

Opinion

Inferring danger with minimal
aversive experienceBlazej M. Baczowski,¹ Jan Haaker,² and Lars Schwabe ^{1,*}

Learning about threats is crucial for survival and fundamentally rests upon Pavlovian conditioning. However, Pavlovian threat learning is largely limited to detecting known (or similar) threats and involves first-hand exposure to danger, which inevitably poses a risk of harm. We discuss how individuals leverage a rich repertoire of mnemonic processes that operate largely in safety and significantly expand our ability to recognize danger beyond Pavlovian threat associations. These processes result in complementary memories – acquired individually or through social interactions – that represent potential threats and the relational structure of our environment. The interplay between these memories allows danger to be inferred rather than directly learned, thereby flexibly protecting us from potential harm in novel situations despite minimal prior aversive experience.

Learning to protect from danger without risk of harm?

Individuals minimize the risk of harm by exhibiting a variety of **defensive behaviors** (see [Glossary](#)) tailored to the recognized imminence of danger [1,2]. Proximal **threats** elicit fixed behavioral responses that prepare us to cope with impending harm (e.g., freezing to a sudden appearance of a dangerous dog). More remote but foreseeable threats, on the other hand, can be surmounted by flexible actions taken when we are in relative safety (e.g., avoiding parks where people walk their dogs without a leash). Although such defensive behaviors are generally adaptive, overly strong defensive behaviors, accompanied by an exaggerated **fear response**, are maladaptive and may contribute to mental disorders characterized by aberrant evaluations of risk [3]. Thus, accurate recognition of danger (i.e., assessing the risk of harm given the current and prospective state of our environment) is crucial for regulating defensive behaviors.

How do individuals learn to recognize that they are in danger? Past life-threatening situations provide useful information about what poses a risk of harm beyond innate threats [1]. For example, you may have learned from a dog attack during a picnic in a park that the sight of this dog entails risk. Learning about threats from first-hand aversive experiences is well described by associative mechanisms of **Pavlovian threat conditioning** [4]. The experience of an unexpected noxious outcome in relation to a biologically neutral cue creates a long-lasting threat memory. Subsequent detection of this cue reactivates the cognitive representation of the threat and mobilizes conserved neuronal circuits in the amygdala and downstream pathways to the brainstem that control defensive responses (e.g., freezing) and aim to prepare for the anticipated imminent harm [5]. Although being a fundamental learning mechanism, the acquisition of Pavlovian threat memory is risky and has a limited scope. It is risky because first-hand exposure to dangerous events may be harmful, if not deadly. Moreover, Pavlovian threat memory is limited to detecting already known (or similar) threats, and fails to guide our defense when such cues are not immediately perceivable but the possibility of harm, even if remote, still exists (e.g., a dog might be hiding).

Highlights

Pavlovian threat conditioning is a fundamental process of how we learn about threats. However, learning from first-hand exposure to danger inevitably involves a risk of harm and is largely limited to detecting known (or similar) threats.

To predict and detect novel threats, Pavlovian threat memory needs to be complemented by other learning systems that operate mainly in safety and represent threats and the structure of our environment in rapid episodic memory, relational (map-like) memory, and conceptual knowledge.

Complementary memories of threats and the structure of our environment may also result from communication with and observation of others, which largely spares exposure to danger.

Building knowledge about the world beyond first-hand Pavlovian threat conditioning allows danger to be inferred rather than learned, which crucially minimizes the risk of harm to the organism.

¹Department of Cognitive Psychology, Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

²Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

*Correspondence:
lars.schwabe@uni-hamburg.de
(L. Schwabe).

Therefore, Pavlovian memories alone are insufficient to fully explain how individuals predict remote threats and detect proximal threats in future situations.

Accumulating evidence shows that Pavlovian threat learning may shape the organization of episodic memory [6–8], and that motivationally significant events change the cognitive underpinnings of goal-directed behavior [9,10]. Based on these recent advances, we discuss here how episodic memory, relational memory, and conceptual knowledge allow an organism to infer danger with minimal aversive experience. These memory processes may capitalize on, but significantly expand, our ability to recognize danger beyond Pavlovian associations. Importantly, the processes relevant for subsequent threat prediction and detection differ in the extent to which they involve prior exposure to danger (Figure 1). First-hand threat conditioning and accompanied episodic memories inevitably entail aversive consequences, but substantial knowledge about the relational structure of our environment whereby one cue is informative about the other (e.g., representing a kennel in relation to a dog) is acquired during safety. Moreover, learning about threats through social interactions (e.g., observing another person being bitten by a dog) builds upon minimal prior first-hand aversive experiences but significantly extends our threat representations without risk of harm. Together, the interaction between these learning and memory processes allows danger to be inferred rather than explicitly learned, which crucially expands our ability to protect against potential harm (Figure 2). In particular, the interaction of Pavlovian learning with episodic relational memory and conceptual knowledge may facilitate defensive behavior by allowing the prediction and detection of novel threats and by mobilizing defensive

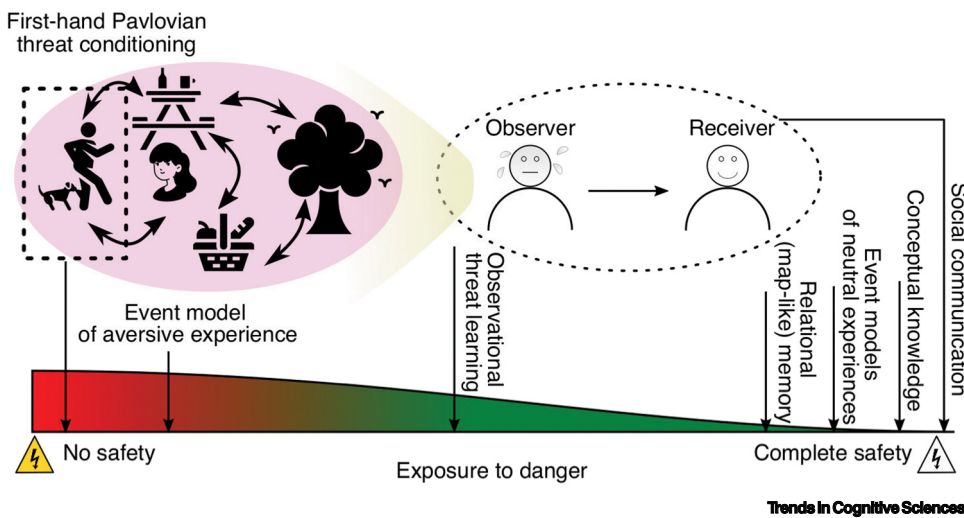


Figure 1. Learning and memory processes operating at different levels of exposure to danger. The degree of exposure to danger varies from actual harm to its absence (complete safety). Pavlovian first-hand threat conditioning inevitably involves harm (e.g., being bitten by a dog; left-hand side of the figure) and is embedded in an individual episodic memory of the aversive event that entails incidental and harm-predictive cues (e.g., being bitten by a dog during a picnic in a park) related to one another (black arrows). Nevertheless, other processes operate during safety (right-hand side of the figure), such as episodic memories of neutral events (e.g., daily commute to work through a park) and the relationships between neutral cues encoded across episodes, that create a map-like structure of our memories (e.g., meeting various people in a park) as well as conceptual knowledge of how similar the objects around us are despite their minimal physical resemblance (e.g., various picnic tables in a park). Moreover, some knowledge does not require any first-hand experience and can be acquired from interactions with others through social communication (e.g., instructions to avoid a dog) or observations (e.g., seeing another person being bitten by a dog; middle part of the figure). These processes ultimately result in knowledge about what constitutes a threat (e.g., the first-hand and observational threat learning) and the relational structure of our environment – how one mnemonic element is informative about the other (e.g., episodic map-like and conceptual knowledge). Icons designed by Freepik, Tempo_doloe, Triangle Squad, Victoruler from [Flaticon.com](https://www.flaticon.com/).

