in laboratory animals as an attempt to better understanding pathological memories in humans, much less attention has been paid on elucidating the biological functions of memory reconsolidation. Here, we discuss the adaptive role that reconsolidation has on memory updating and the main neuronal mechanisms underlying this process.

3.1. Why is memory updating important and how does it happen?

The long-term maintenance of our significant life experiences is an important adaptive process. Based on past experiences, animals can predict future events and optimize their decisions. Previously formed memories can guide animals to repeat behaviors that have resulted in successful outcomes while avoiding those that have previously failed (Dickinson, 1981, 2012; Rescorla, 1988; Zentall, 2013). However, animals live in a dynamic environment and for this reason memories are normally reactivated in situations that differ from those in which the original experience has occurred. Therefore, new pieces of information should be integrated with the original memory in order to keep its current significance. This incorporation of new information is believed to be mediated by reconsolidation mechanisms (De Oliveira Alvares et al., 2013; Hupbach et al., 2008; Lee, 2010; Zinn et al., 2020). Memory retrieval triggers a plastic state in which the original memory trace can be modified and subsequently reconsolidated into an updated form (Lee et al., 2017).

To be updated in terms of content or strength, a previously consolidated memory that is in a stable state (i.e., inactive state) needs to destabilize and enter a labile state (i.e., active state) (Fig. 2A). This memory destabilization process occurs when there is a certain degree of mismatching (i.e., prediction error) between what is expected and what actually happens during the retrieval period (Popik et al., 2020; Sevenster et al., 2013). Studies in both humans and laboratory animals have shown that in the absence of prediction error, retrieval does not induce destabilization and the original memory remains unmodified, reinforcing the idea that retrieval and memory destabilization are two different processes (Barreiro et al., 2013; Bustos et al., 2009; Exton-McGuinness et al., 2015; Merlo et al., 2013; Pedreira et al., 2004; Sevenster et al., 2012; Sinclair and Barense, 2018). These experimental observations make sense if we consider the biological role of memory updating because, when the current situation is closely identical to the original experience, there is no need of updating the memory. Using our previous example again, if during a subsequent visit to that same savory restaurant the food continues as tasteful as before, our memory would not be destabilized and our positive judgment about the restaurant should be integrated with the original memory in order to keep its current significance. This incorporation of new information is believed to be mediated by reconsolidation mechanisms (De Oliveira Alvares et al., 2013; Hupbach et al., 2008; Lee, 2010; Zinn et al., 2020). Memory retrieval triggers a plastic state in which the original memory trace can be modified and subsequently reconsolidated into an updated form (Lee et al., 2017).

Fig. 2. Representative model and synaptic mechanisms of retrieval-induced memory destabilization. A) During memory retrieval, a consolidated memory that is initially in a stable state (inactive state) enters a labile state (active state) if a certain degree of mismatching occurs (i.e., prediction error). Depending on what happens during retrieval, the original memory trace can be updated into distinct forms thereby resulting in a strengthened, weakened, or modified memory. House drawings are used as an analogy for retrieval-induced memory updating. B) The transition of the original memory from an inactive to an active state after retrieval involves a series of postsynaptic mechanisms including: 1) Entrance of calcium following activation of GluN2B-containing NMDA receptor and i-type voltage-gated calcium channels (L-VGCC); 2) Activation of protein kinase CaMKII; 3) Activation of protein complexes (e.g., proteasome); 4) Degradation of scaffolding proteins (e.g., SHANK); 5) Endocytosis of calcium impermeable GluA2-containing AMPA receptors from the postsynaptic density (PSD) followed by autophagy; 6) Insertion of calcium permeable GluA1-containing AMPA receptors in the PSD in part due to the phosphorylation of serine 845 (S845) by protein kinase A (PKA).
would continue the same.

3.2. What are the molecular/cellular processes involved in memory destabilization/updating?

Several important players have been described to participate in the memory destabilization process (Fig. 2B). For instance, the entrance of calcium ions into the cell through the activation of GluN2B-containing NMDA glutamatergic receptors and L-type voltage-gated calcium channels (L-VGCCs) have been shown to be crucial steps for this process. GluN2B-containing NMDA receptors are glutamate gated ion channels that are centrally involved in neuronal signal transduction and synaptic plasticity (Zhang and Luo, 2013), whereas L-VGCCs are transmembrane calcium channels that play a critical role in neurotransmitter release and gene expression regulation (Berger and Bartisch, 2014; Gomez-Ospina et al., 2006). Systemic or intracerebral injection of the GluN2BR antagonist ifenprodil or the L-VGCC blocker nimodipine before memory reactivation prevents memory updating (Ben Mamou et al., 2006; Haubrich et al., 2015; Suzuki et al., 2008). Studies have shown that calcium influx through GluA1-containing AMPA receptor - a subtype of AMPA-type glutamate receptors that is permeable to calcium - is also required for memory updating (Clem and Huganir, 2010; Torquatto et al., 2019). Indeed, during memory labilization, GluA1-containing AMPA receptors are inserted into the postsynaptic density, whereas GluA2-containing AMPA receptors, which are known for rendering the channel impermeable to calcium, are in turn endocytosed (Clem and Huganir, 2010; Hong et al., 2013; Rao-Ruiz et al., 2011b) and degraded through autophagy (Shehata et al., 2018). The postsynaptic density is a specialized region in the membrane of excitatory synapses that is rich in receptors, proteins and signaling molecules, and is known for playing a critical role in synaptic plasticity (Gold, 2012). Increased influx of calcium activates the protein kinase CaMKII, which triggers proteasomes (i.e., large protein complexes) that are responsible for the degradation of synaptic scaffolding proteins (e.g., Shank), an essential step for memory destabilization. (Jarome et al., 2016, 2011; Lee et al., 2008). Accordingly, inhibiting either CaMKII or proteasomes prevent memory destabilization (Jarome et al., 2016; Lee et al., 2008); and the coordinated regulation of protein synthesis and degradation seems to be a critical process for long-term memory maintenance (Park and Kaang, 2019). Other important players that participate in the process of memory destabilization include dopamine (Merlo et al., 2015), cannabinoid receptors (Lunardi et al., 2020), cholinergic muscarinic receptors (Stiver et al., 2015), and NMDA-induced autophagy (Shehata et al., 2018).