Glossary

Cognitive map: originally conceived as an allocentric representation of the relative spatial locations (and objects) in a Euclidean space, it was later treated more broadly as a complex representation of contingencies between multiple stimuli or events (not necessarily tied to physical space) that constitute an internal (episodic-like) model of the environment (or a cognitive task) and thereby guide our expectations about the immediate and distant future.

Cognitive (conceptual) space: a representation of conceptual knowledge where its entities are organized by their relative positions on multiple dimensions of abstract feature values such that similar items are located nearby (e.g., an elephant and a bulldozer, considering weight). Projecting novel stimuli to a particular region in a cognitive space allows infer their properties (also consequences) to be rapidly inferred from a representative exemplar.

Defensive behavior: a class of innate and acquired behavioral outputs (accompanied by physiological changes) comprising reactions and goal-oriented actions that aim to survive (or avoid) a direct (or forecasted) exposure to harmful consequences.

Event model: a spatiotemporal representation of a specific event type (e.g., a picnic in a park) that includes entities such as people, objects, and places, as well as the relationships between them, that is used to comprehend a currently unfolding situation and form accurate predictions about upcoming sensory information.

Fear response: a behavioral response accompanied by physiological changes and subjective feelings induced by a harm-predictive cue.

Latent (hidden) state of the environment: an unobservable, and therefore inferred, state of the environment that is believed to generate sensory observations.

Pavlovian threat conditioning: a fundamental learning process (and an experimental procedure) by which individuals learn to associate a biologically neutral cue (e.g., a tone serving as a conditioned stimulus, CS) with an intrinsically noxious outcome (e.g., an electric shock that serves as an unconditioned stimulus, US) owing to its predictive value.

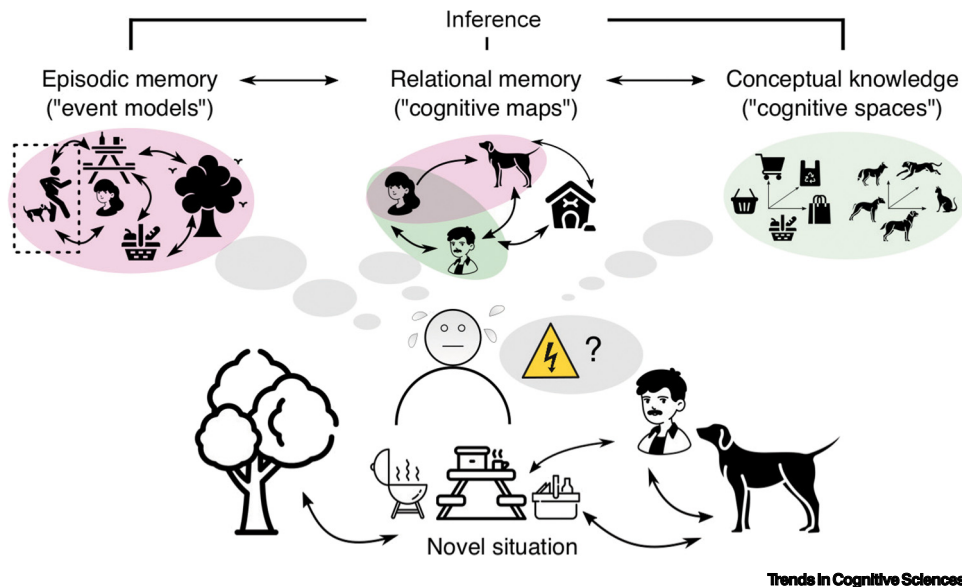


Figure 2. Inferring danger from multiple learning and memory processes. These cognitive processes operate at different levels of prior exposure to danger (Figure 1) and can be jointly deployed to infer current danger. First-hand experiences and social interactions result in knowledge describing the relational (predictive) structure of our environment, in other words how one cue is informative about another (black arrows between mnemonic elements). Such relational knowledge together with Pavlovian threat memories and episodic memories of aversive events can be used to infer the risk of remote or immediate harm in future situations by enabling the prediction and detection of novel threats, respectively. For example, individuals can comprehend the new situation by reusing an existing 'event model' of an aversive episode (being bitten by a dog during a picnic). They can also use their 'cognitive maps' built across neutral and aversive episodes to infer risk from cues that were never experienced in an aversive context, such as seeing a neighbor with a dog (red and green ellipses indicate elements acquired during aversive experience and safety, respectively). Finally, individuals can infer the properties and consequences of novel stimuli based on their similarity to familiar cues quantified by their representational distance in 'cognitive space' (a new vs dangerous dog). Importantly, social interactions with others (seeing another person being bitten by a dog) can replace first-hand exposure to danger and result in a similar representation of threats, and thus also aid the inference process. Icons designed by Freepik, Tempo_doloe, Triangle Squad, Victoruler, loongeeq26, alfan, Smashicons from [Flaticon.com](https://www.flaticon.com/).

behaviors before the actual threat is present (e.g., avoidance of events anticipated to be dangerous), thereby directly reducing the risk of harm.

In the first part of this paper we discuss how individuals may successfully protect themselves from potential harm in future situations by using complementary memories of minimal first-hand aversive experience combined with pre-existing knowledge about the relational structure of our environment formed during safety. Specifically, we illustrate the role of enriched episodic memories of aversive experiences, map-like knowledge built across neutral and aversive learning episodes, and abstract representations extracted from repeated exposures to environmental regularities, which all reflect the relational structure of our environment derived from experiences that gradually progress from being idiosyncratic to more generic. In the second part we briefly portray social learning through communication with and observation of others to illustrate how prior minimal first-hand aversive experience can be capitalized upon to expand our cognitive representations of threats without risk of harm. Throughout, we discuss recent findings on how we encode, retrieve, and combine episodic memories, how we judge the similarity of objects around us, and how we learn from social interactions. Although these cognitive phenomena may engage distinct neuronal substrates and computations, as well as resulting in qualitatively different outputs (e.g., memory recognition, physiological changes, behavioral actions), they all contribute to

Relational (map-like) memory: a memory representation that relates elements of an experience to one another (in spatial and temporal context). Separate event memories can be linked together via their shared elements to create a map-like structure across individual episodes, which in turn enables its flexible recombination and inferences of novel information.