Once memory achieves a labile state, it becomes vulnerable to modification (e.g., enhancement, attenuation, or integration of new information) for a few hours, after which memory is reconsolidated and restabilized. This restabilization process, which is necessary to place memory back into a more steady form, involves RNA synthesis (Sangha et al., 2003), protein synthesis (Nader et al., 2000a), AMPA receptor trafficking (Rao-Ruiz et al., 2011b), and cytoskeleton reorganization (Lunardi et al., 2018). Evidence suggests that the exchange of calcium-impermeable to calcium-permeable AMPA receptors is a hallmark of memory malleability. When high concentrations of calcium-permeable AMPA receptors are located in the postsynaptic density, memory becomes destabilized. Replacement of calcium-permeable by calcium-impermeable AMPA receptors shifts memory back to a stable form (Clem and Huganir, 2010; Torquatto et al., 2019). This switch in memory malleability is accompanied by structural changes mediated by actin remodeling, which results in alterations in the number and morphology of dendritic spines (Roy et al., 2017). Accordingly, actin remodeling in dendritic spines has been described following distinct forms of synaptic plasticity including long-term potentiation (LTP) and long-term depression ( Fonseca, 2012; Ramachandran and Frey, 2009; Szabo et al., 2016). We propose that together with the AMPA receptor trafficking, there is an orchestrated reorganization of actin in which mature dendritic spines become structurally labile through actin depolymerization. Thus, dendritic spines can be rebuilt in an updated form depending on what happens after retrieval: becoming larger in the case of memory enhancement or smaller in the case of memory impairment. Future experiments using two-photon microscopy will test this prediction by longitudinally tracking individual dendritic spines during the time window of memory reconsolidation.

Previous studies have demonstrated that memories with intense emotional content are more resistant to destabilization, and consequently, less susceptible to interference (Bustos et al., 2011; Suzuki et al., 2004; Wang et al., 2009). In short, memory strength is a boundary condition that limits memory from undergoing reconsolidation. For example, during a strong fear conditioning training, activation of locus coeruleus neurons increase the release of noradrenaline in the amygdala, thereby reinforcing the memory trace (Haubrich et al., 2020). Subsequent to retrieval, activation of beta-adrenergic receptors in the amygdala halts AMPA trafficking and down regulates GluN2B-NMDAR, which prevents memory restabilization (Haubrich et al., 2020; Wang et al., 2009). Thus, strong emotional experiences seem to be more resistant to undergoing reconsolidation and as a result less vulnerable to modifications. Nevertheless, studies in rodents and humans have demonstrated that even emotionally-charged memories can be modified through reconsolidation mechanisms under certain conditions that involve the correct length of reactivation, the ideal memory mismatching, and/or the use of drugs to enhance memory labilization (Elsey and Kindt, 2017; Soeter and Kindt, 2012; Zhang et al., 2018a).

3.3. How can memory be updated through reconsolidation mechanisms?

As we have discussed above, memory updating is a critical process that helps to maintain the predictive value of previously acquired information and consequently to avoid inappropriate behavioral responses. Such updating is mediated by reconsolidation mechanisms and may serve as an opportunity for modifying the strength and the content of memories. Memory updating is observed in many species and may have evolved in some organisms to assure that information about resources (e.g. food, water, and mate) and potential threats (e.g., risk of predation) are always up to date. In this section, we explore how reconsolidation can lead to distinct forms of memory updating including: i) the modulation of memory strength, ii) the modification of the original memory content, iii) the formation of false representations, iv) the incorporation of state-dependency into the memory trace, and finally v) the shift of memory valence through counterconditioning.

3.3.1. Memory strength

The first direct evidence that memories can be strengthened through reconsolidation mechanisms was the demonstration that a weak learning event can be reinforced by additional training until it achieves an asymptotic level (Fukushima et al., 2014; Lee, 2008, 2009). In one of these studies, laboratory rats were fear conditioned with foot shocks to achieve a freezing level of around 40 % (freezing was used as a measure of memory retrieval). When the same animals were conditioned again, freezing levels went up to ~70 %. However, different molecular cascades were necessary during these two phases: knocking down the expression of brain-derived neurotrophic factor (BDNF, a critical protein for memory acquisition) before the initial conditioning phase disrupted memory formation, whereas knocking down the expression of zif268 (a critical protein for memory reconsolidation (but see Zalcman et al., 2015)) before the second conditioning session impaired memory strengthening; suggesting that reconsolidation rather than new learning is the main mechanism underlying memory strengthening (Lee, 2008). Other studies using similar experimental designs have drawn the same conclusion by demonstrating that reactivation-induced reconsolidation strengthens the original memory trace (De Oliveira Alvare et al., 2013; Forcato et al., 2014; Inda et al., 2011). Together, these studies indicate that retrieval-induced memory destabilization is an essential step in the strengthening of previously acquired memories, most likely through...
3.3.2. Memory content

In addition to modifying memory strength, another possible function of reconsolidation is to modify memory content. That is, to permit the integration of new information into an existing memory trace during the malleable period that succeeds retrieval (Fernandez et al., 2016; Morris et al., 2006; Winters et al., 2009). For instance, prior studies have shown that rats that were fear conditioned in one context (e.g., context A) showed low levels of freezing when tested in a distinct context (e.g., context B). However, when the conditioned memory was reactivated in a hybrid context that resembled both context A and context B, animals started to exhibit high levels of freezing when tested in context B, suggesting that during the reactivation session the information from context B was inserted into the memory trace previously associated with context A (De Oliveira Alves et al., 2013; Zinn et al., 2020).

3.3.3. False memory

The brain’s capacity to modify memories opens up the possibility of creating a false memory, a process that consists in either changing the representation of previous events by implanting false information upon an existing memory trace or making up a completely inaccurate memory trace by incorporating exogenous misinformation (Loftus and Davis, 2006). The phenomenon of false memory takes place when the information inserted into an existing memory trace is altered until the point that the updated form no longer fit into the original memory (Schacter et al., 2011). This distorted or even fabricated recollection of events is believed to be a side effect of memory malleability that occurs during the time window of reconsolidation.

Memories are known to be formed following the association between biologically salient events and sensory stimuli. Nevertheless, a recent study has demonstrated that mice can create a fully artificial memory without being exposed to a sensory experience (Vetere et al., 2019). In this study, optogenetic activation of olfactory sensory neurons expressing the odorant receptor for acetylcholine was paired with optogenetic activation of lateral habenula projections to the ventral tegmental area (VTA), a circuit known to mediate aversive responses. In a separate group of mice, optogenetic activation of the same olfactory sensory neurons was paired with optogenetic activation of laterodorsal tegmental nucleus projections to the VTA, a circuit known to mediate rewarding responses. After pairings, mice receiving optogenetic activation of the aversive or rewarding circuits spent less or more time in the side of a chamber containing the acetylcholine odorant, respectively, despite the animals never having smelled the acetylcholine odorant before (Vetere et al., 2019). These findings suggest that memories can be artificially formed by entirely bypassing a sensory experience. Although this study demonstrates the genesis of an artificial memory through the modulation of acquisition/reconsolidation mechanisms, a similar process involving the insertion of false information into a previously existing memory trace has been described in other studies (Ramírez et al., 2013; Redondo et al., 2014).

Previous studies using activity-dependent neural tagging have shown that manipulation of neuronal populations that were active during learning can enable previously acquired experiences to become associated with new stimuli (Ramírez et al., 2013; Redondo et al., 2014). In one of these studies, mice hippocampal neurons that were activated by exposure to a neutral context (context A) were subsequently labeled with the light-sensitive cation channel channelrhodopsin (ChR2) (Ramírez et al., 2013). Animals were then exposed to a different context (context B) and the neurons labeled in context A were optogenetically reactivated during a fear conditioning training with foot shocks. In the following day, animals displayed increased freezing responses in context A, even though they had never experienced a foot shock in that context, suggesting that the encoded information about context A was updated and combined with the foot shock information (Ramírez et al., 2013).

3.3.4. State-dependent memory

Besides the inclusion of external cues, reconsolidation also enables the inclusion of internal cues into the mental representation of the environment, which may result in memory updating. Examples of internal cues include the distinct mental states that are experienced by the subject during the retrieval session (e.g., specific mood states, particular arousal situations, or the influence of drugs). In other words, the mental state prevailing during memory reactivation can be incorporated into the neural substrate of the memory (henceforth called engram), thereby making the original memory state-dependent. Afterwards, subsequent memory retrieval will be better achieved if the same internal state is presented again (Gisquet-Verrier et al., 2015; Sierra et al., 2013). A similar state-dependent induction has been also reported during extinction training (Rosa et al., 2014), suggesting that state-dependency is a general memory process that may be established when the memory is in a labile/plastic state (e.g., after acquisition or reactivation), hence incorporating the internal state into the memory engram (for a review see: Gisquet-Verrier and Riccio, 2018).

3.3.5. Counterconditioning

Another way to update a memory during the labile state induced by retrieval is through counterconditioning, a process in which the original memory is modified by the integration of new information containing opposite emotional valence. During counterconditioning, a fearful memory can be updated by presenting hedonic information during the period of reactivation, so that the original negative valence is “replaced” by a positive/less aversive one. This emotional remodeling has been demonstrated in a series of studies in which a conditioned stimulus (CS) previously associated with a foot shock became less aversive after being re-associated with an appetitive stimulus such as chocolate (Haubrich et al., 2015), caffeine (Pedraza et al., 2018), methylphenidate (Arellano Perez et al., 2020) or amphetamine (Toledano and Gisquet-Verrier, 2014). The observation that counterconditioning updating is prevented by the infusion of memory destabilization inhibitors (e.g., GluN2B-NMDA antagonists or L-VGCC blockers) suggests that memory updating during counterconditioning is mediated by reconsolidation mechanisms rather than new learning (Haubrich et al., 2015).

While significant progress has been made in elucidating the neural mechanisms involved in memory updating, there are still several open questions that need to be addressed to clarify how additional information is inserted into an existing memory trace. Are new neurons incorporated into the engram? Are neurons that do not carry relevant information anymore removed from the engram? Which patterns of neuronal activity regulate these processes? Future experiments using activity-dependent neuronal tagging or calcium imaging recordings of neuronal dynamics during memory acquisition vs. reconsolidation are likely to yield more information about how a memory engram activity is modified following reconsolidation; and by what means the insertion or removal of cells from the engram correlates with behavioral changes. Another important question regarding memory reconsolidation is the possibility of translating basic research findings to clinical practice. While some studies have successfully manipulated reconsolidation mechanisms in human patients with the purpose of attenuating emotional memories, others have failed (for a review see: Beckers and Kindt, 2017b; Elsey et al., 2018). This inconsistence of results may be explained by differences in the boundary conditions (e.g., the strength and age of the memories) or in the degrees of prediction error observed in each of these studies, which are important limiting factors for memory to undergo destabilization/reconsolidation (Elsey and Kindt, 2017; Sevenster et al., 2012; Sinclair and Barry, 2018, 2019). Further studies focused on investigating the mechanisms that govern these limiting factors in laboratory animals may provide better insights to elucidate memory destabilization/reconsolidation in humans.
4. Memory extinction: mechanisms and biological functions

In the previous section, we described how memories can undergo a reconsolidation phase after being retrieved, a process by which the encoded information can be updated and re-established. Here, we discuss a different process called extinction, which takes place to attenuate/inhibit the original memory (Bouton, 1993; Pavlov, 1927). Extinction mechanisms are triggered when the retrieval phase surpasses a critical period, and the presence of a previously learned CS does not predict the same outcome anymore (i.e., high prediction error) (Merlo et al., 2014; Salinas-Hernandez et al., 2018) (Box 1). In such cases, multiple exposures to the same CS in the absence of consequences (i.e., unconditioned stimulus) lead to an attenuation of the conditioned response (Quirk and Mueller, 2008). Although the extinction phenomenon has resisted a brief and simple explanation, there is a consensus that extinction is not the same as forgetting or habituation, where habituation is defined as a response decrement to a repeated stimulus (Furlong et al., 2016; Thompson and Spencer, 1966). Extinction can also not be explained by synaptic depotentiation (Hong et al., 2009; Kim et al., 2007), as synaptic inputs that were facilitated during memory acquisition remain potentiated following extinction training (Clem and Huganir, 2010; Kim and Cho, 2017; but see: Park and Choi, 2010). Instead, increasing evidence supports the idea that extinction is a new learning that competes with the initial association (for a review see: Bouton, 2004; Dunsmoor et al., 2015; Luchkina and Bolshakov, 2019; Maren and Quirk, 2004). For example, an extinguished memory can resurface with the passage of time after the extinction training (spontaneous recovery) (Rescorla, 2004), as well as when the subject is tested in a context other than that in which the extinction training occurred (renewal) (Bouton et al., 2006; Goode and Maren, 2014). In addition, an extinguished memory may also reappear when the subject is presented to a simple reminder of the original memory after the extinction training is completed (reinstatement) (Bouton and Bolles, 1979). While many features about the extinction process have been unveiled recently, many basic questions still remain unsolved. For example, why does extinction memory weaken with time? Why is extinction memory context-dependent? How does the single presentation of an unconditioned stimulus retrigger a conditioned response?