Sensory pre-conditioning: learning a (predictive) association between two initially neutral stimuli where the latter subsequently acquires a Pavlovian association and can be mobilized by the former to generate an outcome appropriate response.

Social observational threat learning: a social learning process which entails an observation of a demonstrator who is responding to an aversive outcome. This response is detected by the observer (as a social US) and is predictable by a neutral cue (social CS).

Threat: a harm-predictive cue that is either innate (e.g., the sight of a predator) or acquired through Pavlovian conditioning (or inferred as such) that can elicit (innate) defensive responses (e.g., passive avoidance) and emit learned actions (e.g., active avoidance).

how individuals protect themselves from danger. Together, our perspective aims to stimulate research on how (first-hand and social) mnemonic processes that operate on a continuum from danger to safety enable individuals to flexibly regulate their defensive behavior despite minimal past aversive experience.

Representing the structure of aversive experiences in episodic memory

Our memories facilitate the understanding of current and future situations [11]. To rapidly establish a situational model of ongoing events, individuals constantly infer whether the current sensory observations are generated by known experiences or whether they are entirely new and deserve a separate place in memory [12]. Novel, unexpected, and biologically significant events become units of their own in episodic memory, whereas common experiences are grouped [13]. Although mundane memories transform from detailed to more gist-like representations over time and repetition [14–16], emotionally highly arousing experiences may remain specific and vivid [17,18]. As a result, such memories may be protected from competing memories because new contradictory experiences appear so distinct that they are believed to be caused by a new **latent (hidden) state of the environment** [7,19], rendering aversive memories difficult to modify but adaptive for assessing future risk according to a 'better-safe-than-sorry' strategy. For instance, having been attacked by an aggressive dog in a park, encountering a friendly dog on a leash being walked by a familiar neighbor may not change the representation of the prior aversive event because the current neutral experience is attributed to a separate state of the environment (i.e., a friendly dog on a leash and its familiar owner). Thus, previous exposure to dangerous situations can result in a dedicated **event model** of aversive experiences, which includes entities such as people, objects, places, and the relationships between them. The content and structure of these event models organize our memory, facilitate comprehension of ongoing situations, and allow predictions about future sensory observations [11,12,20,21]. These features of event cognition may facilitate prediction and detection of (novel) threats because event models contain both mnemonic elements and their spatiotemporal relationships that can later be utilized to form accurate predictions about upcoming potential harm.

To effectively infer risk, relevant memories must be easily accessible and long-lasting. Memories of emotionally laden events are prone to neural reactivation even by inherently neutral retrieval cues [22,23]. Episodic memories of aversive events exhibit slow forgetting [24] and are prioritized for their maintenance through increased noradrenergic arousal and a cascade of neuromodulatory changes integrated in the amygdala to modulate mnemonic processes in the hippocampus [25,26]. Fine-grained investigations into the interplay between amygdala-based threat conditioning and hippocampus-based episodic memory recently became possible by incorporating trial-unique (i.e., non-repeating) images of two semantic categories, whereby one remains neutral and the other becomes associated with an aversive outcome [6]. This paradigm revealed that items belonging to the category predictive of the aversive outcome are prioritized in episodic memory [27,28], which involves amygdala activity during encoding [29] and spontaneous post-encoding reactivation of neocortical representations coordinated by the hippocampus [30]. Importantly, the formation of long-lasting memory of what constitutes threat is also established through the surprise elicited by the presence or absence of the aversive outcome [8,31], suggesting that we form memories particularly when our current models of the environment are insufficient or incomplete. Further, harmful events are usually accompanied by incidental cues in a shared spatiotemporal context.

Studies that tested episodic memory for complex emotional (stressful) real-life episodes composed of inherently neutral elements encountered in a sequence showed that, if these elements were informative about the stressor, they were prioritized in memory [32–34]. Only those elements that were entirely uninformative about the stressor exhibited reduced memory, in line

with studies suggesting a detrimental influence of emotion on associative memory [35]. We further note that, even if emotion reduces memory of some elements, this is typically a relative decrease and not a complete absence, and even the relatively reduced memory could contribute to event models useful for the prediction of danger. However, the preferential storage of central elements of an emotional episode indicates that information that is incidental but considered relevant for survival is prioritized in long-term episodic memory. Other studies revealed that memory prioritization of sequentially presented items is graded, reflecting their (spatio)temporal proximity to the biologically significant outcome [36]. Likewise, retrieval of episodic elements recapitulates their sequential presentation (i.e., neural replay) [37]. To create a shared context between the emotional and subsequent neutral information, arousal induced by emotionally relevant material can linger for minutes and generate a prolonged state of increased amygdala–hippocampal network activity that enhances episodic memory for neutral information [38]. More broadly, dynamic fluctuations in pupil-linked arousal tracking of salient environmental changes may reflect shifts in the internal context that either links neural representations of adjacent events or separates them into distinct episodes in memory [39]. Therefore, spatiotemporal context embedding of sequentially encoded episodic elements may provide a mnemonic structure to aversive experiences that allows the past to be rewound and simulate the future to guide our defensive behavior. In line with this view, emotion enhances memory of the location and temporal context of an event [40–44]. Moreover, recent evidence shows that neutral events encoded in the context of an emotionally arousing encounter are better remembered than those encoded in a neutral context, and can become part of an event model that is useful to subsequently predict a threat encounter [35,45].

Together, these findings reveal that the structure of aversive experiences, such as being bitten by a dog during a picnic in a park, is reflected in the structure of our memories. Episodic memories of aversive experiences contain the source of harm that enables threat detection (e.g., a dog), incidental cues encoded in the shared context that can act as additional warning signals (e.g., a tasty sandwich that may unintentionally bait a dog), and their spatiotemporal relationships describing the trajectory of how potentially dangerous situations unfold (e.g., the smell of the sandwich, the sight of the dog, and the attack).

Expanding episodic models of danger with pre-existing relational memories

Episodic memories of aversive events help us to rapidly create models of our environment. However, the full environmental structure is rarely present in a single episode. Instead, memories of related experiences collected across time and space, usually during safety, must be combined to form a complex network of relationships among people, objects, and places such that recognition of one informs about the other [46,47].