Despite the temporary nature of extinction (Vervliet et al., 2013), cognitive behavioral therapies based on extinction learning (e.g., exposure therapy) continue to be considered the gold standard intervention to suppress the expression of emotional memories in humans (e.g., anxiety disorders, substance abuse) (Foa and McLean, 2016; Hollandt et al., 2020; Kieter and Dinter, 2013). The identification of potential cognitive enhancers (also called nootropic drugs) to improve extinction learning, and consequently attenuate fear or substance use disorders, has been the focus of many studies in the last decades (for a review see: Carpenter et al., 2019; Singewald et al., 2015). Additional insights on understanding the mechanisms of extinction have emerged from recent studies centered on investigating both the molecular/cellular processes and the neural circuits underlying extinction learning. Below we describe the main molecular/cellular mechanisms underlying extinction and explore the central determining factors that control the switch between reconsolidation and extinction processes following memory retrieval. Next, we discuss some rodent studies focused on elucidating the neural circuits that regulate the extinction of fear- or reward-associated memories. We draw some parallels between classical studies using lesions/pharmacological approaches and recent studies using optogenetic/chemogenetic tools to investigate extinction circuitry.

---

**Box 1**

The “Retrieval Judge” of memory destinies

To better understand how the destiny of a consolidated memory is determined, here we use a metaphor to compare the process of memory retrieval with a judicial trial court. In this analogy, the judge (memory system) evaluates many aspects of the defendant’s sentence (original memory) and compares it with the new evidence presented during the appeal (retrieval session) in order to offer a new verdict (memory destiny). If a minor degree of mismatching between the new proofs and the original evidence is observed during the appeal (lack of memory destabilization), the judge will resentence the defendant to the same penal conviction (memory persists unaltered). However, when the new pieces of evidence (cues) presented during the appeal differ significantly from the original evidence (wild prediction error), bigger are the chances that the judge will adjust the sentence (memory destabilization) by extending, reducing or modifying the verdict (reconsolidation with enhancement, attenuation, or modification of the memory content). In other cases, when the defense attorney presents new evidence that strongly differ from the original evidence (high prediction error), the judge will temporarily acquit the defendant and a completely different verdict will be pronounced (extinction memory). Finally, if the defendant’s request for an appeal is not sufficient to persuade the court (lack of retrieval), the appeal will be denied and the case closed (forgetting).
4.1. What are the molecular/cellular processes involved in memory extinction?

A large number of molecular and cellular processes are required for extinction memory formation, many of which resemble those required for memory reconsolidation. For instance, the acquisition of both extinction and reconsolidation requires protein synthesis (Milekic and Alberini, 2002; Suzuki et al., 2004) and the activation of several systems of neurotransmission and intracellular messengers in distinct brain areas including the amygdala, the hippocampus, and the medial prefrontal cortex (mPFC) (Baldi and Bucherelli, 2015; Kida, 2019; Lee et al., 2006; Merlo et al., 2014; Wideman et al., 2018). It has been shown that glutamate plays a critical role in extinction learning through the activation of AMPA (Yamada et al., 2009), NMDA (Kwapis et al., 2014), and mGlurR1 receptors (Simonyi et al., 2007). The inhibitory neurotransmitter GABA is also important for extinction formation. Systemic or intracerebral infusions of GABA-A receptor agonists impairs extinction memory (Hart et al., 2009; Laurent and Westbrook, 2009), whereas infusions of GABA-A receptor antagonists facilitates it (Berlau and McGaugh, 2006). In addition, catecholamines such as noradrenaline are important modulators of extinction processes. Augmented noradrenergic transmission has been shown to potentiate the extinction of fear memories (Berlau and McGaugh, 2006; Mueller et al., 2008; Uematsu et al., 2017), whereas the blockade of either alpha-1- or beta-adrenergic receptors has been associated with extinction learning impairment (Bernardi and Lattal, 2010; Cain et al., 2004; Do-Monte et al., 2010a, b; Fitzgerald et al., 2015; Mueller et al., 2008).

Moreover, the endocannabinoid system plays a critical role in fear extinction with increased levels of endocannabinoids being detected following extinction training (Marsicano et al., 2002). Activation of endocannabinoid receptors type 1 (CB1) is indispensable for fear extinction formation as CB1 knockout mice or animals treated with either systemic or intracerebral infusion of CB1 antagonists show a robust impairment in extinction memory (Lisboa et al., 2019; Lutz, 2007; Marsicano et al., 2002; Sachser et al., 2015; Varvel et al., 2005), whereas animals treated with CB1 receptor agonists or inhibitors of endocannabinoid uptake/metabolism show extinction facilitation (Bisby et al., 2020; Bitencourt et al., 2008; de Oliveira Alves et al., 2008; Do Monte et al., 2013; Gunduz-Cinar et al., 2013; Morena et al., 2018; Segev et al., 2018). In contrast to fear extinction, the extinction of drug-associated memories is rather facilitated by the infusion of antagonists/inverse agonists of CB1 receptors (Colombo et al., 2004; Gessa et al., 2005; Hu et al., 2015; Khaleghzadeh-Ahangar and Haghparast, 2015), suggesting that CB1 receptors may have different effects on extinction processes depending on the emotional valence of the original memory. The activation of the abovementioned receptors and others not described here triggers intracellular signaling processes that play an important role in extinction memory formation. For example, calcium influx through NMDA receptors activates the phosphatase calcineurin (Lieberman and Mody, 1994), and its pharmacological or genetic inhibition impairs extinction memory (de la Fuente et al., 2011; Hakeves et al., 2008). Equally implicated in extinction memory formation are a number of protein kinases including ERK1/2 (Herry et al., 2006; Merlo et al., 2018), MAPK (Lu et al., 2001), PI3-K (Kritman and Maroun, 2013), and CaMKII (Szapiro et al., 2003), as well as a series of intracellular pathways that control remodeling of dendritic spines and structural plasticity (Lai et al., 2012; Sananbenesi et al., 2007). For further understanding about the molecular and cellular mechanisms underlying extinction formation, we refer the readers to some comprehensive reviews that have been recently published in this topic (Baldi and Bucherelli, 2015; Pagani and Merlo, 2019).

4.2. How does retrieval trigger extinction?

As described above, retrieval-induced memory destabilization can drive memory in two opposite directions: reconsolidation or extinction. In the former case, memory strength or its content may be updated. In the latter, a new inhibitory memory is created to compete with the original one. Experimentally, the only procedural difference that will determine whether a memory will undergo reconsolidation or extinction processes is the duration of the memory reactivation session. Several studies have shown that increasing the duration of the retrieval session interrupts reconsolidation and activates extinction formation. Whereas a brief exposure to the CS leads to memory reconsolidation, longer exposures trigger mechanisms of extinction (Bustos et al., 2009; Pedreira and Maldonado, 2003; Suzuki et al., 2004), suggesting that the fate of the retrieved memory depends on the length of the retrieval session. However, the specific temporal boundaries and the molecular mechanisms that determine the transition between these two processes remain largely unsolved. The first insight into this topic came from a mice study showing that the post-retrieval course of memory depends on a switch of transcription factors in the hippocampus (de la Fuente et al., 2011). In this study, the authors demonstrated that nuclear factor Kappa B (NFκB) is required for reconsolidation but, during the transition to extinction, calcineurin phosphatase inhibits NFκB expression thereby facilitating extinction memory formation.

Subsequent studies using a fear conditioning paradigm in rats have proposed that reconsolidation and extinction are mutually exclusive processes separated by a transition state in which neither process is recruited (Merlo et al., 2018, 2014). In the first study, the authors demonstrate that manipulating the activity of NMDA glutamate receptors following a reconsolidation (1 CS) or an extinction (10 CSs) protocol affected the original fear memory, but the same manipulation following a transitional state (4 CSs) had not effect (Merlo et al., 2014). Similarly, the second study showed that exposure to either 1 CS or 10 CSs during the retrieval session resulted in increased expression of ERK1/2 in the amygdala, but exposure to an intermediate retrieval session of 4 CSs failed to activate ERK1/2 in this same region (Merlo et al., 2018). Together these findings suggest that, at some point between the first and the fourth CS presentation, the activity of NMDA receptors and the expression of ERK1/2 are arrested to terminate the labilization/reconsolidation period. As the retrieval session progressed and more CSs were presented to the animals (7–10 CSs), both NMDA receptor activation and a second wave of ERK1/2 expression were reestablished and extinction mechanisms were recruited (Merlo et al., 2018, 2014). Increasing the number of CS presentations during the retrieval session also augmented the expression of calcineurin in the amygdala; and blocking calcineurin expression following the extinction protocol, but not following the reconsolidation protocol, impaired extinction formation, suggesting that reconsolidation and extinction recruit mutually exclusive processes (de la Fuente et al., 2011; Merlo et al., 2014). It is worth mentioning that this transition state between reconsolidation and extinction processes has been also demonstrated in other paradigms such as contextual fear conditioning in rodents (Cassini et al., 2017) and differential fear conditioning in humans (Segever et al., 2014).

Inspired by the observation that pharmacological manipulations during the reconsolidation window affect the updating of a destabilized memory trace (Alberini and Ledoux, 2013; Meir Drexler and Wolf, 2017; Nader, 2015; Otis et al., 2015), some laboratories have investigated the effects of post-retrieval extinction training on the persistence of the reactivated memory. In these studies, fear conditioned rodents that were exposed to extinction training session after a single unreinforced CS presentation (i.e., memory reactivation) did not show any return of fear at later retrieval tests, suggesting that the extinction information incorporated during the reconsolidation window was sufficient to permanently attenuate the original memory trace (Graff et al., 2014; Monfils et al., 2009; Pineyro et al., 2013; Ruo-Ruiz et al., 2011a). This reactivation-extinction effect seems to be independent of prediction error or memory destabilization, as behavioral or pharmacological interventions that either reduce prediction error or block memory
4.3. Which neural circuits mediate the extinction of emotional memories?

The recent advent of new techniques to investigate neural circuit function in laboratory animals have led to a significant increase in the number of studies aimed at elucidating the neural circuits of extinction. A better understanding of the neural mechanisms that regulate extinction memory may help to identify: i) genetic markers that are particularly expressed in extinction circuits, ii) distinct patterns of brain activity that are correlated with successful extinction and reduced risk of relapse, and iii) pharmacological targets that may guide the development of new therapies in patients undergoing extinction-based therapies.

It is well accepted that the neural circuit that underlies extinction memory is regulated by a distributed network of brain regions that partially overlap with those recruited during memory acquisition (Fig. 3). Support for this idea comes from several rodent studies demonstrating that manipulation of brain areas that are essential for fear and reward conditioning also interferes with extinction processes. For example, lesions or pharmacological manipulation of the basolateral amygdala (BLA, including both the lateral and the basal subregions) impair not only the acquisition but also the extinction of fear-associated memories (Maren et al., 1996; Miserendino et al., 1990; Sierra-Mercado et al., 2011; Sotres-Bayon et al., 2007; Zimmerman and Maren, 2010). Accordingly, animals exposed to either fear learning or fear extinction sessions show increased expression of the neuronal activity marker cFos in BLA, suggesting that BLA neurons are recruited during both the acquisition and the extinction of conditioned fear memories (Ganella et al., 2018; Holahan and White, 2004; Knapска and Maren, 2009). Additional findings using electrophysiological recordings and optogenetic manipulation of projection-defined neurons in BLA have demonstrated that different subpopulations of BLA cells are recruited during the acquisition and extinction of conditioned fear memories. Whereas a specific subpopulation of BLA neurons show increased CS responses after fear conditioning (fear neurons), a distinct subpopulation of BLA neurons exhibit increased CS responses after fear extinction (extinction neurons) (Herry et al., 2008). These fear and extinction neurons in BLA send dense projections to the prelimbic (PL) and infralimbic (IL) subregions of the mPFC, respectively; and pathway-specific optogenetic activation of BLA-IL projections or inhibition of BLA-PL projections facilitate the acquisition of extinction memory (Senn et al., 2014), suggesting that the balance of activity between these two BLA-mPFC pathways is an important factor during the formation of long-term extinction memories.