Many daily experiences occurring in safety comprise inherently neutral objects, but the relationships between them may later be useful for our defense. **Sensory pre-conditioning** [48] illustrates that separate memories of predictive relationships encoded first in a safe and then an aversive context can be combined through an overlapping element despite their dependence on distinct brain regions [49]. Using the relationships from both experiences, the risk of harm can be inferred from the cues previously encountered only in a safe context because they provide information about the potential presence of a known but so far undetected, threat. For example, by repeatedly seeing neighbors buying dog food, we may learn that they own a dog, and we may infer the next time we meet them that their dog is present somewhere in the vicinity. Linking separately acquired experiences can occur through offline hippocampal memory replay [50,51], which allows knowledge to be built independently from veridical experience. Combining memories offline may be particularly prominent when these are foreseen as relevant for future behavior because post-encoding neural reactivation of memories tends to be enhanced for emotionally

salient experiences [52]. Building networks of relational memories across multiple episodes not only expands our initial memories but also allows us to flexibly extract novel information [53]. The ability to assess risk based on the inference of unobserved links between entities in our environment crucially expands our defense beyond threat detection because it allows potential harm to be foreshadowed before the threat is actually present.

Although aversive memories generally reflect rare events, we constantly form memories of relationships between neutral cues as we navigate the world around us during safety. These so-called **cognitive maps** [54], supported by the hippocampal–entorhinal system [55], are thought to represent not only spatial relationships between locations in physical space but also non-spatial predictive relationships that encode immediate (adjacent) and multistep (long-range) transitions between states of our environment [56]. Predictive episodic-like relationships organized in a map-like format describe what is likely to occur in the near as well as distant future, which vastly supports planning over different timescales [57]. Equipped with such predictive maps acquired during safety and first-hand threat memories, individuals can later plan how to minimize exposure to danger before a threat is near – when a threat is absent but can already be predicted from the current or prospective state of the environment. Indeed, decisions to avoid a threat encounter (e.g., by rejecting an invitation to a house party by a neighbor who owns a dog) and thereby harm can be made ahead of time through mental simulations which reactivate neural representations of the trajectory to a distant goal coordinated by the hippocampus and medial prefrontal cortex [58,59]. Moreover, we can make accurate avoidance decisions even when simulated state transitions have never led to aversive outcomes before but signal this possibility based on their relationships with other previously experienced consequences [59,60]. Finally, the structure of cognitive maps might be adapted by biologically significant events to better support goal-oriented behavior. For example, deformations of spatial maps were observed in the rat medial entorhinal cortex in response to learning new hidden reward locations [9,10].

Together, **relational map-like memories**, built across multiple episodes and acting together with Pavlovian threat memories, significantly expand our defensive capabilities because cues that were never experienced in an aversive context can support risk assessment and planning to avoid threats and ultimately harm. Because memories of aversive events are typically stronger than those of mundane events, and key elements of an emotional episode appear to be prioritized in memory, it is tempting to speculate that cognitive maps might be geared towards predicting threat over and above motivationally neutral information. Direct support for this idea, however, is still lacking. Although we assume that cognitive maps facilitate the initiation of defensive behaviors before threat is present (e.g., avoidance of events anticipated to be dangerous), recent evidence suggests that, under acute threat, the use of cognitive maps and flexible navigation may be impaired [61].

Transferring episodic and relational models of danger to novel situations

Although relational memories collected across multiple episodes expand our models of the environment, they comprise already known events. However, a key challenge is to infer danger in circumstances we have never encountered before. For example, would an outdoor barbecue entail any risk for someone previously bitten by a dog during a picnic? To transfer pre-existing knowledge to novel situations and judge whether a novel stimulus may lead to familiar consequences even when it bears minimal physical resemblance to the past, we must infer that the new sensory input is generated by the latent state of the environment stored in our memory [62,112,113], a process which relies on the orbitofrontal cortex [63]. To perform such inference, and thereby transfer our memories to novel situations, individuals may spontaneously leverage their

conceptual knowledge that is abstracted from individual experience but represents stable relationships among objects according to their multi-feature similarity [64].

Given the complex nature of the objects we encounter every day, similarity can be assessed at many levels. Stimuli that perceptually approximate to known threats via unimodal features (e.g., the size of a dog) can reactivate neural correlates of Pavlovian memories and mobilize defensive behavior [65,66]. However, generalization based on simple perceptual features is not easily applicable to real-world objects that are composed of numerous properties (e.g., size, breed, or facial expression of a dog). The multidimensional structure of our mental representations [64] makes objects appear similar in one dimension but dissimilar in another.

To reduce such complexity and facilitate similarity judgments, individuals direct their attention to only a few dimensions and actively search for the features that are most relevant to current goals [67]. For example, when evaluating risk, participant gaze patterns became biased towards those features that were most predictive of the aversive outcome in the past [68]. Selective attention projects our rich conceptual knowledge into a low-dimensional **cognitive (conceptual) space** where similar stimuli are located nearby, and their similarity can be easily computed using a distance metric [69,70]. The distance metric operating in cognitive space is thought to be supported by the processes we typically recruit when navigating through physical environments [69,71]. For instance, the distance between real-world locations and between stimuli in a conceptual space are both encoded by the hippocampus [72,73], and the way we navigate through either space engages the entorhinal cortex [74,75]. Establishing the proximity of stimuli in the low-dimensional representation allows adaptive generalization because stimuli located nearby in a cognitive space can be inferred to be generated by the same latent cause and therefore bear the same properties [70,76]. All knowledge about a known threat (e.g., a dog) and other cues predictive thereof (e.g., a park) can then be transferred to novel stimuli that are considered to be functionally equivalent, thus leading to consequences similar to those we experienced in the past. If the similarity of stimuli encountered in different events is perceived to be high, these events as a whole may be represented as more similar.

Because relational episodic-like networks are only concerned with already known entities, knowledge transfer based on similarity judgment in cognitive spaces goes far beyond associative inferences. Using cognitive spaces during inference may allow precise estimations of the level of expected harm even in situations we encounter for the first time [77,78]. For example, after being bitten by a dog during a picnic, an individual may form predictions about potential harm when going to another outdoor social event and seeing novel but similar objects such as a blanket, grocery basket, and a dog running around (Figure 2). Notably, emotion has been shown to result in enhanced memory for item details, and this may facilitate the precision of similarity judgments for stimuli encoded under emotional arousal [79,80].

Taken together, goal-oriented and selective attention allows us to reduce the complexity of our mental representations into a low-dimensional cognitive space to ease the similarity estimates between what we learned in the past and what we observe now. In turn, our existing models of the environment with their relational structure can be reused whenever needed to predict and detect novel threats.