In line with the abovementioned findings, lesions and electrophysiological studies have demonstrated that PL activity is correlated with/required for fear memory retrieval (Burgos-Robles et al., 2009; Courin et al., 2014; Dejean et al., 2016; Sierra-Mercado et al., 2011; Sotres-Bayon et al., 2012). In contrast, IL activity is necessary for the formation of fear extinction memory (Bravo-Rivera et al., 2014; Do-Monte et al., 2015a; Laurent and Westbrook, 2009; Santini et al., 2012; Sierra-Mercado et al., 2011). Optogenetic silencing of either BLA projections to PL or the reciprocal connection from PL to BLA attenuates fear memory retrieval (Burgos-Robles et al., 2017; Do-Monte et al., 2015b). In addition, activity in either PL neurons that project to the paraventricular nucleus of the thalamus (PVT) or PVT neurons that project to the central nucleus of the amygdala is critical for the retrieval of well-consolidated fear memories (Do-Monte et al., 2015b; Penzo et al., 2015). On the other hand, fear extinction training increases the excitability of IL-BLA projections (Bloodgood et al., 2018; Cho et al., 2013), and optogenetic or chemogenetic inactivation of IL projections to the amygdala disrupts the formation of fear extinction memory (Adhikari et al., 2015; Bloodgood et al., 2018; Bukalo et al., 2015). While current models have attributed distinct roles for PL and IL in the retrieval and extinction of conditioned fear, a recent study has challenged this anatomical dichotomy view by demonstrating that deep-layer PL glutamatergic neurons that project to IL show increased cFos expression during extinction training, and optogenetically activating or inactivating this specific projection facilitated or impaired fear extinction learning, respectively (Marek et al., 2018). Accounting for the discrepancy between these findings and the previous studies, PL projects to diverse downstream regions and broad manipulation of PL activity could have a range of effects that would cover up the potential role of this region in extinction learning. Future studies using projection-defined and cell-type specific analyses of both PL and IL neurons would help to clarify the precise role of these regions in fear regulation.

Besides its role in fear extinction, BLA neurons have been also implicated in the extinction of reward-associated memories. Increased neuronal activity in a subpopulation of BLA neurons has been correlated with extinction of sucrose-seeking behavior (Tye et al., 2010), and lesions or pharmacological modulation of BLA activity impairs the extinction of appetitive conditioned responses (Balleine et al., 2003; McLaughlin and Floresco, 2007; Portero-Tresserra et al., 2013). BLA neurons send dense projections to the nucleus accumbens (McDonald et al., 2019; Wright et al., 1996), a critical region in the regulation of reward-seeking responses (for a review see: Baldó and Kelley, 2007; Castro and Bruchas, 2019). In coordination with dopamine inputs to nucleus accumbens, BLA projections to nucleus accumbens drive reward-seeking behaviors in response to conditioned cues that predict reward (Ambroggi et al., 2008; Stuber et al., 2011; Wang et al., 2020h). Studies using optogenetics and electrophysiological recordings in vivo have demonstrated that BLA pyramidal neurons respond to both positive and negative valence stimuli; and activity in BLA downstream neurons
Activity in the BLA to nucleus accumbens pathway has been also implicated in the regulation of fear extinction memories, as optogenetic stimulation of this circuit during fear extinction training attenuated the subsequent return of conditioned fear (Correia et al., 2016). Accordingly, a recent study has demonstrated that fear extinction is stored in a genetically distinct subset of BLA neurons (expressing the subunit-1B of the protein phosphatase-1 regulatory inhibitor) which is activated during reward-seeking behaviors, suggesting that fear extinction memory uses the same neuronal ensembles recruited during appetitive memory formation (Zhang et al., 2020).

As previously mentioned, extinction training only reduces the conditioned response to the CS (e.g., fear expression or reward seeking) in the same place where extinction has previously occurred. In other words, extinction is context-dependent and the conditioned response will return (i.e., renewal) when the CS is presented outside the extinction context (e.g., in a novel context or in the original conditioning context) (Goode and Maren, 2014; Podlesnik et al., 2017). The hippocampus has long been involved in the encoding of contextual representations (for a review see: Kubie et al., 2019). Activity in hippocampal neurons is critical to regulate context-evoked fear or reward-seeking responses following extinction. For example, extinction of context fear conditioning results in remapping of place cells in the dorsal hippocampus (Wang et al., 2015); and a series of studies have demonstrated that inhibition of the dorsal or ventral hippocampus interferes with the acquisition and retrieval of extinction for both context-dependent fear-associated memories (Bernier et al., 2017; Corcoran et al., 2005; Corcoran and Maren, 2001; Hobin et al., 2006) and reward-associated memories (Bosser and Stern, 2014; Busse and Schwarting, 2016; Hitchcock and Lattal, 2018). A recent study using activity-dependent neural tagging and optogenetic tools in mice has demonstrated that extinction training suppresses the activity of dorsal hippocampal neurons that were responsive during fear acquisition, and recruits a different ensemble of cells that are both necessary and sufficient for fear extinction retrieval (Lacagnina et al., 2019). This study suggests that a balance between fear and extinction representations in the hippocampus governs the suppression or relapse of fear following extinction.

Other studies focusing on the ventral hippocampus have shown that ventral hippocampal neurons that project to the amygdala or mPFC are activated during the renewal of both fear and reward-associated memories (Anderson and Petrovich, 2018; Jin and Maren, 2015), and pathway specific inactivation of these projections attenuates fear renewal (Vasquez et al., 2019; Xu et al., 2016). Exposure to extinction training increases the expression of BDNF in the ventral hippocampus, and infusing BDNF directly into this region exacerbates the firing rate of IL neurons (Rosas-Vidal et al., 2014, 2018). Consistent with a role of the hippocampus-prefrontal pathway in extinction memory regulation, activity in the ventral hippocampus induces feed-forward inhibition of amygdala-projecting neurons in IL by recruiting local parvalbumin interneurons, and chemogenetic inhibition of ventral hippocampus projections to IL attenuates fear renewal (Marek et al., 2019). Together, these findings emphasize the role of the hippocampus and its reciprocal connections with the mPFC and amygdala in the regulation of extinction memories.

Another important brain region in the acquisition and retrieval of context-dependent extinction memories is the nucleus reuniens, a ventral midline thalamic region that is interconnected with the hippocampus and the mPFC (Hoover and Vertes, 2012; Varela et al., 2014). Inactivation of the nucleus reuniens or its inputs from the mPFC increases conditioned fear responses during both the encoding and retrieval of an extinction memory (Ramanathan et al., 2018; Ram-anathan and Maren, 2019). These findings are in agreement with previous studies suggesting that activity in the nucleus reuniens regulates the specificity of memory attributes for a particular context by processing information from the mPFC en route to the hippocampus (Troyner et al., 2018; Xu and Sudhof, 2013). Understanding the mechanisms regulating the context-dependence of fear extinction has special clinical relevance because attenuating the conditioned response beyond the therapeutic setting is a major challenge during extinction-based therapies.