Inferring danger from the aversive experiences of others

Although learning from first-hand aversive experience allows us to adapt to the environment, it entails harmful exposure to danger. By contrast, social learning – broadly defined as learning from or in interaction with others [81,82] – builds upon the minimal prior first-hand aversive experiences

Box 1. First-hand and observational threat learning in the brain

First-hand and observational threat learning involve a common neural network (Figure 1). Experiments in non-human animals revealed that dorsal anterior cingulate cortex (dACC) and the amygdala (AMY) – two regions that are key for learning about threats from first-hand aversive experiences [1,4] – encode threat-relevant information to learn via observation the responses of others [105,106]. This overlap in the dACC and AMY, together with the anterior insula (AI), secondary somatosensory cortices, and midbrain periaqueductal gray (PAG), in first-hand and observed threats is furthermore implicated by research in humans [87,107]. The involvement of common brain regions suggests that both ways of learning may share similar neurocomputational mechanisms. However, there are differences between observational and first-hand threat learning that relate to information flow within these shared networks: electrophysiological and functional MRI connectivity analyses indicate that socially derived information (e.g., distress of a demonstrator) is preferentially transferred from cortical (ACC or AI) to subcortical regions (AMY) to mediate associative memory formation during observational threat learning. First-hand threat learning underlies a relay from the AMY to higher cortical areas to transfer the motivational threat value [87,105], in addition to projections from the ACC to AMY which allow processing of inferences about future threats [108].

Hence, each direction of these bidirectional ACC–AMY projections modifies different aspects of the information processing (social information and threat inferences vs motivational threat value) to shape threat learning. This would support the idea that inferences about future threats from minimal first-hand experience and social interactions are mediated within similar neural networks. It further aligns with recent research suggesting that previously made first-hand threat memories in the hippocampus (and the basolateral AMY in mice or AI in humans) are reactivated by observational learning [90–92]. It would be interesting to unravel how these shared brain processes are related to circuits that are associated with the integration of socially relevant information and are active during observational learning, such as the superior temporal sulcus (STS) and temporoparietal junction (TPJ) [87]. It will further be important to understand how forebrain representations of threats in observational learning are integrated into sensory pathways of first-hand processing because observational threats are processed in the spinal cord [109] and via endogenous opioids, which are key to defending against first-hand nociception [110,111]. Together, these overlapping and interacting neural processes enable social information to initiate, adapt, and broaden our knowledge about threats, including those that we acquired through first-hand exposure to danger.

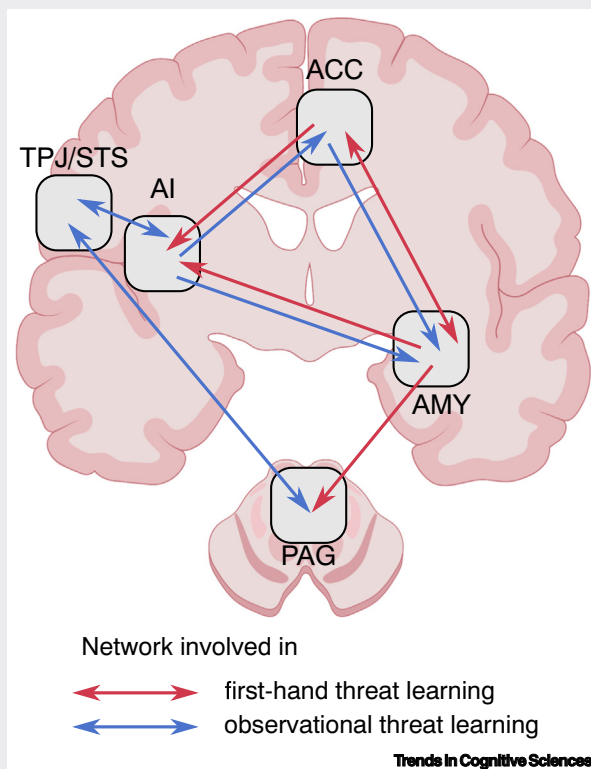


Figure 1. Brain networks involved in first-hand and observational learning about threats. Arrows indicate information flow within the networks of first-hand (red) and observational (blue) threat learning. Background image created with BioRender.com.

but expands our knowledge of what poses a threat without risk of harm. Social communication and observational learning about threats rely on neural processes that are engaged in first-hand aversive learning (Box 1), and thereby may result in comparable cognitive and neural representations of threats that can equally well guide our defensive behavior. Moreover, socially acquired threats may act together with other first-hand memories acquired during safety, such as episodic models of neutral events, map-like episodic memories, and conceptual knowledge. Together, social learning crucially expands our knowledge of threats which may interact with the processes described in the preceding text to aid inference of risk in novel situations (Figure 2).

Knowledge of what poses a threat can be rapidly transmitted through verbal instructions. Threat-related instructions in human experiments are a primary example of how verbal information about the relationships between cues (or contexts) and aversive outcomes elicit defensive responses (i.e., physiological arousal) even though such environmental contingencies were never self-experienced [83,84]. Further, contingencies of self-experienced aversive outcomes can be quickly modified by social instructions that update threat-related neural processing, for example in the dorso-lateral prefrontal cortex but not in the amygdala [85,86]. These findings show that social communication may allow the formation of threat-related knowledge without the experience of aversive outcomes.

Threat-related information can also be transmitted through the expressions and actions of others that are merely observed (**social observational threat learning**). Registering a conspecific receiving a painful electric shock in relation to an inherently neutral stimulus engages brain networks of threat processing and forms threat memories which lead to changes in physiology (e.g., skin conductance response), behavior (e.g., freezing), and action tendencies (e.g., avoidance) comparable to first-hand aversive learning [84,87–89]. Learning about threats from observation of the responses of others to aversive events is facilitated by prior first-hand aversive experience [90,91]. The reactions of others to threats reinstate prior first-hand threat memories in the observer [92] and reactivate so-called memory engram cells in the amygdala that allow the coupling of sensory observations with threat value but without self-experienced aversive consequences [90,91]. These findings suggest that observational learning capitalizes upon the minimal first-hand aversive experiences by interacting with and extending existing threat-related representations.

Observational learning transmits information not only about specific objects but also about time-extended actions that require multistep cognitive representations of the environment. For example, spatial trajectories of a running conspecific in a maze are replayed by hippocampal place cells in the observer [93], which indicates a remote formation of a cognitive map that can later be used for planning routes when pursuing one's own goals, such as avoiding places associated with an increased risk of harm [94]. These findings indicate that the processes that operate at a social level may result in knowledge structures similar to those rooted in first-hand experiences, including cognitive representations of threats and our environment. Moreover, episodic memories acquired from first-hand experiences during safety may interact in a complementary way with memories acquired through social interactions. For example, a plethora of first-hand event models that describe how typical daily-life situations unfold over time may be combined with the socially learned threat supporting the prediction of its encounter (e.g., individual event models of playing in a park where dogs run around and a socially observed attack of a dog on a street). Therefore, knowledge acquired through social interactions with others during safety enables individuals to protect themselves efficiently and flexibly from danger despite minimal prior first-hand aversive experience.

Concluding remarks

Our defensive system is remarkably flexible, which allows us to deal with danger even though our personal aversive experience is minimal. To assess the risk of harm when it cannot be directly

Outstanding questions

How is threat-related information from multiple sources – based on one's own experiences or the experiences of others – integrated? Do we weight personal and social source of information equally, and how is this process implemented in the brain?

How do we deal with discrepant information about danger from multiple sources? Under which conditions do we modify or separate our own experiences from what we have learned from others?

How do different knowledge structures (event models, cognitive maps, and cognitive spaces) support the inference about the level of perceived risk and the anticipated type of aversive outcome? Do we always know what exactly we are afraid of (e.g., if I go there again, something bad might happen)?

How do episodic memories of dangerous situations differ in their relational structure from mundane memories? Is the relational structure of aversive memories distorted towards the highly arousing event? How does it affect sequential retrieval and corresponding predictions about inferred upcoming observations?

How and when do we integrate experiences acquired in safe and aversive contexts? How easily can the integrated knowledge be updated to track changes in environmental contingencies that can be altered when new elements are introduced, and old ones are removed?

Do aversive experiences modify our cognitive (conceptual) space of representations that were acquired during safety? If so, how and through which cognitive and neuronal mechanisms?

How rapidly can the prediction of foreseeable and detection of perceptible threats adapt to entirely novel situations? What circumstances initiate or hamper the inferences about novel threats, and how rapidly are they updated when new information comes in?

How does the inference of remote and immediate harm (threat prediction and

based on first-hand aversive experiences, individuals can use other memories acquired during safety and in interaction with conspecifics. These complementary memories represent what poses a threat and how objects in our environment relate to one another, which together allow danger to be inferred. For example, relational representations not only enable the detection of novel threats (i.e., when they are already present) but also enable prediction of foreseeable threats (i.e., when they are still absent), thus promoting defensive behaviors that directly reduce the risk of harm ahead of time. Such defensive behaviors conducted over an extended time horizon involve an intricate interplay of amygdala, hippocampus, and prefrontal cortex [95].

detection, respectively) translate to defensive behaviors? Are these behaviors tailored to the inferred imminence of danger?

The perspective that we evaluate risk based on danger that is largely inferred rather than learned may have relevant implications for fear-related disorders. Individuals diagnosed with these disorders exhibit avoidance behavior to inherently neutral situations [96] as if they were simulating future threat encounters. In these disorders, avoidance is irrational, and is typically not related to a veridical source of harm, and may emerge without first-hand exposure to danger (e.g., when transmitted by close individuals) [97,98]. The present perspective suggests that some of the fear-related symptoms originate not only from aberrant models of the world, as acknowledged by current cognitive therapies of fear-related disorders [99], but also from biased inferences about risk. Such inferences, in turn, may propel already maladaptive models of the environment through offline simulations, leading to a vicious cycle [100].

In this context, one of the key questions for future research relates to whether the inference of risk modifies pre-existing knowledge structures acquired during safety. To tackle this question, future studies could benefit, for example, from virtual reality techniques that can create arbitrary real-life environments – which have been successfully used to measure threat appraisals [101] or distortions in spatial memory [102]. Virtual reality could lead to novel and enriched experimental designs to measure putative changes in the relational structure of familiar environments, thereby extending recent evidence for how emotional events modify pre-existing knowledge structures acquired during safety [103,104]. Another important question is how cognitive maps may mobilize defensive behaviors whenever needed. Future studies will be necessary to experimentally manipulate the availability and relevance of cognitive maps when studying defensive behavior and, more broadly, to bridge disparate fields of research on cognition and defensive behaviors (e.g., avoidance). Addressing these and related questions that arise from our perspective (see [Outstanding questions](#)) may enhance our understanding of how individuals learn to predict and detect threats in a variety of (novel) situations as well as the aberrancies thereof in fear-related disorders.

Acknowledgments

We thank Andreas Olsson and Christoph W. Korn for helpful comments on a previous version of this manuscript. We further gratefully acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 449640848, GRK 2753/1 to J.H. and L.S.

Declaration of interests

The authors declare no conflicts of interest.

References

1. Fanselow, M.S. (2018) The role of learning in threat imminence and defensive behaviors. *Curr. Opin. Behav. Sci.* 24, 44–49
2. Mobbs, D. et al. (2020) Space, time, and fear: survival computations along defensive circuits. *Trends Cogn. Sci.* 24, 228–241
3. American Psychiatric Association et al. (2013) *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, American Psychiatric Association
4. LeDoux, J.E. (2014) Coming to terms with fear. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2871–2878
5. Roelofs, K. (2017) Freeze for action: neurobiological mechanisms in animal and human freezing. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160206
6. Dunsmoor, J.E. and Kroes, M.C. (2019) Episodic memory and Pavlovian conditioning: ships passing in the night. *Curr. Opin. Behav. Sci.* 26, 32–39