5. Memory forgetting: mechanisms and biological functions

As described above, memory reconsolidation and extinction processes keep our memories updated and useful for predicting the future. However, the vast majority of our daily experiences are forgotten over time. It is believed that forgetting has several important functions, such as emotional regulation (by removing the negative aspects of some experiences), abstraction and generalization (by extracting the rules/gist from related episodes, and forgetting redundant/noisy information), and cognitive economy (by restricting the information that really matters by removing outdated and useless memories) (Hardt et al., 2013; Norby, 2015; Richards and Frankland, 2017). While reconsolidation and extinction mechanisms serve to modify and update our memories, the physiological process of forgetting acts as a filter to remove some memories and prevent unimportant information from being retained. Hence, a system that balances memory maintenance and forgetting is highly adaptive as it retains pertinent information while forgetting unwanted facts. In this section, we explore the neurobiology of forgetting, which is defined here as a failure to retrieve long-term memories that were easily remembered before.

Although the phenomenon of memory forgetting has been studied since the end of the 19th century (for a review see: Medina, 2018; Sachser et al., 2017; Wixted, 2004), until recently no convincing empirical data have demonstrated the mechanisms involved in memory loss over time. Currently, there are three main hypotheses regarding the nature of memory forgetting: interference, retrieval deficit, and memory decay. Below, we discuss each one of these hypotheses.

5.1. The interference hypothesis

Psychologists have posited that the main forgetting mechanism relies on memory interference, a process in which similar memories content ‘compete’ with each other, thereby promoting the incapacity to retrieve memory properly (Wixted, 2004). During interference, a specific memory interferes with items that were learned before memory acquisition (i.e., proactive interference) or after it (i.e., retroactive interference). For neuroscientists, the process of interference relies on the fact that memories undergo a plastic and labile state after learning or retrieval (e.g., during memory consolidation or reconsolidation, respectively, although these terms have rarely been used by psychologists). In retroactive interference, new learning affects previously acquired experiences. Pioneer studies performed in the beginning of the last century (Müller and Pilzecker, 1900) have shown that subjects exposed to a list of syllables immediately after learning a different list of words showed a significant impairment in the retention of the words (Dewar et al., 2007). However, their memories were not affected when the same list of syllables was presented after the time window of consolidation. These findings are consistent with real life situations in which new pieces of information may interfere with memories that are undergoing consolidation/reconsolidation. In contrast, proactive interference can be explained by the fact that our memories are formed over a background knowledge framework called schemas. These schemas may either facilitate or impair the acquisition of new memories. For example, memories can be consolidated faster and more efficiently if subjects are pre-exposed to a situation that reminds one of the acquisition phase (Pedraza et al., 2017; Tse et al., 2007). On the other hand, if the prior experience has distinct meaning, the subsequent learning will be impaired. This happens during latent inhibition, a phenomenon in which prior presentation of the CS alone impairs the subsequent association of the CS with the unconditioned stimulus by establishing a
different meaning for the CS (Lingawi et al., 2017).

Another example of memory interference is called retrieval-induced forgetting (Anderson et al., 1994). During retrieval-induced forgetting, a set of information is acquired but only a particular category of information is retrieved. This selective retrieval process reduces the brain's capability to remember the related non-retrieved information. Thus, the non-retrieved memory is inhibited and less remembered in subsequent retrieval trials (for a review see: Murayama et al., 2014; Pica et al., 2018). Pharmacological inactivation of either the mPFC or the hippocampus abolishes retrieval-induced forgetting, suggesting that these regions play an essential role in this process (Bekinschtein et al., 2018; Wu et al., 2014). Although retrieval-induced forgetting may be attributed to the interference caused by the competition between reminders associated with distinct memories, it may be also discussed as an active inhibition processes (see additional discussion in the next section below).

Memory interference has been also explained in the lens of non-monotonic plasticity hypothesis, which predicts that memories can change as a function of experience and their neural representation can move apart (differentiate) or together (integrate) according to the degree of reactivation (Ritvo et al., 2019; Sinclair and Barense, 2019). According to this view, when two overlapping memories are strongly reactivated the connections between them will be strengthened, thereby resulting in memory integration. In contrast, if only one of these overlapping memories is strongly reactivated, their connections will be weakened, thereby resulting in memory differentiation (Ritvo et al., 2019). Previous studies have demonstrated that the mechanisms underlying the balance between memory integration and differentiation are believed to be mediated by inhibitory interneurons (Barnes et al., 1990; Rashid et al., 2016).

5.2. The retrieval deficit hypothesis

An open-ended question in the memory field is whether forgetting reflects a failure of memory retrieval or impairment in memory storage. Methodological challenges have prevented researchers from adequately addressing this question because the absence of retrieval does not necessarily indicate that the memory trace is lost (e.g., storage deficit). Instead, it is possible that the reminders used to retrieve the memory are not sufficiently accurate to access the memory engram.

Consistent with the idea that memory forgetting results from a retrieval deficit rather than a complete engram erosion, recent studies have demonstrated that the memory disruptive effects of protein synthesis inhibitors administered immediately after training (consolidation phase) or post reactivation (reconsolidation phase) can be optogenetically rescued by artificially activating the engram corresponding to the original memory (Roy et al., 2017; Ryan et al., 2015). In these studies, optogenetic activation of the engram cells, but not the presentation of conditioned cues, was sufficient to induce memory retrieval. Furthermore, in a transgenic mice model of Alzheimer’s Diseases, optogenetic activation of engram neurons was sufficient to rescue memory deficits (Roy et al., 2016), suggesting that although conditioned cues may be unable to trigger memory retrieval, direct optogenetic activation of the 'silent engram' may restore the dormant memory engram. An alternative explanation for these findings is that both the protein synthesis inhibitor and the Alzheimer’s Disease mouse model used in these two studies were not enough to eliminate the whole engram. Thus, the residual part of the memory trace that was left over was sufficient to be restored/reconstructed by subsequent optogenetic activation. In agreement with this, a recent study has shown that optogenetic activation of the memory engram is unable to rescue a conditioned fear memory when the memory trace is erased by autophagy-induced protein degradation (Abdou et al., 2018), suggesting that the memory recovery capacity relies on the extent of interference over the original memory trace.