7. Dunsmoor, J.E. *et al.* (2018) Event segmentation protects emotional memories from competing experiences encoded close in time. *Nat. Hum. Behav.* 2, 291–299
8. Kalbe, F. and Schwabe, L. (2020) Beyond arousal: prediction error related to aversive events promotes episodic memory formation. *J. Exp. Psychol. Learn. Mem. Cogn.* 46, 234–246
9. Boccarda, C.N. *et al.* (2019) The entorhinal cognitive map is attracted to goals. *Science* 363, 1443–1447
10. Butler, W.N. *et al.* (2019) Remembered reward locations restructure entorhinal spatial maps. *Science* 363, 1447–1452
11. Zacks, J.M. (2020) Event perception and memory. *Annu. Rev. Psychol.* 71, 165–191
12. Franklin, N.T. *et al.* (2020) Structured event memory: a neuro-symbolic model of event cognition. *Psychol. Rev.* 127, 327–361
13. Rouhani, N. *et al.* (2020) Reward prediction errors create event boundaries in memory. *Cognition* 203, 104269
14. Brodt, S. *et al.* (2018) Fast track to the neocortex: a memory engraving in the posterior parietal cortex. *Science* 362, 1045–1048
15. Dandolo, L.C. and Schwabe, L. (2018) Time-dependent memory transformation along the hippocampal anterior-posterior axis. *Nat. Commun.* 9, 1205
16. Zeng, T. *et al.* (2021) Tracking the relation between gist and item memory over the course of long-term memory consolidation. *eLife* 10, e65588
17. Atucha, E. *et al.* (2017) Noradrenergic activation of the basolateral amygdala maintains hippocampus-dependent accuracy of remote memory. *Proc. Natl. Acad. Sci.* 114, 9176–9181
18. Krenz, V. *et al.* (2021) Noradrenergic arousal after encoding reverses the course of systems consolidation in humans. *Nat. Commun.* 12, 6054
19. Gershman, S.J. *et al.* (2017) The computational nature of memory modification. *eLife* 6, e23763
20. Bellmund, J.L.S. *et al.* (2020) Sequence memory in the hippocampal-entorhinal region. *J. Cogn. Neurosci.* 32, 2056–2070
21. Shin, Y.S. and DuBrow, S. (2020) Structuring memory through inference-based event segmentation. *Top. Cogn. Sci.* 13, 106–127
22. Bowen, H.J. and Kensinger, E.A. (2017) Recapitulation of emotional source context during memory retrieval. *Cortex* 91, 142–156
23. Kark, S.M. and Kensinger, E.A. (2015) Effect of emotional valence on retrieval-related recapitulation of encoding activity in the ventral visual stream. *Neuropsychologia* 78, 221–230
24. Yonelinas, A.P. and Ritchey, M. (2015) The slow forgetting of emotional episodic memories: an emotional binding account. *Trends Cogn. Sci.* 19, 259–267
25. Hermans, E.J. *et al.* (2014) How the amygdala affects emotional memory by altering brain network properties. *Neurobiol. Learn. Mem.* 112, 2–16
26. Inman, C.S. *et al.* (2017) Direct electrical stimulation of the amygdala enhances declarative memory in humans. *Proc. Natl. Acad. Sci.* 115, 98–103
27. Dunsmoor, J.E. *et al.* (2015) Emotional learning selectively and retroactively strengthens memories for related events. *Nature* 520, 345–348
28. Kalbe, F. and Schwabe, L. (2021) On the search for a selective and retroactive strengthening of memory: is there evidence for category-specific behavioral tagging? *J. Exp. Psychol. Gen.* 151, 263–284
29. de Voogd, L.D. *et al.* (2016) Disentangling the roles of arousal and amygdala activation in emotional declarative memory. *Soc. Cogn. Affect. Neurosci.* 11, 1471–1480
30. de Voogd, L.D. *et al.* (2016) Awake reactivation of emotional memory traces through hippocampal-neocortical interactions. *NeuroImage* 134, 563–572
31. Kalbe, F. *et al.* (2020) Expectancy violation drives memory boost for stressful events. *Psychol. Sci.* 31, 1409–1421
32. Bierbrauer, A. *et al.* (2021) The memory trace of a stressful episode. *Curr. Biol.* 31, 5204–5213
33. Vogel, S. and Schwabe, L. (2016) Stress in the zoo: tracking the impact of stress on memory formation over time. *Psychoneuroendocrinology* 71, 64–72
34. Wiemers, U.S. *et al.* (2013) What we remember from a stressful episode. *Psychoneuroendocrinology* 38, 2268–2277
35. Bisby, J.A. *et al.* (2018) Negative emotional content disrupts the coherence of episodic memories. *J. Exp. Psychol. Gen.* 147, 243–256
36. Braun, E.K. *et al.* (2018) Retroactive and graded prioritization of memory by reward. *Nat. Commun.* 9, 4886
37. Wimmer, G.E. *et al.* (2020) Episodic memory retrieval success is associated with rapid replay of episode content. *Nat. Neurosci.* 23, 1025–1033
38. Tambini, A. *et al.* (2016) Emotional brain states carry over and enhance future memory formation. *Nat. Neurosci.* 20, 271–278
39. Clewett, D. *et al.* (2020) Pupil-linked arousal signals track the temporal organization of events in memory. *Nat. Commun.* 11, 2207
40. Dev, D.K. *et al.* (2022) Negative emotion enhances memory for the sequential unfolding of a naturalistic experience. *J. Appl. Res. Mem. Cogn.* 11, 510–521
41. Mather, M. and Nee, K. (2008) Arousal-enhanced location memory for pictures. *J. Mem. Lang.* 58, 449–464
42. Palombo, D.J. *et al.* (2021) Exploring the facets of emotional episodic memory: remembering 'what', 'when', and 'which'. *Psychol. Sci.* 32, 1104–1114
43. Petrucci, A.S. and Palombo, D.J. (2021) A matter of time: how does emotion influence temporal aspects of remembering? *Cognit. Emot.* 35, 1499–1515
44. Rimmel, U. *et al.* (2012) Memory for time and place contributes to enhanced confidence in memories for emotional events. *Emotion* 12, 834–846
45. Talmi, D. *et al.* (2019) A retrieved context model of the emotional modulation of memory. *Psychol. Rev.* 126, 455–485
46. Eichenbaum, H. (2017) On the integration of space, time, and memory. *Neuron* 95, 1007–1018
47. Park, S.A. *et al.* (2020) Map making: constructing, combining, and inferring on abstract cognitive maps. *Neuron* 107, 1226–1238
48. Brogden, W.J. (1939) Sensory pre-conditioning. *J. Exp. Psychol.* 25, 323–332
49. Wong, F.S. *et al.* (2019) Online integration of sensory and fear memories in the rat medial temporal lobe. *eLife* 8, e47085
50. Gillespie, A.K. *et al.* (2021) Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron* 109, 3149–3163
51. Liu, Y. *et al.* (2019) Human replay spontaneously reorganizes experience. *Cell* 178, 640–652
52. Hermans, E.J. *et al.* (2016) Persistence of amygdala-hippocampal connectivity and multi-voxel correlation structures during awake rest after fear learning predicts long-term expression of fear. *Cereb. Cortex* 27, 3028–3041
53. Barron, H.C. *et al.* (2020) Neuronal computation underlying inferential reasoning in humans and mice. *Cell* 183, 228–243
54. Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208
55. O'Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*, Clarendon Press
56. Stachenfeld, K.L. *et al.* (2017) The hippocampus as a predictive map. *Nat. Neurosci.* 20, 1643–1653
57. Brunec, I.K. and Momennejad, I. (2021) Predictive representations in hippocampal and prefrontal hierarchies. *J. Neurosci.* 42, 299–312
58. Wimmer, G.E. and Büchel, C. (2019) Learning of distant state predictions by the orbitofrontal cortex in humans. *Nat. Commun.* 10, 2554
59. Wise, T. *et al.* (2021) Model-based aversive learning in humans is supported by preferential task state reactivation. *Sci. Adv.* 7, eabf9616
60. Liu, Y. *et al.* (2021) Experience replay is associated with efficient nonlocal learning. *Science* 372, eabf1357
61. Silston, B. *et al.* (2021) Threat impairs flexible use of a cognitive map. *PsyArXiv* Published online July 9, 2021. <http://dx.doi.org/10.31234/osf.io/7p69d>
62. Vaidya, A.R. *et al.* (2021) Neural representation of abstract task structure during generalization. *eLife* 10, e63226
63. Bradfield, L.A. *et al.* (2015) Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. *Neuron* 88, 1268–1280
64. Hebart, M.N. *et al.* (2020) Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nat. Hum. Behav.* 4, 1173–1185
65. Lissek, S. *et al.* (2008) Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behav. Res. Ther.* 46, 678–687