Memories that are not frequently accessed tend to be weakened. In contrast, memories that are periodically reactivated keep their precision and strength over time (Alves Lde et al., 2012; Forcato et al., 2013; Inda et al., 2011). It is believed that, except for memories with strong emotional content that are longer maintained, the natural course of our daily life experiences is to be forgotten in a graded form. Accordingly, memories that are not frequently reactivated tend to be more difficult to be rescued (Fukushima et al., 2014; Inda et al., 2011), suggesting that lack of retrieval may activate mechanisms of forgetting. During an initial stage, the forgetting process is a retrieval deficit caused by the partial erosion of the engram. The following stage involves a disconnection of the engram to a point in which it can no longer support the memory content, thereby leading to a storage deficit. A special case of retrieval deficit concerns the state-dependency of memories. As briefly discussed in the reconsolidation section above, learning occurs under certain mental states that are required for subsequent memory retrieval. Prior studies have shown that memories that cannot be appropriately retrieved under certain conditions can be fully retrieved if the subject achieves the same mental state that was present during the memory acquisition phase. This state-dependency phenomenon has been extensively described with psychostimulants, opioids, benzodiazepines and other drugs and experimental conditions such as high arousal levels and hypothermia (for a review see: Radulovic et al., 2017). These studies suggest that the mental state acts as an important reminder that is "incorporated" into the memory trace during memory acquisition and/or reconsolidation, hence becoming a critical component of the retrieval process.

5.3. The active decay hypothesis

Although forgetting has been viewed as a passive mechanism, a number of studies have proposed that forgetting is an active process that involves the removal of previously consolidated memories (Davis and Zhong, 2017; Hardt et al., 2013). A pioneer study in the early 2000’s provided the first evidence, finding that chronic administration of CPP, an antagonist of the NMDA glutamate receptor, prevents the natural decay of LTP in the hippocampus of rodents (Villarreal et al., 2002). This finding was corroborated by further studies demonstrating that spatial memory is maintained as long as NMDA receptors are blocked (Sachser et al., 2016; Shinohara and Hata, 2014), particularly NMDA receptors containing the GluN2B subunit (Migues et al., 2019; Sachser et al., 2016). It has been proposed that calcium entry through NMDA receptors induces forgetting by activating calcineurin and synaptogamin-3, which ultimately removes GluA2-AMPA receptors from the synapse (Awasthi et al., 2019; Sachser et al., 2016). Consistent with the idea that glutamatergic transmission underlies memory forgetting, inhibiting the endocytosis of GluA2-AMPA receptors from synapses extends the duration of memories with neutral, appetitive, or aversive valence (Migues et al., 2016), and blocks LTP depotentiation (Dong et al., 2015; Migues et al., 2016).

Another potential candidate to modulate the mechanisms of memory forgetting is the enzyme Rac1, which has been implicated in the shrinkage of dendritic spines by regulating actin dynamics. Pharmacological inhibition of Rac1 extends memory life-time, whereas pharmacological activation or its overexpression speeds up memory forgetting in both rodent hippocampus and flies (Liu et al., 2016, 2018; Shuai et al., 2010; Zhang et al., 2018b). Forgetting has been also associated with the formation of functional mature neurons from neural stem cells, as neurogenesis in the dentate gyrus accelerates forgetting of hippocampal dependent memories, mediates infantile amnesia in juveniles, and enables behavioral flexibility in adult mice (Akers et al., 2014; Epp et al., 2016). In addition, the engulfment of large particles in the synapse or cell surface observed during phagocytosis seems to play a critical role in the regulation of forgetting. Either depletion of microglia or the inhibition of phagocytosis by microglia cells prevents memory forgetting, with active synapses being less prone to phagocytosis (Wang et al., 2020a). Thus, synapse elimination by microglia may contribute to the
degradation of memory engrams observed during forgetting.

Taken together, these studies suggest that the molecular bases of memory forgetting rely on both functional and structural reorganization of synapses that have been potentiated during learning. We propose that calcium entrance through the activation of NMDA receptors triggers the endocytosis of AMPA receptors and the initiation of Rac1 intracellular cascades, which will lead to synaptosomal depotentiation and dendritic spine shrinkage, two critical processes for memory forgetting. Further studies are needed to investigate this possibility.

6. General considerations on memory destiny

Here we explored the distinct processes that a previously acquired memory can follow after retrieval: reconsolidation, extinction, and forgetting. We have observed that, depending on the similarities between the information represented by the original memory and what is experienced during retrieval, memories can be conducted to distinct fates. If a high degree of similarity exists between the original memory and the retrieval phase, no changes are observed in the natural course of the memory and the information is maintained in its current form. In contrast, if a mismatching occurs between the retrieval and the original memory, the new piece of information presents during the retrieval phase is incorporated into the original memory, thereby strengthening, weakening, or modifying the existing memory trace. When this mismatching goes beyond certain levels, extinction mechanisms take place and new cells are recruited to create a new memory that will inhibit the existing memory trace without eliminating it. Finally, if the information stored by the original memory is no longer important or is rarely assessed during retrieval, the memory trace is removed by an active process of forgetting. From an evolutionary perspective, this dynamic nature of memory has important adaptive functions including updating our mental representation of past experiences, optimizing our predictions about the future, and properly adjusting our behaviors to make the most appropriate choices.

Funding sources

This work was supported by a CNPq Universal-2018 grant to L.d.O.A. and NIH grants R01-MH120136 and R00-MH105549, Brain & Behavior Research Foundation NARSAD Young Investigator grant, and a UT System Rising STARS Award to F.H.D-M.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

The authors would like to sincerely thank Arturo Romano, Christian Bravo-Rivera, Douglas Engelke, Gabriel Lázaro-Muñoz, John Byrne, Jose Rodríguez-Romaguera, Paul Smolen, and the anonymous reviewers for their valuable comments and suggestions, which helped us to improve the quality of this manuscript.

References


Ben Mamou, C., Gamache, K., Nader, K., 2006. NMDA receptors are critical for unleashing consolidated auditory fear memories. Nat. Neurosci. 9, 1237–1239.


L. de Oliveira Alvares and F.H. Do-Monte


Neuropharmacology 58, 1097–1108.


Neuroscience and Biobehavioral Reviews 125 (2021) 592–607


L. de Oliveira Alvares and F.H. Do-Monte