66. Onat, S. and Büchel, C. (2015) The neuronal basis of fear generalization in humans. *Nat. Neurosci.* 18, 1811–1818
67. Leong, Y.C. *et al.* (2017) Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* 93, 451–463
68. Kampermann, L. *et al.* (2019) Fixation-pattern similarity analysis reveals adaptive changes in face-viewing strategies following aversive learning. *eLife* 8, e44111
69. Bellmund, J.L.S. *et al.* (2018) Navigating cognition: spatial codes for human thinking. *Science* 362, eaat6766
70. Shepard, R.N. (1987) Toward a universal law of generalization for psychological science. *Science* 237, 1317–1323
71. Balkenius, C. and Gardenfors, P. (2016) Spaces in the brain: from neurons to meanings. *Front. Psychol.* 7, 1820
72. Morgan, L.K. *et al.* (2011) Distances between real-world locations are represented in the human hippocampus. *J. Neurosci.* 31, 1238–1245
73. Theves, S. *et al.* (2019) The hippocampus encodes distances in multidimensional feature space. *Curr. Biol.* 29, 1226–1231
74. Constantinescu, A.O. *et al.* (2016) Organizing conceptual knowledge in humans with a gridlike code. *Science* 352, 1464–1468
75. Viganò, S. and Piazza, M. (2020) Distance and direction codes underlie navigation of a novel semantic space in the human brain. *J. Neurosci.* 40, 2727–2736
76. Tenenbaum, J.B. and Griffiths, T.L. (2001) Generalization, similarity, and Bayesian inference. *Behav. Brain Sci.* 24, 629–640
77. Dunsmoor, J.E. and Murphy, G.L. (2014) Stimulus typicality determines how broadly fear is generalized. *Psychol. Sci.* 25, 1816–1821
78. Dunsmoor, J.E. *et al.* (2012) Role of conceptual knowledge in learning and retention of conditioned fear. *Biol. Psychol.* 89, 300–305
79. Szöllösi, Ágnes *et al.* (2020) Behavioural pattern separation is strongly associated with familiarity-based decisions. *Memory* 28, 337–347
80. Kensinger, E.A. *et al.* (2007) Effects of emotion on memory specificity in young and older adults. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 62, P208–P215
81. Boyd, R. *et al.* (2011) The cultural niche: why social learning is essential for human adaptation. *Proc. Natl. Acad. Sci.* 108, 10918–10925
82. Laland, K.N. (2004) Social learning strategies. *Anim. Learn. Behav.* 32, 4–14
83. Mertens, G. *et al.* (2018) Does US expectancy mediate the additive effects of CS–US pairings on contingency instructions? Results from subjective, psychophysiological and neural measures. *Behav. Res. Ther.* 110, 41–46
84. Olsson, A. and Phelps, E.A. (2004) Learned fear of 'unseen' faces after Pavlovian, observational, and instructed fear. *Psychol. Sci.* 15, 822–828
85. Atlas, L.Y. *et al.* (2016) Instructed knowledge shapes feedback-driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *eLife* 5, e15192
86. Costa, V.D. *et al.* (2014) From threat to safety: instructed reversal of defensive reactions. *Psychophysiology* 52, 325–332
87. Lindström, B. *et al.* (2018) A common neural network differentially mediates direct and social fear learning. *NeuroImage* 167, 121–129
88. Lindström, B. *et al.* (2019) Social threat learning transfers to decision making in humans. *Proc. Natl. Acad. Sci.* 116, 4732–4737
89. Jeon, D. *et al.* (2010) Observational fear learning involves affective pain system and Cav1.2 Ca²⁺ channels in ACC. *Nat. Neurosci.* 13, 482–488
90. Finkelstein, A.B. *et al.* (2022) Social reactivation of fear engrams enhances memory recall. *Proc. Natl. Acad. Sci.* 119, e2114230119
91. Terranova, J.I. *et al.* (2022) Hippocampal-amygdala memory circuits govern experience-dependent observational fear. *Neuron* 110, 1416–1431
92. Haaker, J. *et al.* (2021) Observation of others' threat reactions recovers memories previously shaped by firsthand experiences. *Proc. Natl. Acad. Sci.* 118, e2101290118
93. Mou, X. *et al.* (2022) Observational learning promotes hippocampal remote awake replay toward future reward locations. *Neuron* 110, 891–902
94. Lindström, B. *et al.* (2016) Co-evolution of social learning and evolutionary preparedness in dangerous environments. *PLoS ONE* 11, e0160245
95. Korn, C.W. and Bach, D.R. (2019) Minimizing threat via heuristic and optimal policies recruits hippocampus and medial prefrontal cortex. *Nat. Hum. Behav.* 3, 733–745
96. Walz, N. *et al.* (2016) A human open field test reveals thigmotaxis related to agoraphobic fear. *Biol. Psychiatry* 80, 390–397
97. Chang, D.-J. and Debiec, J. (2016) Neural correlates of the mother-to-infant social transmission of fear. *J. Neurosci. Res.* 94, 526–534
98. Marin, M.-F. *et al.* (2020) Vicarious conditioned fear acquisition and extinction in childparent dyads. *Sci. Rep.* 10, 17130
99. Pittig, A. *et al.* (2022) Change of threat expectancy as mechanism of exposure-based psychotherapy for anxiety disorders: evidence from 8484 exposure exercises of 605 patients. *Clin. Psychol. Sci.* Published online October 4, 2022. <https://doi.org/10.1177/21677026221101379>
100. Gagne, C. *et al.* (2018) When planning to survive goes wrong: predicting the future and replaying the past in anxiety and PTSD. *Curr. Opin. Behav. Sci.* 24, 89–95
101. Suarez-Jimenez, B. *et al.* (2018) Linked networks for learning and expressing location-specific threat. *Proc. Natl. Acad. Sci. U. S. A.* 115, E1032–E1040
102. Bellmund, J.L.S. *et al.* (2019) Deforming the metric of cognitive maps distorts memory. *Nat. Hum. Behav.* 4, 177–188
103. de Vries, O. *et al.* (2021) Threat learning impairs subsequent memory recombination with past episodes. *Sci. Rep.* 12, 18878
104. Zhu, Y. *et al.* (2020) Emotional tagging retroactively promotes memory integration through rapid neural reactivation and reorganization. *eLife* 11, e60190
105. Allsop, S.A. *et al.* (2018) Corticoamygdala transfer of socially derived information gates observational learning. *Cell* 173, 1329–1342
106. Smith, M.L. *et al.* (2021) Anterior cingulate inputs to nucleus accumbens control the social transfer of pain and analgesia. *Science* 371, 153–159
107. Olsson, A. *et al.* (2020) The neural and computational systems of social learning. *Nat. Rev. Neurosci.* 21, 197–212
108. Shackman, A.J. *et al.* (2011) The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167
109. Tinnemann, A. *et al.* (2021) Observation of others' painful heat stimulation involves responses in the spinal cord. *Sci. Adv.* 7, eaabe8444
110. Haaker, J. *et al.* (2017) Endogenous opioids regulate social threat learning in humans. *Nat. Commun.* 8, 15495
111. Schenk, L.A. and Colloca, L. (2020) The neural processes of acquiring placebo effects through observation. *NeuroImage* 209, 116510
112. Gershman, S.J. and Niv, Y. (2013) Perceptual estimation obeys Occam's razor. *Front. Psychol.* 4, 623
113. Gershman, S.J. *et al.* (2013) Gradual extinction prevents the return of fear: implications for the discovery of state. *Front. Behav. Neurosci.* 7, 164